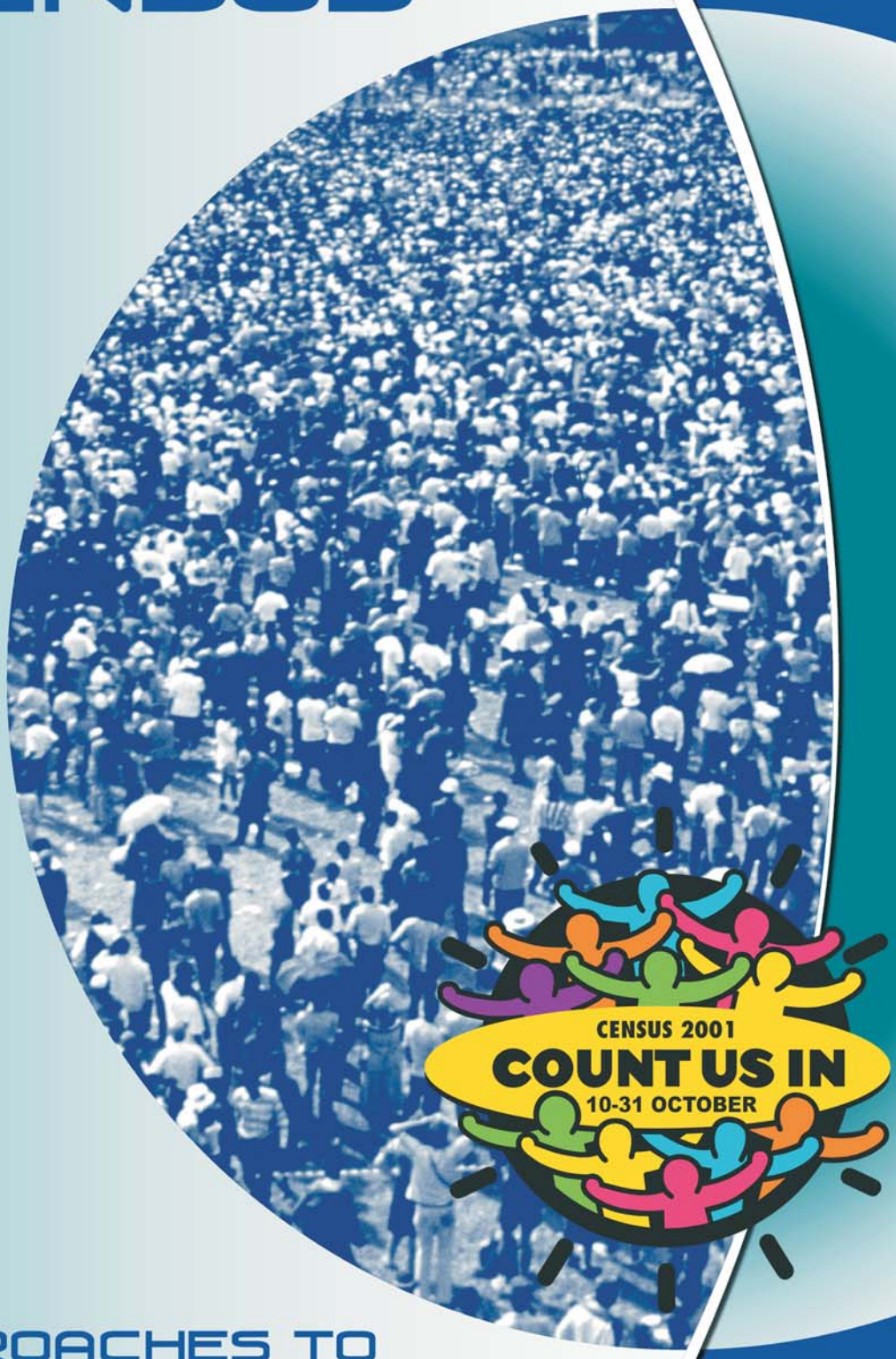
USING THE 2001 CENSUS



APPROACHES TO ANALYSING DATA

Statistics South Africa
Preferred supplier of quality statistics

HSRC
Human Sciences Research Council

Using the 2001 Census

Approaches to analysing data

**A collaboration between Statistics South Africa (Stats SA)
and the Human Sciences Research Council (HSRC)**

Published by Statistics South Africa
Private Bag X44
Pretoria
0001

Using the 2001 Census: Approaches to analysing data / Statistics South Africa. Pretoria: Statistics South Africa, 2007
244 p

ISBN 978-0-621-37259-5

Acknowledgements

The HSRC acknowledges the following researchers for their research work and contribution to the chapters indicated: Craig A. Schwabe (Mapping the 2001 Census); Pieter Kok (Migration studies based on the 1996 and 2001 Census data); Andrew Paterson (The 2001 Census and education in South Africa); Johan Erasmus (The labour force) and Eric O. Udjo (Use of Census 2001 demographic variables in decision making, planning and development).

Stats SA acknowledges everyone who has contributed to the publication, including fieldworkers, data processors, analysts to this publication and members of the South African public whose participation culminated in the consolidation of the census data.

Stats SA Library Cataloguing-in-Publication (CIP) Data

1. Population census – South Africa
2. Demographic variables – South Africa
3. Education – (South Africa)
4. Labour force – (South Africa)
5. Migration – (South Africa)

I. Statistics South Africa
(LCSH 16)

This report is only available on the Stats SA website: www.statssa.gov.za and the HSRC website: www.hsrc.ac.za

Contents

Foreword

I am proud to introduce the census users' handbook, produced through a partnership between Statistics South Africa and the Human Sciences Research Council.

How do we place a value on the information produced through the undertaking of a population census? The price tag of the last two recent censuses conducted, i.e. the cost to South African tax payers, we know to have been R419 million and R987 million respectively. It is obvious – particularly with the information and communication technologies we have at our fingertips today – that while it is extremely costly to *produce* information, it is extremely cheap to *reproduce* information. While information economists might put forward different approaches to how we can quantify the value of information in general, there cannot be much argument against the point that the value of information is derived from it being *utilised*.

Through publishing this book, which is demonstrating how others have used census data, together with the provision of information on how the data were collected, we hope to broaden and deepen the use of census data. Due to funding constraints, the flagship Census '96 data product, SuperCross, which enabled users to tabulate and cross-tabulate census data for the whole or parts of the country, and get this into the format of their choice; could not be given free of charge. We had to charge users R8 000 (for the single-user license / one per province) and up to R400 000 (for the multi-user license, including the complete dataset). This tool provided users with census data, from enumeration to provincial level, including the metadata. The SuperCross software had a built-in charting and mapping functionality. We sold 219 copies of the 1996 SuperCross package and gave away free copies after 1999. With better funding in 2003, the equivalent Census 2001 product was made available free of charge to the public sector and educational institutions – and we supported the installation and training on the use of the product on computers in over 1 400 organisations. And six years on after Census 2001– and particularly in the last 12 months ending 1 July 2007 – over 900 users have downloaded *Census in Brief 2001* from our website. So the utilisation of census data has no doubt increased - but there could still be more who could potentially benefit from using this data. More use means

deriving more value from the information collected through the costly exercise of census taking.

And so, this is why I am proud to introduce the census users' handbook, which itself is a product resulting from collaboration between the producers and key users of census data. I trust that this will encourage the further exploitation of census data – not only as an independent information source, but also to enrich the interpretation and understanding of information from other sources.

Pali Lehohla
Statistician-General (Stats SA)

Preface

The importance of census data in addressing societal needs, developing policy, implementing development strategies and monitoring the progress of government programmes is well documented. In fact, it is the opinion of many international researchers and scholars that census data or the socio-economic information that can be extracted from such data is one of the most important data sets that a country needs in addressing the basic needs of its people and focusing its attention on achieving the objectives stipulated within national policies. From policies such as the Reconstruction and Development Programme (RDP) that was implemented in 1994 by the newly-elected democratic government to improve the quality of life and standard of living of all people in South Africa, through to the New Partnership for African Development (NEPAD), which is an integrated strategy to bring about the holistic socio-economic development of the African continent to Agenda 21, which is a global strategy to achieve sustainable development; the need for socio-economic data emanating from national censuses is seen as being core to the achievement of these policy's goals and objectives.

The intention of this publication is, therefore, to provide a better understanding of what variables are incorporated into census questionnaires, such as the 1996 and 2001 censuses conducted in South Africa; what information can be derived from such data when used to populate various indicators and for what applications the census data can be used. By providing such a reference manual it is hoped that policy and decision-makers, officials at all three tiers of government, planners, non-government organisations, research institutions, private businesses and the communities of South Africa will be better informed about the value of census data and that they will be greatly encouraged to use it in their decision-making. Although the publication focuses mainly on the 2001 census, the chapters on the different thematic components are equally applicable census conducted in the past and future. The approach that has been taken to provide this information was to request the assistance of researchers at the Human Sciences Research Council (HSRC) and Statistics South Africa to provide their specialist insight into the different chapters contained in this manual. This publication is seen as a foundation upon which the knowledge of using census data will be built upon and that in years to come will see an update to reflect on the progress that has been made in the analysis and use of this information for decision-making purposes.

South Africa remains one of the few countries in Africa that has not only conducted regular censuses but has progressed to incorporate the enumerator area boundaries and census attributes into Geographical Information Systems (GIS) for the planning of its future censuses and to disseminate the results. This progressive approach has seen the production of many different products from both the 1996 and 2001

censuses. Furthermore, access to information from the censuses has been dramatically improved by Statistics South Africa (Stats SA) adopting a policy of making the data freely available in relatively easy-to-use cross tabulation software that enables users to analyse and extract a wealth of information. It is hoped that this first census users' manual will do justice to what has been achieved with censuses in South Africa and that a new wave of users will spring up from the insights that are contained in the different chapters.

The manual has 12 chapters that provide insights into how censuses are conducted and what products have emanated from the 1996 and 2001 censuses. Emphasis is also placed on the need to be able to integrate the census data with other layers of information to enable a more holistic perspective of South Africa for better decision-making purposes. One of the mechanisms by which these layers of information can be effectively integrated is by the use of GIS technology. Chapter 4 gives a detailed account of how the geography of the census is established. The next chapter examines the use of GIS as a tool for visualising census data and provides several examples of how best to accomplish this. The idea behind this chapter is to give potential users a clear understanding of what census data is all about and how it can be analysed using a GIS.

The availability of the census data in a variety of different products also enables the use of different statistical methods to analyse it. The use of univariate and bivariate analyses is explored and some of the complexities of using these methods with census data products are identified. From Chapter 6 to 12 the thematic chapters deal with migration, education, labour and unemployment, and the use of demographic variables in planning and development and providing services. The focus of these chapters is mainly on identifying the relevant variables that can be used, providing insights into what indicators can be used and then giving some insight into how this information can be applied.

HSRC

Chapter 1: What is a population census?

1.1 Introduction

The Statistics Division of the United Nations' Department of Economic and Social Affairs defines a population census as "… the total process of collecting, compiling, evaluating, analysing and publishing or otherwise disseminating demographic, economic and social data pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country".[1]

Most countries carry out censuses at regular intervals, usually every five or ten years. The characteristics of all individuals within a defined territory are recorded simultaneously. The data are used to inform government policy making, planning and administration (e.g. distribution of financial resources for administrative and development purposes), for demographic and social research and for research to inform business, industry, labour and the public. Census data also provide a sampling frame for inter-censual surveys that provide further insights into demographic and socio-economic trends for the purpose of monitoring and evaluating the implementation of policies and programmes.

Population censuses provide highly disaggregated data that delineate the characteristics of the population within small geographic areas and within small communities. This information is vital for planning for the unique needs of such areas and population groups.

This chapter focuses on the processes that are undertaken in conducting a population census and highlights the various planning and operational phases. The first major phase is the planning phase. This is followed by: implementation of preparatory activities, main enumeration activities and post-enumeration activities (data processing, analysis of results and release of results).

1.2 The mandate to conduct a census

The Statistics Act (No. 6 of 1999) mandates Statistics South Africa to conduct a population census every five years. The Act covers diverse issues such as the obligation of individuals to provide information about themselves and their households, the topics to be included in the census, the confidentiality of the information that individuals supply and the role of government ministries and departments in census taking. In March 2004 the Cabinet decided that population censuses will hence be

[1] *Principles and Recommendations for Population and Housing Censuses Revision 2*, United Nations, October 2006

conducted in a ten-year cycle. The next population census is therefore scheduled to be carried out in 2011.

The success of a census depends to a great extent on political commitment to, as well as public ownership of, the exercise. Political parties must encourage their constituencies to cooperate with census officials and to provide accurate and complete information.

1.3 The planning phase

This phase entails putting together the plans that need to be implemented, with special focus on activities, timelines and budgetary requirements.

Unlike many other sub-Saharan African countries, the South African government funds the country's population censuses without any external support. This places a considerable fiscal strain on the country that can only be accommodated through careful planning. Censuses therefore cannot be organised and conducted at short notice. International best practices indicate that preparations for a satisfactory census must commence at least three years in advance if a permanent census structure exists. A hurried execution of activities inevitably leads to ineffective operations and wastage of funds.

A census consists of a series of distinct but closely interlinked activities. The close linkage of different activities means that some cannot be undertaken before the completion of others. Strict adherence to sound project management principles is thus required and each activity must be carried out according to a well-planned schedule: the operational plan. This schedule of activities must be factored into the government's funding cycle through a thorough budgeting process. The success of a census depends on how well activities have been integrated and how efficiently they are executed.

1.4 The preparatory phase

This phase constitutes the cornerstone of census taking. The enumeration of the population can only be successful if the various activities in this phase are properly implemented.

1.4.1 Census mapping operations

Census mapping operations are the determination of the geography on which the census is collected and on which census data are disseminated. Two focal points can be distinguished:

- small-area identification; and
- development of a suitable geographic hierarchy.

The key output of a census is the availability of small-area data. The basic geographic unit for which data are collected (and for which data could therefore possibly be reported) is called an Enumeration Area. The most important criteria for the design of Enumeration Areas are:

- They should not overlap, should be compact without pockets or disjointed sections and should cover the entire country.
- They should have boundaries that can be identified on the ground.
- They should be of approximately equal population size to enable an enumerator to cover each one within the census period.[2]

An estimate of the number of dwellings in a particular enumeration area is needed before its boundaries can be finalised. In the past these estimates could only be obtained through an actual listing of dwellings in the field. With the advent of geographic information systems that can capture, store, retrieve, update, manipulate, analyse and display geographically referenced information, alternative approaches that utilise updated address systems and data collected through remote sensing have become available.

If they are large enough to guarantee data privacy, enumeration areas can be used as the geographic basis for reporting census data. The information must, however, also be reported for larger geographic areas. A geographic classification is a hierarchy of statistical areas that is used to report census data. The particular hierarchy that is decided on depends on user needs, the administrative hierarchy of the country and the hierarchy that is already in use for related products. These concepts are further expanded on in Chapter 4.

1.4.2 Development of the questionnaire through user consultations

Census data are a valuable resource that is currently under-utilised. It is very difficult to engage potential users in the census process. An ongoing education and communication effort that involves meetings, seminars and publicity is therefore needed in the years leading up to a census in order to develop an understanding of the nature of decision making and the need for census data to support it.

Census data feed into the National Statistics and Information System. These data are used in conjunction with data from other sources. The coding, editing and data transformation processes must therefore be conceptually sound to guarantee delivery of quality data to satisfy the needs of users.

[2] *Handbook on Geographic Information Systems and Digital Mapping*, United Nations, 2000

A structured consultation programme is followed to determine the expectations of critical users. This consultation typically focuses on the content of the census form and reporting formats. These discussions are informed by research on the effectiveness of alternative formulations of questions and the users' assessment of census products.

1.4.3 Development of logistical support systems

The data collection phase of a census is heavily dependent on the success of vast logistical arrangements. The key issues are:

- establishment of regional offices
- printing of questionnaires
- printing of maps
- delivery of census and promotional material to regional offices
- logistical support such as transport to supervisors and enumerators during enumeration
- other logistical issues such as the payment system.

An effective payment system must be developed to ensure that field staff members are not distracted from their task by concerns about payment. This payment system must ideally link with the management information system of the census.

Government departments have an obligation to provide logistical support to Statistics South Africa where possible. As part of a cost-curtailment strategy the census management must ensure that this resource is tapped.

The following are examples of possible contributions of government departments to the census:

- the use of the police force to ensure the security of fieldworkers
- the provision of infrastructure, such as the use of municipal offices as regional offices and the use of vehicles and drivers during enumeration
- the use of civil servants such as teachers in enumeration
- the sharing of information such as maps
- the provision of specialised services such as printing.

1.4.4 Training

Different training programmes for regional managers, supervisors, enumerators and support staff are needed. Extensive training material such as manuals and videos are developed. The training focuses on the duties of the staff members, their importance, how their efforts fit into the overall census process as well as issues such as confidentiality.

1.4.5 Pilot census

Tests are carried out for all the stages of the census. They are especially useful for identifying weaknesses in the census form or in enumeration procedures.

A year before the census a comprehensive test of all census procedures, called the census pilot, is launched. This test covers the performance of the whole census organisation. It is conducted in selected areas under conditions that resemble the actual census as closely as possible.

1.4.6 Finalisation of the questionnaire and logistics

Based on the lessons learned in the census pilot the questionnaire and logistical arrangements are adjusted and finalised. After this, the questionnaires are translated and sent off for printing.

1.4.7 Publicity

Census publicity aims to increase public understanding of the purposes of censuses and thus ensures the cooperation needed for high-quality responses. Publicity messages cover issues such as when the census is to be held, what is expected of respondents, how the data will be used and how more information can be obtained. The audience is typically reached through mass media such as radio, television and newspapers, but specific interest groups such as community leaders can be approached through meetings. Mass gatherings and publicity material such as T-shirts are also utilised to increase awareness.

Publicity campaigns are very expensive. For them to be effective, special messages must be focused on specific target groups such as hard-to-count sections of the population. Professional communication agencies are typically used in the planning and implementation of the publicity campaigns. The campaigns are ideally extensions of ongoing publicity programmes that are already in place.

1.4.8 Recruitment and training of personnel

In South Africa a population census requires the involvement of close to 100 000 temporary staff members in field and regional offices. Field staff members are recruited within the particular community to ensure their safety and the cooperation of respondents as well as to minimise travelling costs. Recruitment and training of staff start months ahead of the census.

The majority of staff is recruited from the general public. An effective recruitment campaign is needed to ensure that staff members who are capable of undertaking the

various duties are recruited. Training is given as close as possible to the time of enumeration.

1.5 Enumeration

Census night

Census data should be pertinent to a well-defined reference period. Cardinal in the definition of the reference period for a particular question is the concept of the census night. The information that is recorded on the questionnaires must reflect the situation at midnight on that night. For example, even though data collection is only completed some days after the census night and a particular person may have died in the interim, all persons alive on census night would be recorded on the questionnaire.

Enumeration procedures

Different data collection strategies are designed for different types of dwelling units and different types of enumeration areas. In the main, a distinction is drawn between people who are attached to households and those who are found in institutions such as military installations, penal institutions, hospitals and so forth. For these, the unit of enumeration is the institution. The approach in enumerating, for example, the population on a commercial farm also differs from that which is used to enumerate an inner-city block of flats.

Furthermore, specific strategies may be designed to ensure maximum coverage of hard-to-count groups such as homeless or nomadic people. The data collection strategies aim to take cognisance of existing social dynamics as far as possible to ensure respondent cooperation and data validity.

Quality assurance

Procedures to assure quality are designed for all the phases of a census. Of all of these, quality assurance during field operations is by far the most crucial.

Central to quality assurance is the role played by supervisors. Supervisors are required to have a thorough knowledge of enumeration procedures. They are expected to identify enumerators who require retraining; through observation of interviews, checking whether questionnaires are properly filled in and checking of a sample of households already interviewed.

1.6 Post-enumeration phase

This phase concentrates on activities that are undertaken after the enumeration exercise. These activities include receipt of questionnaires from the field, closedown

operations, processing of the data, post-enumeration survey, analysis and evaluation of the data and release and dissemination of the results.

Receipt of questionnaires and other relevant materials

After completion in the field, questionnaires are forwarded to regional offices where they are checked and temporarily stored. With the completion of fieldwork, questionnaires are transported in bulk to processing centres.

At the processing centres all questionnaires are registered. Checks are run to ensure that all enumeration areas in the country and all households within each enumeration area are accounted for. The questionnaires are then submitted to a procedure of preliminary checks to ensure that they are in an adequate condition for data processing.

Closedown operations in the field

Closedown operations are instituted for the temporary regional offices, and material such as training material, office furniture, computers and communication equipment is allocated and transported to new locations or storage.

Data processing

Data processing involves the *coding* and *capturing* of the collected data as well as *editing* and *imputation* procedures.

Coding involves assigning classification codes to all the responses on the census form. Although pre-coded responses with codes printed next to each category are used whenever possible, some responses are in textual format. These verbal responses must be replaced with suitable codes.

During *data capturing* the information on the census form is converted to a format that can be interpreted by a computer. Current technology involves the use of imaging techniques by scanning devices that perform optical mark recognition whereby responses to “tick-box” type questions are detected by the machine, and intelligent character recognition in which complex software “recognises” what was written by the respondent in constrained text fields and translates this into text values.

Editing is the process designed to check on obvious respondent or interviewer mistakes as well as responses that are out of range. *Imputation* procedures go a step further and assign values to missing data through statistical techniques that refer to the probability of a particular value based on the values of other variables.

These complex activities rely on extensive administrative and information technology support. Processing systems based on cutting-edge technology are evolving at a rapid

rate. It is typical for outsourcing to be employed in this phase as many of the activities are highly specialised and require sophisticated technology.

Prior to the handover of data, validation checks are run to ensure that data quality meets specified minimum standards.

1.7 Post-enumeration survey

The data collection phase of censuses may be followed by a post-enumeration survey to measure the extent of undercount or overcount. In these surveys an independent and exhaustive enumeration is conducted in selected enumeration areas to compare against census data. Through this comparison, adjustment factors are derived to remove errors of coverage, such as erroneous omissions and erroneous inclusions. The survey results can also be used to investigate content errors, such as the completeness and validity of the responses given by respondents.

Evaluation of data and results

Time, effort and financial investment in a census have the single objective of making it as successful as possible, with maximum coverage and accuracy. It is, however, an accepted fact that any population census will be subject to some errors. The presence of the errors does not detract from the value, utility and acceptability of the results, provided that the magnitude of the errors is known and does not affect the major users of the data. Therefore, information on the basic measures of quality of the data must be determined and made available to the users of the data. This information will give an indication of the degree of confidence with which the results can be used for various purposes. It is good census practice not only to evaluate the completeness and accuracy of the data but also to use this and other information to evaluate and redesign collection methodologies, systems and procedures.

Dissemination of products

Based on the feedback of users, the products and services of the previous census are reviewed and new products and services are developed. The following are some of the key concerns in product development:

- user needs
- all aspects of data quality
- cost
- respondent confidentiality.

Where applicable, marketing forms part of product development, i.e. a strategic marketing plan is developed to recover some of the costs of operations through charging for products and services. A careful distinction must be made between

commercial products and products that are considered as community service obligations.

Data dissemination typically consists of a staged release of final data. Simple information such as basic demographics is released first. The data for the more complex topics that require more processing resources, such as industry and occupation, are released at a later stage.[3]

A schedule for the release of products takes cognisance of demand, resources and commitments made to users. Large publicity campaigns accompany the launch of census products in order to promote awareness of the products and their uses and as part of publicity for future censuses.

1.8 Conclusion

Censuses are, with the exception of political elections, the largest logistical exercises to be carried out by any country. Careful planning and control during implementation of various activities are needed to ensure effective use of resources and the collection of quality data. The focus of census operations should always be to satisfy users' data needs.

[3] *Handbook on Census Management for Population and Housing Censuses,* United Nations, 2001

Chapter 2: Historical development of census taking in South Africa

2.1 Introduction

Census taking is a momentous and challenging endeavour for any country. It brings on board diverse political, economic and social demands in terms of the magnitude of resources that have to be utilised to collect the requisite information for effective decision making as well as the implementation of development programmes. It is increasingly becoming of paramount importance that countries carry out censuses on a regular basis to ascertain the impact of various programmes on the living standards of their population, particularly disadvantaged groups. In this regard, South Africa is one of the countries that uphold this cardinal commitment.

This chapter gives a historical perspective on census taking in the country. It provides insights into methodological and technological advances as well as the political and social environment, and focuses on the transformation of data needs after 1994. Most of the information contained in the chapter has been collected through interviews, literature research of media articles and reference to the questionnaires that were used in the various censuses.

The value of a historical perspective

It is not possible to understand the challenges that South Africa has faced in undertaking its population censuses without reference to the past. Historically the national statistics agency has been responsible for censuses. Before 1998 this agency was called the Central Statistical Service. In 1998 it was transformed into Statistics South Africa.

The responsibility of a national statistics agency relates to issues that range from ethical standards to prevent misuse of data, independence from political influence and methodological transparency, to others such as open dissemination and professional staffing. The political and social history of the country has impacted vastly on the ability of the agency to meet these responsibilities in the past. During ten years of democracy the organisation has done a great deal to improve its capacity and to fulfil its mandate to ensure effective governance through the collection, processing, analysis and dissemination of pertinent statistics. The most prominent aspects of this history are highlighted in this chapter.

The value of this historical perspective is twofold. First, the institutional capacity of Statistics South Africa is built through an assessment of the past. Past experiences are re-evaluated and the lessons learned are incorporated in planning for the future. Second, the history clearly shows the influence of the political agenda on methodology

in terms of both content and coverage. This illustrates the cardinal relationship between census methodology and the current development agenda, and points to the importance of involving all tiers of decision makers in planning and conducting the censuses.

Census taking prior to 1960

Historically censuses in South Africa have been fragmented, typically covering only parts of the country and sections of the population. The first census was conducted in 1798. Every head of a household in the Cape Colony had to submit a return stating the size of his family and the number of slaves and cattle that he owned. The first South African statistical manuscript, known as *The Annual Blue Book in the Cape of Good Hope*, was released in 1823 and was continued until 1837, but was never published.

Censuses for all races were conducted in the Cape in 1865 and 1875, in the Free State in 1880 and in Natal in 1891. In 1890 only members of the white population took part in the first census conducted in the Transvaal. In 1904 censuses were conducted in the Cape, Natal, the Free State and the Transvaal. Although these censuses were separate they did cover the whole country within the same year.

In 1910, when the Union of South Africa was formed, the government decreed in terms of the South Africa Act of 1909 that a population census for all races be undertaken in 1910 and thereafter whenever the government indicated the requirement. Censuses that covered all population groups took place in 1911, 1921, 1936 and 1951. Additional censuses in which only white people were enumerated were conducted in 1918, 1926, 1931 and 1941.

2.2 Census taking since 1960

The censuses since 1960 are discussed in this section. All of these censuses were *de facto* censuses, i.e. the respondent had to report on all persons, whether a usual resident or not, who spent the night at the particular dwelling. The reference nights of these censuses were 6 May 1970, 6 May 1980, 5 March 1985, 7 March 1991, 9 October 1996 and 9 October 2001.

2.3 Technological and methodological advancements

Census mapping and geographic information systems

Prior to any census the country must be demarcated into census enumeration areas. One enumerator, or person doing the counting, covers each of these areas during the census. An enumeration area should be as homogenous as possible, should consist of a contiguous area and should be of a suitable size. The underlying principle in

demarcation is that all parts of an enumeration area should be within comfortable reach of an enumerator. The enumeration areas also serve as the building blocks for the dissemination of census data.

Historically administrative maps served as the basis for demarcation. Town planning maps were used for large new towns and developments and aerial photographs were utilised to pinpoint new residential units.

During the 1980 census many census enumerators complained that the maps provided to them were insufficient and outdated. New developments were not reflected on these maps.[4] The extensive informal settlements that developed as influx control measures proved unenforceable exacerbated these difficulties. A strategy of listing households in these areas and demarcating on the basis of the lists did not prove viable due to political violence.

In 1991 large parts of the country were not demarcated into enumeration areas. Excluded from the demarcation process were the independent homelands (Transkei, Bophuthatswana, Venda and Ciskei), the rural areas of the self-governing territories that were instead counted by means of a 'sweep enumeration' covering large areas of land, and squatter camps and informal settlements in urban or semi-urban areas for which population estimates were derived from aerial photographs.

For Census '96 about 1 000 persons were recruited and trained for a complete demarcation and listing process. The country was divided into approximately 86 000 enumeration areas. Each enumeration area consisted of about 100 to 250 households. The size of the area was determined by whether it was situated in a densely or sparsely populated area. The information on each enumeration area was brought together in a booklet called the *Census Summary Book*. The booklet included maps, aerial photographs or descriptions of every area and its boundaries. It also included the address or other means of identification of each visiting point and the households within the area. During enumeration the booklet was used to record whether or not each household had been enumerated.

For Census 2001 this methodology was further extended. Geographic information system technology was utilised for the first time in demarcation and map production as highlighted in Chapter 4. Data acquired from government departments and private sector companies were integrated into a comprehensive digital spatial information database. Where such data proved insufficient, further data were collected in the field by means of global positioning systems. The 1996 boundaries of the enumeration

[4] *Die Vaderland*, 9 May 1980

areas could be re-evaluated on the basis of this spatial data. The ability to perform onscreen demarcation saved a great deal of time and effort.

Changes in data collection strategies

In the 1960 and 1970 censuses there was a full count only for white, coloured and Asian people. A different questionnaire was used for black people, for whom a small sample was enumerated. The 1980 census was the first census that aimed to count all population groups through a uniform data collection methodology (although a question on family structure was not asked of the black population). Due to the unplanned and unstructured nature of certain residential areas, and the inaccessibility of others due to political violence and budget limitations, the Central Statistical Service decided to use aerial photographs for obtaining estimates of population counts for 83 urban areas and magisterial districts during the 1991 census.

Census ’96 was the first census in which the whole country was covered and all residents of the country were treated equally. The same methods were used for everyone. Up to Census ’96 questionnaires were only available in Afrikaans and English. In Census ’96 and Census 2001 questionnaires were available in all 11 official languages. Respondents could select the language of their choice when supplying information. The respondents were given a choice: an enumerator could complete the questionnaire for them during a face-to-face interview, or they could opt for self-completion. The vast majority of people chose to be interviewed.

The extent of the growth of field operations is reflected in the number of enumerators recruited for the different censuses. In 1980 there were 27 000 enumerators. In 1985, 35 000 enumerators were used. In 1991 the number of enumerators grew to 41 000. In Census ’96 and Census 2001 close to 100 000 enumerators were used in each of these years.

The utilisation of computer technology in processing data

In 1960 and 1970 the census results were processed manually. In 1980 a total of 1 000 part-time coders were used to capture the results electronically through optical character recognition. The results had to be transcribed to specially designed computer-readable forms prior to capturing. In 1985 and 1991 capturing and editing of the results were done on dump terminals. For the 1980, 1985 and 1991 censuses post-capture processing was done on mainframes.

As a result of the improvement in technology, in the period 1960 to 1985 there was a decrease in time required for data processing. In 1960 the results were released nearly 11 years after the collection of data. In 1970 and 1980 eight and five years respectively were required. In 1985 the final results were available in just over one and a half years. This delivery time has been maintained.

For Census ’96 and Census 2001 processing was done on computer servers. For Census ’96 processing was decentralised to the provinces and only editing was performed at the head office. Capturing was done online on PCs. The processing and editing of the data of Census 2001 were done at a special centralised facility. Intelligent scanning software (able to read free handwriting) was used for the first time. Hot deck imputation, a technique which uses one or more variables to estimate a likely response, was also introduced for the first time to resolve non-responses.

Data validation

Prior to 1996, demographic models that estimated the size of the population were used to validate census results. Pronouncement on the quality of the censuses of 1980, 1985 and 1991 was based on such models and the data for 1980 and 1991 were adjusted accordingly.

Census ’96 was the first census for which a nationwide post-enumeration survey was undertaken immediately after the count, in order to estimate the number of persons or households that were missed during the census. The most skilled fieldworkers revisited 1% of the census enumeration areas and all the households in these areas were listed, revisited and enumerated. The data thus collected were compared with the census data through a laborious process of matching of individuals and households. This exercise yielded adjustments for the undercount. Apart from identifying errors of coverage, a post-enumeration survey also identifies errors of content that pertain to the reported characteristics of the people or the households enumerated. A similar post-enumeration survey was conducted just after Census 2001.

The development of census products and dissemination methods

Up to 1985 census products consisted only of hardcopy reports. In 1991 unit record data were available for a selected few users in ASCII format on magnetic hard tape.

The products of Census ’96 marked a radical departure, in that there were some free publications and there were new electronic products on CD with fast tabulation software (SuperCross) that provided unprecedented flexibility for the users of census data. A 10% sample of unit records was also disseminated in electronic format on CD. Key publications and metadata were placed on the Stats SA website where they could be accessed at no charge by Internet users. In addition, the census products were all deposited in state and other libraries. The focus on ensuring accessibility to support decision making had an added dimension of feedback to the general public and to specific communities. Community profile databases were disseminated in electronic format on CD. Hardcopy reports were limited to a few publications such as *Census in Brief* and *The Count and How it was Done*.

Census 2001 has built further on the successes of Census '96. PC_Axis was introduced to allow users to do interactive cross tabulations on certain combinations of variables on the web. Extensive metadata were also published on the web.

2.4 The political and social environment and its effects on census taking

The homeland policy and its effects on the data series

In 1976 and 1977, respectively, the Republic of Transkei and the Republic of Bophuthatswana were declared independent from South Africa. The same status for the Republic of Venda and the Republic of Ciskei followed in 1979 and 1983 respectively.

In 1979 the Department of Statistics announced that mid-year population estimates would exclude people living in the Transkei and Bophuthatswana. The Transkei was planning to conduct its own head count in 1980 and Bophuthatswana had asked South Africa to conduct an identical census at the same time as the South African count.[5] It was later decided that the Transkei, Bophuthatswana and Venda would each conduct a census on the same day to avoid a double count of migratory workers.[6]

Bophuthatswana conducted another census in 1985, concurrent with the South African census. A sample census, the first of its kind in South Africa, was conducted the same year in the Transkei. The sampled areas showed a massive undercount had occurred in the 1980 census.[7] The results of these censuses were not made public.[8]

Censuses were conducted in all four independent homelands in 1991. The HSRC assisted with those in the Transkei and the Ciskei while the Central Statistical Service assisted with the processing of Venda's data. At this time it was widely accepted that the homelands would be reincorporated into South Africa. The *Diamond Field Advertiser* welcomed the fact that the "… fiction of excluding 'independent' states, which in the past distorted official statistics on everything from unemployment to road accidents to detentions without trail …" was coming to an end.[4]

The effect of the homeland policy on the quality of official statistics for South Africa was devastating. Data for the censuses in the Transkei were never fully recovered. Information for the country as a whole during the period of the independent homelands can only be estimated through computations beset with technical and quality issues.

[5] *Financial Mail,* 23 February 1979
[6] *Financial Mail,* 25 January 1980
[7] *Daily Dispatch,* 25 October 1985
[8] *Diamond Field Advertiser,* 11 March 1991

The political climate and data collection

The success of censuses depends to a large extent on the cooperation of the respondents during data collection. It is very important for the national statistics agency to build a relationship of trust with the general public.

The oppressive governmental policies during the apartheid era ensured that it was not possible to cultivate such a relationship with the greater part of the population in South Africa. The regulations that controlled internal migration forced a status of illegality on a large section of the population. As the political assertiveness of the oppressed grew there were manifestations of open resistance against many governmental programmes, including censuses.

On 25 January 1980 the *Financial Mail* stated with regard to the 1980 census that "… the usefulness of the statistics is clearly going to depend on the degree to which the department is able to persuade Blacks, in particular, that personal details will not be available for scrutiny by other authorities. The *FM* does not rate their chances highly against the background of many years of justified suspicion of government motives". Dr Nthato Motlana, a civic leader of Soweto, said before the census that "… those people living in Soweto who were regarded in South Africa's sick parlance as illegal would never allow themselves to be enumerated in any census"[9].

The government was criticised in 1985 for "… spending R35 million on yet another census at a time when the country is battling through an economic recession".[10] The deteriorating political and economic climate impacted on methodology. During the 1991 census the ANC objected to the methodology of using aerial photographs for obtaining estimates of population counts, as well as to the exclusion of the independent homelands and described the census as an absurd farce.[11] This statement further undermined the cooperation of respondents.

By 1991 a population census as a tool of governance was highly politicised. Under the new dispensation it has been one of the main concerns of Statistics South Africa to change negative perceptions and to gain the confidence of respondents. Both Census '96 and Census 2001 had extensive public outreach and publicity campaigns. The publicity campaigns initially focused on non-urban areas through radio advertising and talk shows. Towards the final stages of the campaigns this focus was extended to television and the print media. The mainstream media campaigns were complemented by grassroots campaigns that included provincial launch shows, visits to local authorities, meetings with civic and traditional leaders and the wide distribution of publicity material. These publicity campaigns have been very successful in convincing

[9] *Rand Daily Mail,* 13 September 1980
[10] *The Star,* 5 March 1985
[11] *Business Day,* 8 March 1991

the majority of South Africans to participate actively in censuses. The democratic state enjoys legitimacy in the eyes of the overwhelming majority and derives its ability to conduct censuses successfully from this. Some hard-to-count sections of the population do, however, still remain. In the more affluent areas characterised by security gates and high walls, for example, enumerators are not always successful in establishing contact with respondents.

The development of census questionnaires

The following table shows the coverage of census topics for all the censuses since 1960.

Table 1: Census topics for all the censuses since 1960

Census topic	1960		1970		1980		1985	1991	1996	2001
	WCA[12]	B	WCA	B	WCA	B				
Usual residence	×	×	×		×	×		×	×	×
Sex	×	×	×	×	×	×	×	×	×	×
Age	×	×	×	×	×	×	×	×	×	×
Population group	×	×	×	×	×	×	×	×	×	×
Relationship to head of household	×	×	×	×	×	×	×	×	×	×
Family structure			×		×					
Language	×	×	×	×	×	×		×	×	×
Education	×	×	×	×	×	×	×	×	×	×
Literacy		×		×	×	×		×		
Religion	×	×	×	×	×	×		×	×	×
Marital status	×	×	×	×	×	×	×	×	×	×
Fertility	×	×		×	×	×			×	×
Mortality		×		×	×	×			×	×
Migration	×	×	×	×	×	×	×	×	×	×
Nationality	×		×		×	×	×	×	×	×
Employment	×	×	×	×	×	×	×	×	×	×
Income	×		×		×	×		×	×	×
Disability					×	×			×	×
Sport and recreation					×	×				
Travel					×	×				×
Dwelling (type and characteristics)			×		×	×		×	×	×
Ownership of dwelling			×		×	×		×	×	×
Number of households			×							
Energy sources									×	×
Water and sanitation									×	×
Amenities					×	×			×	×
Persons employed by household			×		×	×				

[12] WCA: whites, coloureds and Asians; B: blacks

In the 1960 and 1970 censuses there were separate questionnaires for white, coloured and Asian respondents on the one hand and black people on the other. The differences seem to have been motivated by assumptions about the relevance of certain information types to particular population groups. In the 1960 questionnaires, for example, "living together" is listed as a category for blacks but not for whites, coloureds and Asians. Whereas for white, coloured and Asian women the total number of children ever born is asked, for black women the number born alive during the previous 12 months is asked. Also, only for black women is child mortality measured – for the other population groups the recorded vital statistics were assessed as sufficient.

In 1980 the same questionnaire had to be completed for all population groups, but a small section was added in which the family (note, not the household) structure as well as the occupation and income of the "head of the family and wife" was recorded. Since 1985 the same questionnaire has been used for all population groups.

The 1980 census questionnaire was very ambitious. It was reported in the media that the questionnaire was the result of extended consultation with academics as well as with private and public organisations.[13] As can be seen in Table 1, the questionnaire covered a range of topics. For an extended list of amenities the number available at the sole disposal of the enumerated persons was measured. Concern about this was raised before the census. The *Financial Mail* pointed out that while it accepted "… the convenience of using a general census for wider purposes, an argument could be made that the collation of this additional type of information should more properly be the task of private endeavour".[14] The census questionnaire, which was printed on A2-size paper, was difficult to read. In several places on it respondents were referred to a separate instruction sheet.

The 1985 questionnaire was very short. Looking at the questionnaire it is clear that the sole function of the census was to adjust the count that was obtained in the 1980 census.

As in the 1960 and 1970 censuses the questionnaire of Census '91 was concise and to the point. Its layout was a considerable improvement over that of the long census questionnaire in 1980. However, as in that census, the respondent was constantly referred to a separate instruction sheet. Fertility and mortality were noticeably missing as topics in Census '91. The first Demographic and Health Survey was conducted in 1987 and it was believed that these surveys would provide adequate information with regard to fertility and mortality.

[13] *Rand Daily Mail,* 10 October 1979
[14] *Financial Mail,* 15 June 1979

Census ’96 and Census 2001 saw the addition of two new topics: energy sources and water and sanitation. Questions on disability and amenities were also included. Census ’96 also focused a great deal on migrant labour. These topics were clearly included to inform the government’s development agenda. More about this follows in a later section.

The questionnaires of Census ’96 and Census 2001 were very user-friendly. In both cases, questionnaires were printed in book format and all instructions were printed on the questionnaire itself. In the questionnaire of Census 2001 the information of persons was recorded in rows. During processing of the data this proved to have been a mistake, as the information of the same people was seen to have been recorded in different rows on different pages. Apparently recording in columns leads to easier control.

2.5 Data dissemination and the main themes reflected in the media

In the discussion of any subject it is usually the concerns of influential opinion makers and decision makers that feature as the main themes in the media. A study of the media tells one a great deal about current political agendas and how these agendas are informed by census data.

In an article entitled “Migration and Sterility” *Die Volksblad* quoted the large-scale migration, that was measured in the 1980 census, for black people from “white” areas to the homelands as evidence of the government’s success in applying the apartheid policy during the 1970s. At the same time concern was expressed about the decline of the white population in rural areas. Whites were warned that they would not remain agriculturally and militarily prepared if they let go of rural areas.[15] Most of the news coverage of this era focused on the racial and geographic breakdowns of the data and the current growth rates for these breakdowns.

Ironically census numbers were also used to justify reform. In 1992 *Beeld* stated that the uncompromising census numbers give little comfort to those who supported either open or covert white supremacy or partition in any form.[16]

With the release of the final data of the 1985 census *Die Burger* identified the following three serious concerns:

- the high rate of population growth
- the “unhealthy” demographic distribution of the population
- the low percentage of the population that had obtained a degree.

[15] *Die Volksblad,* 12 September 1980
[16] *Beeld,* 11 March 1992

Die Burger stated that overpopulation and too high a rate of urbanisation would create friction that would lead to a revolution.[17] During the late 1980s and the early 1990s the media regularly raised issues around population growth. The government's Population and Development Programme targeted population growth and specifically identified the birth rates of black and coloured people as problematic.

The quality of the 1970 census was never publicly disputed. In the 1980s there was a growing awareness of the difficulties that the political climate imposed on census taking and the uncertainty with regard to the results. This, together with the growing sophistication of the users of the data, put increasing pressure on the Central Statistical Service to communicate with the users of census data and to justify its methodologies.

Census data were adjusted to compensate for undercount for the first time for the 1980 census data. With the release of data of the 1985 census there was an outcry, as the unadjusted 1980 figures that were published with the newest figures in order to permit comparison bore no resemblance to the figures that were used before. Researchers, municipal officials and corporate planners were unaware that they were working with "inflated" numbers.[18]

There were increasing questions with regard to the validity of official data during this period. The quality as well as the relevance of official data was questioned, as the data of the independent homelands were not included in estimates. The Bureau for Market Research at the University of South Africa published independent population estimates at magisterial district level in 1989.[19]

With the introduction of a democratic dispensation in 1994 came a high level of awareness of the government's accountability to the voters and consequently the need for relevant and accurate data to inform service delivery and to monitor the implementation of reconstruction and development. Census '96 by its very inclusiveness, and by producing small-area data, had a high profile as a source of valuable information that would drive development. Just prior to the census, Frank Mdlalose, the Premier of KwaZulu-Natal, stated that "… the statistics of violence show us where we have been. The census statistics will show us where we are going and how we will get there."[20]

[17] *Die Burger,* 12 November 1986
[18] *Weekend Post,* 24 August 1985
[19] *Rapport*, 26 November 1989
[20] *The Star,* 12 July 1996

2.6 The transformation in data needs after 1994

Background to the transformation initiatives

The attainment of democracy in 1994 presented the government with the challenge of implementing the Reconstruction and Development Programme, the basic development agenda that has been elaborated in all post-1994 policies. This programme identifies the following key objectives: meeting basic needs, building the economy, democratising the state and society, developing human resources and building the nation. It requires significant institutional transformation as well as the concurrent introduction of new policies that are in line with a democratic constitution.

The government set out to deliberately and systematically dismantle the social relations that were established by apartheid and to create a democratic society based on the principles of equity, non-racialism and non-sexism. New policies and programmes have been put in place to dramatically improve the quality of life of all people, to address poverty and inequality and to restore the dignity of citizens.

The utilisation of Census '96 and Census 2001 in the transformation

The additions of further topics to inform the government's development programmes were alluded to in Section 2.4. We now further discuss the utilisation of censuses in transformation.

One of the main advantages of censuses over surveys is the fact that censuses provide comprehensive small-area data. This level of reporting allows local government to inform detailed operational planning by means of census data and projections. The 10% sample of unit records is especially helpful in this regard.

Statistics South Africa is currently aligning its products with the needs of the users of official statistics. It does this through the development of a National Statistics System, the standardisation of definitions and the determination of standards for all products. Censuses and sample surveys are designed to complement each other. The role of census information centres on benchmarking the development of sample frames and the provision of basic small-area information. Sample surveys, on the other hand, provide detailed information on selected subjects and monitor the implementation of development initiatives.

Lessons learned in the last two censuses

Three important strategic lessons were learned in the last two censuses:

The relationship of the national statistics agency with the general public is of crucial importance in conducting a successful census. Statistics South Africa has learned the value of focused publicity campaigns combined with feedback to communities. Focus

on user needs in marketing products is critical in enhancing the utilisation of the data collected in the census.

Population censuses are extremely expensive programmes and the principle of utilising scarce resources to establish systems that can be continuously updated and broadly used, instead of using one-off solutions, must always be followed. The development of a geographic frame from the spatial information that was procured for Census 2001 is an example of the development of a permanent system that also services products other than censuses. This frame is now, for example, also used in disaster management. Another example is the Census Administration System, a management information system that will be used in managing surveys.

A permanent structure for conducting population censuses has been established within Statistics South Africa. Core staff members will ensure the retention of expertise and experience within the organisation. A capacity for conducting census research to inform the development of methodologies and content development has also been established.

Chapter 3: Census dissemination

3.1 Introduction

A national census can only be a success if everyone is willing to give up some of their time to complete the census questionnaire. Without this participation a national census cannot succeed. It is important, therefore, to show all citizens that they can benefit, directly or indirectly, from a successful census. They can benefit directly in that they can use the results of the census for their own work, such as developing business strategies or employment equity plans. They can benefit indirectly from the assurance that the government, better informed by a successful census, can provide basic services, such as water and electricity, in a much more efficient way. Therefore, the organisation responsible for conducting a census has the additional task of informing the public how important a census is to the development of the country, and one of the ways it can do so is by letting the public know what results and information are available and how these can benefit the average citizen.

To achieve this in South Africa, Statistics South Africa has conducted the census dissemination phase with a census marketing strategy, the aim of which is to ensure that the results of the census are distributed to those who need them the most. This chapter outlines this strategy: it provides important reasons why a marketing strategy is needed to disseminate census data, outlines how market research is done, covers the different types of markets that are targeted, provides information on what type of census products are distributed and, finally, it outlines how users' perceptions regarding products are gauged on an ongoing basis.

3.2 Why marketing?

The Marketing department views the entire business of Statistics South Africa from the data user's perspective. From this perspective, it can target the data users with appropriate products and services to satisfy specific needs.

A marketing strategy focusing primarily on census data is important for the following reasons:

- *Focused data distribution*: With a marketing strategy in place, Statistics South Africa is able to divide its data users into distinct market segments. A communications strategy that outlines, among other things, which products are useful to that specific segment, is created for each market segment. Instead of waiting for data users to contact Statistics South Africa for data, the Marketing Department takes the initiative to contact each data user first and to inform the user what products are available. This allows for a much more focused dissemination process.

- *Improved awareness of census products*: The Marketing Department is also responsible for creating awareness about census products through data user consultation workshops, presentations and exhibits. Together with one-on-one interaction with some data users, the Marketing Department aims to increase general awareness of census products among those who might find census data useful or vital to their work.
- *Improved awareness of Statistics South Africa*: The marketing strategy aims to generate awareness about Statistics South Africa so that it is regarded as the preferred supplier of easily accessible, quality statistics. It aims to inform the public that Statistics South Africa is the first place where it can search for statistics.
- *Improved statistical literacy*: The State President has, on many occasions, articulated his vision of an economically literate society. One element in the creation of economic literacy is the enhancement of statistical literacy, and Statistics South Africa has to ensure that it has products to promote this enhancement. Moreover, there is great concern in the government that high school learners are avoiding mathematic subjects at school. Statistics South Africa, through the marketing of its census products, can inform the public of why mathematics and statistics are important skills that every citizen needs to acquire.

3.3 Market research

In the previous section, important reasons why a marketing strategy is needed to disseminate census results were covered. The first step in implementing the marketing strategy is to acquire information about users of census data: what data do they require to get their work done and in what format?

Before the dissemination phase of every census, market research is done by holding workshops where all data users are invited to provide input on census products and services. In addition to this, a data user questionnaire is sent to all data users, asking for input on previous products. The feedback from both the workshops and the questionnaires is used to develop future census products.

3.4 Market segmentation

This section highlights the second step taken to disseminate census data, namely market segmentation. Market segmentation is a process by which all users of statistical data are divided into manageable market groups. Each segment or group has different demands and needs with regard to census data; therefore, different products have been created according to the needs of each of these segments. This makes the process of census dissemination much more focused and manageable. Each segment is briefly covered in this section.

Statistics South Africa divides all possible users of its data into the following segments:

- National government
- Provincial government
- Local government
- Parliamentarians
- Non-government organisations
- Constitutional institutions and major public entities
- International agencies
- Research and education
- Private sector
- Media
- General public

National government

There are 37 national government departments in South Africa, all located in Pretoria. These departments are responsible for the implementation of government policy, especially with regard to providing service delivery on a national level. Government departments are also responsible for the implementation of developmental strategies that aim to alleviate the constant problem of poverty among South Africans.

To perform these functions, national government departments need reliable statistics on the state of the nation, statistics that will enable better-informed decision making with regard to government planning and national spending. The State President has mandated Statistics South Africa to be the organisation that collects these statistics, and therefore this market segment is one of the most important to the Marketing Department.

In order to fulfil their mandate, government departments need detailed census statistics for nearly all geographic levels. Therefore, Statistics South Africa provides electronic census products that contain census data for not only large geographic areas, such as provinces or municipalities, but also for small geographic areas, such as suburbs, villages and towns.

In order to reach this group, various presentations with department heads are conducted, highlighting the importance of census data. Once the data products have been developed, members of the Marketing Department from Statistics South Africa's head office install the products on various computers within each department. Training on how to use these products is also provided.

Provincial government

There are 103 provincial government departments in South Africa. These departments are responsible for the implementation of government policy at provincial level. Provincial government departments need census statistics for much the same reasons that national government departments do, the only difference being that provincial government departments are focused more on the provincial than the national level.

The marketing strategy aims to provide the same kind of census products to provincial government, i.e. products that contain detailed statistics on all geographic levels. Because provincial government departments are located throughout the country, the provincial offices of Statistics South Africa, of which there is one in each province, take on the responsibility of installation and training.

Local government

In 1998, the Municipal Demarcation Board demarcated South Africa into 309 local authorities. These authorities are composed of six metropolitan areas and 47 district councils. The district councils are further subdivided into 231 local municipalities and 25 district management areas. These local municipalities have been mandated to deliver services and infrastructure at a local level. This is an extremely important task, as this is the level at which the government's efforts regarding sustainable development and poverty alleviation are centred.

Small-area statistics, provided by census data, are therefore vital to any local authority that needs to make decisions on how to implement developmental plans in areas under its control. Small-area statistics provide information on which suburbs or towns are lacking in various services and where a local municipality needs to focus its money and efforts.

Digital census products together with various print products are developed to meet the data needs of local municipalities. Due to the large number of local municipalities in South Africa, the provincial offices, with assistance from Head Office, conduct training sessions and installations of the products at all municipal offices.

Parliamentarians

National elections are conducted in South Africa every five years. In order to manage the entire election process, the Municipal Demarcation Board demarcates the whole country into electoral wards. Depending on the outcome of the voting process, political parties nominate candidates to each ward. Politicians and parliamentarians are therefore interested in the wards that they are contesting. Therefore, a product that provides statistics on a ward level meets the needs of this market segment.

Non-government organisations

Non-government organisations (NGOs) are normally non-profit organisations that work to uplift the standard of living of the communities. Some assist by implementing development projects at a local level, and others monitor the government's progress in providing service delivery. General census statistics aid this segment in its work.

Due to the large size of this market segment, however, NGOs are encouraged to extract general census statistics from StatsOnline, Statistics South Africa's website. If the need for more detailed statistics arises from a specific NGO, arrangements are made for the installation of, and training on, some of the digital census products that are available.

Constitutional institutions and major public entities

Constitutional institutions ensure that the major elements of the Constitution are applied throughout the country. Major public entities work with the government to provide infrastructure, service delivery and information on a national level (e.g. Eskom, museums etc.).

Census statistics are of great value to this sector as they can be used to determine which communities lack basic infrastructure. They can also provide an indication of where a major public entity can focus its resources.

As in the case of NGOs, members of this market segment can use StatsOnline to meet most of their data needs. However, if more detailed statistics are required, arrangements are made for the installation of, and training on some of the digital products.

International agencies

International agencies (e.g. the United Nations and the World Health Organization) provide regular reports on the progress of developing and emerging market economies. Census statistics are vital because international organisations use them to compare South Africa's developmental progress to that of other countries. The data highlight where donor funds are most needed and allow the international community to make informed decisions about South Africa.

Members of this market segment use StatsOnline to meet most of their data needs.

Research and education

Research and educational institutions include the various universities, schools and universities of technology (formerly technikons) throughout South Africa. They use census data extensively, primarily for research but also for training and development.

Schools require basic census data to advance the skills of data handling and to provide all learners with an early awareness of basic facts regarding their country.

Private sector

The private sector consists of all private companies and businesses that conduct business for profit. Census statistics can be of great value to this sector because they can be used to determine market penetration as well as growth of specific target markets. They also can also be used to assist with employment equity and business plans.

The private sector market segment is extremely large. For this reason, a user-friendly product, which is easily accessible on StatsOnline, is required. This enables many users to access various types of general census statistics at the same time.

Media

The media consist of all forms of mass communication, including newspapers, radio, magazines and television. Census statistics are of great value to this sector because they provide journalists with accurate, understandable statistics that can be used as background material to the stories that they write. Statistics South Africa has realised that a good relationship and effective collaboration with the media are of paramount importance, as most citizens will receive exposure to census results through the media. It is thus vital to communicate results to journalists in a language that they understand. Due to the nature of their work, journalists generally face intense time constraints, thus the reports should be communicated clearly, concisely and in simple language.

General public

As was previously mentioned, the role of statistical literacy in developing economic literacy is of paramount importance. An economically literate society will be able to advance public debate and participate in the democratic processes of the country. For this reason, every individual in the country who does not form part of a previously mentioned group has been placed in this market segment.

The general public requires easily understandable publications and basic information from the census. Due to advances on the Internet, many of these publications can be placed on StatsOnline for easy download. However, for those who do not have access to the Internet, print publications are useful, especially if they can be obtained from libraries and multi-purpose community centres.

3.5 Census products

The development of census products is steered first by the feedback received during the data user workshops and second by the needs expressed by the 11 market

segments that were covered above. Each market segment is targeted with a specific product or range of products in order to meet its needs.

The following is a list of census products:

Printed products

- *The Count and How it Was Done*
- *Post-Enumeration Survey: Results and Methodology*
- *Key Results*
- *Census in brief*
- *Age Tables*
- *Primary Tables*
- *Key Municipal Data*
- Poster – *Did You Know?*
- Poster – *Thematic Maps*

Digital products

- *Community Profiles* databases
- *Sample Database (10%)*
- *Digital Census Atlas*
- *Ward Profiles 2003*
- PX-Web databases
- *GIS Spatial CD*
- *Urban/Rural Investigation Discussion Document*
- *Census 2001: Concepts and Definitions*
- *Census 2001: Computer Editing Specifications*
- *Census 2001: Metadata*

The next sub-section deals with the brief discussion of the contents of each of these printed materials.

The Count and How it Was Done

This report focuses mainly on the methodology that is used in South Africa to conduct the national census. It provides detailed information on how the census was conducted and provides a summary of three-and-a-half years of census planning, which includes demarcation, enumeration, processing and data checking. The document also highlights the major problems that are encountered during each census and what challenges need to be addressed when census planning is considered.

The Count and How it Was Done has been written for research institutions that use census data extensively and that need to know and understand how the census data were compiled. The Census 1996 and Census 2001 versions of this publication are available as printed documents or as Acrobat Reader documents on StatsOnline.

Post-Enumeration Survey: Results and Methodology
After each national census, a post-enumeration survey is conducted to determine the level of undercount. It is an important survey as the results indicate how many people were not counted during the census, and the final census results are eventually adjusted to reflect this. This publication provides an overview of how the post-enumeration survey is conducted in South Africa, with particular emphasis on the methodologies used to ascertain the level of undercount.

This publication is also available for download on StatsOnline, and print copies are available. It is intended for research institutions that use census data extensively and is also available for Census 1996 and Census 2001.

Key Results
The *Key Results* publication is aimed at the general public. It is a small A6-size pamphlet that contains some basic results from the census in the form of tables and charts. It provides basic information on how the census was conducted and includes a few census variables, such as education, language, household goods and provincial population figures.

Census in Brief
The A6-size *Census in Brief* document is much like the *Key Results* pamphlet but more detailed. This is an accessible and easy-to-use report consisting of over 80 tables and graphs at national and provincial levels. It provides an extensive range of statistics pertaining to individuals and households: including population, gender, language, education, employment and household services.

Census in brief is the most popular printed census product and is widely distributed around the country as well as being available from all Statistics South Africa offices and from StatsOnline. It is aimed at all market segments although the general public and educational institutions benefit the most from using this publication. It is also available for both Census 1996 and Census 2001.

Age Tables
This printed publication is aimed at researchers and provides the national age profile with provincial breakdowns. These tables are presented in both single years and five-year age groups, broken down into urban and non-urban areas, gender and

population groups. Actual numbers and percentages are shown for each table. At present, the *Age Tables* publication is only available for Census 1996.

Primary Tables

This A4-size printed publication for researchers contains detailed tables showing gender and population group tables, all cross-tabulated with 23 variables from the census. The Census 2001 version of this publication also contains comparative tables for Census 1996. In addition, provincial versions are also available as separate publications.

Key Municipal Data

This A4-size publication, consisting of 345 pages, is aimed at local government as it contains various census results for every one of the 309 local authorities. Together with maps of the various municipalities, each page covers a single local authority and includes tables on important variables such as population, education, household services and employment. The document is ring-bound, so photocopies of specific pages can be made with ease.

Poster – *Did You Know?*

Schools, libraries and multi-purpose community centres have been targeted with various census-related posters. The aim of the posters is to raise awareness of census in the general public. Moreover, the posters can help learners at schools to learn basic facts and figures about their country. The *Did You Know?* posters are large and colourful and display interesting facts and figures from the census.

Poster – *Thematic Maps*

The second type of posters is intended mostly for schools, and these posters display provincial population data from the census in the form of thematic maps. The maps do not only assist in the data-handling skills of learners but they also assist in their geographic knowledge of South Africa.

Community Profiles databases

The first digital product covered in this chapter is Statistics South Africa's flagship product. The *Community Profiles* package consists of 12 CDs that contain a range of summarised databases with almost all variables from the census. SuperTABLE cross-tabulation software is included and allows the user with a PC to extract a wide range of census tables for sub-places (villages or suburbs), main places (towns or tribal areas), wards, municipalities, district councils, magisterial districts and provinces. Created tables can be exported into various formats, such as Excel and CSV as outlined in Annexures 1 and 2.

Charting and mapping software are also included, and these are linked directly to SuperTABLE. One can use the mapping function to create interactive thematic maps directly from the tables designed in SuperTABLE.

Many market segments use the *Community Profiles* package although it is primarily aimed at all three government market segments. Training and installation sessions are also included to provide a user with an introduction on how to use the package. A training document has been developed and is attached to this chapter as Appendix 1. Due to its size, the *Community Profiles* package is not available on the Stats SA website.

Sample Database (10%)

The *Community Profiles* package contains summarised databases that can be used by many organisations and departments to extract census figures for a wide range of geographic levels. However, this might not suffice for the serious researcher who might want to perform his or her own statistical analysis on census data. For this reason, Statistics South Africa has published the 10% sample of unit records which consists of various ASCII databases on a single CD. The 10% sample can only be used by those with access to a complex statistical package, such as SAS or SPSS, and the CD only contains geographic data down to local municipality level. Like the *Community Profiles*, the 10% sample is not available on the Internet.

Digital Census Atlas

For those interested in census statistics for local municipalities and looking for a much simpler package than the *Community Profiles* or the 10% sample, Statistics South Africa has made the *Demographic Atlas* available on StatsOnline and as a single CD. The *Digital Census Atlas* is a user-friendly tool that allows a user to select pre-created thematic maps, tables and charts from a wide range of census themes which can be selected for South Africa, the nine provinces and all the local municipalities. Only Census 2001 figures are available. The *Digital Census Atlas* can be run straight from the website or from the CD using Microsoft Explorer. No installation is required.
Although many market segments use this product, it is primarily aimed at international agencies, NGOs and the private sector.

Ward Profiles 2003

Due to the excitement over national elections in 2004, Statistics South Africa embarked on a project to adjust all its Census 2001 figures to the electoral ward boundaries of 2003. Hence this product was produced and it is almost identical to the *Digital Census Atlas* in that it is a user-friendly selection tool that allows the user to select maps, charts and figures and that can also be run from a single CD and off the website using Microsoft Explorer. However, *Ward Profiles* is different in that it provides statistics for all the electoral wards in the country with comparisons to Census 1996.

This product is used by many market segments, but it is primarily targeted at parliamentarians, political parties and local municipalities.

PX-Web databases

The main advantage of the *Digital Census Atlas* and *Ward Profiles 2003* is that one can easily obtain basic tables, charts and maps for wards, local municipalities and provinces on the Internet. However, some users may prefer to create their own tables and cross-tabulations if they find the basic pre-created tables insufficient. Personalised cross-tabulations can be performed using the *Community Profiles* package, but this is not available on the Internet.

To provide cross-tabulation functionality on the Internet, Statistics South Africa has developed an online database utility that allows anyone with an Internet connection and Microsoft Explorer to create and download custom-made census tables. One can perform basic cross-tabulations online for provinces, district councils and local municipalities. Tables can be downloaded in various formats, such as Excel or Lotus. However, although there is a charting facility available in PX-Web, there are no maps.

This product is aimed at all market segments, but primarily at international agencies, NGOs and the private sector. A training document has been developed and is attached to this chapter as Appendix 2.

GIS Spatial CD

The national census is based on a geographic framework. To make use of the census results, one has to express the figures in the form of some geographic structure. For this reason, census results are available for various geographic boundaries, such as provinces, district councils, local municipalities, magisterial districts, towns, suburbs, villages and electoral wards, to name a few. Within Statistics South Africa, these different boundary systems are stored in the form of a Geographic Information System (GIS). The *GIS Spatial CD* contains all the various boundary systems in ArcView shapefile format, which allows a GIS analyst to map the census results with other mapping data. The *GIS Spatial CD* can only be used with compatible GIS software, such as ArcView or GeoMedia, and is primarily aimed at GIS analysts in the government, research institutions and the private sector.

Urban/Rural Investigation Discussion Document

The legal definition of ‘urban’ and ‘rural’ was scrapped when local municipalities were demarcated in 1998. Since then, many have expressed a need for a new, non-legal definition of urban and rural. That is what this discussion document, available in Acrobat Reader format, attempts to explore. The paper identifies differences in how areas were categorised as urban or rural in 1996 and 2001. It experiments with two

different classifications of urban and rural, the old legal definition used for 1996 and a newer population density classification, and then it examines census results in the light of these two classifications. This paper is available for download from StatsOnline.

Census 2001: Concepts and Definitions
The national census is an extremely large project. Its size and complexity have spawned an entire plethora of acronyms, definitions and concepts. This document, available in Acrobat Reader format, explores the various definitions and concepts related to the census and provides explanations for the many acronyms used. This document is ideal for the user who is using census information on a regular basis, and it is available for download from StatsOnline.

Census 2001: Computer Editing Specifications
After the initial enumeration period is complete, the task of capturing the actual data from millions of questionnaires begins. This 370-page document outlines the various computer editing specifications that were applied to the raw data in Census 2001 so that meaningful results could be reached. It provides information on how derived variables were calculated as well as specifying how different imputation methods were used to allocate values for unknown and incorrect answers. This document can be downloaded from StatsOnline in Acrobat Reader format.

Census 2001: Metadata
Census 2001: Metadata files are available for download from StatsOnline in Acrobat Reader format. These cover information on all the variables that were used during the census. They explore each question in the questionnaire and provide information on how the question was asked and the code lists used later on in the capturing phase. The various geographic boundaries that were used during the census are also explored. The following metadata documents are available:

- Information on geography
- Information on personal variables
- Information on household variables
- Information on mortality

3.6 Monitoring and evaluation

The Marketing Department of Statistics South Africa has committed itself to constantly monitoring and evaluating the changing needs of data users as well as their perceptions regarding the utility of the products listed in the section above. Some of the ways in which it monitors the various market segments are by:

- providing an online questionnaire for all users who log on to use its Internet products
- gathering feedback through an annual user needs survey
- inviting users to attend workshops where they can provide feedback
- recording the number of telephonic and electronic queries received through an electronic user inquiry database
- measuring market penetration on a quarterly basis through user workshops, Internet questionnaires etc. tracking the number of users trained on the manipulation and interpretation of census data.

3.7 Conclusion

A national census is a massive undertaking that requires the involvement of almost every citizen. It is important, therefore, to ensure that every citizen knows and understands why a national census must be conducted and how the census results can benefit the life of the average person. Census statistics are vital to the development of the country, and if the results are not used by those who need them the most, the resources spent on conducting a census are wasted. For this reason, Statistics South Africa has implemented a marketing strategy to ensure that all relevant data users know and have access to census data. In order to do this, the Marketing Component has divided all data users into 11 market segments, and various products have been developed to meet the needs of these segments. Census data *must* be used, and the data *must* be taken to the people. The marketing approach ensures that this is achieved.

Chapter 4: Census geography

Glossary

Census geography	▪ Describes the geographic breakdown of a country to facilitate the census processes, field operations and logistics, processing and analysis of results and reporting of these results. ▪ Describes the subdivision of the country into geographical areas for the purpose of the census (Martin, 2000). ▪ Refers to the land area and how it is organised for purposes of facilitating census operations in terms of collection and dissemination of statistical data.
Geography	▪ Refers to the land area and how it is organised for purposes of collection and dissemination of statistical data. ▪ The geography is essentially a hierarchical system of nested areas which varies according to the level of information that is required. The lowest level of the hierarchy consists of enumeration areas.
Geographic unit	Refers to the geographic elements (or areas) of a certain geographic level in the geography hierarchy.
Geography hierarchy	A hierarchy of different geographic levels nested together. These include the country's administrative and statistical areas.
Census geography frame	The general outline of the manner in which geographic levels nest hierarchically from the lower level to the uppermost level, used for census purposes.
Digital map product	Pre-prepared static maps that can be displayed on the computer monitor (i.e. JPEG, PDF etc.).
Geography frame	The general outline of the manner in which geographic levels nest hierarchically from the lower level to the uppermost level for administrative purposes.
Enumeration area	The smallest unit area of the census geography frame, of about equal population size with other enumeration areas, that one person visits during a census to interview members of households within a specified time period.
Place name	The most easily recognisable small-area geographic entity at a local level, e.g. a community, village or suburb.
Municipality	A generic term used by the Municipal Demarcation Board (MDB) to refer to metropolitan areas, local municipalities, district councils and district management areas.
Metropolitan area	Metropolitan areas are conurbations featuring high

or Category A municipality	population density, intense movement of people, goods and services, extensive development and multiple districts and industrial areas. Other features include a complex and diverse economy, a single area where integrated development is desirable and strong interdependent social and economic linkages among the area's constituent units.
Local municipality **or** Category B municipality	Local municipality boundaries were determined according to settlement type, the rationalisation of municipalities, manageable size and functionality. Including the DMA, they make up the district councils.
District council **or** Category C municipality	District councils were created for better management of local municipalities and the DMAs that constitute them.
DMA	DMAs are areas with both district and local municipality features where the establishment of a local municipality is not appropriate (does not meet a set of requirements). These are areas of special interest, e.g. deserts and semi-arid areas, state-protected and conservation areas and special economic areas.
Geography type	Geography types are defined based on at least four broad settlement types characteristic of South African communities (urban formal and informal, tribal and farm).
Enumeration area type	EA types profile land use and human settlement within an area using two principles: first, the specific geographical location of the EA (e.g. urban, adjacent to urban or within local tribal area) and second, the land use and the kind of dwellings predominant within the EA.
MDB	Municipal Demarcation Board

4.1 Introduction

Census geography is fundamental to census taking by organising the way information is collected and published. Two geographies were developed for the 2001 Census. The first facilitates the management of census-taking operations, primarily through balancing the workload among enumerators. The second enables the publication of small-area data in order to provide sensible results for decision making and planning.

South Africa's census geography has undergone several changes. These changes were necessitated by administrative boundary alterations over the last century. Census data are collected in accordance with administrative boundaries. Changes in these boundaries usually imply adjustments of the lower-level geographical units: statistical and enumeration boundaries.

This chapter provides an overview of changes in South Africa's census geography over time. It gives guidelines for the utilisation of the census geography and explains

the reasons for the different geographic frames. Practical guidance on the analysis of census data with spatial elements is also given through examples of relevant digital map data products.

4.2 Background

The smallest unit area of the census geography frame is the enumeration area. These areas of approximate equal population size are so designed that an enumerator can visit and interview members of households within a specified time period.

Coombes (1995) argues that the usability of census information, particularly in a geographic and GIS environment, can be determined by answering the following seven questions:

- Are the building blocks (enumeration areas) the **smallest** that are deemed to be possible given **confidentiality restrictions** (to allow for maximum flexibility of aggregation)?
- Is each of the areas in a set of building blocks, or other set of areas, **defined on a consistent basis** across the country?
- Do the enumeration area boundaries represent **real-world entities**, such as settlements, which can thus be recognised using these boundaries?
- Do the enumeration areas **cover the whole country?**
- Can the enumeration area boundaries be made available in **digital format**?
- Do the enumeration area boundaries create consistent boundaries at that level or for some minimal grouping of areas to allow for **comparisons with previous censuses**?
- Can the enumeration area boundaries be readily and accurately **linked** through location coding to the areas used in **non-census outputs**?

The geographic frame should also **adhere to administrative, political and statistical boundaries** and their change over time. These criteria provide a good measure of the flexibility of the geography of census outputs.

4.3 Adherence to administrative, political and statistical boundaries

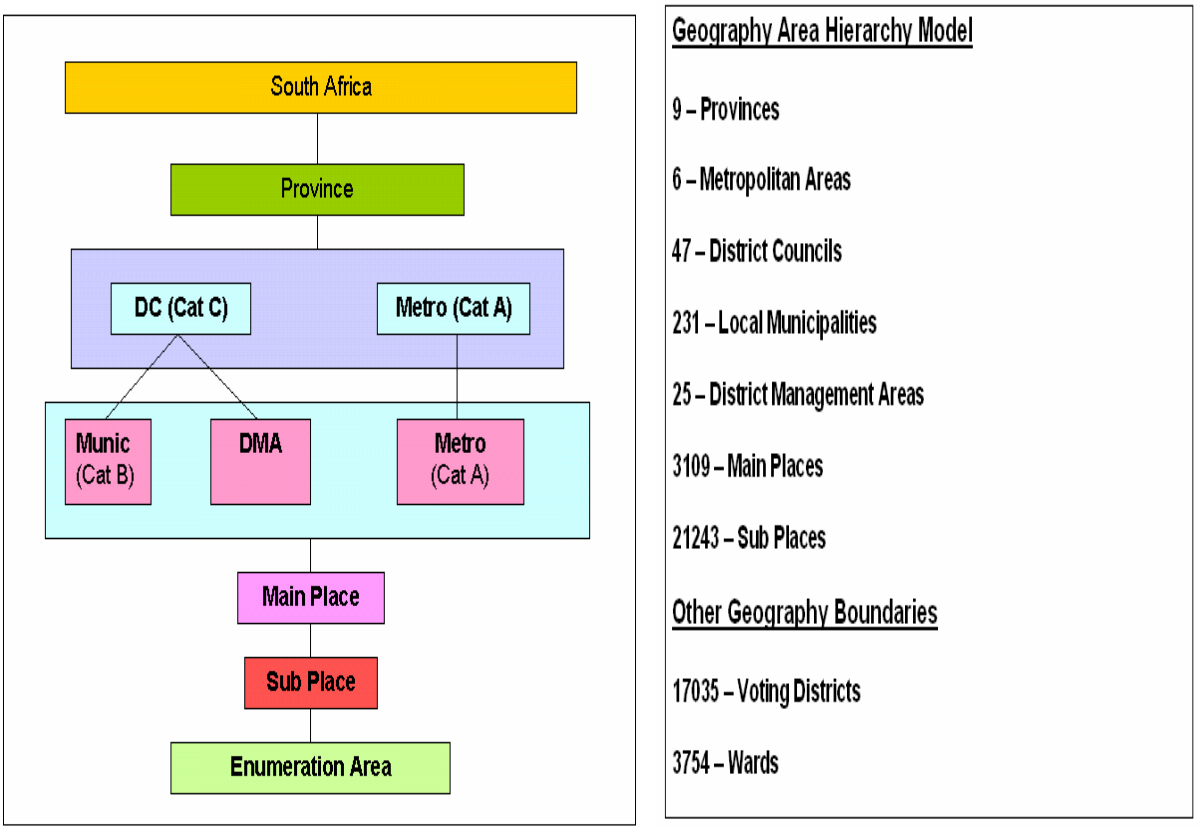
Figure 1 shows the geographic frame for the planning and management of data collection.

4.3.1 *Enumeration area*

Each enumeration area has a unique eight-digit identifier called the geo-code. The geo-code is used to link the enumeration area geographic units with higher-order geographic units such as place names, municipalities, district councils and provinces.

This enables the information of the enumeration areas to be aggregated and represented at any of the administrative boundaries.

Figure 1: Geography Area Hierarchy Model (For organisation and management of Census 2001)



4.3.2 Place name

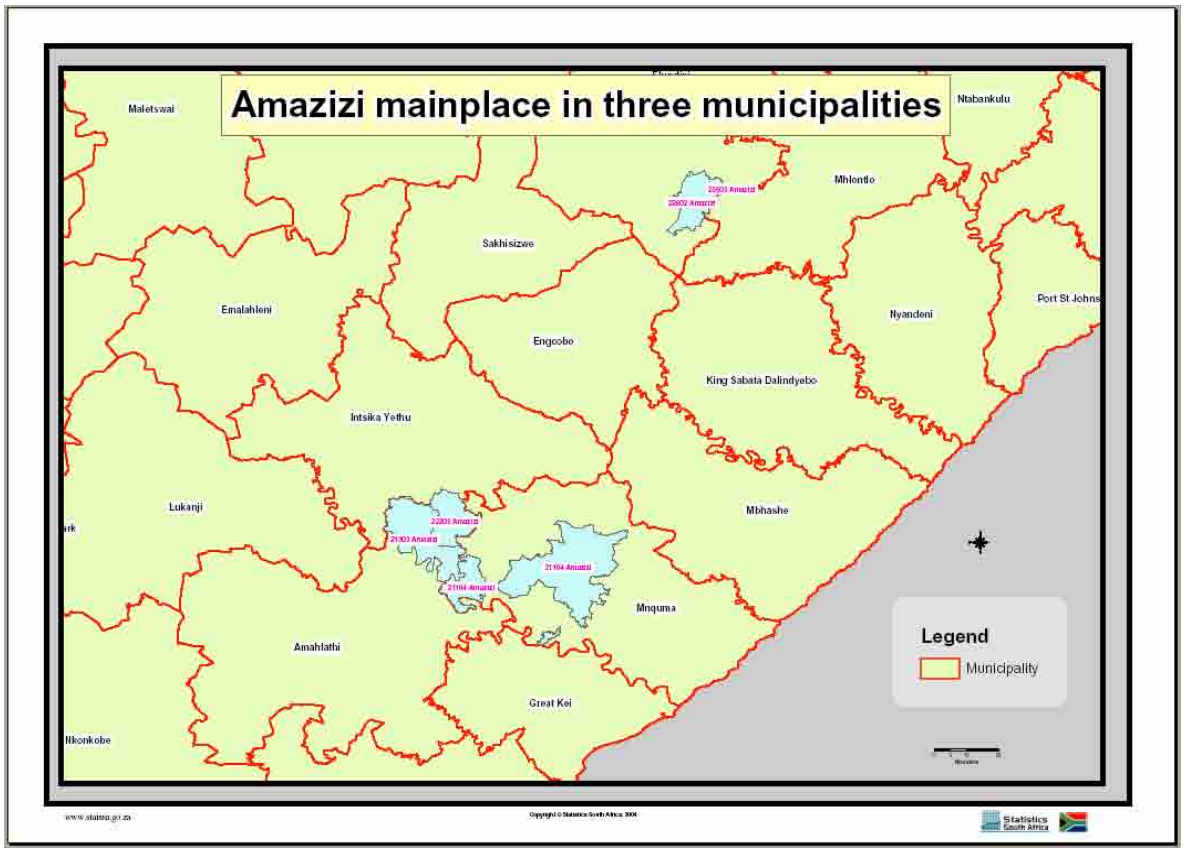
The most easily recognisable small-area geographic entity is a place name that signifies a particular community, village or suburb. Place names were first coded in Census '96. In Census 2001 a more complex hierarchy of two place name levels (the main place name and the sub-place name) was used. In order to distinguish among different settlement types, enumeration areas within the same locality (main place name) with the same enumeration area type were merged to create a sub-place name. In cases where sub-places are not defined the main place name has the suffix SP.

Main place

A main place name is the equivalent of a small city, a suburb with subdivisions in a large city or a tribal area in communal or trust land areas. Municipality names were used where main place names were not supplied. A total of 3 109 main place names were coded although there are only 2 674 main places (Geography Metadata).

Place names that straddle more than one municipality were split into a corresponding number of entities with the same name but different codes. For main place names a suffix such as “Part 1” or “Part 2” was added to the name. Figure 2 shows Amazizi, a main place name that occurs in four municipalities.

Figure 2: Amazizi – an example of a main place name that occurs in more than one municipality

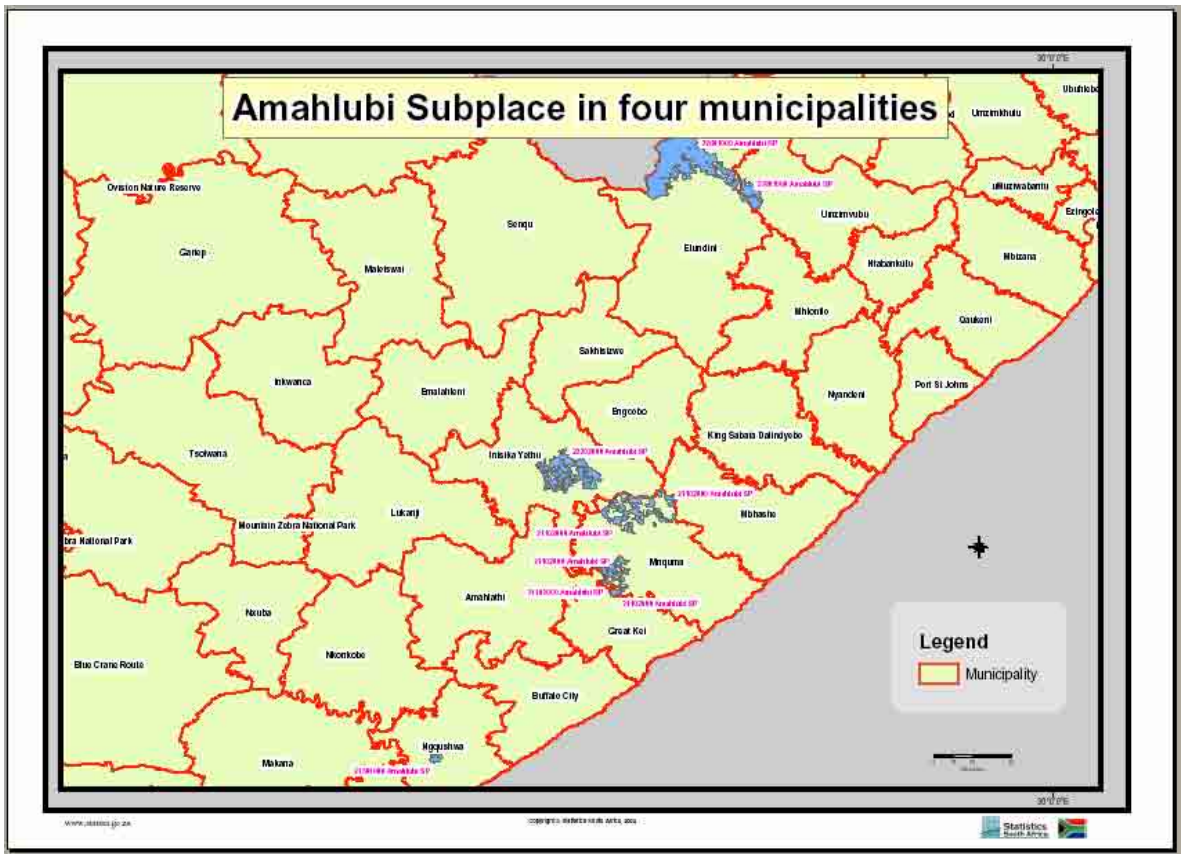


Sub-place

A sub-place is equivalent to a split suburb or merged suburb in urban formal areas, a locality in the informal areas and a village in the traditional areas. There are 15 966 unique sub-place names. Due to cross-boundary sub-places, as well as the same sub-place names that occur in different municipalities, a total of 21 243 sub-place names were coded.

Figure 3 shows Amahlubi, a sub-place name that occurs in four municipalities.

Figure 3: Amahlubi – an example of a sub-place name that occurs in more than one municipality



The data for place names that are in different municipalities or even in different provinces (the sub-place name Mandela Park, for example, is found in the Western Cape, the Free State, Gauteng and Mpumalanga) must be accessed by their codes.

4.4 Municipality

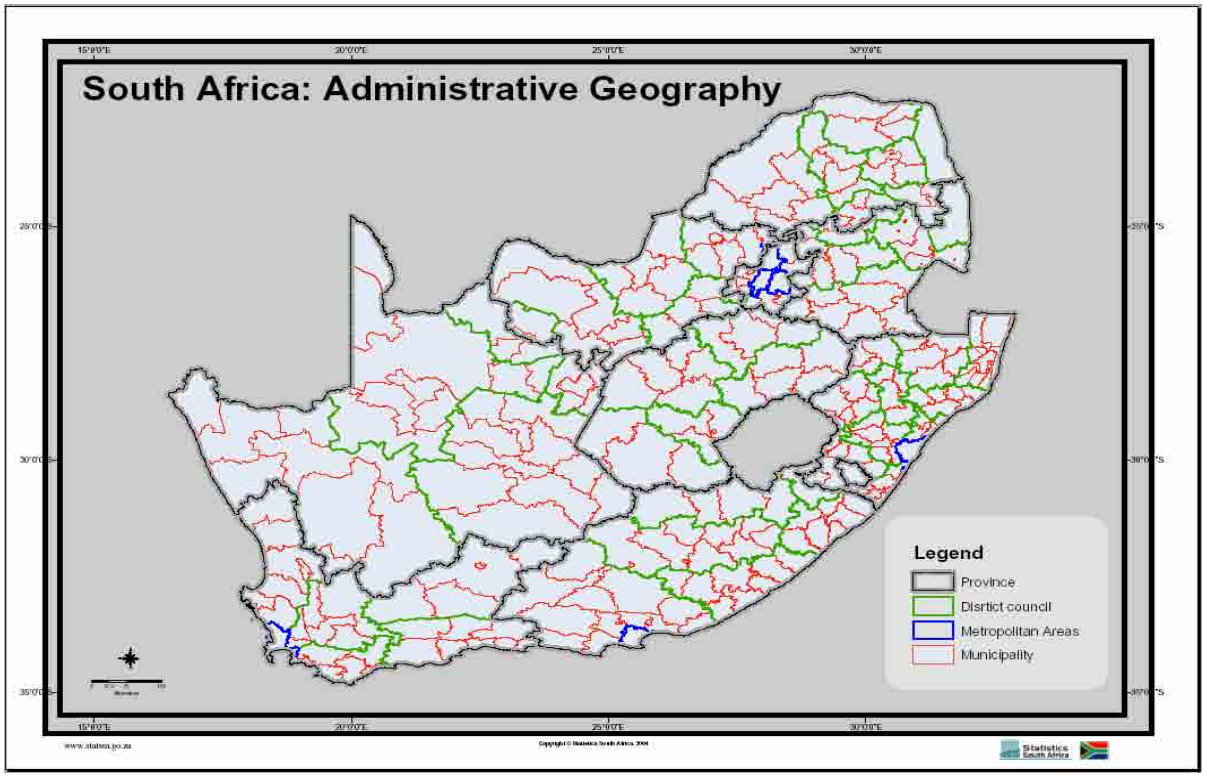
The term *municipality* refers to local municipalities, district councils and district management areas. The local municipalities and district management areas constitute the larger district councils. District management areas were created where the requirements of a local municipality are not met, such as deserts and semi-arid areas, state-protected and conservation areas and special economic areas.

Metropolitan areas are conurbations featuring high population density; intense movement of people, goods and services; extensive development and multiple districts and industrial areas. Other features of a metropolitan area include a complex and diverse economy, a single area where integrated development is desirable and strong interdependent social and economic linkages among its constituent units.

In total there are six metropolitan areas, 47 district councils, 231 local municipalities and 25 district management areas. Figure 4 is a diagrammatic representation of these administrative boundaries.

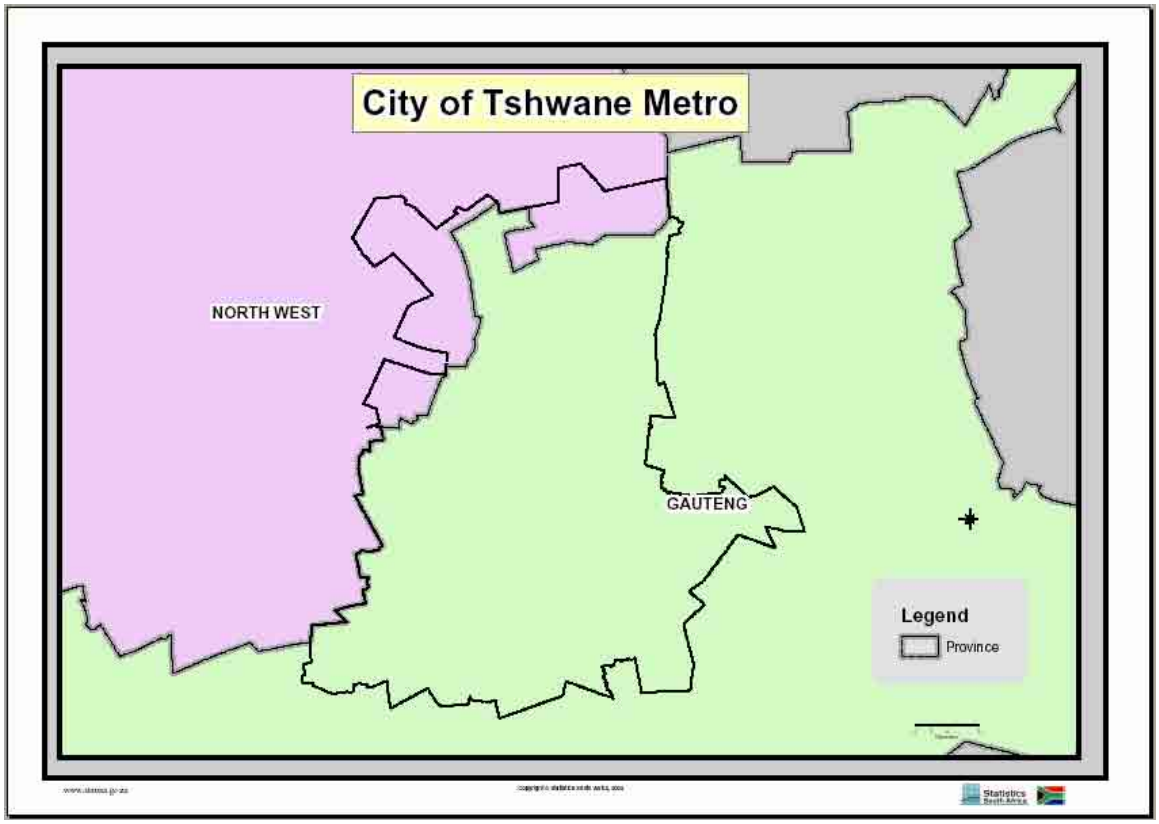
The Municipal Demarcation Board defined the boundaries for these areas in 2000 when a new municipal structure replaced the old structure. Key considerations in determining these boundaries were the need for rationalisation, the need to create areas of manageable size as well as settlement type and functionality. The resulting district councils ensure better coordination with other spheres of government, better planning and better resource allocation across the local municipalities of which they are made up.

Figure 4: Provinces and their administrative geography



Provincial boundaries were obtained from the Department of Land Affairs. It should be noted that there are some cross-boundary local municipalities, district councils, district management areas and metropolitan areas (see Figure 5). This must be taken into account when cross-tabulations are compiled.

Figure 5: An example of a cross-boundary metropolitan area



More details on municipalities and their coding structure are available on Statistics South Africa's website (www.statssa.gov.za). Further information on municipalities and other geographical entities can be obtained from the Constitution of South Africa (Act 108 of 1996), the Local Government Municipal Structures Act (Act 117 of 1998) and the Organised Local Government Act (Act 52 of 1997). Information can also be found on the web at www.local.gov.za, www.salga.org.za and www.demarcation.org.za

4.5 Confidentiality restrictions

Confidentiality issues shape census geography. The location coding that is used in census geography is so detailed that households or individuals may be identified through a distinctive combination of characteristics.

For previous censuses the smallest commonly used area from which census information could be accessed and displayed on a map was the enumeration area. Due mainly to confidentiality considerations it was decided that the Census 2001 data would be released only at the sub-place level. A number of standard products based on this output geography have already been published.

The trade-off for this decision to safeguard confidentiality is that the linking of any data other than the data collected in Census 2001 to their geographic frame must be done through special data re-aggregations at Head Office. The procedures that are used for this are described in the next chapter.

The use of the sub-place as the lowest level in the output geography did not resolve all the confidentiality issues. The splitting of localities that overlap two or more municipalities resulted in the creation of very small sub-places, some made up of a single enumeration area. This further possibility for a breach of confidentiality was eliminated through the randomisation of results for all cells with five or less counts when cross-tabulations are run in SuperCross, the software through which census data are accessed by end users.

Martin (1997) advised that the creation of output (or dissemination) areas might prove useful as census outputs ideally require some degree of statistical control over the size and characteristics of small-area populations. Statistics SA is currently researching an output geography based on dissemination areas for Census 2001.

4.6 Consistency of building blocks

Confidentiality was not the only reason why enumeration areas were perceived as unsuitable small areas for data publication. Standardisation of enumerator workloads resulted in an unacceptably high variation in population size among enumeration areas. Furthermore, some enumeration areas lacked social homogeneity as they were of mixed type.

The enumeration areas consisted of approximately 80 to 250 dwelling units, depending on the geographic and enumeration area type. Their geographic types were based on the four broad settlement types characteristic of South African communities: urban formal settlements, urban informal settlements, tribal settlements and farms.

Table 1: Geographic and enumeration area types

Geographic type code	Geographic type	Enumeration area type	Enumeration area type code	Number of dwelling units
1	Urban formal settlements	Vacant	0	
		Commercial farm	2	80 to 120
		Smallholding	3	80 to 120
		Urban formal settlement	4	120 to 250
		Informal settlement	5	120 to 170
		Recreation facility	6	80 to 120
		Industrial area	7	
		Institution	8	
		Hostel	9	
2	Urban informal settlements	Vacant	0	
		Informal settlement	5	120 to 170
		Recreation facility	6	80 to 120
		Institution	8	
4	Tribal settlements	Vacant	0	
		Tribal settlement	1	80 to 120
		Recreation facility	6	80 to 120
		Institution	8	
		Hostel	9	
5	Farms	Commercial farm	2	80 to 120
		Recreation facility	6	80 to 120
		Institution	8	
		Hostel	9	

Apart from these geographic types, ten further enumeration area types were defined in demarcation. Table 1 shows the resultant combination of geographic and enumeration area types that profiled land use, human settlement and the predominant type of dwelling.

Enumeration areas in informal settlements are usually small because of high population density. In rural and farm areas they tend to be large.

Census 2001 used 80 787 enumeration areas versus the 83 126 demarcated in Census ’96. Each enumeration area used in Census ’96 was verified, assessed and modified where necessary. The modifications involved the splitting and merging of enumeration areas to obtain results in line with the demarcation rules explained above.

4.7 Boundaries as real-world entities

Boundaries of each element of a geographic entity should represent real-world boundaries. This must be for both the data collection geography and the publication geography. Real-world or physical boundaries are important in assisting fieldworkers in navigating to visiting points. Lower-level geographic units that constitute output boundaries (EAs and place names) were demarcated so that they preserve social homogeneity. Social homogeneity is crucial for local service planning and resource allocation.

4.8 Coverage

After Census '96, Statistics South Africa launched Project Eagle. This project established the census geography frames for 1996 and 2001 as illustrated in Table 2 in which each geographic unit fits within a comprehensive hierarchy (See Figure 1 for Census 2001 geography frame.). A geographic area unit at any level therefore consists of a grouping of smaller areas at the lower levels. Any geographic area in the hierarchy is made up of whole units at the lower level, the lowest level being the enumeration areas. This spatial enumeration area database, that is linked to a detailed attribute database and that covers the whole of South Africa, was used in Census 2001.

4.9 Digitisation

The digitised EAs from Project Eagle formed the foundation for the creation of a digital census geography frame for 1996 census dissemination. These EAs were modified to provide the base for the development of the 2001 census geography frame.

Up to Census '96 administrative maps served as the basis for demarcation. Town planning maps were used for large new towns and developments. Aerial photographs were used to pinpoint new residential units. Many maps were insufficient and outdated, as new developments were not reflected on them. The map-based publications of census results were difficult to produce. The first steps to address this situation were taken when the Human Sciences Research Council captured the enumeration areas of Census '91 digitally.

Census '96 represented a transition from traditional demarcation and mapping methods toward an electronic spatial database. Due to budgetary and time constraints the traditional methods were used to demarcate the enumeration areas and to manage the census. A digital census geography frame that provided a spatial backdrop for the reporting of the census results was then constructed. This reporting frame was based on the old municipal structure and magisterial districts. It also enabled Statistics South Africa to monitor progress and coverage of enumeration areas in the census and the post-enumeration survey. The collaboration of Statistics South Africa, the Department of Land Affairs, the Independent Electoral Commission

and the Municipal Demarcation Board in Project Eagle led to the conversion of the analogue enumeration area boundaries into digital format in 1997.

Census 2001 was the first census in South Africa in which an established geographic information system was utilised in census management and map-based product production.

4.10 Comparability across censuses

The following factors complicate the work of researchers and academics who wish to do comparative studies using data collected in South African censuses:

- the continuous and comprehensive changing of administrative boundaries
- the revision of the set of enumeration areas that was used in Census '96
- the decision not to release the Census 2001 data at enumeration-area level.

Coombes (1995) notes that time and space are the two fundamental organising principles for any population census information. Population censuses that covered all the population groups in South Africa date as far back as 1911. However, the spatial dimension that is used to collect, compile, assess, analyse, disseminate and publish information has been fragmented by the Bantustans and self-governing territories imposed by apartheid. Efforts in the last decade to reintegrate and rationalise administrations have had a similar effect.

Coombes (1995) also points out that an enumeration area definition is unique to a particular census. This has severe disadvantages for the study of change and complex issues such as poverty that require merging of census data with other datasets. Where possible, changes are limited to facilitate comparability.

The data of different censuses are compared through the re-allocation of the data according to a common set of boundaries. Historical boundaries such as the magisterial districts are best suited for this. It is done either through re-aggregation techniques based on proportional enumeration area overlap (area-based interpolation) or the centroid location of enumeration areas (point-in-polygon location). Statistics SA is currently working towards releasing 1996 census data for the new municipal boundaries.

Note that comparisons between Census '96 and Census 2001 cannot be done at place name level due to the change in place name structure and the magnitude of the changes in the place name dataset. Furthermore, comparisons are extremely difficult for place names that are too small for re-apportioning enumeration area data using the centroid or area-based interpolation method.

It is important to maintain historical geographic frames in order to make comparison over time possible. Table 2 shows geographic frames for three time periods.

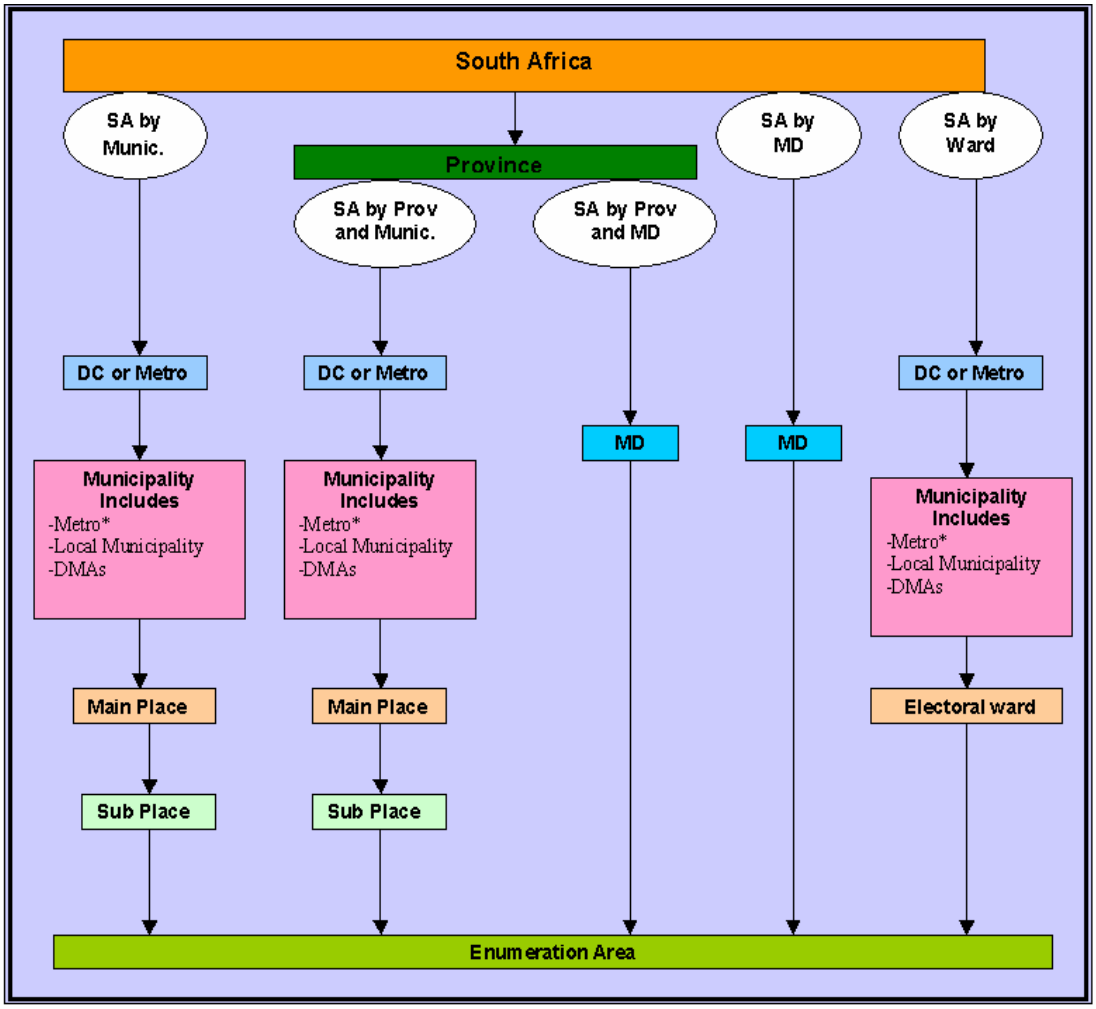
Table 2: Previous geographic frames

Geographic frame prior to Census '96	Geographic frame for Census '96	Geographic frame for Census 2001
▪ Provinces ▪ Bantustans (Transkei, Bophuthatswana, Venda and Ciskei) ▪ Magisterial districts ▪ Municipalities ▪ Suburbs ▪ Tribal authorities ▪ Administrative areas ▪ Wards	▪ Provinces ▪ Magisterial districts ▪ Transitional district councils o Regional local councils o Transitional metropolitan councils ▪ Municipalities o Metropolitan local councils o Metropolitan substructures o Transitional local councils o Transitional rural councils ▪ Place names o Cities o Towns o Suburbs o Tribal authorities o Administrative areas o Wards o Villages ▪ Enumeration areas	▪ Provinces ▪ Metropolitan areas ▪ District councils o Local municipalities o District management areas ▪ Main places ▪ Sub-places ▪ Enumeration areas
	Census '96 hierarchy of metropolitan areas or services councils	**Other geography boundaries**
	o Metropolitan councils o Municipal substructures o Town suburbs	o Magisterial districts o Wards

4.11 Linkage of census geography to non-census outputs

Martin (2000) noted two major requirements for a census geography system: it should facilitate the organisation and management of the census itself and it should provide an appropriate framework for the publication of small-area census statistics that meets users' needs. Due to new administrative boundaries and changes to enumeration areas during demarcation for Census 2001, these requirements led to the decision to create separate geographies for the data collection and output of Census 2001. Figure 6 shows output geography used in the publication and dissemination of census data.

Figure 6: Census 2001 geography: dissemination routes in SuperCross

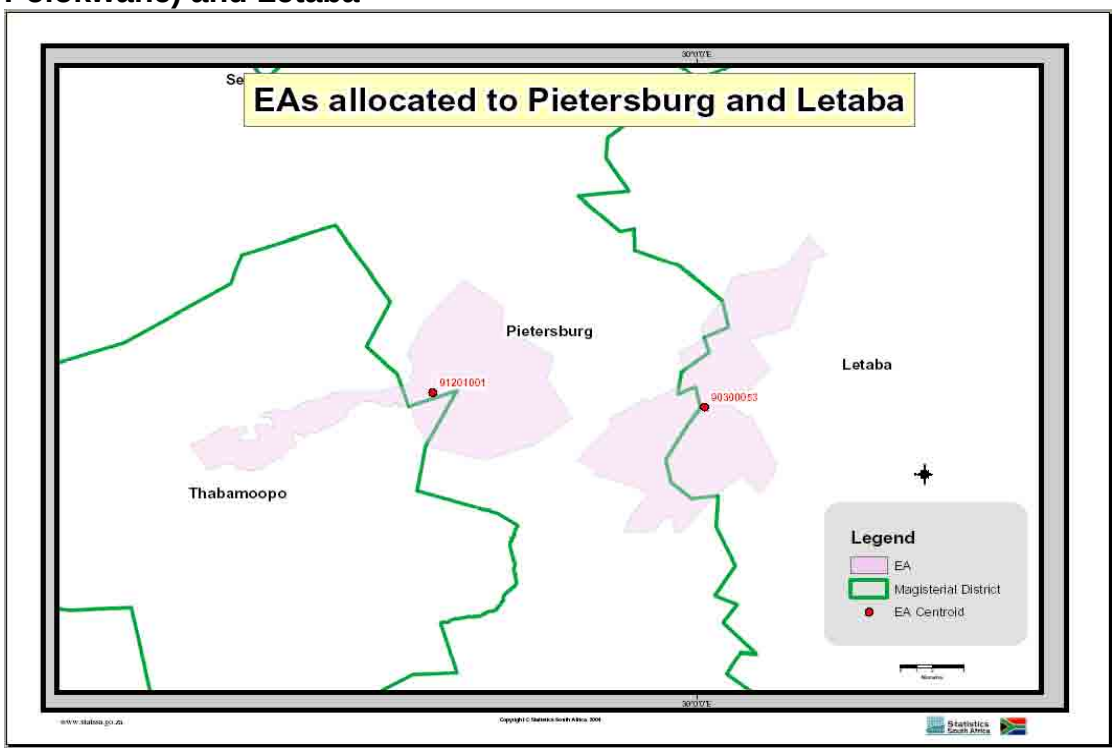


Magisterial districts and electoral wards feature in the output geography but are not accounted for in the planning and management frame.

A magisterial district is a judicial area under the jurisdiction of a magistrate. Although the demarcation of enumeration areas for Census 2001 was based on the new municipal structure, data had to be released by magisterial district too as several government departments still use this unit. Figure 7 shows how the point-in-polygon method was used to assign those enumeration areas, that were split by the borders of magisterial districts, to a specific magisterial district. This situation arose as the new municipal boundaries to which the enumeration areas had to conform were not aligned to the magisterial districts. The link that was thus established between enumeration areas and magisterial districts enabled the creation of the two dissemination hierarchies in SuperCross based on magisterial district. (Refer to Figure 7).

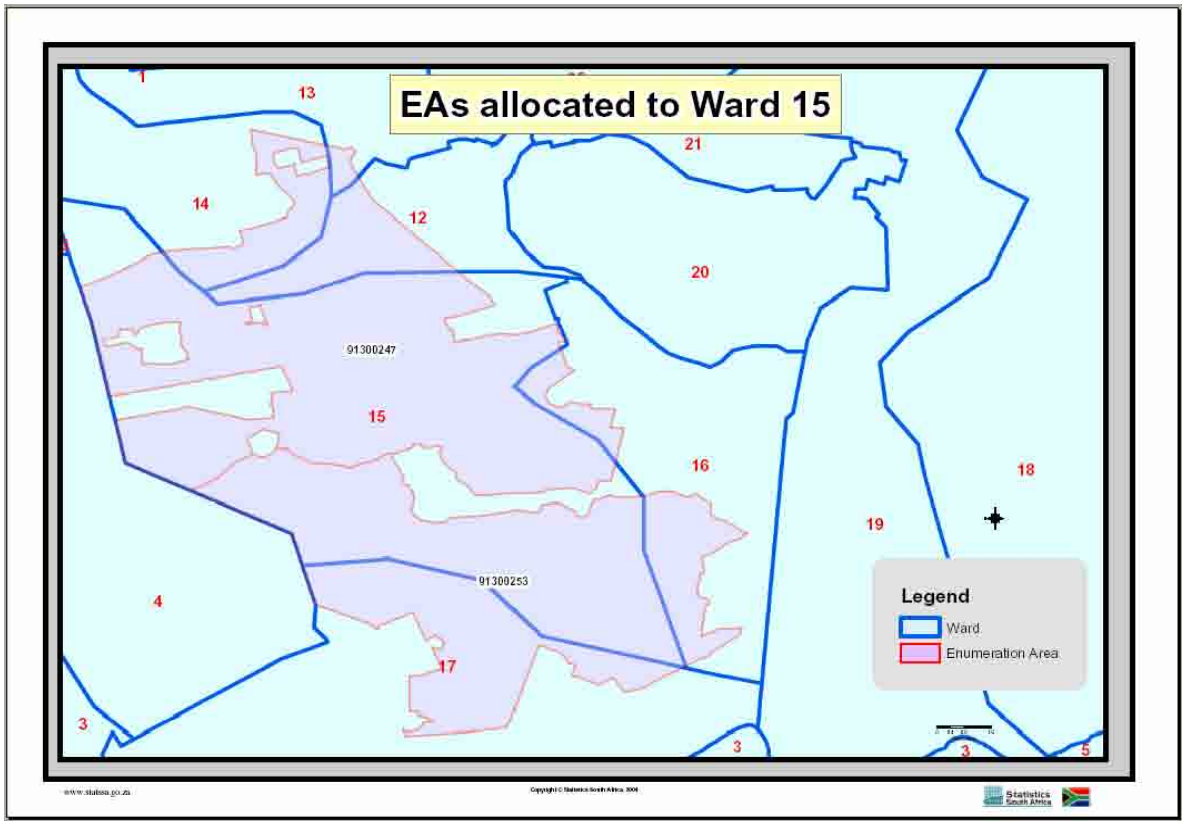
Also see *Report 03-02-25* (metadata document) on the Statistics SA website]. Note that the magisterial districts were not linked to the new place name set that was used in Census 2001.

Figure 7: Enumeration areas allocated to Pietersburg (now known as Polokwane) and Letaba



Enumeration areas were linked to electoral wards by assigning them to the ward with which they shared the most area. Settlement locations within enumeration areas were manually checked by means of aerial photographs and, where appropriate, enumeration areas were allocated according to population concentrations (Refer to Figure 8.). A ward was also linked to the municipality to which it belonged. A ward that overlapped more than one municipality was assigned to that municipality in which most of its area fell. This process ensured that each ward belonged to one municipality, enabling the creation of the right-most dissemination route indicated in Figure 8.

Figure 8: Enumeration areas allocated to Ward 15 in Lebowakgomo



4.12 Standard geography outputs for Census 2001

The geographical products that can be found in digital format on Statistics South Africa's website are listed below:

- hierarchies and metadata/geography concepts
- atlases (hardcopy/digital)
- digital boundary CD
- ward atlas (My Constituency)
- urban/rural reports and discussions
- web products
- about place names
- expanding GIS technologies into sampling
- enhancing map production (automation tools)

4.13 Conclusion

Reference to Coombes' criteria shows that Statistics South Africa has in the past years made extensive progress toward the establishment of appropriate and flexible geography for census taking. Further innovations that are currently being planned are described in Chapter 5.

Bibliography

Census Mapping Manual v1.5

Collaborative Maintenance Document by Stats SA, DLA, MDB, IEC

Coombes, M. (1995). *Census User's Handbook,* edited by Stan Openshow, Bell and Brain Glasgow

Geographical Hierarchies and Attributes (2003)

GIS Closure Report (January 2002)
GIS Database Design Document
GIS Workflow for 2001 Census

Investigations into Appropriate Definitions for Urban and Rural Areas in South Africa (2003)
Martin, David (1997a). *From Enumeration Districts to Output Areas: Experiments in the Automated Creation of a Census Output Geography, Population Trends 88: 36-42.*

Martin, David (2002a). *Geography for the 2001 Census in England and Wales, Population Trends 108: 7-15.* Also available at
http://www.statistics.gov.uk/downloads/theme_population/PT108.pdf

Statistics South Africa (2003). *Investigation into the Density Criteria as Part of Defining Urban and Rural Areas of South Africa*

Stats News (December 2002). *Expanding the Extent of the Geography Division*

Stats News (June 2004). *Understanding and Accessing Geographies Through Intuitive Spatial Products*

The Count and How it Was Done (1996)

The Count and How it Was Done (2001)

Chapter 5: Mapping the 2001 Census

5.1 Introduction

The mapping of census information is a relatively new phenomenon in South Africa because a national geographical information systems (GIS) database for the entire country has been available only since the release of the 1996 census GIS database information in 1999. It was only then that users throughout the country were exposed for the first time to a comprehensive database that could be mapped at many different spatial levels. The enumerator area (EA) is the spatial unit at which socioeconomic information is collected during a census. It is also the smallest spatial area at which census variables can be mapped. The small area level (SAL) is the smallest spatial unit at which the 2001 census data can currently be mapped.

Numerous challenges face users of census data, one of which is developing a thorough understanding of the many census variables present in the 2001 data and using these variables appropriately in their planning and decision making. Another major challenge for users of census data is developing the skill to map the census variables so that they make sense. Users in South Africa should be encouraged to use census data in a mapped format so that they can become familiar with the data and the use of the data in a spatial format. Users should endeavour to understand the meaning of each map produced at the different spatial levels, especially in respect of the location of the high and low concentrations of any census variables mapped, the possible meaning of this distribution, the existence of any unique features of the illustrated data and the shown relationships between areas.

This chapter focuses on the different ways census information can be displayed so that users can optimise the use of such information. The chapter will also discuss new techniques for mapping the 2001 census information.

5.2 Mapping census information

Census data are usually mapped on a computer screen or on a paper map. In creating a map, several questions are often asked:

- What problem needs to be addressed using the census data, and can an indicator (e.g. population density, sex ratio, per capita income) be used?
- What census variables are required to provide relevant information on the problem being investigated or to populate the indicator, and in what format

should the variables be represented (e.g. raw numbers, proportions, percentages, shares[21])?

- At what spatial level should the census variable be mapped?
- What is the purpose of the map (e.g. scientific paper, report publication, poster, book, graphic picture)?
- What type of map would best display the particular census variables (e.g. thematic map, dot density map, three-dimensional surface map, contour map)?
- What standard map features should be used (e.g. title, legends, scale bar, source, north arrow)?

Selecting census variables

The selection of a census variable is not always a simple task because users often have limited knowledge of all the questions asked in the census and are not always familiar with the way the data associated with a variable should be modified to address a particular problem or to populate an indicator. Census information specialists and researchers who use census information frequently in specific research areas are generally the only people who properly understand what variables to use and how best to use them, especially in the form of maps. Users can develop such knowledge by reading up on the subject and by eliciting the help of organisations such as Statistics South Africa (Stats SA) and the Human Sciences Research Council (HSRC) to gain access to the right census variables.

Access to census information in desktop GIS has greatly facilitated variable selection and map creation. Care should, however, be taken as “mapping can be extremely informative but it can also mislead” (Rhind 1983). Consequently, this handbook is being published with various chapters written by experienced researchers who discuss what census variables can be used in certain development planning areas and how the variables can be applied.

Format of the variable

Raw numbers should be used to map a census variable at the different spatial levels (e.g. enumerator area, ward, place, municipality and province). However, the resultant map may not facilitate a proper understanding of the distribution of the variable or of the spatial relationship between different geographic areas. An example would be the mapping of the population that speak Afrikaans as their first language at a municipality level. This mapping would show that the populations with the highest number of Afrikaans speakers live in the metropolitan areas of Johannesburg, Pretoria, Cape Town and Port Elizabeth.

[21] A share index generally examines the value of a particular variable in a specific geographic area in relation to the national average. It looks at how much an area “shares” in the national picture of a particular census variable.

An examination of the proportion of the population that speak Afrikaans as their first language in relation to the total population would show that more than 80 per cent of the population in most districts in the Western and Northern Cape speak Afrikaans. An examination of the ratio of the number of Afrikaans-speaking people to the national average of speakers of the rest of South Africa's official languages would show that Afrikaans is dominant in districts in the Western and Northern Cape, northern Gauteng and parts of Limpopo. The format of the variable used to create a map, therefore, has a marked effect on the distribution of the variable and the spatial relationships illustrated. Users will have to decide, on the basis of experience, what variable format best illustrates the reality in society.

What spatial level to use

Census information can be provided at various spatial levels in South Africa from the smallest unit, which is the EA, to the largest, which is the whole of South Africa. What spatial level to use is often determined by who the end user of the map is, the purpose of the map and the part of the country that has to be shown. If the total population of the whole of South Africa had to be mapped, any spatial level below municipality would be inappropriate as the spatial units would be so small that information on metropolitan areas would not be visible and the information would be lost. To show a more detailed breakdown of the total population in certain areas, map insets can be used where the spatial unit is smaller than a municipality.

EAs, wards, places and suburb boundaries could be used to illustrate census variables for municipalities. As the user begins to focus on smaller areas of interest, the size of the spatial unit to use in mapping census variables may also become smaller. Again, this would depend on the purpose of the map and for whom the map is intended. Planners and decision makers focusing on national and provincial issues would generally require census information primarily at the provincial and municipal levels.

Although provincial governments and municipalities might want to compare information on their areas with that on neighbouring areas, their focus would usually be within the provinces and district councils. Consequently, they would want the census information to be presented at spatial levels that are small enough to enable them to see the variations within their areas. Metropolitan areas and municipalities would, in most cases, require census information at a ward, place or EA level to enable users to understand the variations in demographic patterns within their areas.

5.3 Purpose of the map

Maps are produced to illustrate information and can be presented in several different published formats such as reports, wall posters, books, hardcopy maps and published

maps. Where the map will be used will determine the spatial level, the layout, the colours and text fonts, the size and quality of the paper, whether the map is plotted from a GIS or printed by a publisher, and the degree of effort that should be put into making the map graphically pleasing.

Some important principles in producing a map:

- The correct variable must be used, and the illustrated information must be accurate.
- The map must portray the inherent patterns associated with the particular census variable.
- All supplementary information must be provided to ensure the correct interpretation of the map.
- The map must be kept as simple as possible to facilitate its readability.

Maps for clients and for inclusion in reports should be designed in such a way that the reader can easily understand what is being represented on the map. As the reader is in close proximity to the map when viewing it, thematic colours can be more graded and single toned (e.g. light to dark red to show differences in the size of the population) while text fonts can be smaller but generally not less than ten points. Such maps tend to be fairly simple in their layout. Because wall posters are often viewed across a room and are intended to attract people, colours on the posters tend to be much bolder and multicoloured. The text font of the title informs the reader about the topic of the poster and is, consequently, usually quite large and bold. Once the reader is drawn closer to the poster, additional information about the map, such as the legend and source of the information, can then be seen.

Where maps are published in books or as high quality posters, the information is usually transferred from GIS to publishing or graphical software (e.g. Aldus Freehand or CorelDraw). Fortunately, most GIS packages can transfer geographical and textual information to these publishing packages in standardised formats such as JPEG, TIFF and PDF. Once transferred, the full functionality of graphical packages can be used to create high quality maps. However, this process can take some time and can be expensive. A user should, therefore, do a cost-benefit analysis before following this route.

5.4 Types of maps

The presentation of spatial information in maps has changed in recent times. Initially, with the advent of desktop GIS, spatial information was presented in the form of thematic or choropleth maps. With advances in GIS technology and theory, new methods of representing spatial information have been developed. These methods have often incorporated information from sources other than censuses. For example,

three-dimensional models and isolines have been used to illustrate land terrain and temperature gradients respectively. Research on techniques for representing census information is ongoing worldwide. The following section explores some of the methods that are being used to display spatial information in GIS.

Area thematic maps

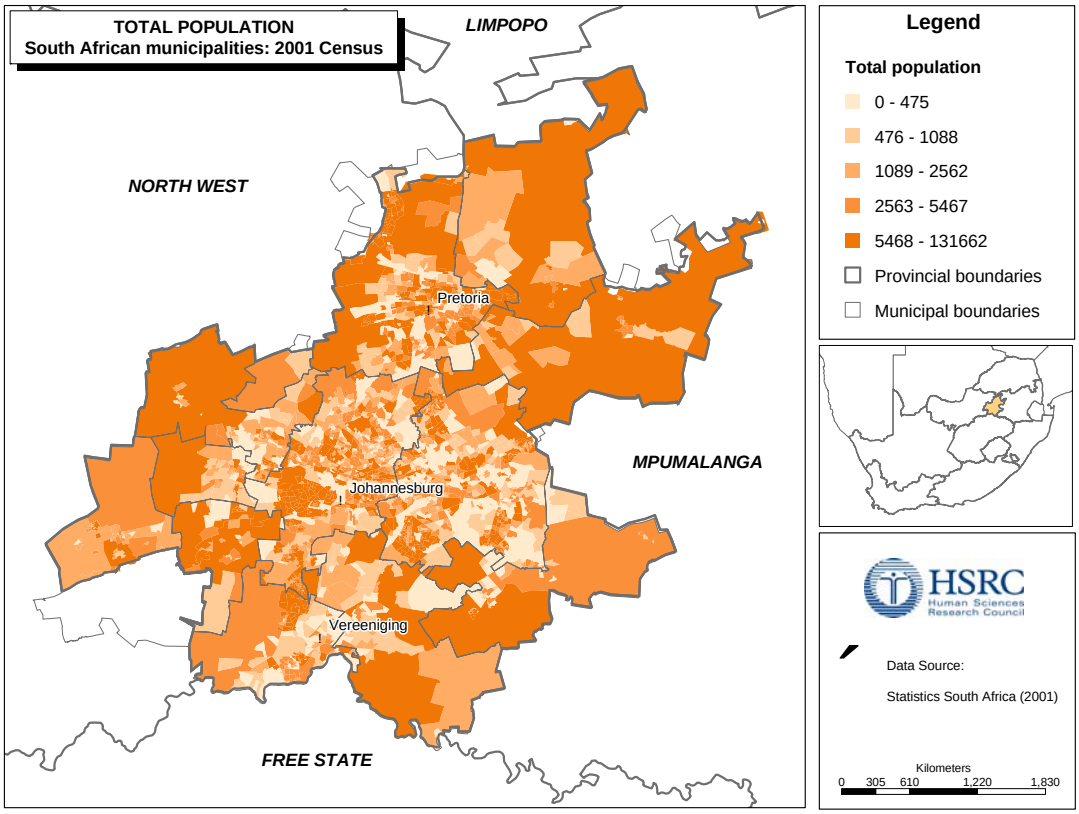
The most common way to present census information in GIS is to use area thematic maps. Commercially available desktop GIS software can design and produce high quality area thematic maps. The mapping of census variables associated with the different spatial levels by shading polygons is also known as choropleth mapping. Although much criticism has been raised against the use of choropleth maps in illustrating census variables because of the use of "artificial" boundaries (i.e. EA boundaries) to demarcate the location of the population, these maps can help users in South Africa to develop the skill of mapping census variables and gain a preliminary understanding of why particular patterns in the population exist.

Desktop GIS software offers several options for dividing the census data into classes represented by colours and patterns on the area thematic map. Care should be taken when deciding on which classification method to use as many of the census variables do not have a normal distribution.

Common classification methods:

- Quantiles: divides the number of spatial features (e.g. EAs) equally among the number of chosen classes (Figure 1).
- Equal size: creates classes of the same range (e.g. 0–100, 101–200, 201–300).
- Natural breaks: identifies distinct breaks in the distribution of the data in order to define classes.
- Standard deviation: creates classes of one standard deviation in size around the mean.

Figure 1: Total population of South Africa at a municipality level, classified according to quantities



Some of the complexities associated with area thematic maps are that the impression is created that the population exists in distinct homogeneous groups and that neighbouring groups may vary markedly from each other. In reality, such distinct groups are not a fair reflection of the population as people with different socioeconomic characteristics may be living in the groups. Similarly, people in different groups may have similar socioeconomic characteristics. Defining people in groups and drawing inferences about the individuals/households is known as the *ecological fallacy*.

A classic example of this fallacy is the mapping of poverty at a municipality level. Such mapping may suggest that in metropolitan areas such as Johannesburg less poverty is encountered than in the more rural municipalities of South Africa. Users should be aware of this problem when using area thematic mapping. The problem can be overcome by creating spatial areas that not only reflect the reality of community groupings in society but also cater for the changes that occur in population characteristics. This problem is known as the *modifiable areal unit problem*. A new technique to minimise the problem is the use of continuous gradients across spatial boundaries.

Bivariate thematic maps

Bivariate mapping allows the mapping of two variables for each spatial feature. The first variable mapped is usually a colour area thematic map and the second variable mapped is usually a patterned area thematic map or a symbol map. A practical example of bivariate mapping would be first to map the 0 to 2-year-old population and then to map either the per capita income or the number of households. The value of such a bivariate map would be the ability to determine where high concentrations of 0 to 2 year-olds reside and to show populations with high incomes or many households so that optimal markets for commercial infant products can be identified.

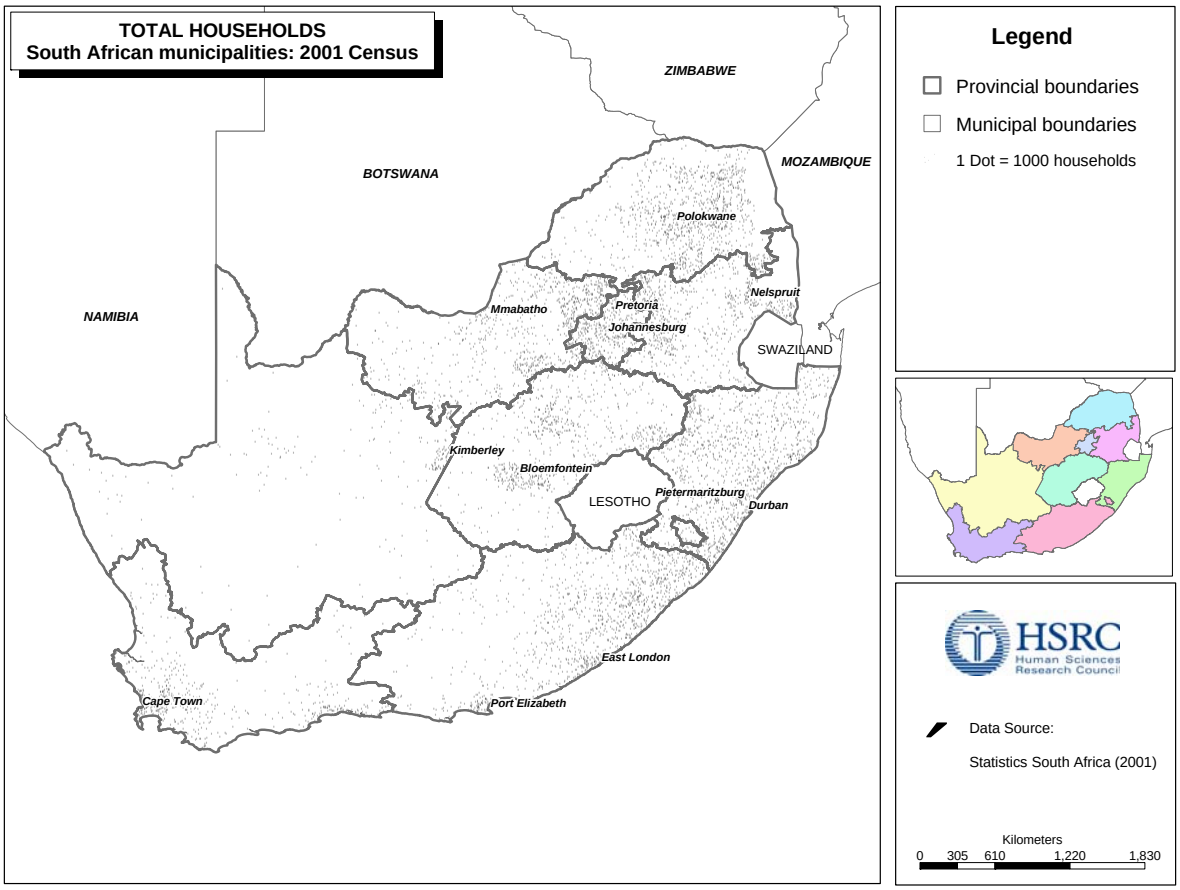
Dot density maps

Another method of representing census data on a map is dot density mapping. In dot density mapping, a dot representing a certain number of people (e.g. 1 dot = 1 000 people) is randomly located within the spatial area (e.g. municipality) to which the census data are linked. For example, if an area has 100 000 people, 100 regularly spaced dots will be seen in the area representing that population. This method is the most appropriate for showing the distribution of the population (Figure 2) and economic variables such as gross geographic product. It is also useful when mapping two different variables at a time, the one variable (e.g. total population) being thematically mapped and the other (e.g. total males) being dot density mapped. This method is very useful for determining whether any spatial correlation exists between areas with a large population and a large number of males.

Point thematic maps

A method of population mapping that has gained some favour over choropleth mapping in recent times is the use of points positioned where the highest concentration of people live within an area (the epicentre of the population). The centroid of an EA would in many instances, especially in urban areas, be sufficient because the demarcation process used during the 2001 census attempted to group equally distributed and homogeneous populations together. Even in rural areas, the centroid of an EA would probably be the epicentre of the irregularly spaced dwellings located in that geographic area. However, in some cases, the centroid of the EA would have to be geographically shifted to where the concentration of the population is in that EA.

Figure 2: An example of a dot density map showing Gross Geographic Product in South Africa



In dasymetric mapping, the population concentrations are represented by points in the form of symbols on the map. To show the different classes of a census variable, symbols of different types, sizes or colours can be used. The size of the symbol can be defined proportionally to the values of the census variables being mapped (Figure 3).

Figure 3: An example of a map showing dasymetric mapping

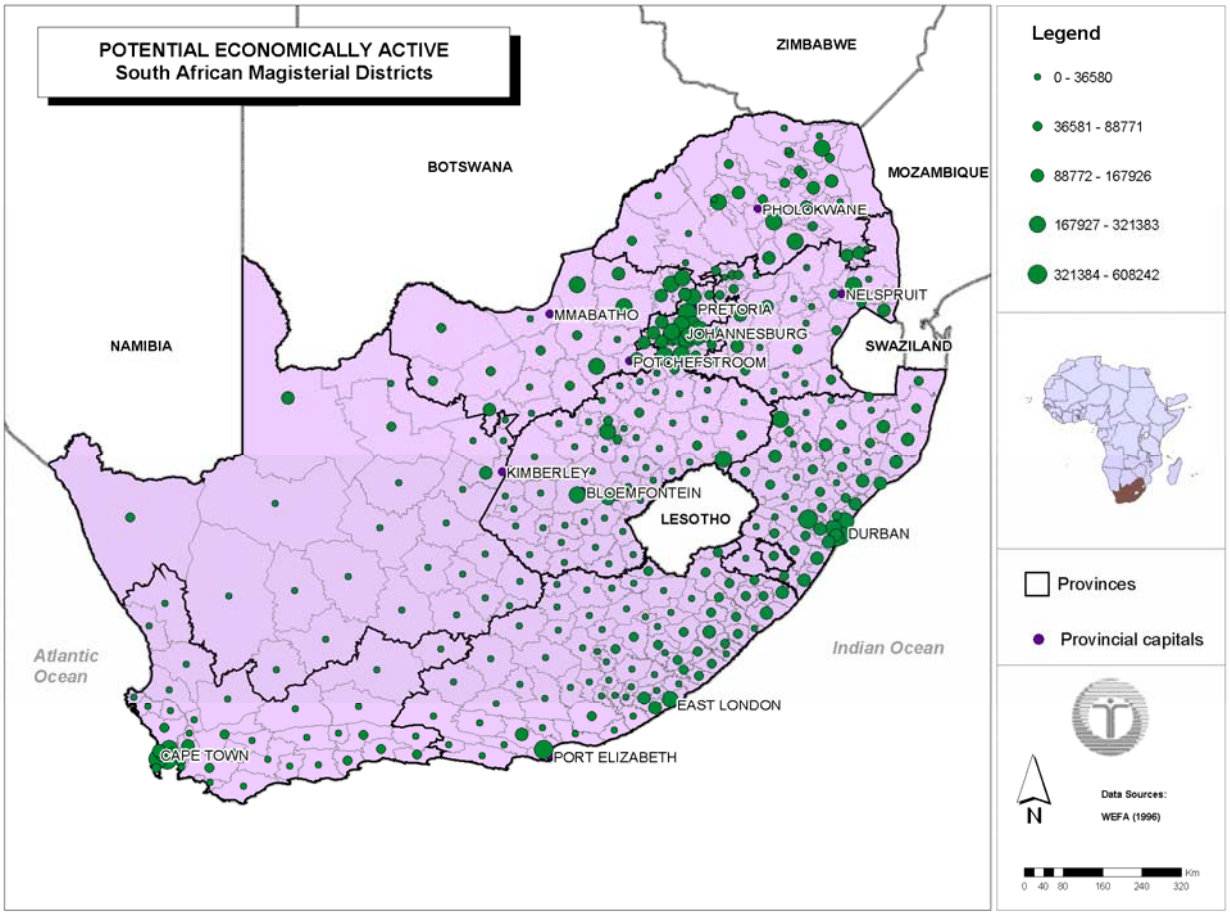
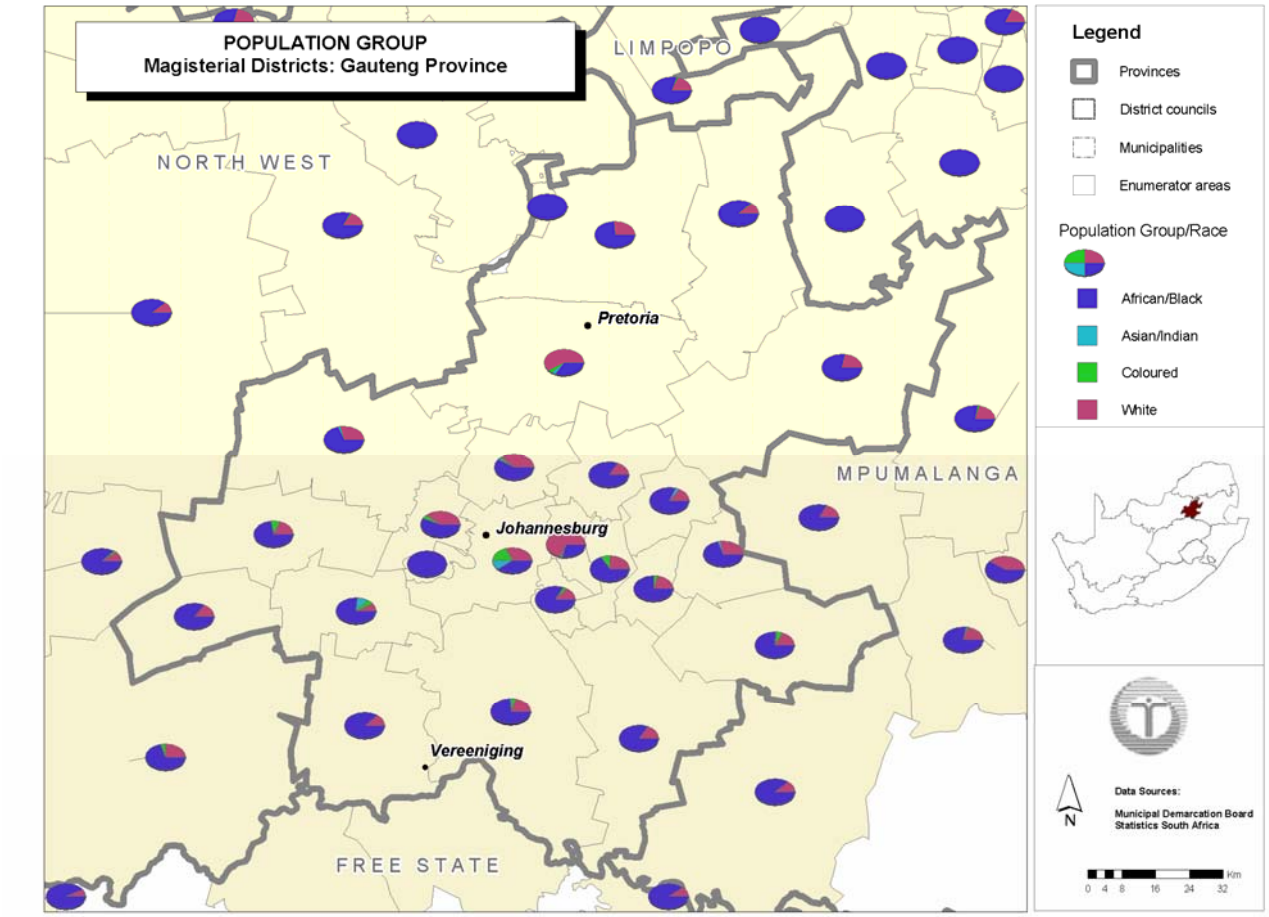


Chart thematic maps

The use of pie and bar charts to represent census variables associated with spatial areas such as EAs has gained popularity in recent years, especially as such charts can be shown with commercial desktop GIS. The value of these charts is that the mapping of many census variables can be done for each spatial area. Thus, pie or bar charts could be used to show the racial composition of the population at a particular spatial level (Figure 4). However, GIS specialists believe that most census users will find chart thematic maps difficult to interpret and that too much information is presented at one time, which means that the value of the census information may be lost. Another disadvantage of chart thematic maps is that the bars or pieces of the pie have no values associated with them. If charts have to be used, bar charts should be used before pie charts.

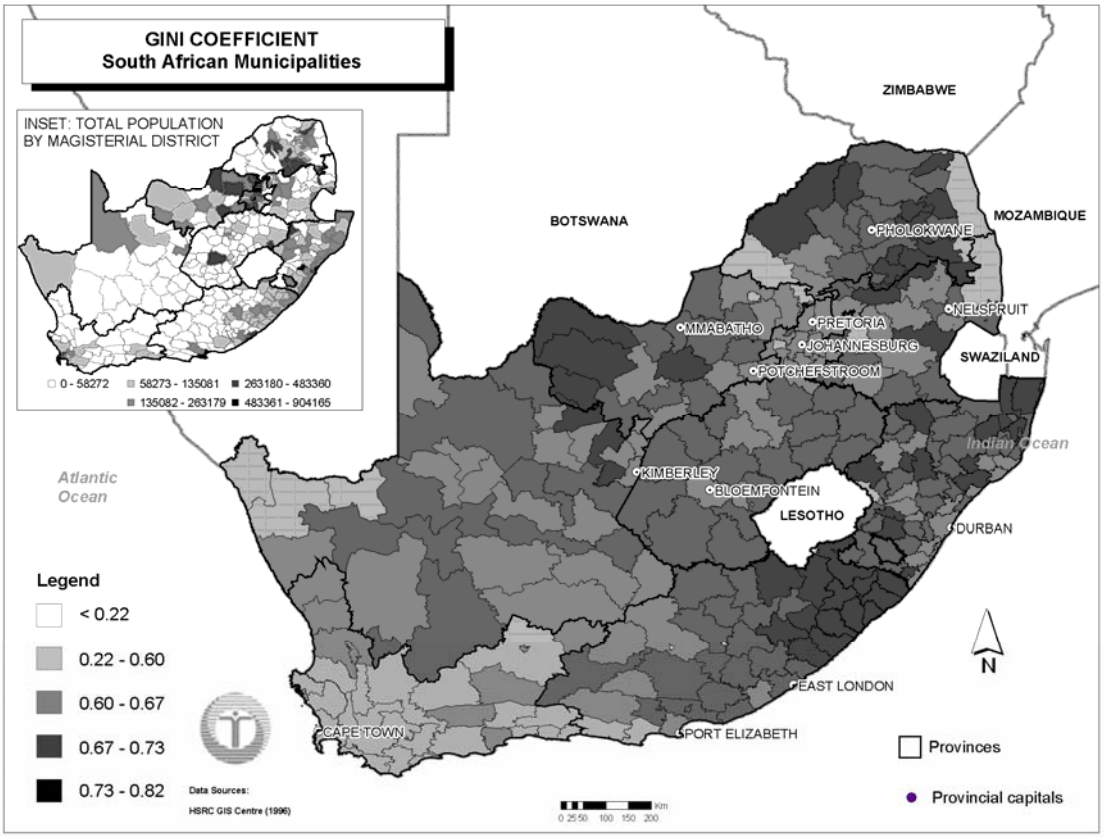
Figure 4: An example of how pie and bar charts can be used to show information



Map insets

Maps insets can be used to represent several different census variables (Figure 5) or to show a particular census variable at different spatial levels. The different maps that are shown on the same screen or piece of paper can then be compared. This technique is considered by some to be better than techniques such as bivariate maps or chart thematic maps because users are not flooded with too much information that is difficult to comprehend and assimilate. The user can concentrate on one inset at a time and see how the values associated with specific spatial features relate to each other. Comparisons can then be made between the maps and individual spatial features. However, map insets allow only the simplest of correlations between census variables in the different insets.

Figure 5: Example of an inset in a map



5.5 Modified spatial boundaries

The concern that delineated census boundaries do not reflect the true demographic characteristics of a population has led to the modification of spatial boundaries in an attempt to represent the reality of a population in society more accurately. Census boundaries can be modified in several different ways. The simplest way is to aggregate the boundaries (i.e. EAs) to create a new set of spatial boundaries (e.g. suburbs) based on attributes associated with each census area. Clustering techniques can also be used to group the census areas into new spatial features based on a set of census attributes. An example would be the aggregation of census areas to create small marketing areas (SMAs) based on a set of predefined census variables. The voting districts created by the Independent Electoral Commission (IEC) are a form of modified spatial boundaries. Essentially, the IEC combined EAs according to the size of their populations. Through consultation, the voting districts were modified further to take into consideration administrative boundaries, access to suitable polling stations and so on. Techniques such as artificial neural networks (ANNs) are also being applied to census data to create new groupings of the census boundaries. This topic will be further discussed in the section dealing with the development of geodemographic databases.

Thiessen polygons or Voronoi regions

Thiessen polygons or Voronoi regions are basically catchment areas that are created around a set of point features and used to model continuous surfaces. Each polygon or region is centered on a point, and any position within the polygon or region is closer to its central point than any of the neighbouring points. According to Dorling (1995), Thiessen polygons or Voronoi regions create a more "continuous" picture of the population and are one step closer to mapping the true social landscape of an area. Thiessen polygons or Voronoi regions essentially combine dasymetric mapping and modified spatial boundaries.

The centroid of an EA is first created and then shifted geographically to represent the concentration point of the population in that area. Thiessen polygons or Voronoi regions are then created around these shifted centroids or points to create new or modified spatial boundaries. The census attributes associated with each EA or its centroid are then attributed to each Thiessen polygon or Voronoi region. Figure 6 shows Thiessen polygons or Voronoi regions created from EA centroids in Bloemfontein.

Surface modelling

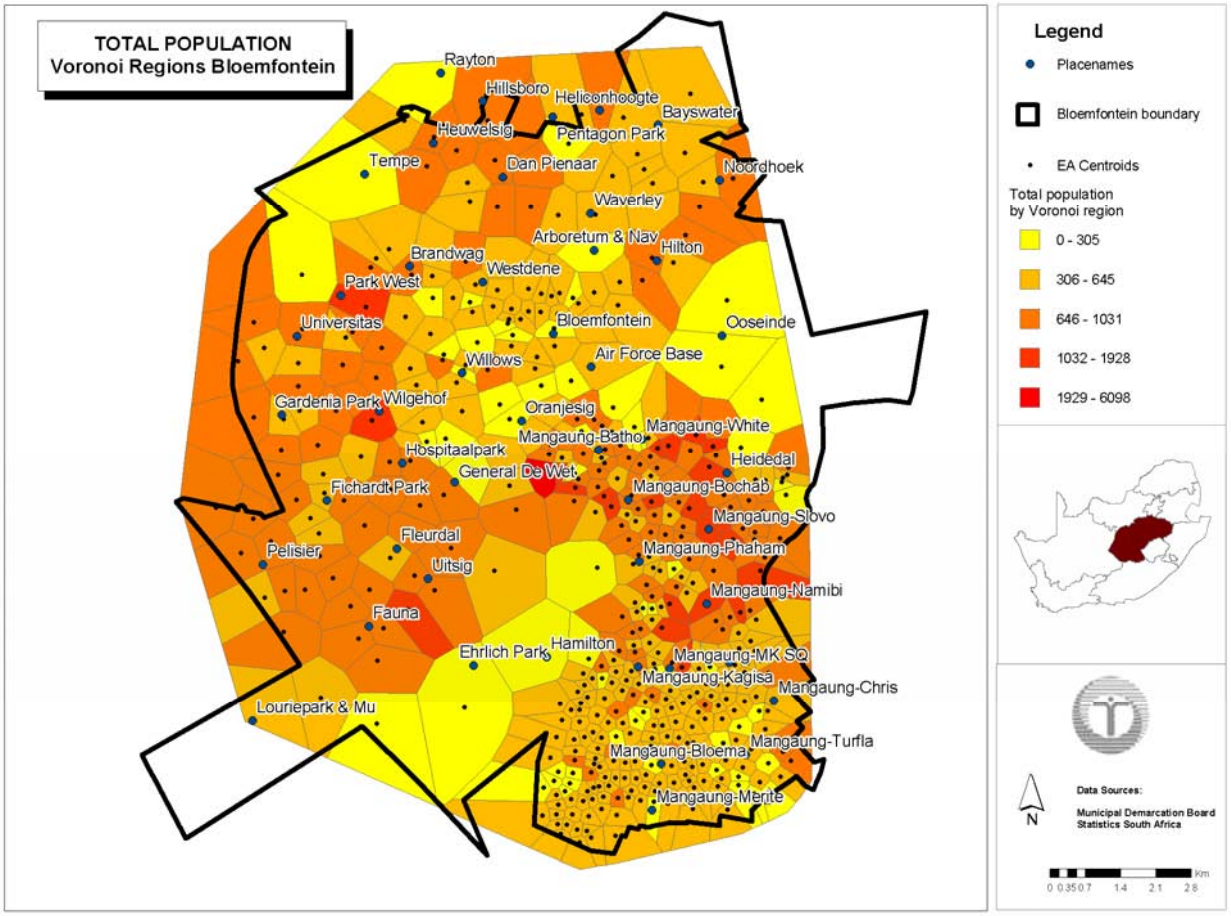
Several different surface modelling techniques can be applied to represent the "continuous" flow of the demographic characteristics of a population in a two-dimensional or three-dimensional space. Surface modelling applies different interpolation[22] algorithms to spatial features (e.g. points, lines, polygons, rasters) with particular values to create a new spatial layer. The centroid points created from EAs and shifted to represent the concentration of the population and the census variables associated with each point form the baseline information needed for the surface modelling.

The following types of surface modelling techniques can be used:

- Surface fitting: uses an attribute associated with a point feature to create a three-dimensional elevation surface. Various methods can be used to interpolate the data between the vector points (e.g. kriging);
- Profiling;
- Triangulation (e.g. Delauney);
- Contouring (e.g. linear interpolation);
- Digital elevation models;
- Three-dimensional models.

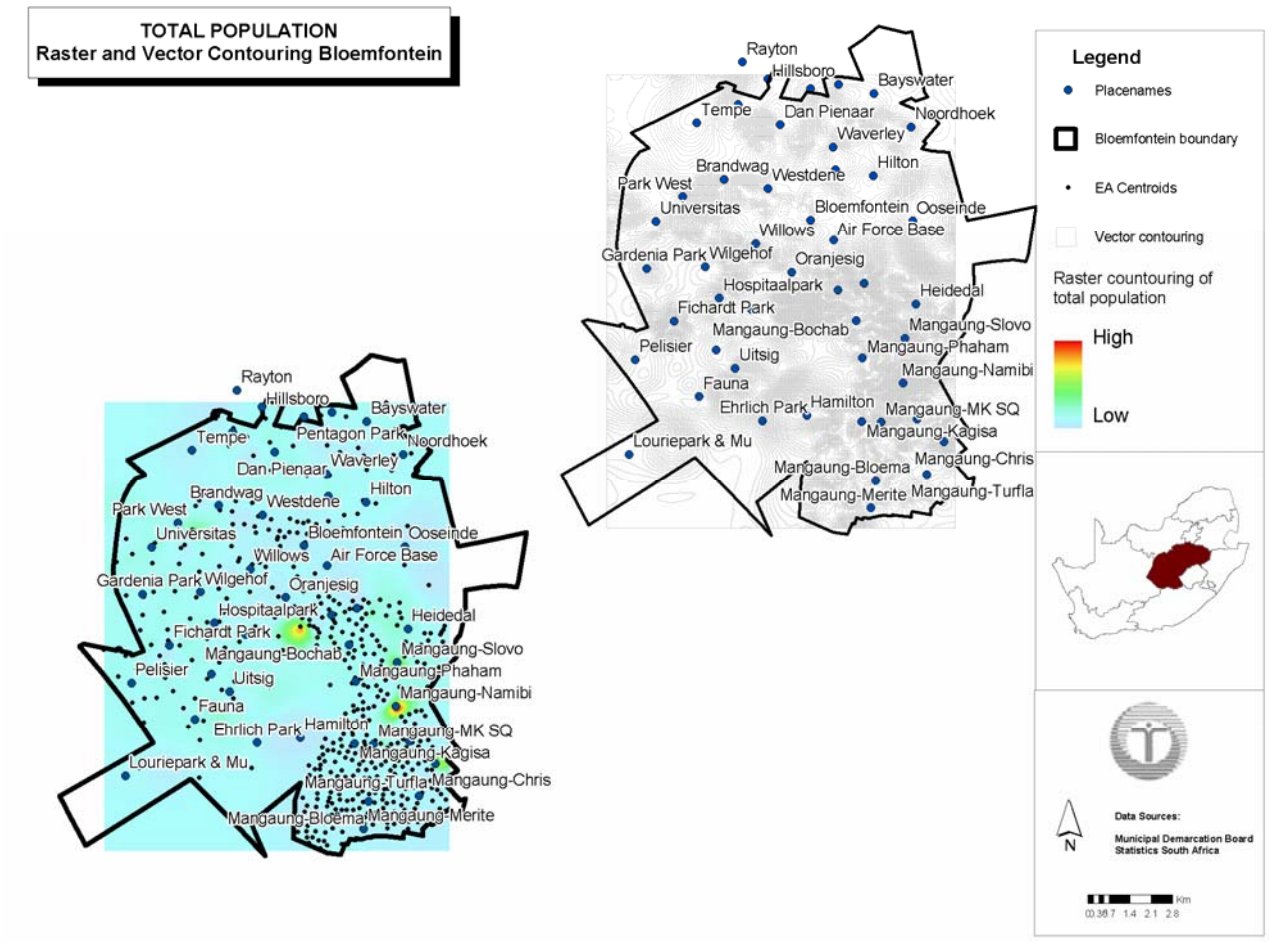
[22] Interpolation is a mathematical technique used to determine the approximate values between given points that have a particular value (e.g. total population) and distance apart.

Figure 6: An example of a Voronoi regions map showing EA centroids in Bloemfontein



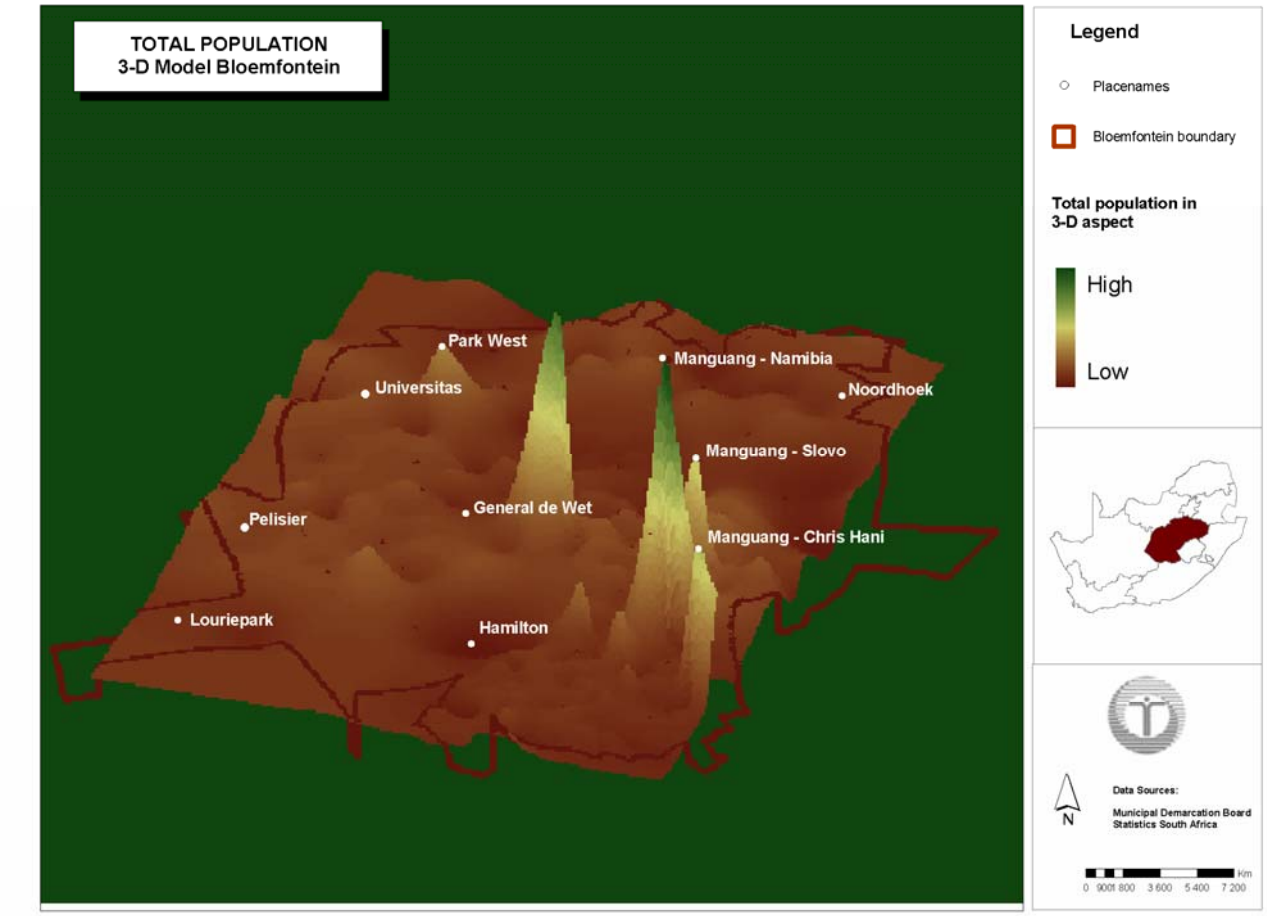
Surface modelling provides an alternative method for mapping, analysing and interpolating census information. However, the surface modelling of census variables tends to create maps that are difficult to interpret mainly because most users do not think in three dimensions, generally have poor depth perception and have a short visual memory. Denham and Rhind (1983) suggest that users should “avoid contour mapping of census data” and go so far as to state that no surface modelling techniques should be used with census data. Nevertheless, these techniques still offer an important avenue for more sophisticated analysis and interpretation of census information for researchers in the field of GIS and demography. Surface modelling is also a method that can be used to fill in gaps in data.

Figure 7: An example of a contour map showing raster and vector contouring of the population of Bloemfontein



The contour map of Bloemfontein shows the interpolated values for the total population, which were generated from EA centroids (Figure 7). The contouring identifies the extent of zones of population size that could influence planning in the city. The high concentration areas are those with numerous flats near the city centre and the townships on the southern outskirts. Zones of low population are identified around the industrial areas of eastern Bloemfontein. This type of contour modelling of the population can be used effectively to examine changes in the distribution of the population from one census year to another (e.g. 1991–1996), especially where the census boundaries do not match completely, as is the case in South Africa.

Figure 8: An example of a three-dimensional model showing the relative size of the population in different areas and the distance between such areas



The displaying of census variables in a three-dimensional model has many advantages, for example, being able to see the extent of the variation in the size of the population from one area to another as well as being able to see the distance between such areas (Figure 8). The draping of a colour ortho-photograph over the three-dimensional image of a census variable also allows for additional interpretation of the census information in terms of what the main land cover is, or land uses are, in the different areas. The ability to rotate and tilt a three-dimensional image means that census information can be viewed from different perspectives. The more modern GIS software allows one to “fly-through” the three-dimensional model, which gives a virtual reality feel to viewing the census information.

Raster and grid maps

Census information can also be displayed in the form of raster maps. Usually, census data are stored in vector GIS formats and, then, to create raster maps, the census boundaries with their associated census data have to be converted to a raster through the process of rasterisation. Within a raster GIS, raster maps can subsequently be

analysed using complex techniques for interpreting the census information. Raster GIS is ideal for undertaking detailed modelling of many different spatial layers of information and for understanding the correlations between features in these layers. A good example would be the construction of land use suitability maps from the integration of many different layers of information including census data such as population density.

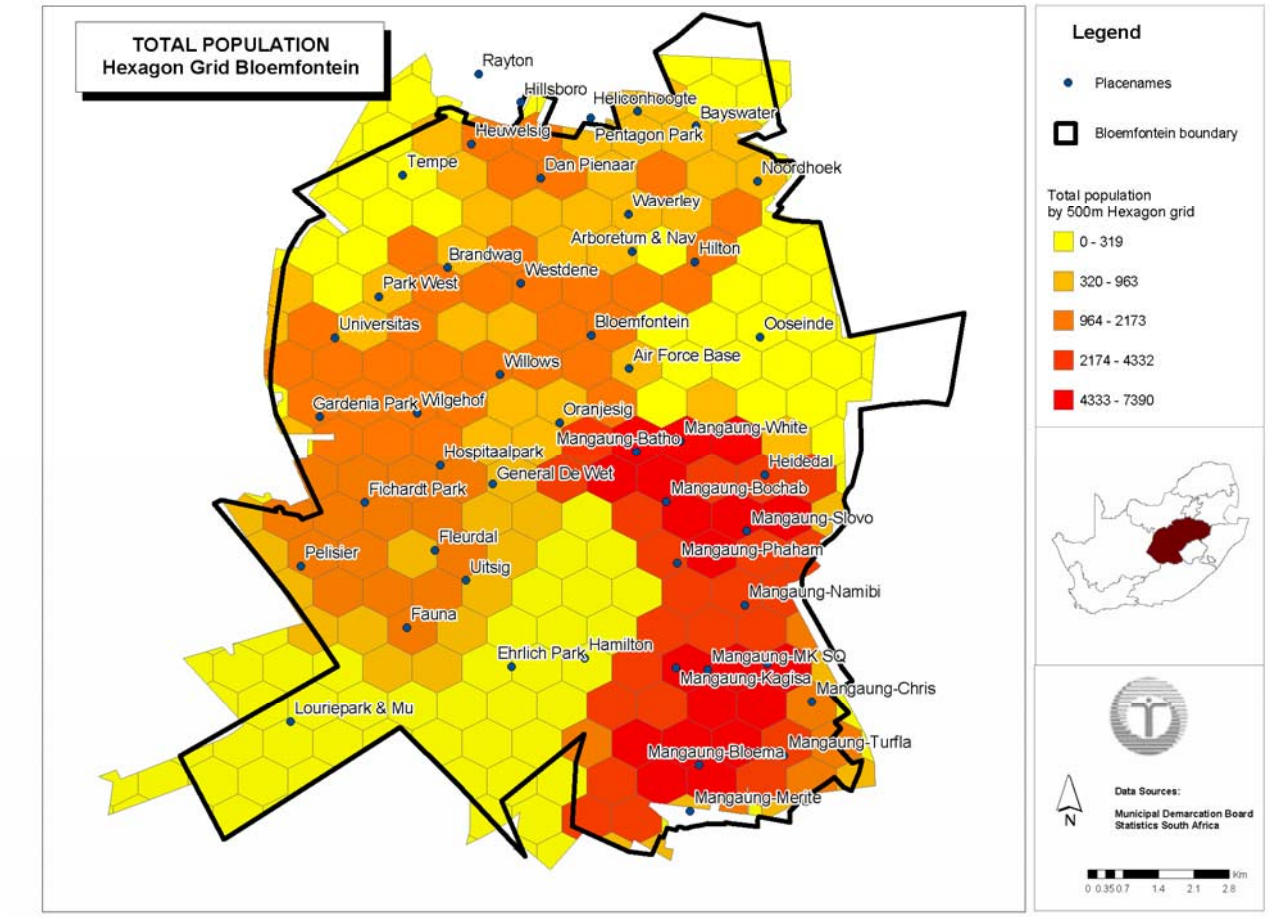
The use of grids, whether they are shaped as hexagons, squares, rectangles or diamonds, is a technique that has gained great popularity in recent years. The value of attributing census data to grids with a particular area, circumference, width and height is that sophisticated modelling of the information can be accomplished. For example, one could create grids of a similar size to the footprint of a cellular antenna and attribute total population figures from the census EA boundaries to these grids (Figure 9). This technique allows the identification of highly populated areas and the calculation of the distance between population concentrations.

Using the grids, modelling such as gravity flows, shortest path analysis, site location analysis and catchment area analysis can be undertaken with a number of variables from the census. The Council of Scientific and Industrial Research (CSIR) in conjunction with the University of Utrecht in The Netherlands, has developed a product called AccessMap that allows these types of modelling and other accessibility analyses to be done with desktop GIS software. Because of the limited financial resources in South Africa for the provision of basic services such as schools and clinics, the use of census variables with grid modelling must surely become an integral part of planning in the country.

Geodemographic or lifestyle segmentation mapping

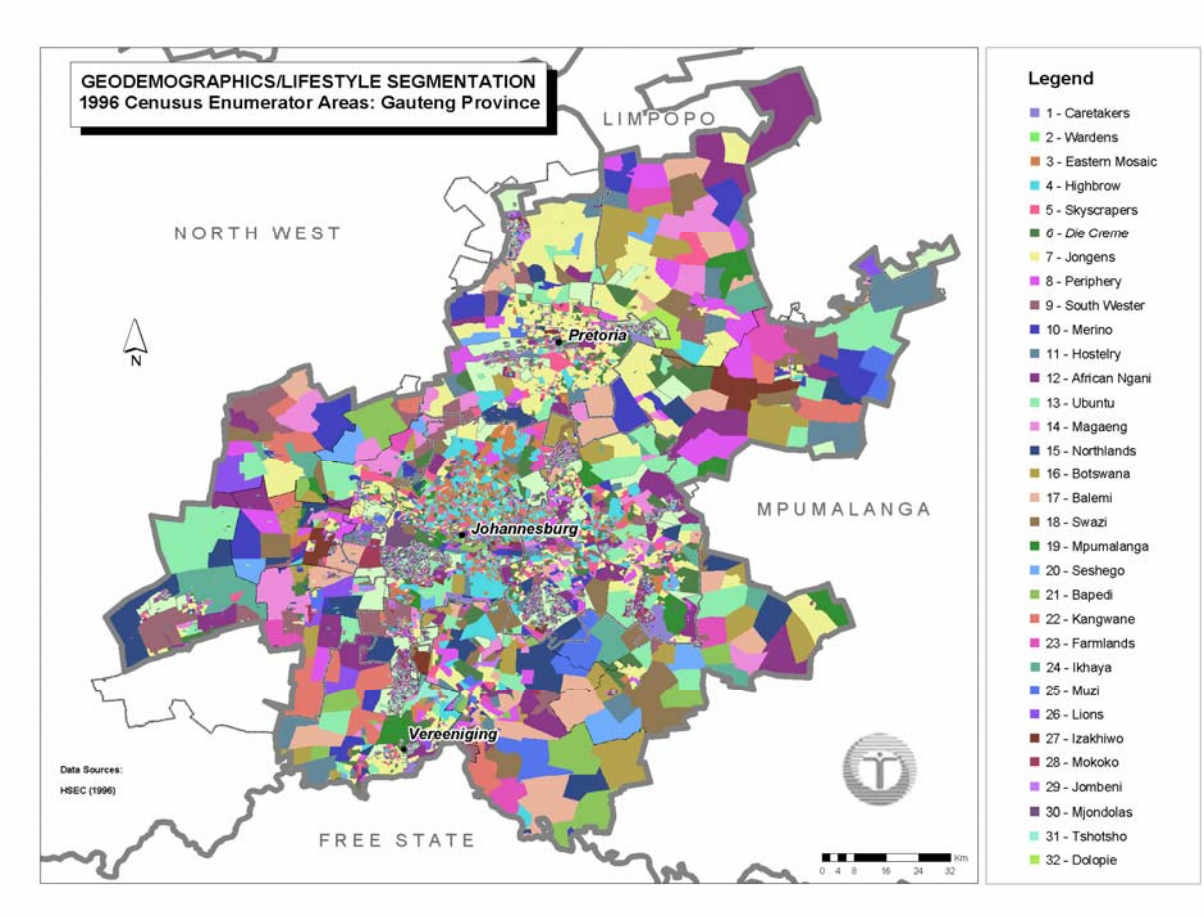
With over 6 000 possible combinations of variables from the questions asked in the 2001 census, the possibility of data overload is great when users start to use the census information. It is also virtually impossible to understand the many relationships that exist between the different census variables in a multidimensional space. Therefore, what is required is a method to reduce these variable combinations to a few easily understandable classifications. Clustering and ANN are techniques that can be used to categorise all EAs into a few classes. Linking these classes to the spatial information creates what is more commonly known as a geo-demographic database (Martin 1991).

Figure 9: An example of a raster map showing population density in different areas using a hexagon grid



Lifestyle segmentation databases can be developed by combining consumer behaviour, buying patterns and attitudinal information with the different classes in the geo-demographic database. This information can be of use in understanding potential markets and in the development of marketing strategies for commercial products in the business sector. Geodemographic information is also very useful in the development of sampling strategies for surveys. As users become more aware of geodemographic and lifestyle segmentation databases, the use of these databases in research and planning will no doubt increase. Figure 10 shows the 35 geo-demographic categories at an EA level in Gauteng.

Figure 10: An example of a lifestyle segmentation map



5.6 Map layout

When deciding on a map layout, the KISS principle – Keep It Short and Simple – should be observed. After going through the process of selecting a particular variable and a particular format, deciding on the purpose of the map and the spatial level at which the information should be presented and deciding what type of map should be used, the next stage is designing the map layout. When Atlas GIS from ESRI is first launched, a default map layout is generated, which is a good example of what a simple map layout should look like. The basic components of a meaningful map are the spatial information, map title, layer legend and thematic legend. Other secondary information can then be added to the map for cartographic purposes.

Map title

Every map should have a title that is grammatically correct and that states clearly what information is being illustrated. Depending on the purpose of the map, the title should be in a text font that is sufficiently large and bold to complement the spatial information. The title is usually at the top of the map but may be positioned elsewhere. The title should ideally indicate the spatial level and area that the information covers (e.g. municipalities in South Africa) so that the user can quickly know the spatial coverage and spatial unit of analysis.

Spatial information

The spatial information should form the focus of the map and, therefore, should occupy as much of the map page or screen as possible. The spatial information can be positioned anywhere on the map as long as the focus of the map is not lost. As has been described above, different types of maps can be used in addition to different colours and shades when using thematic maps. Secondary information can be added to the map such as labels and reference information (e.g. main roads and town names).

Layer legend

The layer legend provides information on which spatial layers are being shown on the map. The symbols representing the different layers of spatial information should not only be clear but should also adhere to cartographic standards. For example, a double red line represents a national road and a red cross in a circle represents a hospital. The layer legend is usually on the left or right-hand side of the spatial information. Town plans usually have the layer legend in a column on the right-hand side of the map while topocadastral maps have the legend below the spatial information. A simple title should be at the top of the legend (e.g. Legend), and the legend box is usually outlined.

Thematic legend

The thematic legend displays the colours representing the different classes on the spatial information and their associated values. The thematic legends are created automatically by most desktop GIS software but have to be placed on the map page. Appropriate text fonts for the legend title and class values should be selected.

Secondary information

Other information that can be indicated on a map:

- Scale bar or scale
- North arrow
- Source of the information (e.g. Statistics SA 1996)
- Geographic coordinate system
- Map producer's name or logo
- Additional explanatory text, tables, charts or images

5.7 Conclusion

This chapter has dealt with a number of issues associated with the use of census information in GIS. The main challenges that will be faced by most users will be the incorporation of the 2001 census information into GIS, understanding what the different spatial patterns associated with each census variable mean and deciding on

what types of maps to use within the GIS environment. As users become more experienced, they will tend to use new techniques for mapping the census information and to apply more sophisticated analysis methods. The 2001 census will become an indispensable source of information for research, planning and the provision of much-needed basic services to the people of South Africa.

Bibliography

Charlton M., Rao L. & Carver S. (1995) GIS and the census, *Census User's Handbook.* GeoInformation International: Cambridge.

Coombes M. (1995) Dealing with census geography: principles, practices and possibilities. In *Census User's Handbook.* GeoInformation International: Cambridge.

Davies H. (1995) Accessing the data via SASPAC91, *Census User's Handbook.* GeoInformation International: Cambridge.

Denham C. & Rhind D. (1983) The 1981 Census and its results. In Rhind D. (ed.) *A Census User's Handbook*. Methuen: London. pp. 55-87.

Dorling D. (1995) *A New Social Atlas of Britain*. John Wiley & Sons: Chichester.

Martin T.J. (1991) *Geographic Information Systems and their socioeconomic applications.* Routledge: London.

Rees P. (1995) Putting the census on the researcher's desk, *Census User's Handbook.* GeoInformation International: Cambridge

Chapter 6: Migration studies based on the 1996 and 2001 Census data

6.1 Introduction

For some considerable time, migration research in South Africa was hampered severely by the lack of useful data. Sample surveys cannot provide the level of information on detailed movement patterns between areas that is needed to obtain a satisfactory understanding of the nature of the phenomenon. Censuses are potentially the only appropriate sources of information for this purpose. Unfortunately, past censuses in South Africa failed to produce the required data for the study of migratory patterns in South Africa. The 1996 and 2001 censuses were, however, welcome exceptions to the rule.

Although all South African censuses have provided data on country of birth, only two of the previous censuses generated data on migration within the borders of the country. In the census of 1980, respondents were asked to furnish information on place of residence five years before, and, in the 1991 census, a question on duration of residence at the current address was included. Both these questions have inherent flaws when used in isolation. Perhaps more problematic, though, is that only a part of the country's population was covered. During the 1980 census, the migration data on the former (constitutionally independent) homelands of Transkei, Bophuthatswana and Venda were not available, and in the census of 1991, no migration data were available either for the later addition, Ciskei. Only the 1996 and 2001 censuses yielded migration data for the entire country.

The details of the 1996 and 2001 census variables for migration studies are given here. The author hopes that this contribution will enable those interested in migration studies to use the census data effectively. It is important, though, first to gain a clear understanding of the issues involved in migration definitions. As will be seen later, the 1996 and 2001 censuses use a specific definition of migration, and this conceptualisation needs to be placed in proper perspective.

6.2 Issues of definition

Migration

Migration is best defined (in general terms) as the *crossing of a spatial boundary by one or more persons involved in a change of residence*. Although this definition appears to be rather simple, complex issues and a lengthy theoretical debate underlie it. It will not serve any purpose to dwell here on all these issues, but the rationale for this specific conceptualisation needs to be explained. The two main components included in the definition – those dealing with "spatial boundaries" and "changes of

residence" – will be discussed below as will the components that have *not* been included in the above definition, namely the "time" and "social" dimensions.

It is generally accepted that migration has "distance", "time" and "social" dimensions. The time and space dimensions were built into the concept from the outset, but *this is where consensus among migration researchers ends*. Many migration analysts have indicated the numerous problems that arise in any deeper conceptual analysis of these two dimensions. Morrison (c. 1980) highlighted the main problems with the *distance* dimension, and he criticised the longstanding practice of movers being defined as migrants if they move some arbitrary distance, or if they cross some boundary that reflects a distant move. Neither are the *time* and *social* dimensions free from serious problems (as will be seen later).

The "spatial boundary" component

The concept "spatial boundary" refers to the borders of a particular study's "migration-defining areas". These could be political areas (e.g. countries, provinces or local governments) or administrative units (such as magisterial districts or census enumerator areas). Every study should, therefore, have its own (unambiguous) definition of an appropriate "spatial boundary".

This restriction has been severely criticised by many prominent migration researchers. The distance of a move is clearly important to distinguish migration from the more local moves (which are covered in the migration literature under the terms "circulation" or "residential mobility"). The problem is how to find an unambiguous definition of what Shaw (1975:6) called a "significant" distance. A change of residence varies along a distance continuum. Highly localised moves – from one apartment to another in the same building, from one house to another in the same neighbourhood or town – are forms of residential mobility that should, according to the United Nations' manual (1970), not be considered migratory moves. The concept of distance may be a key spatial aspect of mobility, but it is not a straightforward matter of mere kilometres. Distance incorporates certain elements that help to contextualise and inform one's understanding of the concept such as geographical, economic and social distance (Standing 1984:32).

The reality is that the use of actual distance in migration research has been somewhat limited, but contemporary technology, such as that offered by GIS, can help to overcome many of these problems. If fairly precise information can be obtained on the point of origin of the most recent move, for example, an analysis of moves in terms of the actual distance covered will be possible. As Shryock, Siegel and Associates point out, however, a definition of a "migrant" that is phrased in terms of a minimum distance moved will remain arbitrary "unless there were some natural break in the continuous distribution of moves" (1976:373–374). The actual distance of a move is,

therefore, not a clear enough indicator to distinguish migration from residential mobility.

Accordingly, one is almost compelled to return to the need for an area's boundary to be crossed before migration can be said to have taken place. The question that begs answering is consequently: What constitutes an "'area"? Standing (1984) points out that the limits often placed on the concept "area" can be largely arbitrary or merely expedient. These limits are usually determined by the administrative unit identified in censuses or surveys – to the detriment of scientific inquiry: "Somewhat remarkably, most demographers and other social scientists have let statisticians and survey administrators determine the areas between which moves are classified as 'migration'. In principle, this surely cannot be generally acceptable. Indeed, it has been said that areas between which moves count as migration are first defined by bureaucrats and later rationalized by social scientist researchers" (Standing 1984:32).

Standing suggests that one should rather use a change in "activity space" as a criterion. However laudable this suggestion may seem, it is fraught with practical implementation problems. Inappropriate census and survey data usually prevent analysts from using activity change as a criterion in the practical definition of migration. Unless the research or census questionnaire specifically makes provision for the collection of appropriate information, it is virtually impossible to determine whether a change in activity space has taken place or not. The theoretical debate on the matter has also not subsided yet. A difficult issue that still needs resolution is defining "activity space" and "activity change". In terms of Standing's suggestion, the definition of migration should exclude "changes in activities" within the same "activity space". But how does one deal with changes in "activity space" that do not involve any "activity change"? Should the definition include "activity changes" outside employment and education, and, if so, where does one draw the line? Is a change of employer as important as a change in occupation? How "activity change" can be defined unambiguously is, consequently, not at all clear and needs further debate. At this point, it cannot serve as a criterion for classifying a move as migration.

Shryock, Siegel and Associates (1976) refer to an earlier suggestion that migrants should be defined as those persons moving relative to labour market areas. Accordingly, the migration-defining area's boundary should perhaps be set at the distance line where commuting to work becomes so time consuming and expensive that a change of residence is required (373–374). This suggestion certainly has some appeal from a purely logical point of view, but it has not caught on in practice. The lack of interest can conceivably be explained in terms of Standing's conclusion about those economists who have analysed migration in terms of movements between labour markets: They "would surely have to admit that the criteria distinguishing the boundary lines of such markets are conceptually weak and empirically somewhat arbitrary"

(1984:33). This problem is, of course, particularly relevant in the context of the very long commuting distances travelled daily by many workers in a developing country such as South Africa (Gelderblom & Kok 1994:105). Such commuting distances complicate the scientific demarcation of labour market areas as possible migration-defining areas.

The "change of residence" component

The "change of residence" requirement implies that not all movements constitute migration (as in the case of daily commuting, for example) since only those moves that involve "moving house" would meet this basic requirement. This concept dispenses with the widely criticised conceptualisation of "usual place of residence", which has been found to lack sufficient applicability in both less and more developed countries. Debate is, of course, still ongoing about the issue of "residence" or "home". Restricting migration to "changes in residence" is seen as excluding homeless persons without any residential address, but, then, who said anything about the need for an address?

Also criticised is the underlying notion that people have only one place of residence when they may well have two (or even more) homes. This problem is discussed by, among others, McHugh, Hogan and Happel in their 1995 article on the flaws in the official definition of "usual place of residence" used in the USA. They conclude that "today, the assumption of a single, fixed place of (usual) residence is untenable for a substantial number of Americans" (251). This comment also applies to South Africa with its history of migrant labour and expected incidence of migrant labour in the future. However, again, the author's response would be: Who said a change in any (one or more) of these places of residence cannot be regarded as migration?

The criticism about multiple residences does become highly relevant when one deals with *de jure* counts of the population. A *de facto* population count refers to the number of people who were actually enumerated. A *de jure* population count, on the other hand, takes into account the fact that a significant proportion of the population were not "home" on the night of the census because of business, study, holiday or other reasons. Therefore, information on a person's "usual place of residence" is usually needed to be able undertake a *de jure* count of the population. In general, it is difficult to obtain totally reliable information on people's "usual places of residence" when a significant proportion of the population resides in two or more places. Such people do not have a single place called "home" in the traditional sense of the word.

Standing suggests that the notion of "usual" is also inherently problematic: "Does it refer only to a relative amount of time spent in one place rather than another, or does it mean a place that an individual considers being his or her "home"? Even if the former is taken as the yardstick, an unambiguously specified reference period, either retrospective or anticipated, should be given. Furthermore, asking where people lived

“most of the time” over the past six months or past year is not necessarily the same as asking where they regard as their “home” (1984:34–35). Despite these potential difficulties, the vast majority of people can be expected to identify their “usual place of residence”, provided that it is defined as that place where they spend most of their time (for example, at least four nights a week).

The ‘time’ component

What has not been included in the above definition is the temporal dimension because of its limited utility in helping analysts to distinguish migration from other forms of mobility. Morrison (c. 1980) refers to problems with the “permanence” of spatial mobility because the concept ignores evidence of the existence of a whole range of impermanent moves reflecting wide variations in length of stay when people are classified as “migrants” or “non-migrants”. Some moves may consequently be excluded quite unjustifiably.

Another point to consider when discussing the temporal notion of ’permanence‘ is the issue of ’intended time‘ (Standing 1984:36). The lack of an underlying logic for restricting migration to a minimum time period is probably best reflected in the different preconditions of an intended or actual minimum length of stay for someone to be counted as a migrant in the Scandinavian population registers. According to Bretz (1996), these preconditions range from three months in Denmark to six months in Norway to 12 months in Sweden and Finland. The same problem exists in other countries with such registers. In Indonesia, the criterion is six months or “less than six months but intended to stay there permanently”, while the Japanese criterion is only three months (Skeldon 1990:18). It should be clear, therefore, that the ’time‘ component is not particularly useful in migration studies and should consequently not be used as a criterion in the definition of migration.

The ‘social’ component

Usually, the social component in the definition of migration requires that a significant change in social context must have taken place during the move. The formulation of a suitable definition of migration, which incorporates the social component, is not a simple matter. Statistics South Africa is, however, planning a migration publication for the 2001 census, and the formulation of appropriate definitions is expected to be a key part of the exercise. Incorporation of the social component into the definition will probably be informed by the research conducted by the Agincourt Demographic and Health Surveillance System (DSS) located in the Agincourt subdistrict of the Bushbuck Ridge region, (now) Mpumalanga Province (see, for example, Collinson et al. 2001).

6.3 Other concepts and definitions

Other concepts used in migration studies need clarification here.

Origin and destination

Every residential move has an *origin* (the place from where the person moves) and a *destination* (the place where the specific move ends). The origin and destination of a residential move can be in the same country/area or in different countries/areas.

International versus internal migration

If the migratory move involves the crossing of a national boundary, it is known as *international migration*. The person involved in such a move is simultaneously called an *emigrant* (from the perspective of his/her country of origin) and an *immigrant* (from the perspective of the country of destination).

If the origin and destination of a specific migratory move are both in the same country, the move constitutes *internal migration*. If the origin and destination are in the same country, the person who migrates from a particular place is called an *out-migrant* from that area, and at the same time he/she is an *in-migrant* into the area of destination.

Residential mobility, circulation and commuting

If the origin and destination of a residential move are in the same country and in the same (migration-defining) area, the move is not regarded as migration but rather as residential mobility. The term “residential mobility” is reserved for those moves that involve a change of residence to distinguish such moves from circulation.

Box 1: Migration questions asked during Census 1996

Section A: In respect of each household member:

1.1 Was (the person) born in South Africa? (Include the former Transkei, Bophuthatswana, Venda, Ciskei—TBVC states)

1 = Yes

2 = No

1.2 (If NO) In what country was the person born?

Write in the name of the country. ……..

2 Is (the person) a migrant worker? (Someone who is absent from home FOR MORE THAN A MONTH each year to work or to seek work.)

1 = Yes

2 = No

3.1 Is this DWELLING (e.g. house, room, shack, flat) the place where (the person) usually lives, i.e.

where (the person) spends at least four nights per week?

1 = Yes

2 = No

3.2 (If NO) Where does (this person) usually live?

Name of

suburb/village/settlement:….....……..…….....

Name of city/town/farm/tribal authority: .....……..…….....

Name of magisterial district: .....……

If not South Africa, please state name of country: .....……

If no usual address, circle 3.

4.1 In which year did (the person) move to the DWELLING (e.g. house, room, shack, flat) where he/she usually lives?

Write in the year that he/she moved 19.....…

OR

1 = The person has never moved. (Lived in the dwelling since birth)

4.2 (For the person who has moved)

From where did (the person) move? (Before moving into the dwelling where he/she usually lives)

Name of suburb/village/settlement: ……

Name of city/town/farm/tribal authority: ….....……..……..

Name of magisterial district: .....….....

If not South Africa, please state name of country: .....

Section B: In respect of the entire household:

5. Are there any persons who are usually members of this household, but who are away for a month or more because they are migrant workers? (A migrant worker is someone who is absent from home for more than a month each year to work or to seek work).

1 = Yes

2 = No

(If YES) indicate the person's particulars:

Age in years

Gender

Relationship to head of household

Where is (the person) living:

Name of suburb/village/settlement: ……..……..……..

Name of city/town/farm/tribal authority: .....……..……..

Name of magisterial district: .....……..……

If not South Africa, state name of country: .....

Box 2: Migration questions asked during Census 2001

Ask for all people in the household:

1a. Was (the person) born in South Africa?

Y = Yes

N = No

[If NO, go to 1c.]

1b. If YES: In which province was (the person) born?

1 = Western Cape

2 = Eastern Cape

3 = Northern Cape

4 = Free State

5 = KwaZulu-Natal

6 = North West

7 = Gauteng

8 = Mpumalanga

9 = Northern Province [Limpopo]

[Go to 2a.]

1c. If NO: In which country was (the person) born?

Write the present name of the country. ……

2a. Is (the person) a South African citizen?

Y = Yes

N = No

[If YES, go to 3a.]

2b. If NO: What is the name of the country of citizenship? ……

3a. Does (the person) usually live in this household for at least four nights a week?

Y = Yes

N = No

[If YES, go to 4a.]

3b. If NO: Where does (the person) usually live?

IF IN THE SAME PLACE as the place of enumeration, dot the S box "S"

IF NOT the same place, write the PROVINCE, MAIN PLACE (city, town, tribal area, administrative area) and SUB-PLACE (suburb, ward, village, farm, informal settlement).

IF ANOTHER COUNTRY, write the name of the country. ……

4a. Five years ago (at the time of Census '96), was (the person) living in this place (i.e. this suburb, ward, village, farm, informal settlement)?

Y = Yes

N = No

B = Born after October 1996

[If Y or B, go to the next section.]

4b.	If NO: Where did (the person) move from? If more than one move, give details of the last move. Write the PROVINCE, MAIN PLACE (city, town, tribal area, administrative area) and SUB-PLACE (suburb, ward, village, farm, informal settlement). IF ANOTHER COUNTRY, write the name of the country……
4c.	If NO: In which year did (the person) move to this place? 1 = 1996 4 = 1999 2 = 1997 5 = 2000 3 = 1998 6 = 2001 If more than one move, write the code for the year of the last move.

If a move involves no change of residence, it is neither migration nor *residential mobility* but some or other form of *circulation*. If such a circulatory move takes place repeatedly between the person's place of residence and his/her place of work, it is known as commuting. If the commuting normally takes place on a daily, weekly or monthly basis, this form of mobility is called *daily, weekly or monthly commuting* respectively.

The questions included in the 1996 census questionnaire are listed in Box 1. These questions cover the entire battery of migration-related questions in the census. Box 1 shows that five fields of study in migration analyses are covered by the 1996 census questions. The questions are:

- "Lifetime migration" (see Items 1.1 and 1.2);
- "Migrant labour" (see Items 2 and 5);
- "Place of usual residence" (Items 3.1 and 3.2);
- "Duration of residence" (Item 4.1); and
- "Origin of the most recent move" (Item 4.2).

The entire battery of migration-related questions included in the 2001 census questionnaire is given in Box 2. The fields of study in migration analyses covered in the 2001 census are:

- "Lifetime migration" (see Items 1a and 1b)
- "Place of usual residence" (Items 3a and 3b)
- "Duration of residence (in terms of the 1996–2001 interval)" (Items 4a and 4c)
- "Origin of the last move (if moved since 1996)" (Item 4b)

A five-year migration interval (1996–2001) was evidently used, but this interval was supplemented by details of the last move (if it occurred after the previous census), which naturally gives the analyst a particularly rich source of migration data, especially with regard to the recent past. Together, the questions from these two censuses form

a potentially powerful battery of questions dealing with migration. For the very first time in the country's history, spatially detailed as well as historically comparative migration studies have now become possible. The next sections contain information on how these census data can be used in migration studies.

6.4 Potential uses of migration data

Lifetime migration

Little needs to be said about lifetime migration. Whereas in the 1996 census, only the country of birth is known, Census 2001 also made provision for information on the (South African) province of birth. The study of lifetime migration can, for both censuses, take the form of an analysis of a table of frequencies listing the proportions of foreign-born people in South Africa, for example according to continent, subcontinent and country. The 2001 data also allow analyses of lifetime migration between provinces.

Migrant labour

The extent of migrant labour should be calculated *only* on the basis of Item 2 in Box 1. Item 2 deals with the *de facto* presence of an individual at a certain place of enumeration on census night (i.e. between 9 and 10 October 1996). If both Items 2 and 5 in Box 1 are used, an overestimation of the magnitude of migrant labour will undoubtedly occur. If only Item 5 is used, however, there will be a much greater risk of errors in the reporting of migrant labour. Item 5, which is meant to cater for the *de jure* situation, constitutes indirect reporting on people, presumably belonging to the household, who are physically absent at the time of the census. There are simply too many possible sources of error to rely solely on Item 5.

A slim chance will still exist of an overreporting of migrant labour due to the enumeration of the same individual in different places even if only Item 2 in Box 1 is used. This can be attributed to the relatively long period of time over which the census took place. Household spokespersons might not have taken sufficient care to restrict the information to only those who had actually been present in the household on census night. They might also simply have forgotten who had been present on census night (and who had not been present). Some of the more mobile people, such as the migrant labourers in particular, could therefore have been counted twice. From the research reported by Kok, O'Donovan, Bouare and Van Zyl (2003), it is clear, however, that the extent of (single-person) labour migration is probably underestimated when analysts use the 1996 census data.

The 2001 census did not generate specific data on labour migration (see Box 2), which makes it impossible to compare data on labour migration from the two most recent censuses.

Place of usual residence

The questions on place of usual residence (see Items 3.1 and 3.2 in Box 1, and Items 3a and 3b in Box 2) were needed to obtain a *de jure* count of the population. The rationale behind the *de jure* count is that people should be counted where they normally live, not where they happened to be on census night. These *de jure* localities are the places where people spend most of their time and for whom planning should be done. (As suggested before, the notion of "usual place of residence" is flawed, but it does help analysts to place population figures in their proper spatial perspective.) The 1996 census data released in respect of Question 3.2 in Box 1 were coded to the level of magisterial district, and in Census 2001, the coded data for Question 3b (Box 2) were released down to the level of "main place" (city, town, tribal area or administrative area).

Duration of residence

Duration-of-stay data should be linked to spatial migration data (e.g. data on previous area of residence). This was not done in the case of the 1991 census, and, in the absence of any spatial information, the data on duration of stay had very little practical value. In the 1996 and 2001 censuses, the error in the 1991 census was rectified and, as is clear from Items 4.1 and 4.2 (Box 1) and Items 4a–4c (Box 2), provision was made for the important linkage of temporal and spatial information.

Origin of the most recent move

In the 1980 census, a question was asked about place of residence five years before. The responses were coded at the suburb level, which theoretically made it possible to study residential mobility. The problem was, however, that this was so costly an exercise that only five per cent of the questionnaires was fully coded down to this level of spatial detail.

Information on residential mobility can be obtained quite cost-effectively through localised sample surveys, if and when required. Consequently, it is the author's opinion that Statistics South Africa *has the responsibility to concentrate its scarce resources on the more general migratory moves in the country as a whole, rather than to provide details of localised moves.* What is needed, therefore, is for the (full) census to provide data that will allow analyses of migration patterns (across political and administrative boundaries) because these data cannot be obtained in any meaningful way from sample surveys.

Fortunately, Statistics South Africa rose to the challenge in the 1996 and 2001 censuses. The 1996 census included a question on the previous place of residence together with a question on duration of residence. The data on previous place of residence were then released to the level of magisterial district. This enabled analysts to link the information on "usual place of residence" (Item 3.2 in Box 1) with the

information on “previous place of residence” (Item 4.2 in Box 1). Combined with Item 4.1 (Box 1), this battery of questions provided an exceptional source of migration information. In Census 2001, the details of the last move (if the person had moved since the 1996 census) were released to the level of main place.

However, a conceptual flaw is apparent in the migration questions posed during the two censuses. This flaw relates to the absence of predetermined migration-defining areas:

(a) Census 1996

The magisterial district was clearly (and correctly at the time) seen as the most appropriate “migration-defining area” for the 1996 census, but, although Statistics South Africa also coded the origin information down to “place”, only the data coded to the level of magisterial district were publicly released. Strictly speaking, Item 4.1 (in Box 1) should have been phrased differently to accommodate this. The words “this DWELLING” should rather have been “this MAGISTERIAL DISTRICT”. The last part of Item 4.1 (Box 1) should then have read as follows:

“1 = The person has *never lived in another magisterial district.*
(Lived in the *same magisterial district since birth*)”

(b) Census 2001

In the case of the 2001 census, the question on last migration was asked down to subplace (suburb, ward, village, farm or informal settlement) level (Item 4b in Box 2), but so far the data have been released only to the level of main place. It would have been better to exclude the reference to the subplaces (suburbs, wards, villages, farms or informal settlements) in Question 4a (Box 2), indicating the need to report moves to the current suburb, ward, village, farm, informal settlement. Item 4a (Box 2) should rather have read as follows:

“Five years ago (at the time of Census 1996), was (the person) living in this place (i.e. this *city, town, tribal area or administrative area*)?”

The words in italics in the two rephrased questions above indicate the changes that would have been necessary to meet all the theoretical requirements. The problem with the actual phrasing of the census question was that it did not necessarily cover all the last migratory moves. For example, a data problem would arise if the most recent move was made from an origin within the same magisterial district or main place, and the individual concerned had *previously* moved there from another district/main place. In this example, the earlier move – *which is the only valid move in terms of the adopted approach, namely to use magisterial districts or main places as migration-*

defining areas – would not be accounted for at all. This problem would inevitably lead to an underestimation of migration and a concomitant relative overestimation of residential mobility.

Although this is a significant problem from the perspective of a "purist" in migration analysis, in the greater scheme of things it is a relatively minor issue. In fact, it might have been more problematic to use "magisterial district" instead of "dwelling" in the 1996 census. The potential problem with the concept of a "magisterial district" is that many people in South Africa may not be aware of the district-defined location of their usual place of residence. By forcing a response on district location in the field (as against a classification in the office afterwards), the quality of most of the information could actually have been compromised. The concept of "main place" as denoting city, town or tribal area, as used in Census 2001, would have been more understandable to respondents than "magisterial district". However, the undefined spatial category "administrative area", included under "main place", would probably have been even more confusing than "magisterial district".

The data from Census 2001 will also allow analysts to dispense with the restriction of "place of usual residence", if they so wish, by using "place of enumeration" instead. This also has the benefit of analysts not having to contend with the difficulties associated with "usual" in the context of residence. However, this *de facto* approach would not appropriately deal with the need for *de jure* migration analyses, and therefore the use of "place of usual residence" is suggested in most cases.

6.5 Analysing the available census data on migration

Some examples of how the 1996 and 2001 migration data can be used in analyses are given. The idea is to introduce principles of comparative migration analysis, migration selectivity and migration indicators, and also to explain their interpretation.

Interprovincial migration

Up till the time of writing, only interprovincial migration data have been released for the full 2001 census. When the spatially more detailed migration data become available, analysts will be able to study migration patterns at lower levels of spatial detail (e.g. "magisterial district" for both censuses or "main place" for analyses of the 2001 census data).

Table 1: Number of interprovincial migrants during the periods 1992–1996 and 1996–2001

(a) Interprovincial migration (1 January 1992 to 10 October 1996)

		Destination province: Interprovincial migrants (1992–1996)									
Province of origin	**Provincial population 1996**	**Western Cape**	**Eastern Cape**	**Northern Cape**	**Free State**	**KwaZulu-Natal**	**North West**	**Gauteng**	**Mpumalanga**	**Limpopo**	**Total**
Western Cape	3 956 876		11 022	6 041	3 429	4 263	1 535	14 782	2 196	600	43 868
Eastern Cape	6 302 524	110 416		1 883	18 017	34 023	11 908	59 487	6 822	1 339	243 895
Northern Cape	840 323	15 094	1 783		5 529	854	5 615	6 656	1 070	210	36 811
Free State	2 633 503	6 599	3 606	5 279		3 746	16 778	34 208	5 324	901	76 441
KwaZulu-Natal	8 417 020	11 423	6 226	737	7 210		2 700	79 721	12 243	1 025	121 285
North West	3 354 824	2 552	852	10 072	9 034	1 409		101 252	4 638	5 108	134 917
Gauteng	7 348 425	37 492	12 728	4 186	25 946	26 462	44 198		42 172	19 671	212 855
Mpumalanga	2 800 710	2 092	935	591	3 584	5 145	6 070	59 839		9 677	87 933
Limpopo	4 929 365	1 308	523	398	2 270	1 609	17 478	119 214	32 826		175 626
Total	40 583 570	186 976	37 675	29 187	75 019	77 511	106 282	475 159	107 291	38 531	1 133 631

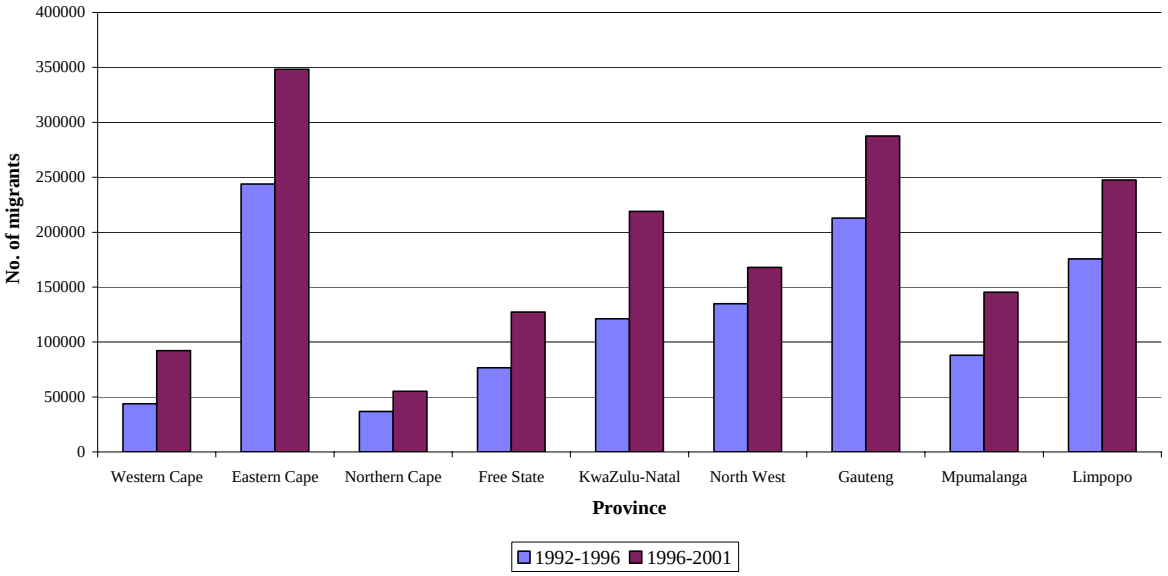
Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996)

(b) Interprovincial migration (11 October 1996 to 10 October 2001)

		Destination province: Interprovincial migrants (1996–2001)									
Province of origin	**Provincial population 2001**	**Western Cape**	**Eastern Cape**	**Northern Cape**	**Free State**	**KwaZulu-Natal**	**North West**	**Gauteng**	**Mpumalanga**	**Limpopo**	**Total**
Western Cape	4 524 341		26 522	9 635	5 144	9 232	3 713	32 279	3 133	2 515	92 173
Eastern Cape	6 436 764	141 386		4 113	16 686	59 114	21 093	89 487	9 983	6 358	348 220
Northern Cape	822 729	21 140	2 948		7 581	1 856	7 428	11 033	1 421	1 716	55 123
Free State	2 706 774	12 942	8 679	6 335		8 479	19 986	59 588	6 936	4 396	127 341
KwaZulu-Natal	9 426 015	24 474	18 160	1 839	8 849		7 905	132 020	18 673	7 076	218 996
North West	3 669 353	6 989	4 264	16 278	10 225	4 380		107 992	6 303	11 504	167 935
Gauteng	8 837 174	57 540	28 743	6 742	24 840	43 928	52 854		34 151	38 747	287 545
Mpumalanga	3 122 991	5 956	3 231	1 474	5 716	11 163	11 467	88 367		17 877	145 251
Limpopo	5 273 639	5 139	2 735	1 394	4 146	5 159	21 175	170 205	37 529		247 482
Total	44 819 780	275 566	95 282	47 810	83 187	143 311	145 621	690 971	118 129	90 189	1 690 066

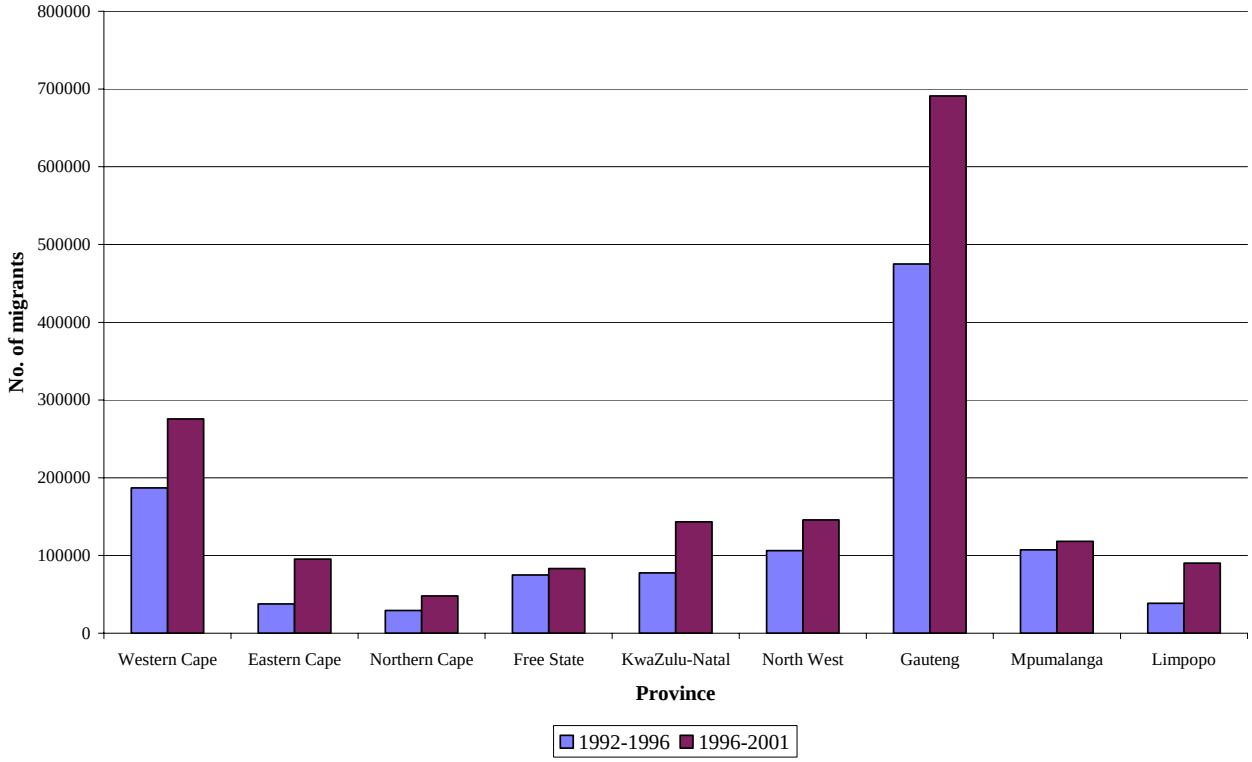
Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 2001)

Graph 1: Inter-provincial migration, number of out-migrants (1992–1996 and 1996–2001)



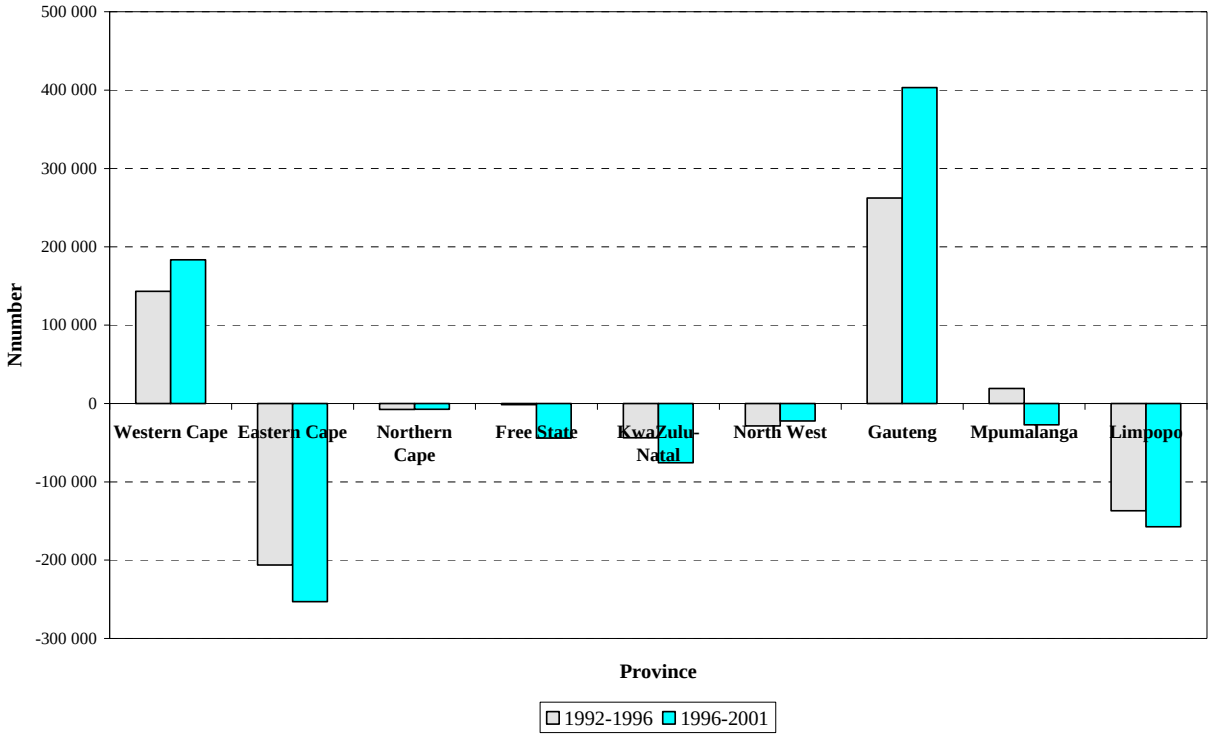
Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996 and 2001)

Graph 2: Inter-provincial migration, number of in-migrants (1992–1996 and 1996–2001)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996 and 2001)

Graph 3: Inter-provincial migration, net migration volumes (1992–1996 and 1996–2001)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996 and 2001)

Table 1 shows the absolute numbers of interprovincial migrants during the periods 1992–1996 and 1996–2001. Since the two migration intervals are not exactly the same, with the first period (1 January 1992 to 10 October 1996) being almost three months shorter than the second period (11 October 1996 to 10 October 2001), the migrant numbers for the two periods are not entirely comparable. A large increase was, however, evident in the numbers of interprovincial migrants over the two periods, not all of which can be explained in terms of the shorter first period. Migration therefore seems to be on the increase.

Table 1 indicates that while Gauteng attracted most of the interprovincial migrants during the period 1992–1996, the Western Cape attracted the largest proportions of interprovincial migrants from the Eastern Cape (45%) and the Northern Cape (41%). Although slightly less pronounced in the time interval 1996–2001, the same proportional patterns of out-migration to the Western Cape continued from the Eastern Cape (41%) and the Northern Cape (38%).

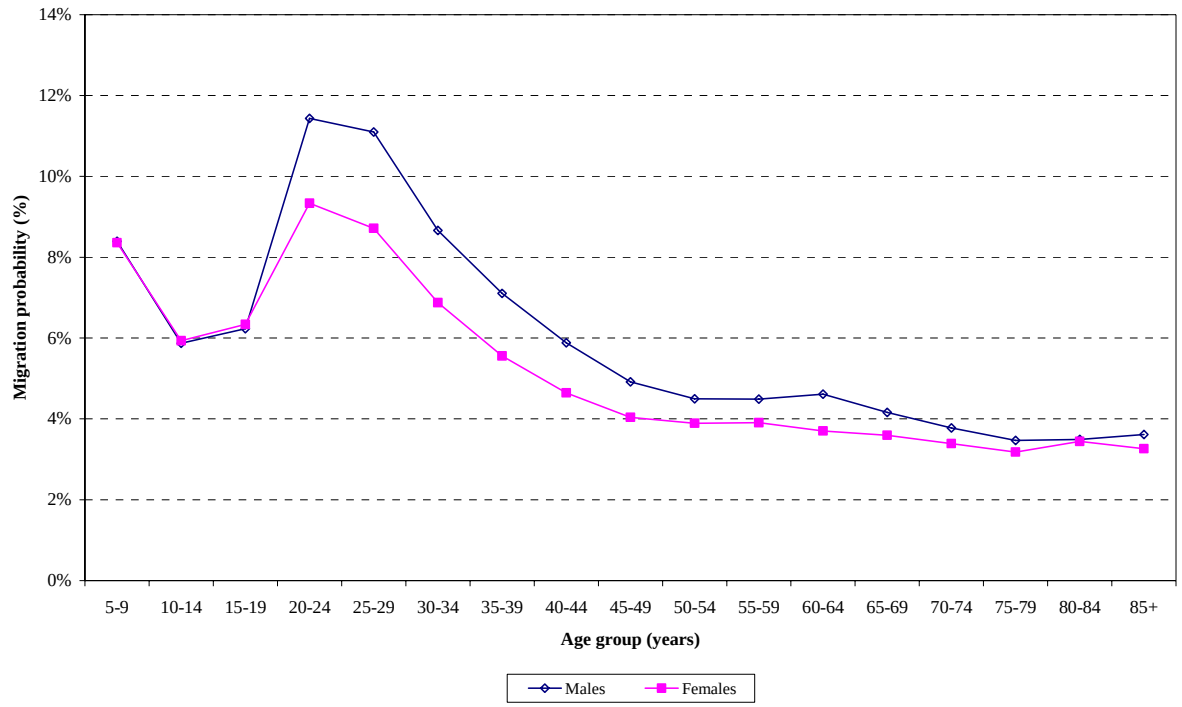
The provincial differences in numbers for the two periods are illustrated in Graphs 1–3. Graph 1 illustrates the out-migration volumes per province, Graph 2 depicts the volumes of in-migration per province and Graph 3 shows the provincial net migration volumes. Graph 3 indicates that only the Western Cape and Gauteng experienced a

consistent positive net migration. Mpumalanga, which experienced a slight net migration gain during the period 1992–1996, showed a net loss in migration during the period 1996–2001.

Migration selectivity

"Migration selectivity" usually relates to the distinguishing characteristics of migrants from the perspective of the area of origin.[23] Graph 4 illustrates the interprovincial migration selectivity in terms of age and sex for the period 1992–1996.

Graph 4: Inter-provincial migration selectivity in terms of age and sex (1992–1996)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996)

The migration volumes for persons aged 0–4 years are not available in respect of Census 2001. This age group is therefore also excluded for the purposes of the 1992–1996 profile (shown in Graph 4).

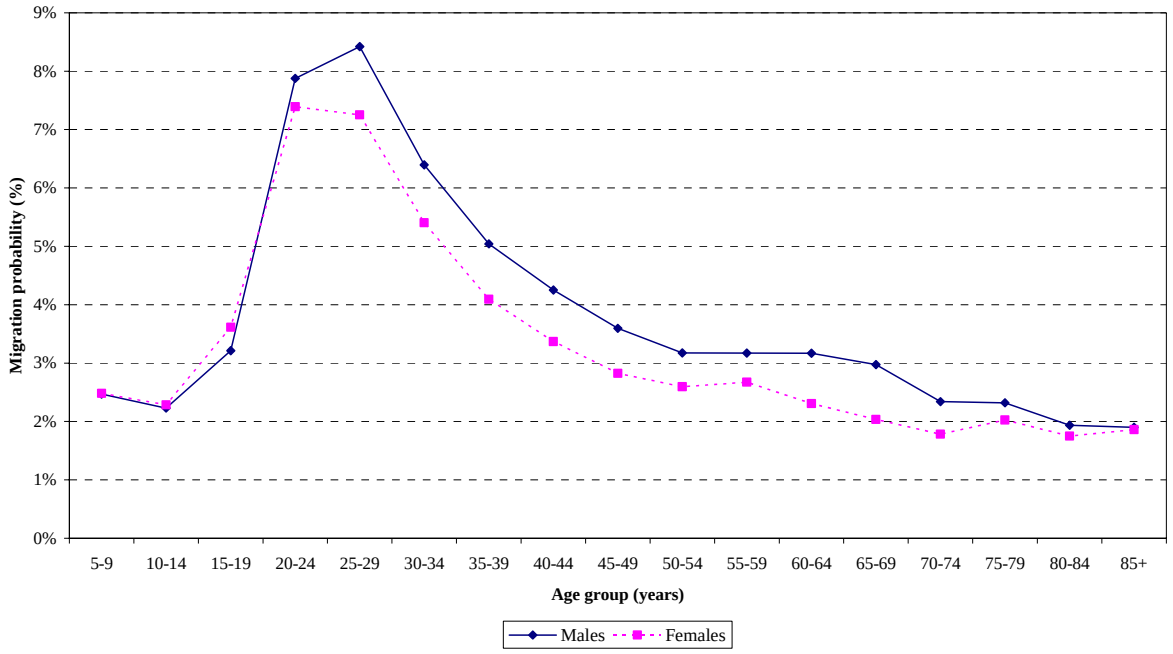
The general interprovincial migration selectivity profile in Graph 4 differs in one important respect from the overall 1992–1996 migration selectivity profile reported by Kok et al. (2003:56), which corresponded with the sample-based profile for 1975–1980. If one bears in mind that interprovincial migration probabilities should be (and are) much lower than overall migration probabilities, the relatively high migration

[23] The concept of "migration differentials", on the other hand, is commonly used to describe the characteristics of migrants from the point of view of the area of destination.

probabilities for persons in the age groups 5–9 and 10–14 years in Graph 4 are noteworthy. It appears, therefore, as if adults involved in interprovincial migration may be more likely to be accompanied by young children (under the age of 15 years) than migrants in general.

Graph 5 shows the interprovincial migration selectivity in terms of age and sex for the period 1996–2001.

Graph 5: Inter-provincial migration selectivity in terms of age and sex (1996–2001)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 2001)

The notion raised above concerning adults involved in interprovincial migration tending to be accompanied by young children to a greater extent than migrants in general is not borne out by the profile given in Graph 5. The graph for 1996–2001 (Graph 5) shows much lower migration probabilities for persons aged 5–9 and 10–14 than those depicted in Graph 4. The peak for young adults (aged between 20 and 34 years) in Graph 5 is also notably more pronounced than that in Graph 4. Whether these inconsistencies are due to actual changes in migration selectivity over time or to data inaccuracies remains to be seen and can be determined only from analyses of data from future censuses.

Migration indicators

Migration indicators are, at best, problematic. This is so because, firstly, with them one is dealing with low occurrence levels (i.e. the so-called rare events). A second problem relates to data quality. Unlike most other demographic events, a person's last migration might have taken place a long time ago, and this could give rise to reporting problems that are associated with memory lapse. Thirdly, one should also remember (at least in the context of internal migration) that the boundaries of spatial entities (e.g. magisterial districts or local government areas) change fairly often. Since "migration" must be defined (in "distance" terms) on the basis of "migration-defining areas", changing boundaries could be a particularly difficult problem to deal with in the development of appropriate indicators.

Nevertheless, no purpose would be served by being overly concerned about the anticipated problems. All the implications of these data issues may not be overcome, but indicators remain key instruments in summarising and presenting data on migration. The most common migration and related indicators deal with *levels* and *rates* (of change).[24] The generally used indicators that would probably be appropriate for the purposes of most migration-related studies are the (a) "urbanisation level" and (b) "migration rate".

6.6 Urbanisation level

Although the indicator "level of urbanisation" is frequently used, scholarly debates are being conducted on the appropriateness of such indicators denoting urbanisation level. The main problem lies with the definition of "urban areas" (as distinct from "rural areas"). Some academic literature criticises the use of the dichotomy "urban" versus "rural" when so many areas exist that cannot unambiguously be defined as belonging to either the one or the other category. An influential school of thought consequently insists on the use of a rural-urban continuum instead of the rural/urban dichotomy. The demand has accordingly arisen in South Africa to classify certain former homeland settlements as either semi-urban (if they are located "deeper" in predominantly rural environments) or peri-urban (if they are located relatively close to urban centres) or similar categories.[25]

However, sometimes even those in favour of a more nuanced approach feel the need for a simple distinction between rural and urban areas. This debate is often undermined when the pro-continuum advocates themselves fall back on the dichotomy interpretation when talking about urbanisation (and this author is no exception). The tendency seems rather to attempt to achieve a proper classification

[24] A conceptual problem in this regard is that demographers often use the term "migration rate" when they actually mean "migration level", but this is not a particularly serious matter for the purposes of this discussion. (This distinction does become important, though, when the rate/tempo of *urbanisation* is being investigated.)

[25] See, for example, the work by Graaff (1986), the Urban Foundation (1990), and Gelderblom and Kok (1994).

of “urban areas” by looking at a number of variables (size, density, dominant employment sector, urban-like services, local government status and so on instead of a single variable such as “local government status”) than to develop separate indicators (or highly complex and interrelated indices) for the rural, semi-urban/ peri-urban and urban (“proper”) populations.

The “urbanisation level” is probably the best-known migration-related indicator. It is given by the formula below.

$$PU = \frac{U}{P}k$$

where PU = proportion of the population living in urban areas (i.e. the urbanisation level) at a particular point in time (for the purposes of the analyses that follow, census day in 1996 and 2001, i.e. 10 October, is used);

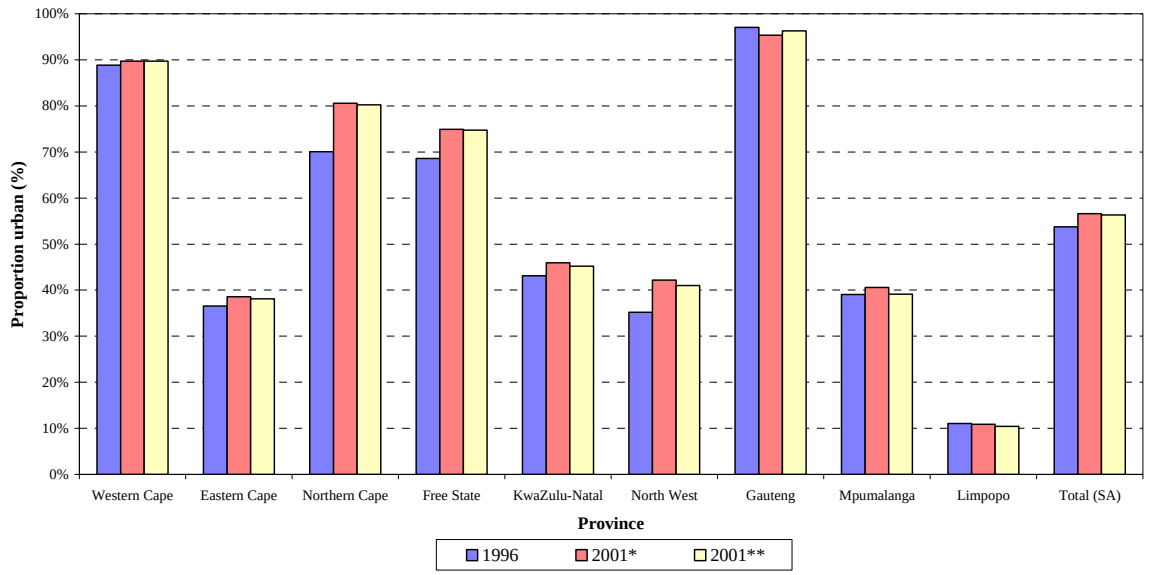
U = population living in urban areas at the given point in time (in this case, census day in 1996 and 2001);

P = total population at the given point in time (in this case, census day in 1996 and 2001), and

k = a constant (usually 100 so as to express the quotient as a percentage).

The calculation of this indicator by province yielded the results illustrated in Graph 6. Since the 2001 urbanisation figures were not provided when the community profiles were (first) released, two estimates of the 2001 urbanisation level are reflected in Graph 6. The first estimate is based on the author’s classification of enumerator area (EA) type in the descriptive community profile, and the second estimate is Statistics South Africa’s own classification based on 1996 EA types in the 10 per cent sample. Only when the latter (Statistics South Africa) classification of urban areas for the full census (community profiles) is released will it be possible to draw firm conclusions about changes in urbanisation levels over time. The advantage of this approach is, of course, that the urbanisation figures for the two censuses are at least spatially comparable.

Graph 6: Provincial and national urbanisation levels (1996 and 2001)



* The classification for 2001 was done by the author and based on enumerator area (EA) type. "Urban settlements", "informal settlements", "industrial areas", "institutions" and "hostels" were categorised as "*urban*" while "sparsely populated enumerator areas (with 10 or fewer households)", "tribal settlements", "farms", "smallholdings" and "recreational areas" were classified as "*rural*".

** The Census 1996 definition of what constituted an urban EA (i.e. whether or not it fell within a municipality or local authority) was used to make a final allocation of Census 2001 EA types to the category "urban" while EA types that did not meet the selection criteria were regarded as rural. Census 2001 EAs were spatially matched and linked to the 1996 types.

Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996 and Census 2001) and the data from the 10 per cent sample for Census 2001 (also provided by Statistics South Africa)

The comparatively high urbanisation levels of Gauteng and the Western Cape are clear from Graph 6. It also seems that the Northern Cape, North West and Free State have experienced notably higher rates/tempos of urbanisation than other provinces since 1996, but this needs to be confirmed when the urbanisation data from the full census are available.

6.7 Migration rates

The most frequently used migration indicator is the so-called *migration rate*, which refers to the level of in-migration, out-migration or net migration (in-migration minus

out-migration) compared to the size of the population concerned. The generic formula for the calculation of a migration rate is as follows:

$$m_{ij} = \frac{M_{ij}}{P} k$$

where m_{ij} = rate of migration from *i* to *j* during a specified time interval;
M_{ij} = number of migrants moving from *i* to *j* during the given time interval;
P = the population concerned (i.e. either in *i* or *j*) at a particular point in the given time interval, and
k = a constant (usually 100 so as to express the quotient as a percentage).

An interesting theoretical debate is taking place on the issue of what *P* should be. The general agreement seems to be, however, that *P* should be "the population at risk of migration" (which could be interpreted as either "the population at risk of migrating" or "the population at risk of receiving migrants").

Another topic of debate, also relating to *P*, is what the most appropriate "point in the given time interval" should be. The consensus seems to be that it should, for various reasons, preferably be the mid-point of the time interval. If, however, the population at the mid-point cannot be directly obtained or reliably estimated, the population at either end point of the interval can be used provided that a clear motivation for selecting that particular end point is given.

The 1996 and 2001 South African census data will be used here to "populate" the rates. For the calculation of the 1992–1996 migration rates, only the end-point (1996) population can be used as *P* because the 1991 census excluded the populations of the formerly independent homeland areas. Since the full 2001 migration data have so far been released only at the provincial level, only interprovincial migration rates can be calculated. Although it would have been possible to use an estimated mid-point population for the period 1996–2001, it is suggested that the end-point (2001) population is used for *P* here as well so as to ensure the necessary comparability between the periods 1992–1996 and 1996–2001.

To populate the migration indicators from the 1996 census data, the period 1 January 1992 to 10 October 1996 (census day in 1996) will be used. Similarly, the period 11 October 1996 to 10 October 2001 will be used for the 1996–2001 interval. These periods represent five-year intervals, and the times between the beginning and end of the intervals are short enough to warrant the use of the end-point population instead of the (ideal) mid-point population sizes.

Out-migration rate

The formula for calculating the *out-migration rate* can be given as follows:

$$m_{i\cdot} = \frac{M_{i\cdot}}{P_i} k$$

where $m_{i\cdot}$ = rate of migration from *i* to all other destinations during a particular time interval;

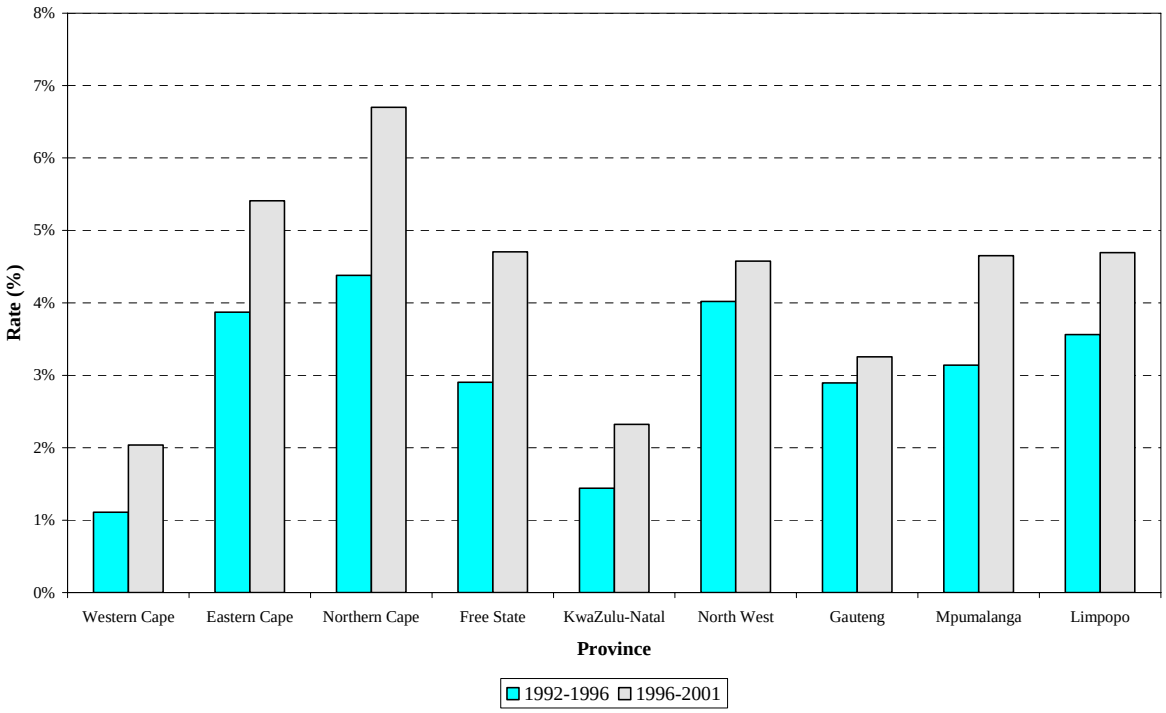
$M_{i\cdot}$ = number of migrants moving from *i* to all other destinations during the given time period;

P_i = the population concerned (in *i*) at the end of the given time period (10 October 1996/2001), and

k = constant (100).

The provincial out-migration rates for the periods 1992–1996 and 1996–2001 are illustrated in Graph 7.

Graph 7: Inter-provincial migration – provincial out-migration rates (1992–1996 and 1996–2001)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996/2001)

Graph 7 shows that the Northern Cape experienced the highest out-migration rates in both periods (4,4% during 1992–1996 and 6,7% during 1996–2001). The Eastern Cape, which had the third highest out-migration rate (3,9%) during the period 1992–1996, lost the second highest proportion of migrants during the period 1996–2001 (5,4%). The other provinces with comparably high out-migration rates were North

West (4,0% during 1992–1996 and 4,6% during 1996–2001), Limpopo (3,6% and 4,7% respectively) and Mpumalanga (3,1% and 4,6% respectively).

In-migration rate

For the *in-migration rate*, the following formula applies.

$$m_{\cdot j} = \frac{M_{\cdot j}}{P_j} k$$

where $m_{\cdot j}$ = rate of migration from all other origins to the destination *j* during a particular time interval;

$M_{\cdot j}$ = number of migrants moving from all other origins to the destination *j* during the given time period;

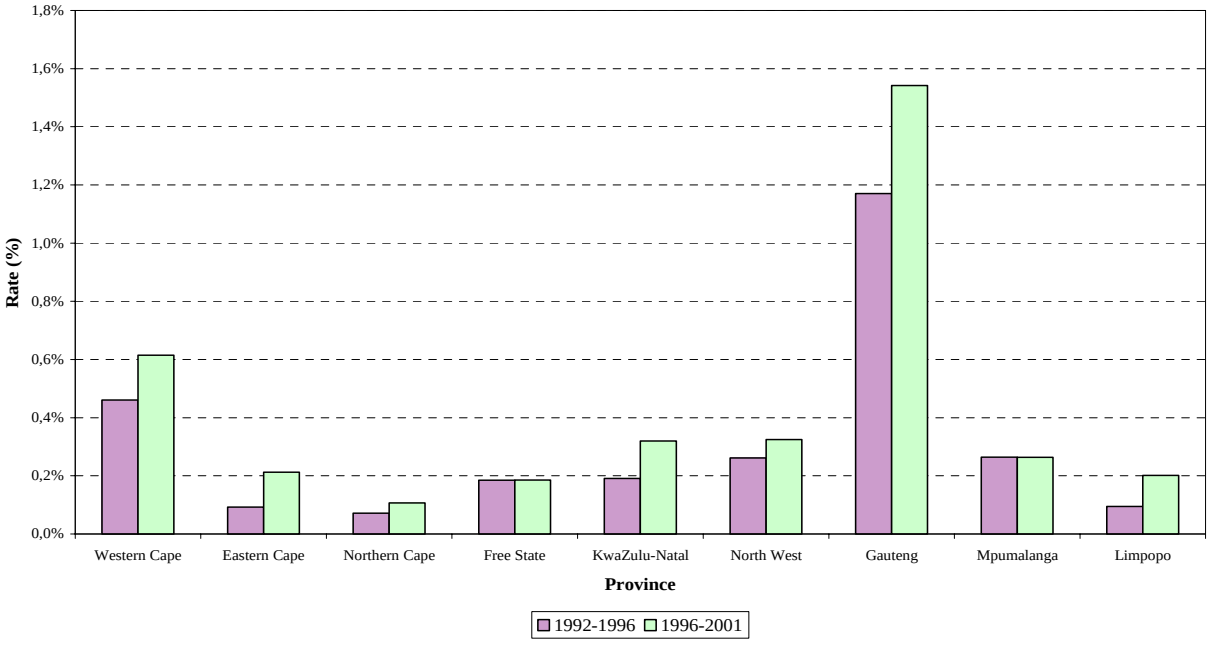
P_j = the population concerned (in *j*) at the end-point of the given time period (in this case, 10 October 1996/2001), and

k = constant (100).

Graph 8 below illustrates the provincial in-migration rates for the periods 1992–1996 and 1996–2001.

The relatively large impacts of in-migration on the total population of Gauteng in the two periods 1992–1996 and 1996–2001 are clear from Graph 8. It should, however, be borne in mind that the absolute impacts were not all that high (1,2% and 1,5% respectively for the two periods). Nevertheless, compared to the other provinces, the population impact of in-migration into Gauteng was by far the greatest.

Graph 8: Inter-provincial migration – provincial in-migration rates (1992-1996 and 1996-2001)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996/2001)

This conclusion should, however, be treated with caution. Only when one considers the overall net migration rate can the true impact on the province's population be properly evaluated.

Net migration rate

The formula for the *net migration rate* is as follows:

$$nm_i = \frac{M_{\cdot i} - M_{i\cdot}}{P_i} k$$

where nm_i = net migration rate in respect of area *i* during a particular time interval;

$M_{\cdot i}$ = number of migrants moving from all origins to *i* during the given time period;

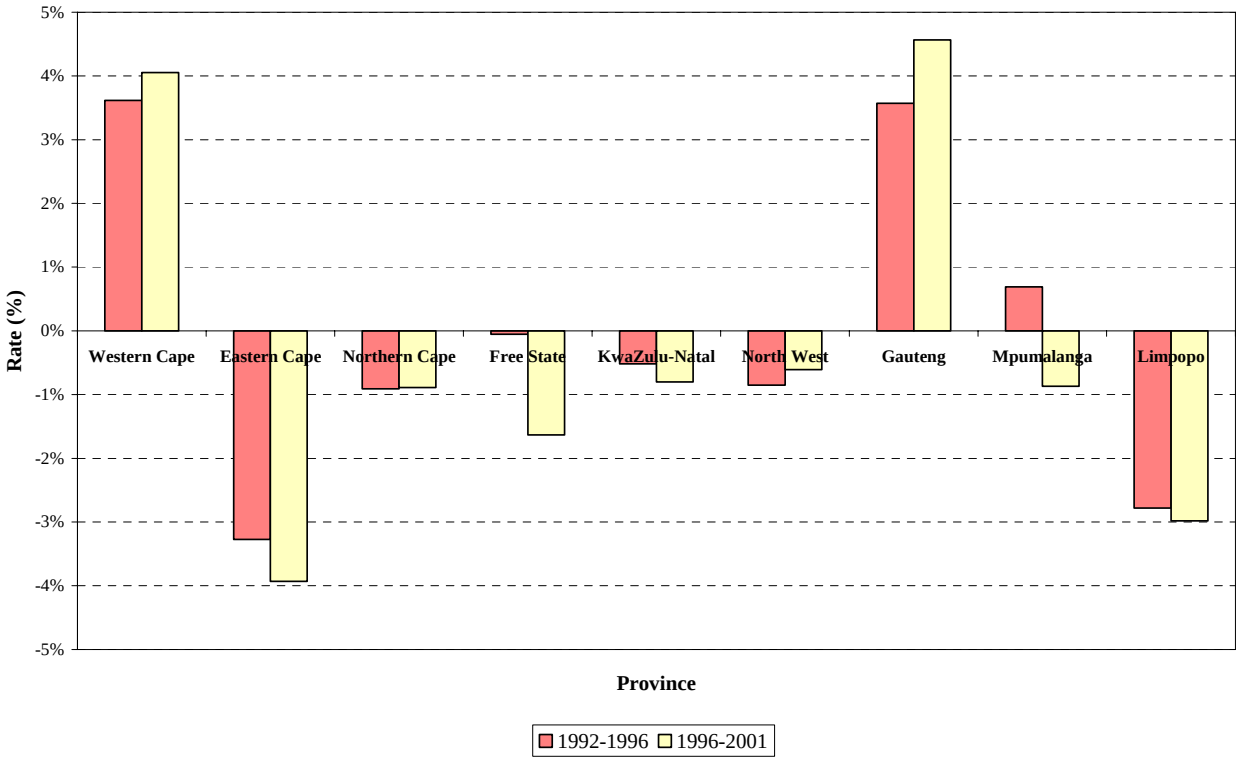
$M_{i\cdot}$ = number of migrants moving from *i* to all destinations during the given time period;

P_i = the population concerned (in *i*) at the end point of the given time period (in this case, 10 October 1996/2001), and

k = constant (100).

Arguably, the most significant of the migration rates, the net migration rates for the different provinces, illustrate the differential net impacts on their populations. Graph 9 illustrates these provincial impacts for the periods 1992–1996 and 1996–2001.

Graph 9: Inter-provincial migration – provincial net migration rates (1992–1996 and 1996–2001)



Source: Calculated from Migration Community Profile data provided by Statistics South Africa (Census 1996/2001)

Graph 9 clearly shows the net effect of migration from other provinces into Gauteng. In 1996, Gauteng accommodated a total population of more than seven million. By 2001, this number had increased to nearly nine million. Over five years, Gauteng's total population therefore rose by just under 1,5 million people, a massive growth that strains government delivery capacity and puts stress on the province's economy. Gains from in-migration (403 000 persons) accounted for more than a quarter of this total increase.

The net effect of in-migration into the Western Cape, which during the period 1992–1996 experienced a net migration rate (3,6%) that was slightly more than Gauteng's (3,6%), was somewhat less (4,0%) than that of Gauteng (4,6%) during the later period (1996–2001). These differences are, however, virtually negligible, and suffice it to note that Gauteng and the Western Cape both experienced notable net in-migration during both periods.

6.8 Summary and conclusions

The 1996 and 2001 censuses offer a unique opportunity to study internal migration in South Africa. The problems of inherently flawed census questions and the non-availability of information for some areas no longer exist. The migration studies that can be undertaken on the basis of the 1996 and 2001 census data are numerous and potentially crucial. Although a slight redefinition of migration is needed theoretically to analyse the census data correctly, opportunities exist for migration research, including the analysis of migration trends, that have never been possible before. Some examples of the analyses that can be carried out on the basis of the data generated by these two censuses have been provided. Unfortunately, these analyses could be undertaken only at a provincial level due to the unavailability (at the time of writing) of the detailed migration data from Census 2001. As soon as the detailed migration data become available, better spatial analyses can be undertaken.

The only remaining issues are those related to possible quality problems in the 2001 data. An evaluation of the 1996 migration data showed acceptable reliability in terms of levels and propensities (Kok et al. 2003), but no such detailed evaluation has so far been done in respect of the 2001 data. If notable problems with the quality of the data arise, one can only hope that these problems will be more or less uniformly spread throughout the country. If some areas have worse problems than others, however, the data quality issue could become quite serious, especially for more localised or detailed analyses. But on a positive note, and until proven otherwise, it must be accepted that the 2001 census data on migration will also be above suspicion. If that proves to be the case, Statistics South Africa has done researchers and the population at large a particularly great favour. At long last, the hope is high that some understanding of the migration processes in South Africa will be achieved!

Migration research in South Africa has received a major boost since the data from the 1996 and 2001 censuses have become available. Nevertheless, migration researchers will still have to undertake complementary empirical and other types of single-purpose migration studies. Such studies are needed to obtain a clearer understanding of what caused the processes that are highlighted by the censuses and also to indicate and predict what the most likely implications are for the country and its people.

Bibliography

Bretz, M (1996) Migration statistics in Germany: Sources, concepts and selected results *Espace Populations Sociétés*, No. 2-3: 247–356.

Collinson, M, Mokoena, O, Mgiba, N, Kahn, K, Tollman, S, Garenne, M, Herbst, K, Malomane, E & Shackleton S (2001) Agincourt DSS, In Sankoh, OA, Kahn K, Mwageni, E, Ngom, P & Nyarko, P (eds) *Population and Health in Developing Countries (Volume 1: Population, Health, and Survival at INDEPTH Sites*, Chapter 16. [See URL http://web.idrc.ca/en/ev-43018-201-1-DO_TOPIC.html; last accessed on 30 June 2004.]

Gelderblom, D & Kok, P (1994) *Urbanisation: South Africa's challenge (Volume 1, Dynamics),* Pretoria: HSRC.

Graaff, JF de V (1986) *The present state of urbanisation in the South African homelands and some future scenarios* (Occasional Paper No. 11), Stellenbosch: Research Unit for Sociology of Development.

Kok, P, O'Donovan, M, Bouare, O & Van Zyl, J (2003) *Post-apartheid patterns of internal migration in South Africa,* Pretoria: HSRC.

McHugh, KE, Hogan, TD & Happel, SK (1995) Multiple residence and cyclical migration: A life course perspective, *Professional Geographer*, 47(3):251–267.

Morrison, PA (circa 1980) Introduction and overview, In *Population movements: their forms and functions in urbanization and development* PA Morrison (ed.), Liège: International Union for the Scientific Study of Population (IUSSP).

Shaw, RP (1975) *Migration theory and fact (*Bibliography Series No. 5), Philadelphia: Regional Science Research Institute.

Shryock, HS, Siegel, JS & Associates (1976) *The methods and materials of demography* (condensed edition by Edward G Stockwell), Orlando: Academic Press.

Skeldon, R (1990) *Population mobility in developing countries: a reinterpretation* London: Belhaven.

Standing, G (1984) Conceptualising territorial mobility, In *Migration in low income countries: guidelines for survey and questionnaire design* Bilsborrow, RE, Oberai, AS & Standing, G (eds.), London: Croom Helm.

United Nations (1970) *Methods of measuring internal migration* (Manuals on methods of estimating population, Manual VI, Population Studies, No. 47), New York: UN Department of Economic and Social Affairs.

Urban Foundation (1990) *Population trends* (Policies for a New Urban Future, Urban Debate 2010, No. 1), Johannesburg.

Chapter 7: The 2001 Census and education in South Africa

7.1 Social interest in the 2001 census as an education data source

Internationally, public provision of educational services is viewed as an indicator of the accountability and the capacity of government to see to the welfare of its citizens. Education is a major sector within which the state commits financial and human resources in the interests of social equity and economic development. Through census information, education issues and dimensions of educational change can be identified and tracked. Census information related to education in any national context is accordingly of considerable interest.

In South Africa, the publication of census data for 2001 is of particular interest for historical reasons. Education was a primary vehicle for systematic, racially based unequal treatment of people under the apartheid order. After 1994, with the advent of democratic government, access to education was enshrined in the Constitution as a basic human right, stated as follows: “Everyone has the right to a basic education including adult education and to further education which the state through reasonable measures must make progressively available” (Constitution of the Republic of South Africa Act No. 108 of 1996). Education is regarded as a key element in the strategy of government to bring about social transformation. This implies improving access to educational opportunities for the historically disadvantaged and implementing policies that focus on redressing the inequities of the past. Policy makers, planners and managers will therefore be interested in the census data in order to shed light on the dimensions of the ongoing education challenges facing government in meeting the needs of the population, in ameliorating inequity and in facilitating economic growth.

7.2 Census data on education and South Africa’s political history

Census data are not necessarily value free. The methods used to obtain census data and the categories according to which the data are collected and analysed will reveal the values of those who design the census instruments.

In the past, census statistics codified racial differences and thereby served as a form of knowledge that could be used to justify unequal treatment of people labelled as “black” or “coloured” or “Indian”. The history of South African census activity shows that data were collected by colonial governments on a racial basis from as far back as the 19th century and that this information was employed in the move to segregate schools formally in terms of the Smuts Act in the Transvaal and the Cape School Board Act in the Cape Colony of 1903 and 1905 respectively (Paterson 1992). Although racial categories were the creation of colonialism and apartheid, these

labels must still be used for the foreseeable future in order to monitor the pace of social change in fields such as education.

The spatial, administrative and political structure of apartheid distorted the collection and organisation of national statistical information thereby casting considerable doubt on the reliability of statistics gathered in that period. Territorial segregation in the form of the homeland system under apartheid had divisive effects on data gathering and analysis. For example, the 1991 census had serious limitations with respect to information on the former TBVC (Transkei, Bophuthatswana, Venda and Ciskei) states, which were not included in the administration of the census by the former South African Central Statistics Services (CSS). These territories conducted census operations independently. Data for education from the TBVC territories were incomplete and inconsistent. Furthermore, the process of adjusting for undercount in the TBVC states was a matter of controversy with regard to previous South African censuses. In 1991, no adjustments were made for the TBVC states (Krige et al. 1994:7). Reliable territorial analysis based on 1991 census data was consequently highly problematic for education analysts.

The 1991 and older data sets were, in effect, the historical reflection of a quite different political and spatial arrangement, namely the old provincial and homeland spatial systems of territorial jurisdiction. A new system of nine provinces was implemented in South Africa in 1994 with the advent of democracy. The 1996 census thus presented the first opportunity to examine the educational characteristics of South African society from within the new provincial governance framework. Researchers who seek to examine changes in the trajectory of educational development in the South African population by using pre-1996 census data will need to take account of the discontinuities identified above.

For the purposes of this chapter, the focus will mainly be on the 1996 and the 2001 censuses, which do not reflect the kind of data unevenness that derived from the spatial fragmentation and disjunctures of the apartheid era. The 2001 census data therefore present an opportunity to investigate change over time based on the 1996 benchmark. The following censuses of the new millennium will be closely monitored in order to assess the impact of equality, democracy and redress on South African society.

7.3 Education and training legislation and the 2001 census

Major legislation related to school education and skills development has been enacted on an almost continuous basis since 1995. Complementary layers of legislation have addressed the structure of the curriculum, an integrated qualifications framework, institutional governance and funding, equitable access to education and the employment of educators (Department of Education 1995a,

1995b, 1996a, 1996b, 1996c, 1997a, 1997b, 1998a, 1998b, 1998c). Policy has given particular attention to the challenges of stimulating skills development and training activities in the workplace through the implementation of a compulsory levy grant scheme for employers (Department of Labour 1998,1999,2001).

However, policy does not automatically translate into social and economic change. Intervening factors can constrain implementation and restrict the success of policy while unintended consequences may also flow from policy prescriptions. Furthermore, policy is necessarily subject to revision on the basis of assessments regarding its effectiveness in implementation. A degree of caution is therefore called for when making judgements about the impact of policy on education and training outcomes or about the impact of educational improvements on other facets of the social reality such as employment.

Nevertheless, many observers will be keen to assess the success or otherwise of education policy in the democratic period as recorded in the census. Two instances only will be cited here:

- The proportion of the population aged 20 and older with Grade 12 or higher levels of education rose from 22,6% in 1996 to 28.8% in 2001 (Statistics South Africa 2004a:143–144; Statistics South Africa 2004b:35)
- The proportion of the population aged 20 and older with higher education increased from 6.2% to 8.4% (Statistics South Africa 2004a:143–144; Statistics South Africa 2004b:35)

These data reveal that the broad educational characteristics of the South African population have improved since 1996. However, we should realise that national education systems change slowly in their institutional forms, practices and effects. Moreover, improvements registered on an aggregate level may mask deepening disparities in conditions experienced between social groups or between spatial categories (e.g. provinces, rural-urban). The unit of analysis chosen by the researcher will influence the nature of his/her findings, which need to be understood in relation to the bigger – provincial or national – picture.
This chapter will now turn to examine the range of education-related data collected in the censuses of 1996 and 2001.

7.4 Education-related information in the 2001 census

The extent to which the census offers insights into the education-related characteristics of the population of South Africa depends on the set of questions asked in the questionnaire. The 1996 and 2001 censuses are dealt with in some detail here as they cover the post-1994 period.

The groups of variables covered in the 1996 census are first described and then changes in the 2001 census are detailed.

Data describing the following education-related characteristics are contained in the 1996 census. The list below is not meant to be exhaustive – it seeks merely to identify the main population characteristics usually associated with education-related analysis.

Where the population is located:

- Where the population is concentrated or dispersed, for instance in urban or rural areas.
- Changes in the distribution and dispersal of the population over time.
- Distribution of population characteristics such as age, gender, language and race group.
- Attendance at education institutions.
- Mode of educational delivery – part time or full time.
- Distribution of education in the population by level achieved.
- Distribution of income in the population.
- Disabilities in the population.
- Religions in the population.
- Employment in the population.

Education-related questions absent in 1996 but added in 2001 are listed below (Cronje & Budlender 2004:71) (comments in brackets are by the author of this chapter).

Travel to work or school. (Information on travel is important because travel time will affect the amount of time the learner has available to spend on homework, and travel costs will be taken from household disposable income.)

- Household goods such as radios, refrigerators, televisions, telephones in the dwellings as well as computers and cellphones in working condition. (The presence of these goods can be used as indicators of socioeconomic status. In particular, home exposure to computers can influence the subject choice, career choice and employment opportunities of individuals.)
- Deaths in the household in the past 12 months. (This information can be used as an indicator of the impact of HIV/AIDS on households. Mortality will affect households with school-going learners and may be linked to the existence of child-headed households and so on.)

Education-related questions asked in 1996 but omitted in 2001 are listed below (Cronje & Budlender 2004:72) (comments in brackets are by the author of this chapter).

- Whether an individual worked full or part time. (The employment status of the employee is important because employers tend to provide much higher rates of training to full-time than to part-time workers.)
- Main duties or activities performed at work. Respondents were required only to provide their occupational title. This information alone does not provide sufficient information to enable accurate coding of occupations. Consequently, a second question – on the main duties or activities performed at work – is needed to enable more accurate post coding but was omitted in the 2001 census thus reducing the quality of the occupational information. As a result, any analysis of the links between educational achievement and sub-occupation category, for example, will be less reliable.

In addition to the removal and insertion of entire questions, some questions were expanded or reduced in terms of their coverage. Cronje and Budlender (2004:73) identify a number of instances where education-related questions were expanded or reduced. In 1996, the questionnaire enquired not only about the language spoken most often in the household but also about other languages spoken at home. In 2001, the questionnaire enquired only about the language spoken most often in the household. (This means that the extent of multilingualism in households in relation to educational characteristics cannot be directly ascertained.)

- Questions on postschool education were changed between 1996 and 2001. On this issue, Cronje and Budlender (2004:73–74) observe:

> In Census 1996, respondents were asked to name the(ir) highest qualification and the question was a 'write-in' one which had to be post-coded. Those responsible for the post-coding experienced considerable difficulties because of the ways in which respondents reported qualifications. The 2001 questionnaire asked respondents to mark off which of 22 learning learning areas (was) the relevant one or best describe(ed) the person's highest post school qualification. This approach (had) the advantage of avoiding post coding. The disadvantage (was) that while this categorization (was) used by all universities in their annual reporting to the Department of Education, it (was) not widely known by the public. Further, the names of the learning areas are not always intuitively easy to understand. One indication of the difficulties in this respect (was) that some of the translators of the questionnaires were not able to find appropriate translations in their language and left the names in English. The change in the way this question was asked might have created difficulties for respondents and enumerators in completing the question in Census 2001.

(As a result of this change, data on highest post-school qualification, disaggregated by learning area, must be treated with circumspection.)

7.4 Analytic limitations of the 2001 census

Of relevance here is the Report of the Census Sub-Committee to the South African Statistics Council on Census 2001, which indicated that the 2001 census probably reflected an

- underestimation of the number of children below age five;
- overestimation of the number of teenagers aged between 10 and 20;
- underestimation of the number of men relative to the number of women;
- underestimation of the size of the white population;
- overestimation of the extent of unemployment. (Statistics South Africa 2004b:3)

These likely inaccuracies are noted as they may have some effect on researchers' analysis of education-related questions.

This chapter will now examine various options for analysis open to researchers, education managers or any other users of census data.

7.5 Description of demographic trends that have an impact on education

The demographic characteristics of a population are key to understanding education access. In the first instance, the researcher would want to examine the extent of growth or decline in the proportion of the population between birth and school-going age in order to predict gross increases or decreases in school enrolments over time. In this regard, population pyramids give a graphic picture of the prospective demand for schooling at the compulsory levels that must be taken into account (Statistics South Africa 2001:30). In order to predict longer term changes, the trajectory of population growth can be extrapolated by modelling trends observed in previous census reports.

Overall increases or decreases in the population need to be described spatially in order to plan the expansion and location of school facilities properly within the norms for accommodation prescribed by government. Information on the distribution of the population and on changes in the distribution of the population is therefore needed. For instance, maps can be drawn of particular age cohorts to indicate where future pressure on school accommodation is most likely to occur.

Change in population distribution is the consequence of migration or of changes in the ratio of births and deaths. The former phenomenon was marked historically by

apartheid policies, which by various means set out to limit African urbanisation and which contrived to locate African population concentrations in the predominantly rural former homelands. Because of apartheid policy, distinct racial patterns of population distribution are evident, but this situation is changing. Constitutional protection of freedom of movement in the democratic era since 1994, along with the trend towards urbanisation that is typically associated with developing societies, has made urbanisation a dominant feature of current demography.

The overall effect of these trends is a rapidly growing and increasingly younger urban African population. The combination of rapid population growth and in-migration to urban areas, which Census 2001 seems to confirm, presents significant challenges in terms of educational provision. Concentrations of population in urban and rural areas have different effects on the cost of providing educational facilities and opportunities of equal quality. Fine-grained analysis of these complex phenomena will be of value in the planning and monitoring of educational access.

A question added to the 2001 census requested data on ”Deaths in the household in the past 12 months”. This question attempted to elicit information that could be used – indirectly – in understanding the impact of HIV/AIDS at the household level. HIV/AIDS is an extremely difficult and sensitive phenomenon to investigate given the nature of the syndrome, the stigmatisation of those infected, the sensitivity surrounding disclosure, the research ethics issues and the hotly contested views on how government could and should be responding to the HIV/AIDS crisis. For these reasons, other studies (e.g. Shisana 2002) should be consulted for a better understanding of the relationship between HIV/AIDS and education. However, considerable methodological and analytic circumspection should be exercised in relating census data to data on HIV/AIDS from any other source.

In this chapter, the concern is with the impact of HIV/AIDS (Vass 2003:201; Crouch & Perry 2003:494) on

- the quality and quantity of the human resources supply (including teachers) to the South African labour market;
- teachers in the classroom who have to cope with high learner-teacher ratios and with the pastoral needs of learners affected by AIDS;
- learners who may have lost parents, are AIDS orphans, are child household heads or are themselves HIV positive (Brookes, Shisana & Richter 2004).

Clearly, HIV/AIDS can be demographically ”measured” in terms of mortality, but its impact is far wider as will be seen in the themes discussed below.

Changes in the educational characteristics of the population aged 20 years and older

Census 2001 enables the user of the census data to establish the extent to which access to various education levels (basic schooling, intermediate education and higher education) has percolated through to the population as a whole. Given the historical legacy of apartheid discrimination, it is important to track developments and to redress imbalances and inequities.

We know that, overall, the proportion of the population aged 20 and older with Grade 12 or higher levels of education rose by 6,2% between 1996 and 2001 and that the proportion of the population aged 20 years and older with higher education increased by 2,2% in the same period (Statistics South Africa 2004a:143–144; Statistics South Africa 2004b:35). Racially disaggregated data will reveal differences in the proportion of members of different population groups who have achieved Grade 12. It will be useful to create maps showing the education levels achieved by the population. For example, higher education can be treated by generating maps to show the number of people with degrees per magisterial district or per thousand of the population. Again, disaggregation of the population by race will help identify disadvantaged groups.

Information on levels of education can be analysed by gender to determine inequalities between males and females in respect of access to education (Statistics South Africa 1998:2.26). The 2001 census data reveal a striking pattern where 51,9% of those aged 20 and older with a higher education qualification are women and, at the same time, where 59,9% of those aged 20 years and older without any schooling are women (Statistics South Africa 2003a:46). It will be interesting to explore this pattern with reference to the race and socioeconomic class characteristics of these populations.

The 2001 census data can be tapped to develop a picture of the proportion of the adult population that has had limited access to education. Disaggregation and mapping of these data according to race will facilitate the identification of groups with low levels of education and highlight areas where significant numbers of adults require education provision in terms of adult basic education and training (ABET). For example, the percentage of the population aged 20 years and older with no education was 17,9% in 2001 at the national level. However, the percentage varied significantly by province from 33,4% and 27,5% in Limpopo and Mpumalanga respectively to 5,7% in the Western Cape (Statistics South Africa 2003a:44). Clearly, ABET needs to be looked at especially – though not exclusively – in the former two predominantly rural provinces. Further disaggregation of these data at the magisterial district level will facilitate the targeting of ABET programmes.

Capacity to support education

The capacity of households to support members engaged in education is likely to affect directly the extent of educational access in the population and also the extent to which those with access to education can take advantage of this access.

Demographic analysis can add value to our understanding of the household context in which school-going learners live. Some approaches to the question of household support are presented below.

First, we can approach the question of household support on the basis of gender. By generating maps that show males as a percentage of the total population by magisterial district, areas with relatively low proportions of males can be identified. Areas that exhibit high levels of male absenteeism because of migrant labour and other factors are characterised by significant numbers of female-headed households that have to shoulder the responsibility of ensuring household sustainability and the education of children.

Secondly, we can address the question of household support on the basis of the educational background of the household head (Statistics South Africa 2001:31–33). The educational background of the household head has important implications for her/his ability to support students in the household, especially where these students have achieved higher educational qualifications than the household head.

Thirdly, we can address the question of household support on the basis of the economic resources the household has at its disposal to support educational activities and to absorb the attendant opportunity costs. Pressure is increasing on parents or guardians to contribute to the cost of education such as fees and other direct costs such as books and uniforms. Simple measures of the capacity of households to support learners financially can be extracted from the 2001 census. One such measure is the annual per capita income. It may be useful to examine racially disaggregated data for this measure given the disparities between white and black South Africans in terms of income levels. Maps of income distribution are likely to reveal areas where low per capita incomes and poverty are most prevalent. It is in these areas, of course, that the educational chances of learners are most at risk.

Another simple measure of household economic sustainability could be derived by mapping the proportion of the population aged 18 to 65 years to ascertain the number of potentially economically active employed people in an area who could support a child at school. More sophisticated measures of the capacity of households to support learners could be developed on the basis of poverty indicators such as the minimum living levels (MLLs) calculated by the Bureau for Market

Research and applied by Krige et al. (1994:48) to the 1991 census data. It should be possible to update this form of analysis with inputs from the 2001 census data.

Education and labour market linkages

A critical index of educational functioning is the link between education and the economy although the positive effects of increased access to education do not necessarily translate into improved social and economic conditions. McCord and Bhorat (2003:137) observe: “Education is a particularly important determinant of success in the labour market process in the context of employment rationing. The accumulation of education, however, is not in itself a sufficient condition for improving employment prospects”. Technical and structural changes in the South African labour market over some decades have produced rising unemployment with particularly poor employment growth for unskilled workers in the primary sectors where African labour is mainly concentrated (McCord & Bhorat 2003:137). A major challenge for South Africa is to create jobs for the growing population.

The question is how can we use census data to achieve a better understanding of the apparent mismatch between the skills demands of the economy and the supply of appropriately skilled workers. At the outset, we should acknowledge the significant difference between the 2001 census measure of unemployment at 42% and unemployment as estimated at 29,5% in the 2001 Labour Market Survey, which is also produced by Statistics South Africa. This difference is attributed to a more precise – and internationally standardised – definition of unemployment adopted in the Labour Market Surveys, which excludes the survivalist economic activity of people who, some would argue, are effectively unemployed. The choice as to which data source is used for analysis will fundamentally affect the results achieved. Therefore, the researcher should be acutely aware of the consequences of the choice of data source.

Having made the decision to pursue analysis using the census, the researcher may elect to address a variety of questions – a few options are entertained below.

Data relating to employment and unemployment levels among people who possess particular educational qualifications can be drawn from the 2001 census. These data can be used to analyse the capacity of the labour market to absorb skilled labour or to analyse the capacity of the education system to equip students with appropriate skills. Disaggregation by race can assist in determining the extent to which labour market segmentation on the basis of race has taken place. This type of analysis can be performed by broad occupational category and by economic sector.

The census data also give us the opportunity to deepen our understanding of the training and employment prospects of people who have had no formal education.

Observation across a range of contexts has shown that working adults who possess higher level qualifications are much more likely to participate in training than those with lower level attainments (O'Connell 1999:2; Ok & Tergeist 2003:3–15). This means that "workers who are disadvantaged by poor educational backgrounds" may be exposed to a "vicious cycle of lack of opportunity and neglect" in respect of their own workplace training needs (Paterson & Du Toit 2004:62) – a serious problem given the history of disadvantage of black South Africans.

The census data show that the proportion of the population aged 20 and older with no schooling declined from 19,3% in 1996 to 17,9% in 2001 (Statistics South Africa 2004a:143–144; Statistics South Africa 2004b:35). International experience suggests that those in this group who are employed are less likely to receive training opportunities than their skilled counterparts. However, Paterson and Du Toit (2004:62) show that in South Africa some progress is being made in improving the skills levels of formerly disadvantaged groups where the training ratio among African workers is higher than that for other race groups. Africans in managerial and elementary occupational categories, in particular, are the main beneficiaries of increased access to workplace training. In 2003, 26% of Africans received some form of training while 23% of coloureds, 18% of Indians and 23% of whites received some form of training. Although South African employers are responding to the training needs of workers who have no schooling, it may be useful to consider the general distribution of those with no schooling and how ABET or other forms of training can be planned for this group.

Labour markets and work environments are changing under the multiple influences of technology, global competitiveness and macroeconomic shifts towards service industries. In some economic sectors, workers are under pressure to reconfigure their skills to remain competitive. Pillay (2003 108–109) has recently observed that in South Africa, skills requirements are rising in sectors as diverse as banking and forestry. The Census 2001 data could be used to explore skills profiles by occupational categories in order to better understand sectoral labour markets.

Finally, a central pillar of South African education and training policy is the concept of "lifelong learning", which stresses the importance of continuous learning in the workplace. Further analysis of the levels of enrolment for part-time as opposed to full-time studies could reveal interesting trends in this growing area of educational provision (Statistics South Africa 1998: 2.27–2.28).

7.6 Themes

Education is an extremely broad field because of its links to social and economic change. Accordingly, the researcher could not identify, let alone address in any meaningful way, the multiple approaches to education analysis on the basis of

census data. Three themes that can, however, be investigated using the 2001 census data are briefly noted below.

Special education needs

In line with government recognition of specific inequalities including access to education experienced by particular disadvantaged groups such as women, youth and those with disabilities, the 2001 census specifically provides for the collation of data on disabilities by age, race and gender (Statistics South Africa 2003a:38–40). Data on special education needs can be related to information on facilities currently available. The transformation of centres of learning and the infusion of ”special needs and support services” constitute an important area of research and planning (Department of Education 1998b). Although the Department of Education (DoE) gives special emphasis to the mainstreaming and breaking down of barriers to learners with ”special needs”, specialised centres will nevertheless still be necessary. Research could also be conducted on levels of access to education and the educational achievement of the disabled so that more effective targeting of educational resources can be undertaken.

Language

In terms of the South African Schools Act (1996:8), school governing bodies may determine language policy subject to the Constitution and with the provision that no form of racial discrimination may be practised in implementing such policy. Language plays a crucial role in the teaching-learning process. In particular, the language of learning has implications for the educational progress of learners depending on whether or not the language of learning is their mother tongue. Access to education where the mother tongue is the medium of instruction is therefore crucial.

Maps can be constructed from the census to identify the numerically dominant and minority languages in a particular province or census district. Although in most provinces one language is dominant, often more than one minority language is spoken by sufficiently large numbers of people to affect planning for the provision of education in the mother tongue. Accordingly, information on the spatial distribution of languages and the proportions of the languages spoken is needed to cater for multilingualism and the possibility of instituting dual-, parallel- and multi-medium schools.

Religion

The 2001 census contains information on the religious affiliations of members of the population, which may have educational implications. The Constitution provides for freedom of conscience and religion at public schools, and the South African Schools Act (1996:8) makes provision for the expression of religious observances at public schools subject to the Constitution ”if such observances are conducted on an

equitable basis and attendance at them by learners and members of staff is free and voluntary". A Ministerial Committee on Religious Education entitled "Religion in Curriculum 2005" presented its report in 1999. However, because religious beliefs are often closely linked to notions of appropriate educational forms, religious freedom and the historical association between religious denominations and denominational schools remain sensitive issues.

7.7 School enrolment and "out of school" youth

School enrolment is affected by a number of factors that include fertility and mortality rates in the population, the quality and efficiency of schools, policies on age-grade norms and promotion, and the ability of the household to support school age learners. These factors, in turn, influence the rates of repetition, drop-out and drop-in, which contribute to the overall flow-through or throughput rates of the schooling system.

"Out of school" youth – or children of school-going age who are not in school – was an endemic problem, particularly under the various apartheid educational structures that administered black education. Monitoring the "out of school" population served as an index of the holding capacity of schools.

Two standard education indicators of educational enrolment are commonly used (see Perry and Arends (2003) for an application of these indicators):

- Gross enrolment ratio (GER) is a measure of access and coverage – it indicates the proportion of the population concerned that has access to the school system. The GER thus reflects the percentage of children in a particular age range who are accommodated in schools.
- Net enrolment ratio (NER) is a measure of the efficiency of the system – it indicates only those learners who are the appropriate age for the relevant grade.

The Census 2001 data enable the researcher to calculate the GER and also the size of the "out of school" population. A comparison between 1996 and 2001 showing the proportion of the population attending school by single-year age categories indicates an interesting pattern. Basically, the proportion of the population aged 5 to 13 attending an educational institution increased between 1996 and 2001 whereas the proportion of the population aged 14 to 18 attending an educational institution decreased in the same period (Statistics South Africa 2004b:38, Figure 9). This suggests that although the DoE has been successful in increasing the "capture" of school age learners in school up to the age of 13, from age 14 onwards, the proportion of appropriately aged youth in the post-compulsory further education and training phase is visibly declining. The reasons for such anomalies are of particular

importance to education managers as is the opportunity to compare the Census 2001 figures with enrolment data emanating from the DoE itself.

7.8 Department of Education data

It is important to look beyond analysing the census data on their own and to consider what other sources of education data may be available to enrich analysis. The census serves as a data source from which policy-relevant statistics can be extracted and combined or linked with other data to reveal the performance of the education system or its subsystems.

Other than the census, the primary source of data on education is, of course, the national Department of Education (DoE). The DoE depends on the collection of comprehensive information on the key characteristics of its students, educators, curricula, institutions, facilities, management, and support and administration infrastructures. The DoE needs such information to plan, manage and monitor implementation as well as assess the impact of policy in the education sector. To meet these information needs, the DoE has developed a number of strategic surveys of which two are particularly valuable for use in conjunction with Census 2001 data:

- Annual Schools Survey
- School Register of Needs

The Annual Schools Survey, as its name suggests, is conducted annually in all of the approximately 27 500 ordinary public and registered independent schools in South Africa. Through this survey, the DoE has produced data on the number and distribution of schools, learners and educators overall and according to grade and school phase. The DoE has also developed key indicators such as the learner to educator ratio, the learner to school ratio, the gross enrolment ratio (GER) and the Gender Parity Index. These data are described at the provincial level in the publication *Education Statistics in South Africa at a glance in 2001*" (Department of Education 2003), published in the same year as the 2001 Census.

In 1996, the School Register of Needs (SRN) was launched by the DoE in order to obtain an empirical basis for the rational planning of educational infrastructure based on principles of equity and redress. The SRN culminated in the development of a comprehensive geographically referenced database in which all schools and educational institutions in the country were mapped together for the first time. The database included information on the geographical/spatial location of each school, the nature and condition of facilities and equipment, teaching resources and the services provided by each institution. In keeping with the commitment of the DoE to update the SRN database, the exercise was repeated in 2000. The SRN offers detailed information on the quantity and quality of school resources (e.g. classrooms and educators) and infrastructure (e.g. from provision of water and electricity to

computers) on a national and provincial basis. Although the SRN 2000 was conducted the year before Census 2001, it offers a wide range of infrastructural information for analysis purposes. These data are available in the Report on the School Register of Needs 2000 Survey (Department of Education 2002).

The Annual Schools Survey of 2001, and to a somewhat lesser extent the SRN 2000, provides comprehensive data that are comparable with the Census 2001 data. However, in the census, the basic unit of data collection is the household, and the basic unit of analysis is the enumerator area, whereas in the Annual Schools Survey and the SRN, the basic unit of analysis is the school. This difference does not impede analysis at provincial or national levels, but, at lower levels of disaggregation, the researcher will have to decide how to relate the census demographic data to the DoE's education data. Using Geographic Information Systems (GIS), the researcher can relate layers of education data (by school location or education district/circuit/region) to demographic data (by enumerator area or magisterial district).

For example, for planning purposes, information on adults who have no schooling per enumerator area can be linked to data that show the location of the nearest school providing adult literacy classes. Alternatively, the researcher can focus on individual constellations of schools in particular localities in order to develop plans for improving education access such as the location of new schools or the allocation of teachers offering scarce subjects.

7.9 Conclusion

The limitations of this chapter are such that a full treatment could not be given of the complex ways in which education and development are linked. Much of the analysis discussed here is based on relatively simple descriptive statistical techniques in order to provide a basic overview of the key issues and possible relationships between key variables and concepts.

After acknowledging the census as a major quantitative data resource, it is appropriate also to acknowledge the value of other forms of social scientific data collection and analysis that can enrich data drawn from the census. It is important to refer to other research approaches such as qualitative methodologies and to institutional studies that assist the social scientist to understand the education reality.

The 2001 census data on education reflect the democratic South African dispensation and its achievements. Perhaps the key achievement in educational terms are the higher levels of enrolment that have been attained in South African schools and higher education institutions in the period since 1994. The most critical factor to look out for in the next census will be the rate of employment creation and

the evolving relationship between employment creation and skills development in the South African population.

Bibliography

Brookes, H., Shisana, O. and Richter, L. (2004) *The national household HIV prevalence and risk survey of South African children*, Human Sciences Research Council, Pretoria.

Chisholm, L. (ed) (2004) *Changing class: Education and social change in post-apartheid South Africa,* Human Sciences Research Council, Pretoria.

Cronje, M. and Budlender, D. (2004) Comparing Census 1996 with Census 2001: An operational perspective, Southern African *Journal of Demography* 9(1) 67-90.

Crouch, L. and Perry, H. (2003) Educators In: Human Sciences Research Council (ed) Human Resources Development Review 2003: Education, employment and skills in South Africa, Cape Town: HSRC Press. 476-497.

Department of Education (1995a) White Paper on Education and Training No. 1. Education and Training in a Democratic South Africa, *Government Gazette* Vol. 357 (16312), Pretoria: Department of Education.

Department of Education (1995b) South African Qualifications Authority Act, *Government Gazette* Vol. 364 (16725), Cape Town: Government Printer.

Department of Education (1996a) White Paper on Education and Training No. 2. The Organisation, Governance and Funding of Schools, *Government Gazette* Vol. 357 (16312), Pretoria: Department of Education.

Department of Education (1996b) National Education Policy Act, *Government Gazette* Vol. 370 (17118), Cape Town: Government Printer.

Department of Education (1996c) South African Schools Act, *Government Gazette* Vol. 377 (17579), Pretoria: Department of Education.

Department of Education (1997a) *Curriculum 2005 Discussion Document*, Pretoria: Department of Education.

Department of Education (1997b) Education White Paper 3. A Programme for the Transformation of Higher Education, *Government Gazette* Proclamation No. 1196, Pretoria: Department of Education.

Department of Education (1998a) Education White Paper 4. A Programme for the Transformation of Further Education and Training, *Government Gazette* Vol. 399 (19281), Pretoria: Department of Education.

Department of Education (1998b) Quality of Education for All: Overcoming Barriers to Learning and Development. Report of the National Commission on Special Needs

in Education and Training and the National Committee on Education Support Services, Pretoria: Department of Education.

Department of Education (1998c) Employment of Educators Act No. 76 of 1998, Pretoria: Department of Education.

Department of Education (1999) Religion in Curriculum 2005. Report of the Ministerial Committee on Religious Education, Pretoria: Department of Education.

Department of Education (2002) Report on the School Register of Needs 2000 Survey, Pretoria: Department of Education.

Department of Education (2003) Education Statistics in South Africa at a glance, Pretoria: Department of Education.

Department of Labour (1998) The Skills Development Act 97 of 1998, Pretoria: Department of Labour.

Department of Labour (1999) The Skills Development Levies Act, Pretoria: Department of Labour.

Department of Labour (2001) National Skills Development Strategy (final version), Pretoria: Department of Labour.

Krige D., Dove S., Makalima B. and Scott D. (1994) The Education Atlas of South Africa, Durban: Education Foundation.

Leibbrandt, M. and Woolard, I. (2001) The labour market and household income inequality in South Africa: Existing evidence and new panel data, *Journal of International Development* 13 671-689.

McCord, A. and Bhorat, H. (2003) Employment and labour market trends In: Human Sciences Research Council (ed) Human Resources Development Review 2003: Education, employment and skills in South Africa, Cape Town: HSRC Press. 112-141.

Minister of Finance (2003) Statement by the Minister of Finance on the 2001 Census Results, National Assembly 11 September 2003.

O'Connell, P.J. (1999) *Adults in training: An international comparison of continuing education and training* CERI/WD (99)1. Centre for Educational Research and Innovation, Paris: Organisation for Economic Co-operation and Development.

Ok, W. and Tergeist, P. (2003) *Improving workers' skills: Analytic evidence and the role of the social partners,* Employment, Labour and Social Affairs Committee, Directorate for Employment, Labour and Social Affairs, DELSA/ELSA/WD/SEM(2003)10 Annex Table A1, pp. 60-61, Definition and Measurement of CET of Adult Workers, Box 1, p. 11.

Paterson, A. (1992) Contest and Co-option: The struggle for schooling in the African

Independent Churches of the Cape Colony: 1895-1920. PhD, University of Cape Town.

Paterson, A. and Kruss, G. (1998) Education migration and its effect on schooling in South Africa, *South African Journal of Education* 18 (3). 149-155.

Paterson, A. and Du Toit, J. (2004) Training rates and training expenditure in small, medium and large private enterprises In: Paterson, A., McGrath, S., Badroodien, A. (eds) (2004) *South Africa: National Skills Survey, 2003,* Pretoria: HSRC. 43-74.

Perry, H. and Arends, F. (2003) Public schooling In: Human Sciences Research Council (ed) Human Resources Development Review 2003: Education, employment and skills in South Africa, Cape Town: HSRC Press.

Pillay, P. (2003) The skills requirements of specific economic sectors In: Human Sciences Research Council (ed) Human Resources Development Review 2003: Education, employment and skills in South Africa, Cape Town: HSRC Press. 86-111.

See, L., Gould, M.I., Carter, J., Durham, H., Brown, M., Russell, L. and Wathan, J. (2004) Learning and teaching online with the UK census, *Journal of Geography in Higher Education* 28(2). 229-245.

Shisana, O. (2002) *Nelson Mandela/HSRC study of HIV/AIDS Household Survey 2002 South African National HIV Prevalence, Behavioural Risks and Mass Media: Household Survey 2002,* Pretoria: Human Sciences Research Council.

Statistics South Africa (1998) The People of South Africa Population Census, 1996: The Count and how it was done, Pretoria: Statistics South Africa .

Statistics South Africa (2001) Education in South Africa: Selected findings from Census '96, Pretoria: Statistics South Africa.

Statistics South Africa (2003a) Census 2001: Census in Brief Statistics South Africa, Pretoria:
http://www.statssa.gov.za/census01/html/CInBrief/CIB2001.pdf

Statistics South Africa (2003b) *Investigation into appropriate definitions of urban and rural areas for South Africa Discussion document,* Report Number: 03-02-20 (2001), Pretoria: Statistics South Africa.
http://www.statssa.gov.za/census01/html/UrbanRural.pdf

Statistics South Africa (2004a) Stats in brief: Ten years of democratic governance, Pretoria: Statistics South Africa.

Statistics South Africa (2004b) Primary tables South Africa Census '96 and 2001 compared, Report No. 03-02-04, Pretoria: Statistics South Africa.
http://www.statssa.gov.za/census01/html/C2001PrimTables.asp

Statistics South Africa (2004c) Census 2001: Digital Census Atlas,
http://www.statssa.gov.za/census2001/digiatlas/index.html

Vass, J. (2003) The Impact of HIV/AIDS In: Human Sciences Research Council (ed) *Human Resources Development Review 2003: Education, employment and skills in South Africa,* Cape Town: HSRC Press. 186-207.

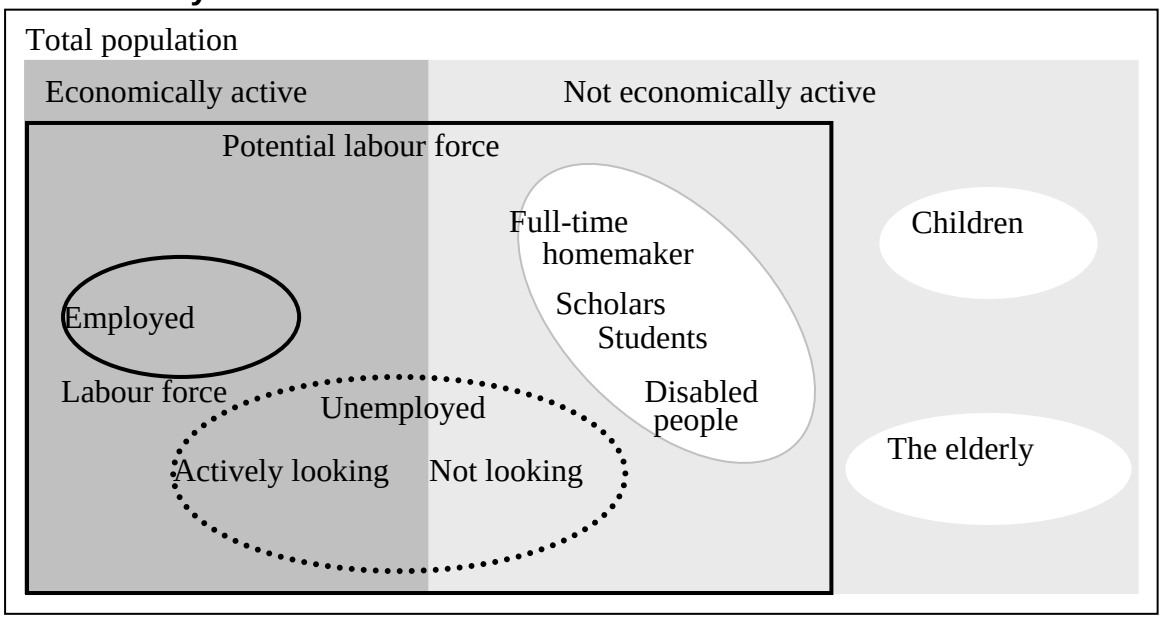
Chapter 8: The labour force

8.1 Introduction

In discussions on labour issues – whether about labour market structures, industrial relations and conditions of work, macroeconomic policies or development strategies – the challenges are normally expressed in numbers. An important source of data on the South African labour force is the population census carried out, approximately once every five years, by Statistics South Africa (Stats SA). The purpose of this chapter is to introduce the variables – captured and made accessible for analysis by Stats SA – that can be used to describe the labour force in South Africa. The focus will be on major labour force concepts, such as employment and unemployment, which are constantly referred to in the media. The key variables relating to employed and unemployed persons are age, education level, gender, race, occupation and industry group.

In order to explain the term 'labour force', various related terms such as 'economically active population', the 'employed' and the 'unemployed' are used to describe the type of involvement in the economy or economic activity. In the diagram in Figure 1, the population is divided into various levels of social and economic activity. The diagram illustrates the relationship between terms frequently used in discussions on labour and employment.

Figure 1: Division of the population into the economically active and the not economically active



The total population of a country can be divided into two main groups. The first group, the potential labour force, comprises all people in the age group 15–65 years, and the second group consists of children younger than 15 years and elderly people

older than 65. The terms labour force and economically active population refer to all people over 15 years of age who are either employed or actively seeking work (Ehrenberg & Smith 2003; Stats SA 2003a; ILO 2000). Because not everyone in the 15–65 age group is working, some people in this group are not economically active. For example, housewives or full-time homemakers, scholars or students, disabled people and people not wanting to work, that is, unemployed people not actively looking for work, are not economically active. Children and the elderly are also not economically active.

These terms are defined below and linked with population census variables most frequently used to describe employment and unemployment.

8.2 Census variables related to the labour force

Work status

The *employed* comprise all persons 15 years and older who, during the seven days before 10 October 2001, worked for one hour or more for pay (in cash or in kind) or for profit or family gain (Stats SA 2003a). The term employment also covers work on a farm or the land, whether for a wage or as part of the household's farming activities, and therefore includes people in paid employment as well as self-employed people. Information is also collected on whether the employed individual works on a full or part-time basis or in the formal or informal sector of the economy.

Unemployment is a situation where members of the labour force are without work (not in employment as defined above), are currently available for work and are seeking work (Barker 1999). Stats SA recently changed the official definition of unemployment in line with 80 per cent of other developed and developing countries.

The *strict definition* of unemployment (used by Stats SA as the official definition) covers those people within the economically active population who

- had not worked during the seven days prior to the specific survey interview;
- wanted to work and were available to start work within a week after the interview;
- had taken active steps to look for work or to start some form of self-employment in the four weeks prior to the interview.

The *expanded definition* of unemployment includes all economically active people who

- had not worked during the seven days prior to the specific survey interview;
- wanted to work and were available to start work at some time in the future but

- had not taken active steps to look for work or to start some form of self-employment in the four weeks prior to the interview.

The *unemployment rate* is the number of unemployed people as a percentage of the total economically active population (Barker 1999). For the official unemployment rate, the strict definition of unemployment is used and therefore excludes discouraged work seekers who are not actively looking for work.

Some economists question the use of narrow unemployment statistics by the government to denote the official unemployment rate. According to McCord and Bhorat (2004), the exclusion of discouraged workers under the narrow definition leads to an underestimation of the actual level of unemployment. Kingdon, Knight and Nattrass (cited by McCord & Bhorat 2004) suggest that in the context of persistent and elevated unemployment, particularly in rural areas, discouraged workers should be included among the unemployed since failure to look for work does not necessarily imply lack of desire to work.

Stats SA (2003a) warns that the 2001 census produces lower estimates of labour force participation than the September 2001 Labour Force Survey (Table 1). According to Stats SA, underreporting of employment may occur in the informal and subsistence agriculture sectors, particularly among those working only a few hours a week. The Labour Force Survey questionnaire includes more prompts to clarify these issues, which is not possible during census enumeration. The United Nations and the International Labour Organisation note that labour force surveys are expected to produce more reliable estimates of labour market variables than censuses. The results from the Labour Force Survey of September 2001 are accepted as the official labour market statistics at the time of Census 2001.

Table 1: Population of working age (15–65 years) by labour market status – Labour Force Survey and census data compared

	Census 2001	**LFS, September 2001**
Employed	33,7%	39,6%
Unemployed	24,0%[a]	16,5%[a]
Not economically active	42,3%	43,9%
Total	**100,0%**	**100,0%**

Note: [a]This is the unemployed as a percentage of the entire working age population and not the unemployment rate.
Source: Statistics South Africa, 2003b

The *labour absorption rate* is calculated as the percentage of the working age population that is employed (Stats SA 1998). The so-called *labour absorption capacity* of the economy is an indication of the capacity of the formal economy to absorb new entrants into the labour force (Barker 1999).

New entrants to the labour force include people (most likely to be in the age group 15–24 years) who have completed their compulsory education or passed matric, and adults who have completed their post-school education and want to work.

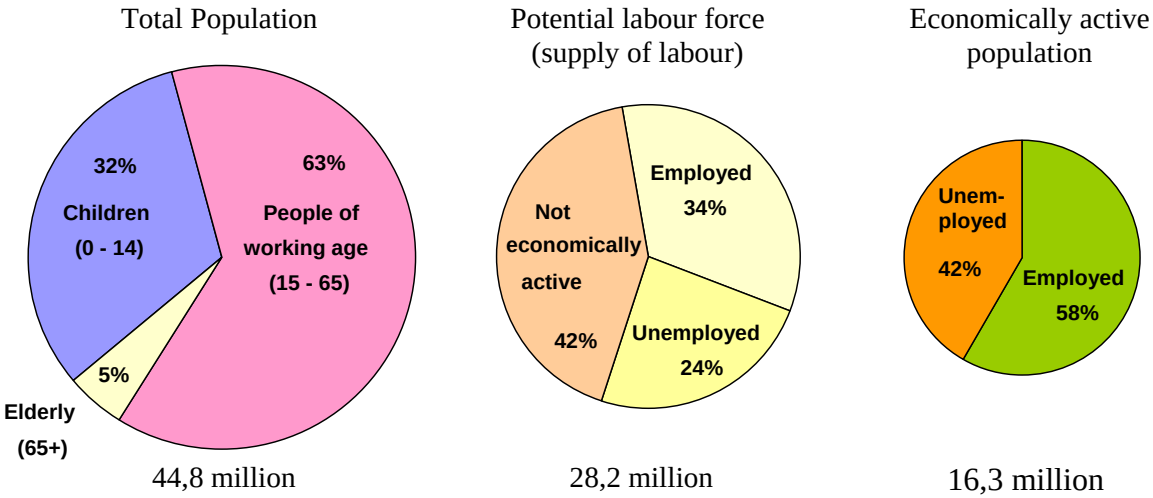
Age

The total population and the population growth rate are important determinants of labour supply. Not all inhabitants physically present in a country at a particular time can be regarded as workers as some may only be future or potential workers and dependants. Stats SA usually publishes the age distribution of the population in five-year intervals.

Children younger than 15 years and the elderly (people older than 65 years, not working and not self-reliant) who are dependent on working parents or relatives and social welfare for subsistence can be considered not economically active. The total population of South Africa was calculated at 44,8 million in the October 2001 population census (Stats SA 2003b) (Figure 2). Children under the age of 15 and the elderly constituted 37% (or 16,6 million) of all the people living in South Africa.

Those who will enter the labour market over the next decade or two have already been born. Forecasts of future labour supply can therefore be made by calculating the total number of 14–year-old children in the year of reporting (or in the five-year intervals: 10–14 years, 5–9 years and 0–5 years).

Figure 2: Total population, potential labour force and economically active population, 2001



Notes:
[1]This is the unemployed as a percentage of the entire working age population and not the unemployment rate.
[2]Unemployment rate: official or strict definition.
Source: Statistics South Africa, 2003b

The supply of labour is calculated on the basis of the total population of working age (15–65) people and is defined as the number of people potentially available for work (Griffiths & Jones 1980) or the potential economically active population (Roukens de Lange 1992). In 2001, 28,2 million people (63%) fell in the age group 15–65 and were potentially available for work (Stats SA 2003b) (Figure 2). Of the potential labour force, an estimated 11,9 million people (42%) were considered not economically active (housewives/homemakers, scholars/full-time students, pensioners/retired people, disabled people and people not wishing to work) (Stats SA 2003b).

The labour force or economically active population comprises those people of working age who present their labour for remuneration on the labour market. It includes workers formally or informally employed, the self-employed and the unemployed who wish to work (Barker 1999). The actual size of the economically active population depends on how the employed and the unemployed are defined and measured. According to the results of the 2001 census, more than half (58%) of the potential labour force in South Africa was economically active at the time (Figure 2). Of these 16,3 million economically active people, 9,5 million (58%) were employed while 6,8 million people (42%) wanted to work (or were actively seeking work) but could not find a job (Stats SA, 2003b).

People who are out of the labour market are considered ‘not economically active’. This term includes people of working age such as housewives or homemakers and full-time students or scholars. Pensioners, retired people, the severely disabled and unemployed people who are not actively looking for a job are included among those who are not economically active (Stats SA 2001).

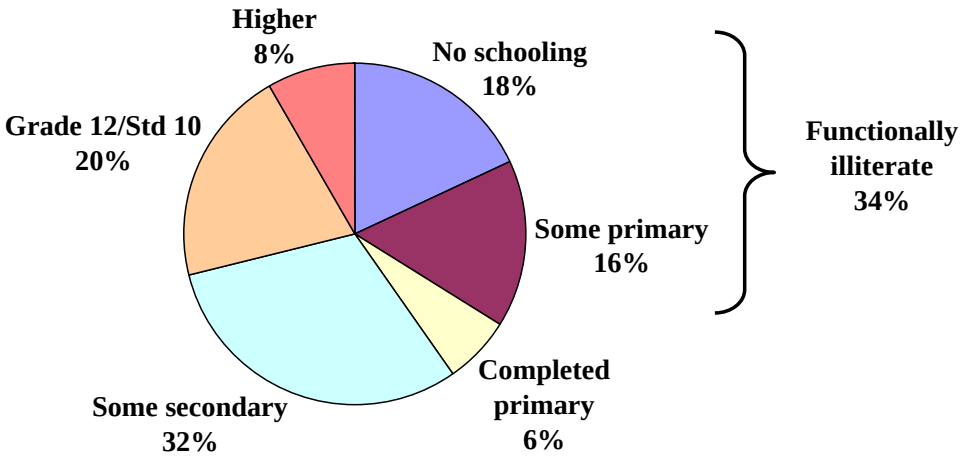
The labour force participation rate is the ratio of the labour force to the labour supply (Griffiths & Jones 1980). It is calculated by dividing the number of people presenting their labour for remuneration on the labour market (economically active population/labour force) by the population of working age people (labour supply) and expressed as a percentage (Barker 1999).

Education levels

Human capital refers to the skills, knowledge and other acquired (usually through education and training) characteristics of workers that make them more productive (Gottheil 2002; Barker 1999). It is generally accepted that the rate of economic progress in a country is closely linked to the stock of human capital it can draw on (Griffiths & Jones 1980; Van Dyk et al. 1997). According to Miyamoto (2003), a high level of human capital is a key factor in attracting foreign direct investment (FDI) to a country.

Stats SA provides data on the level of education of the population either in broad categories (none, primary, secondary or higher education) or in every grade or standard. It also provides data on those who have obtained diplomas or certificates and the number of university graduates. The level of education of a worker will play an important role in the type of job he or she can obtain. According to the 2001 population census, almost a fifth (18%) of South Africans aged 20 years or older had received no education, and a further 16% had not completed their primary school education (Stats SA 2003b) (Figure 3). According the Joint Education Trust, people who have not had at least seven years of formal schooling are regarded as functionally illiterate (SAIRR 1997:152). Almost half (38%) of the above group had completed primary or acquired some secondary school education. Only 20% of the population 20 years and older had obtained Grade 12, and only 8% had a post-school qualification (Stats SA 2003b).

Figure 3: Highest level of education among those aged 20 years and older, 2001



Source: Statistics South Africa, 2003b

Higher education refers to all post-school education undertaken by those who have completed their matric at accredited institutions. Certificates, degrees and diplomas issued by an accredited institution are regarded as examples of higher education qualifications (Stats SA 1998).

Secondary education for a person normally starts at the age of 13 years and extends over a five-year period. It consists of Grades 8–12/Standards 6–10/Forms 1–5 and National Trade Certificates (NTCs) 1–111. This broad category can be divided into two groups: 'some secondary' and 'Grade 12 or Standard 10' (Stats SA 2003a). The first group refers to those who have completed Grades 8 to 11 or obtained a diploma or certificate but have not completed Grade 12. The second group refers to those who have matriculated (Grade 12 or Standard 10 or Form 5 or matric or NTC111).

Secondary level education appears to be the minimum level sought by relatively high value-adding, efficiency-seeking multinational enterprises (Miyamoto 2003).

Primary education for a person normally starts at the age of six and extends over seven years. It consists of Grades 1–7 or Sub A and B and Standards 1–5. This broad category can also be divided into two groups: 'some primary' and 'complete primary' (Stats SA 2003a). The first group refers to those who have completed Grades 1 or Sub A to Grade 6 or Standard 4. According to the Joint Education Trust, people who have had less than seven years of formal schooling are regarded as functionally illiterate (SAIRR 1997:152). The second group refers to those who have completed Grade 7 or Standard 5.

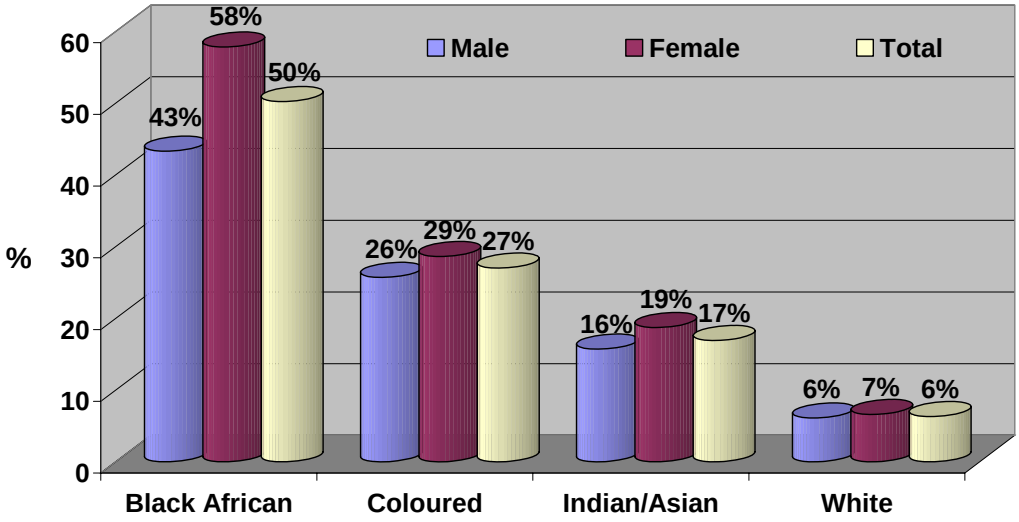
None/No formal education: These are people who are part of the labour force but who have never had the opportunity to receive formal education or training.
The prospects for human capital development can be deduced from the current trends in education among children and adolescents (learners) and the training programmes of enterprises (Miyamoto 2003). A significant number of undereducated children (as indicated by net primary school enrolment ratios among primary school aged children) lowers the average level of human capital among the working age population. Entry rates in post-basic schooling (upper secondary and tertiary) indicate how a country is doing in fostering a more highly skilled workforce. An example would be new entrants to each schooling level as a percentage of the total population at the typical age of entry.

Race and gender

Economies and labour markets have to adapt to changes in society. During apartheid, blacks – those classified as African, Indian and coloured – women and people with disabilities did not have a fair chance of securing employment in all occupational categories and levels in the labour force. The demise of apartheid saw the scrapping of the old laws and ways of working. New laws, such as the Employment Equity Act, 55 of 1998, were introduced to achieve equity in the workplace.

An analysis of the racial and gender composition of the labour force in South Africa reveals that inequalities still remain. As illustrated in Figure 4, the unemployment rate for blacks is considerably higher than for whites, and female unemployment across all races is higher than male unemployment (Stats SA 2003b).

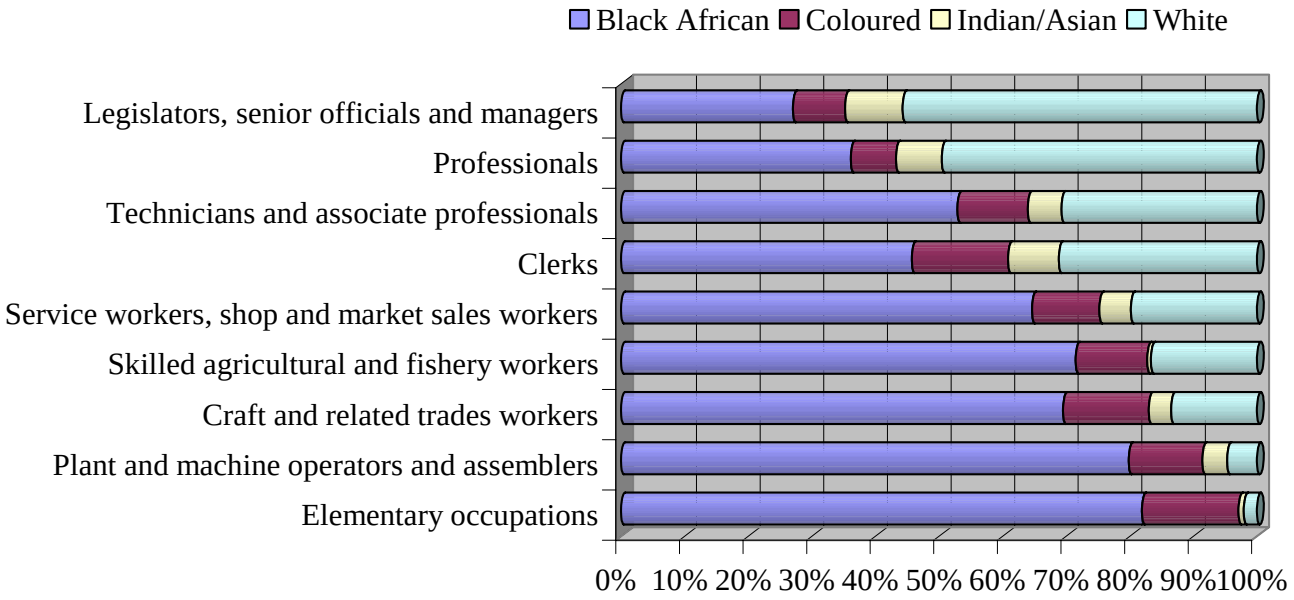
Figure 4: Unemployment rate (strict definition) among those aged 15–65 years by gender and population group, 2001



Source: Statistics South Africa, 2003b

An analysis of the racial composition of different occupational levels shows an inverted pattern – whites predominate in the few top-level jobs while mainly blacks fill the large number of positions at the lower end of the job spectrum (Figure 5). Male workers predominate in most of the occupational categories. The exception is clerical and sales personnel, where more female than male workers are employed.

Figure 5: Occupation by population group among the employed aged 15–65 years



Source: Statistics South Africa, 2003b

Occupation

Occupation refers to the actual work a person does, for example, professional, clerical or elementary work (domestic work, for instance, is usually classified as elementary work). Based on the second edition of the South African Standard Classification of Occupations (SASCO), which is in turn based on the International Standard Classification of Occupations (ISCO-88), Stats SA uses nine broad occupational categories at the top level of aggregation. The following brief outline of ISCO-88 major groups will facilitate interpretation of the classifications.

Legislators, senior officials and managers include people in occupations whose main tasks are determining and formulating government policies, laws and public regulations, overseeing their implementation, representing government and acting on its behalf or planning, directing and coordinating the policies and activities of enterprises and organisations or departments.

Professionals include people in occupations whose main tasks require a high level of professional knowledge and experience in the fields of physical and life sciences or social sciences and humanities. The main tasks consist of increasing the existing stock of knowledge, applying scientific and artistic concepts and theories to the solution of problems, and teaching the foregoing in a systematic manner (ILO 1990). Most occupations in this major occupational group require workers with higher education qualifications.

Technicians and associate professionals include people in occupations whose main tasks require technical knowledge and experience in one or more fields of physical and life sciences or social sciences and humanities. The main tasks consist of doing technical work connected with the application of concepts and operational methods in the abovementioned fields and teaching at certain educational levels (ILO 1990).

Clerks include people whose main tasks require the knowledge and experience to organise, store, compute and retrieve information. The main tasks are performing secretarial duties, operating word processors and other office machines, recording and computing numerical data and performing various customer-oriented clerical duties, mostly in connection with mail services, money-handling operations and appointments (ILO 1990).

Service workers and shop and market sales workers include people whose main tasks require the knowledge and experience to provide personal and protective services and to sell goods in shops or at markets. The main tasks consist of providing services related to travel, housekeeping, catering, personal care, protection of individuals and property, maintaining law and order and selling goods in shops or at markets (ILO 1990).

Skilled agricultural and fishery workers include people whose tasks require the knowledge and experience to produce farm, forestry and fishery products. The main tasks consist of growing crops, breeding or hunting animals, catching or cultivating fish, conserving and exploiting forests and, especially in the case of market-oriented agricultural and fishery workers, selling products to purchasers, marketing organisations or at markets (ILO 1990).

Craft and related trades workers include people whose tasks require the knowledge and experience of skilled trades or handicrafts which, among other things, involve an understanding of the materials and tools to be used as well as of all stages of the production process including the characteristics and the intended use of the final product. The main tasks consist of extracting raw materials, constructing buildings and other structures and making various products and handicraft goods (ILO 1990).

Plant and machine operators and assemblers include people whose main tasks require the knowledge and experience to operate and monitor large-scale, and often highly automated, industrial machinery and equipment. The main tasks are operating and monitoring mining, processing and production machinery and equipment as well as driving vehicles and driving and operating mobile plants or assembling products from component parts (ILO 1990).

Elementary occupations are those that require the knowledge and experience to perform mostly simple and routine tasks involving the use of hand-held tools and, in some cases, considerable physical effort. With few exceptions, only limited personal initiative or judgement is required. The main tasks are selling goods in streets, door keeping and property watching as well as cleaning, washing, pressing and working as labourers in mining, agriculture and fishing, construction and manufacturing (ILO 1990).

These broad occupational categories are subdivided into sub-major groups (second level of aggregation), minor groups (third level of aggregation) and unit groups (representing the most detailed level of the occupational structure). The subdivisions of the broad occupational groups provide successively finer detail and are based on skill specialisation defined by reference to the field of knowledge required, the tools and machinery used, the materials worked on or with, as well as the kinds of goods and services produced. The same occupation may be encountered in various economic sectors or industry groups.

Industry

Economic sector or *industry* refers to the type of organisation in which a person works, for example agriculture, forestry and fishing, and manufacturing and construction (Stats SA 2003a). Nine economic sectors have been identified:

- agriculture, hunting, forestry and fishing
- mining and quarrying
- manufacturing
- electricity, gas and water supply
- construction
- wholesale and retail trade
- transport, storage and communication
- financial, insurance, real estate and business services
- community, social and personal services (including domestic work)

The composition of each sector is based on The International Standard Industrial Classification of All Economic Activities (ISIC) (ILO 2001). The nine broad sectors can be divided into three subsectors, namely primary, secondary and tertiary. The *primary* sector includes agriculture, hunting, forestry and fishing, and mining and quarrying. The *secondary* sector includes manufacturing, electricity, gas and water supply, and construction. The *tertiary* sector includes wholesale and retail trade; transport, storage and communication; financial, insurance, real estate and business services; and community, social and personal services.

Shifts occur in the economy of a country as part of developmental processes. In all advanced economies, the nature of economic growth has been shifting away from primary production towards secondary and tertiary production (Neef 1999). Consistent with this trend, the pattern of activity in the South African economy has also changed. While the contribution of agriculture to South Africa's gross domestic product (GDP) dropped between 1920 and 1969, the contribution of the manufacturing sector rose in the same period (Malherbe 1975). According to the South African Reserve Bank (2000), the relative share of value added (measured in terms of GDP) by the tertiary sector increased over the last decade while the relative share of the primary and secondary sectors declined.

These shifts in the economic output of the different sectors give rise to changes in the employment structure. Parallel with the shift in emphasis away from the primary and secondary sectors to the tertiary sector is a movement within individual sectors away from unskilled and semiskilled work towards more highly skilled work. The past two decades have seen declining employment in the primary (agricultural and mining) and secondary (manufacturing) sectors of the South African economy and substantial increases in employment in the tertiary sector (transport, trade, finance and services).

An increase or decrease (from a previous reporting period to a current reporting period) in the number of people employed in a specific industry (and in different

occupations) is an indication of the *demand for labour* in that industry (and of the types of qualifications that are in demand). The demand for labour can also be calculated by subtracting the number of unemployed workers from the labour force (Roukens de Lange 1992).

Other variables

The variables mentioned earlier are central in describing the labour force. They can be cross-classified with other population variables supplied by Stats SA such as province of residence, place of birth, marital status, monthly income, and so on. Researchers can thus describe employment and unemployment with various degrees of detail from a broad overview of the labour force in South Africa to detailed analyses of

- the provincial and sectoral distribution of job opportunities;
- the extent of unemployment in a province among the youth and among women;
- the demand for specific types of workers;
- estimates and projections of the labour force to enable a comparison of levels of employment and trends in the labour force in different provinces and in the country in past and future years.

Bibliography

Barker, FS (1999) *The South African labour market: critical issues for renaissance* 3rd ed. Pretoria: Van Schaik.

Ehrenberg, RG & Smith, RS (2003) *Modern labour economics: theory and public policy* 8th ed. San Francisco: Addison Wesley

Gottheil, FM (2002) *Principles of economics* 3rd ed. USA. South-Western.

Griffiths, HR & Jones, RA (1980) *South African labour economics.* Johannesburg: McGraw-Hill.

International Labour Organisation (1990) *ISCO-88: International standard classification of occupations.* Geneva.

International Labour Organisation (2000) *Sources and methods: Labour statistics.* Geneva.

International Labour Organisation (2001) ISIC: International Standard Industrial Classification of all Economic Activities. Geneva

McCord, A & Bhorat, H (2004) Employment and labour market trends, in Human Sciences Research Council (2003) *Human resources development review 2003: Education, employment and skills in South Africa.* Cape Town: HSRC Press and East Lansing: Michigan State University Press.

Malherbe, E.G. 1975. *Education in South Africa. Volume 1, 1652–1922*. 2nd ed. Cape Town: Juta. p.521.

Miyamoto, K (2003) *Human capital formation and foreign direct investment in developing countries.* Paris: OECD.

Neef, D (1999) *A little knowledge is a dangerous thing: understanding our global knowledge economy* 2nd ed. Boston: Butterworth/Heinemann.

Roukens de Lange, A (1992) *Manpower demand and supply in South Africa: A study of trends and interactions with the economy.* Stellenbosch: Institute for Futures Research, University of Stellenbosch.

South African Institute of Race Relations (1997) *South African Survey 1996/97.* Johannesburg: p910.

South African Reserve Bank (2000) *Annual economic report 2000.* http://www.resbank.co.za/Economics/Annual2000.html (accessed 8 January 2001)

Statistics South Africa (1998) *Census in brief*: *Report no. 1: 03–01–11*(1996). Pretoria.

Statistics South Africa (2003a) *Census 2001: metadata (information on persons).* http://www.statssa.gov.za/census2001/HTML/persons.pdf (accessed 28 May 2004)

Statistics South Africa (2003b) *Census 2001: Census in brief.* http://www.statssa.gov.za/census01/html/CInBrief/CIB2001.pdf (accessed 28 May 2004)

Van Dyk, PS, Nel, PS, Loedolff, PVZ & Haasbroek, GD (1997) *Training management: a multidisciplinary approach to human resources development in South Africa.* Johannesburg: Thomson International.

Chapter 9: Estimation of fertility and mortality rates from Census 2001 data

9.1 Introduction

The aim of this chapter is to provide the user with essential information on demography (with particular emphasis on estimation of fertility and mortality rates) to help the user in understanding why it is important to have questions on the demographics of a country in a census. The issues related to migration have already been covered in chapter 6. This chapter first looks at the importance of population censuses as a source of demographic information. It then shifts its focus to fertility and mortality and describes the most important indicators for these processes that were obtained from Census 2001.

9.2 A population census as a source of demographic information

Demography is concerned with population size, distribution and structure (composition) as well as the processes that lead to change in these aspects. Population size is captured as the number of units (persons) in the population. The population distribution is the arrangement of the population in a geographic area, while the population structure is the distribution of the population among its gender and age groupings. The three demographic processes are fertility, mortality and migration. Any factors that affect the number and distribution of people must operate through one or more of these variables.

The statistical measurement of fertility, mortality and migration forms the core of formal demographic analysis (Kammeyer, 1969). However, the study of demography is not confined to measuring, counting, collecting and analysing statistical data. It is evident that fertility, mortality and migration are not independent variables. They are to a large extent socially and biologically determined. The number of births, deaths and migrants is therefore affected by a whole host of physical, biological, social and psychological factors. In turn, population changes have far-reaching effects on the social organisation and the economic system of the societies in which they occur.

A census is an important source of information on demography because it obtains demographic information from all persons in a country unlike other sources such as surveys. In countries without complete registration of vital events, however, a population census or survey may include questions about births or deaths of household members during a specified period preceding the census. In addition, even when vital statistics of good quality exist, the census or survey may include questions on fertility (children ever born, children still alive and date of birth of each child) because the distribution of women by number of children ever born and by

interval between successive births cannot be discovered from birth certificates (Shryrock, Siegel & Associates, 1976).

Direct demographic estimation is used when data collected are perfect, meaning the data are complete and accurate, but because censuses yield far from perfect data, indirect estimation is used. Censuses suffer from two types of errors: failure to enumerate all the members of the relevant population or to over count, which happens occasionally; and poor age reporting on the part of the population enumerated.

9.3 Fertility

As one considers the basic elements in the population formula, it is difficult not to come to the conclusion that fertility and mortality are the most crucial factors in determining the demographic future of individual countries, and South Africa is no exception. Future fertility trends will certainly influence the nature of human life on earth. In South Africa there is general agreement that fertility began to decline among all major population groups prior to the end of apartheid. Since the recognition of this decline, and after the 1994 Cairo Conference, there has been a shift from an emphasis on fertility reduction with the sole aim of reducing the population growth rates of countries to a concern with reproductive health and with enhancing the ability of women and couples to achieve their fertility goals (Finkle, 2002). This section focuses on fertility as a variable that provides information that is used to plan social programmes.

Socio-economic indicators are often required, among other uses, to measure development progress, monitor intervention programmes arising from government policies and inform future policies. The knowledge and quantification of the link between fertility and socio-economic demands thus enable decision makers to allocate resources more efficiently and to do proper planning regarding resources (Udjo, 1997, 1998, 2003). The census tries to capture fertility and related information, although not comprehensively.

Census 2001 questions on fertility

In the 2001 census, the following lifetime and fertility questions were included for women aged 12–49 in the census questionnaire:

1. How many children, if any, that were born alive has the person ever had?
   How many of these were boys?
   How many of these were girls?
2. If the person has ever given live birth, when was the person's last child born? (Month and year of birth)

Using the information collected on fertility, there are several indicators that can be directly computed which would give information about the situation regarding births in the country, and by doing comparisons of different periods, information can be obtained on patterns and trends in fertility over time. There are two different methods demographers use to compute fertility indicators: direct and indirect methods of estimation. Direct methods are used when the data are available, but if there are inconsistencies or incompleteness, indirect methods are used (United Nations, 1983) . In this chapter we look at the following indicators: age-specific fertility rate; total fertility rate which is very important in measuring fertility levels in our country and gross reproduction rate. These are just some of the few examples of how the data collected in the census can be used.

Age-specific fertility rate

Within the female population the frequency of childbearing varies markedly from one age group to another. In fact, there is a characteristic age pattern to fertility, which is quite similar all over the world. Computing age-specific fertility rates reveals this age pattern. The age-specific fertility rate is defined as the number of births to women of a specific age group (aged 15 to 49), per 1 000 women in that age group.

Total fertility rate

The total fertility rate is calculated by cumulating the age-specific fertility rates for women of all ages (15–49 years) as shown in Table 1. The total fertility rate is a measure of the total number of births a woman would have if she were to produce at the prevailing age-specific fertility rates. The total fertility rate is perhaps the most widely used indicator of fertility levels and changes in these levels. However, although the total fertility rate is a useful and very commonly used measure of fertility, it is not a perfect measure of a real situation as it combines the fertility experiences of women of all reproductive ages for a particular year even though experiences relate to different time periods. It is thus a measure of what a hypothetical group of women would experience at the end of their childbearing, given the prevailing age-specific fertility rates.

Table 1: Calculation of age-specific fertility rates and total fertility rate for South Africa in 2001

Age of women	Female population	Number of births	Age-specific fertility rates
(1)	(2)	(3)	(4) = (3)/(2) * 1 000
15–19	2 528 642	190 047	75,2
20–24	2 195 230	288 428	131,4
25–29	2 035 814	274 445	134,8
30–34	1 746 412	196 776	112,7
35–39	1 630 264	119 822	73,5
40–44	1 385 832	46 342	33,4
45–49	1 119 776	13 640	12,2
		Sum	573,1
		Sum * 5 / 1 000	2,9
The total fertility rate in South Africa for 2001 was 2,9 births per woman on average.			

Source: Statistics South Africa, Population Census 2001

Table 2: Age-specific fertility rates derived from direct-estimation techniques, Census 1996 and Census 2001

Age group	RSA 1996	RSA 2001
15–19	47	75
20–24	100	131
25–29	116	135
30–34	115	113
35–39	95	73
40–44	72	33
45–49	57	12

Source: Statistics South Africa, Population Census 2001

The above tables show one of the uses of the census data and the indicators that could be computed using census data. Table 1 shows how age-specific fertility rates and total fertility rates can be calculated using the female population by age groups and number of births. Table 2 shows a comparison of age-specific fertility rates estimates for the country as a whole for 1996 and 2001. From the tables we observe that there has been a fertility increase for women in age groups 15 to 29 years and a significant decline for those aged between 40 and 49 years.

Gross reproduction rate

The gross reproduction rate measures the number of daughters a cohort of women will have. It assumes that all females survive to the end of the childbearing period. It

is thus derived in the same manner as the total fertility rate but uses a set of age-specific fertility rates based on female births only. Calculating the gross reproduction rate follows the same method as calculating the age-specific fertility rate except that it takes only the births of females into account. Multiplying the total fertility rate by the proportion of all births that are female can also give an acceptable approximation.

The gross reproduction rate by population group based on data from Census 2001 is shown in Table 3. For all population groups the gross reproduction rate based on direct estimates from Census 2001 is below 2.

Table 3: Direct estimates of gross reproduction rate by population group in 2001

Population group	Gross reproduction rates
African	1,525
Coloured	1,425
White	0,860
Indian/Asian	0,985
South Africa	1,435

Source: Statistics South Africa, Population Census 2001

Social factors that impact on fertility

It is important to note that fertility is affected by social variables such as population group, education, marital status and employment status. Research over the years has proved that the above-mentioned variables affect the level of fertility in South Africa. For example, the census data show that the African population has the highest fertility rates when compared with whites, coloureds and Indians/Asians. Another factor is that women who are educated are likely to have fewer children than women who are less educated. The data on these social variables are found in the census, running simple cross-tabulations between them can show the relationship between fertility and these variables.

9.4 Mortality

Death statistics are needed for purposes of demographic studies and for public health administration. The most important uses of death statistics include: analysis of the present demographic status of the population as well as its potential growth; fulfilment of the administrative and research needs of public health agencies in connection with the development, operation and evaluation of public health programmes; determination of administrative policy on health programmes; determination of administrative policy and action in connection with the programmes of government agencies other than those concerned with public health; and fulfilment

of the need for information on population changes in relation to numerous professional and commercial activities (Shryrock, Siegel & Associates, 1976).

Census 2001 questions on mortality

Statistics South Africa carried out the second post-apartheid census in October 2001. The data for this study are taken from the 2001 census and conclusions are based on the final edited data.

For estimates on childhood mortality, the questions on survivorship in the 2001 census were:

1. How many children, if any, that were born alive has the person ever had?
   How many of these were boys?
   How many of these were girls?
2. If the person has ever given live birth:
   How many boys are still alive?
   How many girls are still alive?

For estimates of adult mortality, the 2001 census questions were similar to those asked in the 1996 census. These questions were asked of each person listed in the household. The answers were then used to calculate indirect estimates of adult female and male mortality. For 2001, the questions were:

1. Is the person's own biological mother still alive?
2. Is the person's own biological father still alive?

The census also included a direct question on deaths occurring in the household: each household was asked whether a death had occurred in the 12 months preceding the census. If so, the age, sex and cause of death of the deceased were recorded. The questions were:

1. Has any member of this household died in the past 12 months?
2. If yes, how many?

Other questions about deaths in the household were:

1. What was the month and year of death?
2. Write down the sex of the deceased.
3. Did the person die from an accident or through violence?
4. If the deceased was a woman under 50 years, did the person die while pregnant or within six weeks after delivery?

Crude death rate (CDR)

Crude death rate is defined as CDR = (D/P) * 1 000 where D is the number of deaths in a population during a specified period and P is the number of person-years lived by the population during the same period. It is expressed as deaths per 1 000 of the population. The crude death rate for a single year is usually calculated as the number of deaths during the year divided by the mid-year population.

Child and infant mortality rates

Infant mortality rate is defined as DO/B * 1 000 where DO represents deaths of infants during a year and B represents live births during the same year. It is expressed as the number of infants who have died in their first year of life per 1 000 of the population. On the other hand, child mortality rate refers to the number of children who have died while older than one year but younger than five years of age. It is defined as the number of dead children under five divided by the number of live children under five.

Age-specific mortality rate

Age-specific mortality rate is defined as the number of deaths occurring during a specified period of persons (usually specified by gender) of a specified age or age group. When the age-specific mortality rate is calculated for a calendar year, the number of deaths of persons of a specified age is usually divided by the mid-year population of persons of that age.

Table 4: Calculation of age-specific mortality rates for South Africa in 2001

Age	Population	Number of deaths	Mortality rates
(1)	(2)	(3)	(4) = (3)/(2) * 1 000
Less than 1	908 406	41 723	45,93
1–4	3 541 409	23 648	6,68
5–9	4 853 555	6 769	1,39
10–14	5 061 917	4 900	0,97
15–19	4 981 721	11 035	2,22
20–24	4 294 523	25 240	5,88
25–29	3 934 939	38 709	9,84
30–34	3 340 901	39 232	11,74
35–39	3 071 770	35 908	11,69
40–44	2 619 465	29 795	11,37
45–49	2 087 380	26 147	12,53
50–54	1 638 020	23 241	14,19
55–59	1 205 266	19 575	16,24
60–64	1 065 294	22 556	21,17
65–69	787 927	20 895	26,52
70–74	631 469	21 539	34,11
75–79	367 537	16 295	44,34
80–84	270 945	16 030	59,16
85+	157 333	18 545	117,87
Total	**44 819 778**	**441 784**	**9,86**
Crude death rate in South Africa for 2001 was 9,86 per 1 000 people.			
Infant mortality rate in South Africa for 2001 was 45,93 per 1 000 infants.			

Source: Statistics South Africa, Population Census 2001

Adult mortality and HIV/AIDS

It is not a sound practice to look at the levels of adult mortality without taking into account the effects of HIV/AIDS. The census does not capture information on the prevalence of HIV/AIDS; information on the disease is obtained from other sources and is used together with census data. It is difficult to know the exact number of HIV/AIDS-related adult deaths in South Africa. This results largely from the widespread lack of complete and accurate vital registration systems and the fact that people infected with HIV die of many different immediate causes that are often recorded as the primary cause of death even though HIV is the primary contributing cause. Moreover, the use of HIV antenatal clinic data as a source of HIV data has been debated as the data only involve women who are pregnant, and therefore the reliability of applying such information to the whole country is questioned. Consequently, all figures describing the number of HIV/AIDS-related deaths in South Africa are just estimates.

Causes of death

The census does not usually capture information on causes of death, but for the 2001 census a question was asked on whether the person died from an accident or through violence. Another question asked was whether a woman under 50 years of age who died did so while pregnant or within one month after delivery. Statistics South Africa undertook a study to investigate the causes of death in South Africa during the period 1997–2001. It was based on a 12% stratified random sample of deaths occurring during the study period, obtained from the Home Affairs population register. Causes of death were coded by utilising guidelines contained in the tenth revision of the *International Classification of Diseases (ICD-10).* The results of this study indicate changes in mortality patterns over time. These changes have tended to affect South Africans differently, depending on population group, gender and age.

The five leading causes of death among South Africans between 1997 and 2001 were unspecified unnatural causes (e.g. suicide, drowning and motor car accidents), ill-defined causes, TB, HIV/AIDS and influenza and pneumonia, accounting for 40,9% of deaths in the sample. Mortality due to unspecified unnatural causes declined significantly during the study period. This decline seems to have been offset by a steep rise in mortality due to HIV/AIDS, TB and influenza and pneumonia. For example, the proportion of deaths due to HIV/AIDS nearly doubled from 4,6% in 1997 to 8,7% in 2001 whereas the proportion of deaths due to unspecified unnatural causes declined from 15,3% to 8,2% during the same period.

Social factors that impact on mortality

Like fertility, mortality is also affected by social variables such as population group, education, marital status and employment status. For example, the census data show that people who are educated and know more about hygiene tend to live a healthier life than those not educated, improving the chances of survival. It is also a known fact that the mortality rate of people in urban areas is lower than that of people in rural areas, but because of HIV/AIDS, there is some shifting. Furthermore, women who are educated are more likely to attend antenatal clinics and have fewer children than women who are less educated, decreasing maternal mortality and improving maternal health. The data on these social variables are found in the census, and running simple cross-tabulations between them will show the relationship between mortality and these variables. For advanced analysis, statistical methods such as regression analysis can be used.

9.5 Conclusion

It is important to have information on fertility and mortality in the census questionnaire, as this information will be used for planning of programmes by the government and other interested parties. Generally, it is known that fertility and

mortality in South Africa have been declining since the 1970s, but with HIV/AIDS to take account of, fertility can be expected to decline further and mortality to increase further. This chapter shows that important indicators can be calculated directly from the census data without having to do advanced analysis or use indirect methods. However, indirect techniques are used to deal with several content and coverage errors. These errors are not unique to South African data. The errors in the 2001 Census are well known and have been extensively discussed in the relevant literature. The fertility and mortality data in the 2001 Census should not be used to estimate fertility and mortality rates using direct methods. The data require that indirect estimates be used to calculate fertility and mortality rates.

Bibliography

Finkle, J, (2002). *Impact of the 1994 International Population Conference on Population and Development*, A background paper for the Expert Group Meeting on Completing the Fertility Transition, 11–14 March, New York.

Kammeyer, K. C. W. (1969). *Population Studies: Selected Essays and Research*. University of Kansas.

Shryrock, H. S., Siegel, J. S. & Associates. (1976). *The Methods and Material of Demography.* (Condensed edition by Edward G. Stockwell). Orlando: Academic Press.

Udjo, E. O. (1997). *Fertility and Mortality Trends in South Africa: The Evidence From the 1995 October Household Survey and Implications for Population Projections*. Pretoria: Statistics South Africa.

Udjo E. O. (1998). *Additional Evidence Regarding Fertility and Mortality in South Africa and Implications for Population Projections*. Pretoria: Statistics South Africa.

Udjo, E. O. (2003). *Recent evidence of levels, trends, and differentials in fertility in South Africa.* Journal of Biosocial Science.

United Nations. (1983). *Manual X: Indirect Techniques for Demographic Estimation*. New York: Department of International Economic and Social Affairs.

Chapter 10: Use of Census 2001 demographic variables in decision making, planning and development

10.1 Introduction

The United Nations (1958) defines a population census as "the total process of collecting, compiling and publishing demographic, economic and social data pertaining, at a specified time or times, to all persons in a country or delimited territory". Although censuses attempt to collect demographic, economic and social data pertaining to all persons in a country, such data collection is never achieved in practice as an element of undercoverage always obtains. The undercoverage of the population in censuses, however, varies from country to country depending on the level of statistical sophistication of the country concerned. The above definition of a population census provided by the United Nations is accordingly somewhat idealistic.

Although population enumeration in South Africa dates back to 1652, the first reliable census was conducted in 1865 in the Cape Province. In 1904, a census covered what was later to become known as the Union of South Africa. The modern census began in South Africa in 1911. Censuses in South Africa since 1911 have not consistently covered the entire Union of South Africa – the last census covering the whole country was conducted in 1970 during the apartheid era. The apartheid censuses of 1980, 1985 and 1991 excluded the former homelands – Transkei, Bophuthatswana, Venda and Ciskei (TBVC). (See Khalifani, Zuberi, Bah & Lehohla: in press.)

Following the new political dispensation in South Africa and the groundbreaking democratic elections of 1994, the first nationwide census since 1970 was carried out in October 1996. The 2001 population census was the second post-apartheid census in South Africa that attempted to canvass the entire country.

The United Nations' census definition provides a broad framework for any country to plan its census although this is rarely an easy operation. For example, while some may see the main objective of a census as the determination of the number of inhabitants in a country or territory (Shryock, Siegal & Associates 1976), others may see the main purpose as the provision of information for the guidance of public policy (Crone 1990). Census planners are often caught between these two historically evolving views when designing the content of a census. South Africa's post-apartheid censuses of 1996 and 2001 questionnaires may be seen as an expression of the latter view as they are among the longest contemporary census questionnaires in the world. The census is a primary source of demographic information, especially in countries where registration of vital events is non-existent or where it is of limited coverage and deficient. South Africa has included a variety of

demographic questions in its censuses. The demographic questions in the 1996 and 2001 post-apartheid censuses are largely similar.

10.2 Importance of demographic information

Demographic information constitutes the bedrock of all socioeconomic planning in any country. The collection, analysis and dissemination of accurate demographic information enable policy makers to plan for the future development of a country. Issues such as the future size of the labour market, unemployment, job creation, poverty and environmental degradation are intrinsically linked to demographic processes. Unfortunately, this link is often not obvious, and the analysis of various socioeconomic phenomena relevant to planning, decision making and development in a country is frequently done in isolation of demographic processes.

The objective of this chapter is (1) to identify the demographic variables in the 2001 census and (2) to describe ways these demographic variables can be applied in decision making, planning and development.

10.3 Demographic variables in the 2001 Census

The demographic variables in the 2001 census can be broadly classified as (1) age and sex, (2) fertility, (3) mortality, and (4) migration variables. In combination, these variables determine the future development of any population with regard to the size and structure (composition) of the population. Planning in the private and public sectors is often aimed at meeting present and future “demands” of the population in order to improve citizens’ standard of living. Knowledge of the age-sex distribution, fertility (including nuptiality), mortality and migration patterns in any population is critical to effective planning to meet these demands. Insights into the magnitude of future demands in various sectors are usually obtained from population projections. Population projections, however, require population parameters as inputs, which include the age-sex composition of the population and fertility, mortality and migration rates.

The impact of these variables determines the rate of growth of a population, and this rate has implications for planning in all sectors of the economy. Each of the variables also has planning implications for the different sectors of the economy. The demographic questions in the 2001 census within the broad classification noted above are identified in the following section. The section also illustrates ways the demographic variables can be applied in decision making, planning and development.

10.3.1 Age-sex variables

The age-sex distribution of a population is a cross-cutting variable in planning in the sense that planning in all sectors needs to take into consideration, directly or indirectly, the age and sex composition of the population.

The following questions in the 2001 census provide information on the age-sex distribution of the population.

What is (the person's) date of birth?

Is (the person) male or female?

Information on the age-sex distribution resulting from the above questions is invaluable for planning and decision making as it is intrinsically linked to all aspects of the life cycle including childhood, education, marriage, childbearing, entry into the labour market, retirement, ageing, morbidity and mortality (Udjo: in press). As already noted, the age-sex distribution is a key factor in population projections. On the basis of the projection results, policy makers gain some idea of the probable future size of the population at national and subnational levels. This information, in addition to other information, enables policy makers to share resources among competing demands. The sectoral importance of the age-sex distribution is discussed below.

Age-sex distribution and planning for education

The age-sex distribution can inform planning in education. For example, the number of persons of school-going age, among other factors, provides a yardstick for gauging the magnitude of the current and future demand for school infrastructure, teachers and so on if certain standards are to be maintained or improved.

If the age-sex distribution is tabulated against educational attainment, planners can get an idea of the extent of sex inequality in educational attainment at a given age in the general population. This information can inform decision making on appropriate interventions to address educational attainment imbalances in a country.

Age-sex distribution and labour market planning

The age-sex distribution has implications for labour market planning in various ways. The age-sex distribution in South Africa gives an indication of the size of the population that will move in and out of the labour market in successive years. The current size of the labour market can be gauged from the age-sex distribution of the population. This distribution provides some indication of the number of additional jobs that need to be created to absorb new entrants into the labour market taking into consideration existing jobs. In general, therefore, policies aimed at reducing unemployment should take cognisance of the size of the labour market as determined by the age-sex distribution.

It should be stressed, however, that the size of the labour market is partly determined by fertility at the individual level. Put differently, the ability of policy makers to create jobs is partly dependent on the fertility of couples since the rate of childbearing to a large extent determines the number of new entrants into the labour market in future years.

An issue related to labour market considerations is economic dependency. Age dependency is often used as a proxy for economic dependency because of problems in defining economic dependency. An age dependency ratio is the ratio of children (0–14 group) and the elderly (65 years and older) to the working age group (15–64 years old). This ratio is thus a proxy measure of the number of people not working who are dependent on the working population. This information can be obtained from the age-sex distribution in the 2001 census. The main determinant of the magnitude of the dependency ratio in any population is the number of children aged 0–14 relative to the working age group. Thus, the age dependency ratio is a strong indicator of the magnitude of the "burden" of children on the working-age group. A large burden of children potentially reduces the saving capacity of the labour force, which, in turn, potentially reduces investment and capital formation in the population. Economic policies designed to stimulate the economy would be enhanced if they were formulated with some knowledge of the extent of the burden of children on the working age group.

Age-sex distribution and planning for social welfare services

Information about the magnitude of the "burden" of children is also important for planning in other ways. For example, policy makers could use the information to assess the adequacy and distribution of existing facilities (e.g. recreational centres) for children. This information could inform planning for the provision of additional facilities, if necessary.

The age-sex distribution revealed by the 2001 census provides a yardstick for measuring the current size of the elderly population by sex. The future size of the elderly population can be estimated on the basis of the current age-sex distribution combined with other demographic variables. This information is critical for appropriate policy decisions on resource allocation to meet the present and future needs of the elderly population including old age benefits, old age homes and other pertinent social welfare services for the elderly. This may be important where the traditional support systems for the elderly are eroding due to increasing modernisation and urbanisation.

Age-sex distribution and planning for housing

Information on the 2001 census age-sex distribution is important in planning for the provision of housing units. Combining the age-sex distribution with the headship

information (based on the relationship question) can be useful in estimating future housing requirements. Such information is valuable for planning in the housing sector as it has implications for resource allocation.

10.3.2 Fertility variables

Fertility relates to the number of live births a woman has had. Fertility has two components: lifetime fertility and current fertility. The former is measured by parity in each reproductive age group (children ever borne by women in each reproductive age group). The latter is measured by the total fertility rate (the number of live births a woman would have if she survived to age 50 years and experienced the current fertility rate in each reproductive age group).

The following questions constituted the fertility variables in the 2001 census.

(For women aged between 12 and 50 years at the time of the census):
How many children, if any, has (the person) ever had that were born alive?
How many of these were boys?
How many of these were girls?

If (the person) has ever given live birth, when was (the person's) last child born?

Teenage fertility (fertility among women under the age of 20) and nuptiality (the demographic study of marriage and its dissolution) are two aspects of fertility.

The 2001 census provides indirect information regarding current levels of teenage fertility based on the question:

(For women under 20 years of age at the time of the census):
How many children, if any, has (the person) ever had that were born alive?

Direct information on nuptiality was collected in the 2001 census through the following question.

What is the (person's) PRESENT marital status?

The following paragraphs illustrate how the fertility variables in the 2001 census can be used for decision making and planning.

Fertility and planning in respect of socioeconomic demands

Fertility is a direct determinant of population growth and age-sex distribution. The future size of the population and its age and sex composition are determined partly by the levels of fertility in the population. Consideration of the levels of fertility at the national and subnational levels is therefore important for planning in all sectors.

Sustained high levels of fertility can increase various future socioeconomic demands since high fertility fuels population growth. By the same token, declining fertility may reduce the future demand for certain kinds of services. Awareness of this fact is important for planning so that the allocation of resources to some sectors can be reduced and diverted to other sectors. Knowledge and quantification of the link between fertility and socioeconomic demands thus enables decision makers to allocate resources more efficiently. Apart from this and other considerations, fertility levels have more obvious implications for planning in specific sectors. The following paragraphs illustrate this point.

Fertility and planning for education

Sustained fertility decline decreases the proportion and number of persons of school-going age in the population thus, in turn, decreasing the cost of education provision. The 2001 census fertility information in combination with other information can help in the assessment of present and future demands for schools and, in so doing, inform planning in the sector.

Fertility and the environment

As already noted, high fertility fuels population growth, which can put pressure on the environment. Sustainable development therefore requires environmental policies to take cognisance of fertility levels and trends.

Before policy makers can provide adequate sanitation and safe water in the future, they need information on the number of people who will require these services. The 2001 census provides valuable baseline information on fertility that can help in such as assessment.

Fertility and planning in reproductive health

The level of fertility and teenage fertility in a population are among the most commonly used indicators of reproductive health. The following goals and objectives are included in the Year 2000 health goals and objectives for South Africa.

1. Reduce the proportion of births among girls under 16 and 16–18 years to 5% and 10% respectively.
2. Increase clinic attendance for contraceptive and family planning services.
3. Reduce infant and child mortality and morbidity.
4. Reduce maternal mortality.

With regard to the reduction of teenage pregnancy, information from the 2001 census can be used to gauge the extent of teenage pregnancy in South Africa. This information can be used as baseline information for monitoring the effectiveness of programmes designed to achieve a reduction in teenage pregnancy.

Indirect insights into the other three health objectives can be obtained from the 2001 census fertility information as a certain level of fertility implies certain levels of contraceptive prevalence (Frank 1983). Also, certain levels of fertility usually imply certain birth intervals and, hence, certain levels of infant and childhood mortality and morbidity as well as maternal mortality.

Information from the 2001 census on teenage pregnancy and fertility levels can therefore provide direct and indirect insights into the extent of certain reproductive health issues. On the basis of these insights, new policies can be formulated and/or existing policies revised.

In addition, risk groups can be identified regarding some of the reproductive health issues outlined above by relating teenage pregnancy and fertility to other socioeconomic variables (such as education and income) in the 2001 census so that policies and intervention programmes can be more specifically targeted.

Fertility and planning for the elderly

As fertility declines, the proportion and also the number of elderly people in the population increases. This situation generally leads to an increase in geriatric diseases and the demand for health services for the detection, management and treatment of such diseases. Appropriate analysis of the fertility variables in combination with other demographic variables (mortality and migration) from the 2001 census can help predict the proportion and number of elderly persons in the future so that adequate health care services for the elderly can be planned for.

Importance of marital patterns for planning

Marital patterns (nuptiality) indirectly impact on some of the planning issues relating to population growth noted above since these patterns are one of the direct determinants of fertility. Policies designed to have an effect on fertility levels are usually formulated within the framework of the direct determinants of fertility (including marital patterns). Marital patterns may also have implications for social welfare support at the level of the household. A vulnerable group in this context partly due to marital breakdown may be households headed by females who have low levels of education compared with other forms of household headship. Female-headed households may need more social welfare to alleviate poverty in such households. Information on marital patterns together with other socioeconomic variables such as those contained in the 2001 census can be used in assessing the need for such support.

10.3.3 Mortality variables

Mortality is the demographic study of death in human populations. The level of mortality is one of the indicators of the well-being and health status of a population. It

is also an indicator of the level of human development hence its inclusion in the construction of human development indices. The multidimensional approach to poverty recognises that the level of mortality is an indicator of poverty in a population.

There are several indices (measures) of mortality. The most commonly used measures include infant mortality rate, mortality in the first five years of life (under five mortality rate), life expectancy at specific ages (including life expectancy at birth) and maternal mortality rate (maternal deaths per 10 000 women in the reproductive age group). Most governments implement policies to lower the level of mortality in their countries. In countries (including South Africa) that have explicit or implicit policies relating to mortality or health goals, indices of mortality provide a yardstick for monitoring the success of programmes aimed at reducing mortality.

Because mortality is one of the components that determine population growth, it should be considered with other demographic variables in the planning process since it impacts on age-sex distributions, fertility and migration (see next section).

The 2001 census included questions on mortality. These questions can be classified broadly into two categories: *childhood* and *adult* mortality questions. The following were the childhood mortality questions.

(For women aged between 12 and 50 years at the time of the census):
How many children, if any, has (the person) ever had that were born alive?

How many of these were boys?
How many of these were girls?

If (the person) has ever given live birth:
If boys: how many boys are still alive?
If girls: how many girls are still alive?

Indices of childhood mortality can be computed by sex on the basis of these questions. The adult mortality questions can be further categorised into indirect and direct adult mortality questions.

Indirect questions:

Is (the person's) own biological mother still alive?
Is (the person's) own biological father still alive?

Direct questions:

Has any member of this household died in the past 12 months?

Of those deceased:

What was the month and year of death?

What was the sex of the deceased?

What was the age of the deceased in years at death?

Did the person die from an accident or through violence?

If the deceased was a woman under 50 years, did she die while pregnant or within six weeks after delivery?

The indirect questions are referred to as “orphanhood questions” for estimating adult male and adult female mortality respectively. Although some researchers have misleadingly used the orphanhood questions to estimate the number of orphans aged less than 15 years in the context of HIV/AIDS, the questions were not designed for that purpose. They were designed for the estimation of adult mortality. Brass, who developed the orphanhood questions, cautions about the reliability of the orphanhood information pertaining to younger persons (especially those under 15 years of age) and recommends that this part of the information should not be used for estimation purposes (Brass 1971) as it is likely to be biased by the “adoption effect” (Blacker 1977; Hill & Trussel 1977; Timaeus 1991; Udjo 1997).

Appropriate indirect techniques of analysis of the childhood and adult mortality questions can provide information about the levels and trends in mortality in the population and, on the basis of this information, future levels of mortality in the population can be projected. The direct questions can also be used to estimate the levels of adult and maternal mortality but only for the reference period for which the questions were asked. The following paragraphs illustrate the usefulness of mortality information in planning.

Mortality and planning for health

As noted above, the Year 2000 health objectives for South Africa include a reduction in maternal mortality and a reduction in infant and childhood mortality and morbidity. Certain indices of mortality can be computed from the mortality questions in the 2001 census and, among other things, used as baseline information for planning with regard to the above health objectives. For example, the 2001 census data suggest that childhood mortality has been increasing in recent years. Further analysis of the data may help researchers identify vulnerable groups in the population and, on the basis of this information, policy makers can introduce appropriate intervention programmes to reverse the trend. The information from the 2001 census about trends in mortality could be the platform for exploring other data sets to attempt to isolate factors that may be responsible for the trends. Knowledge of these factors could then form the basis for formulating intervention programmes.

Mortality and life assurance

The 2001 census mortality information could inform decision making in the insurance industry. Assessment of mortality risks in the population is central in setting insurance premiums. Insurance premiums are not static; they are revised from time to time depending on new information regarding mortality risks. The 2001 census provides the most recent data on mortality at national, provincial and smaller geographical area levels.

10.3.4 Migration variables

Migration usually refers to a process of moving from one administrative area to another with the intention of staying in the new area. Thus, migration has two broad components: international and internal migration. The former involves moves across international boundaries, and the latter involves moves across administrative boundaries within a country. Information on migration from the 2001 census is based on the following questions.

International migration:

Was (the person) born in South Africa? (including the former homelands)
In which country was (the person) born?
Is (the person) a South African citizen?

Note that the above questions provide information about immigration only. Information about emigration is usually collected from administrative records.

Internal migration:

(For persons born in South Africa):

(1) In which province was (the person) born?
(2) Five years ago (at the time of Census '96), was (the person) living in this place?
(3) If five years ago (the person) was not living in this place, where did (the person) move from?
(4) If five years ago (the person) was not living in this place, in which year did (the person) move to this place?

Migration has socioeconomic consequences for the state, community and individuals at the sending and receiving areas. An important consideration in respect of the consequences of migration in planning is the volume of the "net flows" (balance between "inflows and outflows": immigration and emigration in the case of international migration; in-migration and out-migration in the case of internal migration). The volume of immigration, in-migration and out-migration by province in

2001 can be estimated from the 2001 census migration questions. This information, combined with administrative records on migration, can inform policy in respect of the mitigation of the consequences of migration.

Migration and population growth

Migration is one of the components of population growth that affects the size and composition of the population at the national and subnational levels. This component is usually less important than the other demographic components (fertility and mortality) at the national level in less developed countries. However at the subnational level (e.g. provincial level), it is an important determinant of population growth. The 2001 census can provide insight into the additional number of persons due to migration at the national and provincial levels and in the other smaller geographical areas that planning focuses on.

Migration and labour force planning

The size and distribution of the labour market are partly determined by net migration. Some areas in the country are characterised by inflows of people from neighbouring areas for employment-related reasons. These inflows can compound problems of unemployment (if any) in the receiving areas. Policy makers should be able to gauge and monitor the volume of such movements so that appropriate policies can be put in place to cope with the resultant growth in the labour force. The 2001 census migration information in combination with the age-sex distribution information is useful in this regard.

Migration and environmental planning

Rural-urban migration accelerates the growth of urban areas. Increased urbanisation encroaches on the natural environment. Squatter settlements at the fringes of urban centres arising from rural-urban migration may deplete non-renewable resources and accelerate environmental degradation. The 2001 census migration information can provide a baseline for gauging and monitoring rural-urban migration and can inform environmental policies aimed at preventing undesirable consequences of migration.

Migration and planning for social services

Migration often leads to increased demand for housing and education as well as health, sanitation, water, electricity, safety and security services at the point of destination. Migration may exacerbate pressure on available infrastructure and social services, especially in urban areas. Policy makers should therefore take into consideration additional demands due to migration when allocating resources for the provision of infrastructure and social services so as to maintain and improve the standard of living of residents and avoid social unrest. The 2001 census migration

information in combination with other variables can help officials estimate the additional demand in various sectors due to migration.

10.4 Conclusion

The above discussion shows that demographic variables are intricately linked to all forms of planning in the private and public sectors. While some of the linkages are obvious in some sectors, they are less so in others. Lack of consideration of demographic variables in planning in one sector may have undesirable consequences in other sectors. For example, if the housing sector does not take into consideration the increased demand for housing and electricity due to natural increase or migration in some areas, this may accelerate environmental degradation in those areas. Consequently, the importance of a multi-sector approach in planning cannot be over-emphasised.

Appropriate analysis of demographic variables in combination with other variables can provide direct and indirect information that is crucial to planning in various sectors. An integrated national statistical database that incorporates the results of such analysis and other indicators can be a key facilitating instrument for planning in all sectors.

Bibliography

Brass W. (1971). *Methods for estimating fertility and mortality from limited and defective data*. The University of North Carolina at Chapel Hill.

Blacker J. G. C. (1977). The estimation of adult mortality in Africa from data on orphanhood. *Population Studies*, vol. 31, no. 1, pp. 107–128.

Crone M. D. (1990). *Counting heads: a guide to census taking in developing countries*. Overseas development administration, London.

Frank O. (1983). Infertility in sub-Saharan Africa: estimates and implications. *Population and Development Review*, vol. 9, no.1, pp. 137–144.

Hill K. & Trussel J. (1977). Further developments in indirect mortality estimation. *Population Studies*, vol. 31, no. 2, pp. 313–334.

Shyrock H. S., Siegel J. S., & Associates (1976). *The methods and materials of demography*. Academic Press, New York.

Timaeus I. M. (1991). Estimation of mortality from orphanhood in adulthood. *Demography*, vol. 28, no.1, pp. 231–217.

Udjo E. O. (1997). *Fertility and mortality trends in South Africa: the evidence from the 1995 October household survey, and implications on population projections*. Central Statistical Service, Pretoria.

Udjo E. O. (in press). *An evaluation of age-sex distributions of South Africa's population within the context of HIV/AIDS*. Development Southern Africa.

United Nations. (1958). Principles and recommendations for national population censuses, *Statistical papers*, series M, no. 27, p. 3.

Khalifani A. K., Zuberi T., Bah S., & Lehohla P. (in press). Population statistics in South Africa. In, Zuberi T., Sibanda A. & Udjo E. O. (eds). *The demography of South Africa*. M. E. Shape, New York.

Chapter 11: Gender

Since South Africa's first democratic election in 1994, the government has deliberately set out to dismantle apartheid-era social relations and create a democratic society based on the principles of equity, non-racialism and non-sexism (PCAS, 2003). Mechanisms put in place through new policies and programmes were meant to dramatically improve the quality of life for all South Africans. The defining point for this process was the Reconstruction and Development Programme (RDP), which has been elaborated in all post-1994 policies. The key objectives of the RDP were to meet basic needs, build the economy, democratise the government and society and develop human resources and nation building. Inherent in government priorities since then has been the need to achieve gender equality in South Africa.

The purpose of this chapter is to contribute to an understanding of gender statistics in an era of democratic change that proposes to eradicate gender inequality. For South Africa to achieve gender equality, it is necessary to recognise that every policy, programme and project affects men and women equally. While this does not mean that men and women are the same, it does carry an explicit assumption that women's empowerment would be central to gender equality. Achieving gender equality requires that specific objectives with specific measurable indicators to measure progress be put in place.

The objectives of this chapter are three-fold. First, it sets the scene for an understanding of the uses of gender indicators derived from census data in informing policy formulation and planning. Second, it identifies gender disparities and gaps in the demographic and socio-economic indicators captured in the census, which will enable a more complete compilation of gender statistics.

11.1 Key concepts and definitions

Gender-specific statistics constitute a new field. Unlike the simple disaggregation of data by sex, gender-specific statistics are built on concepts and definitions designed to detect gender-differentiated conditions and characteristics and gender interactions.

More often than not the word *gender* is incorrectly used as synonymous with the word *sex*. At first glance this may appear harmless, yet it has significant implications for interpretations made and for the measurement of progress towards gender equality. It is important to point out the difference between sex and gender because a population census does not necessarily set out to establish gender indicators or measure gender relations in South African society. The census questionnaire rather determines certain variables on the basis of sex. The analysis and interpretation of key variables on the basis of differences between men and women lead to a gender

analysis of population census data. *Sex* therefore does not have the same meaning as *gender*, and the two terms should not be used interchangeably.

Sex is the biologically determined state of being male or female. It defines the biological differences between boys and girls. Sex is genetically determined at conception.

Gender is a socio-cultural construct of society that determines the identity, roles or functions, entitlements and deprivation of women and men in society. Gender does not prescribe a focus on women only but emphasises the differences between men and women with respect to entitlements, resources and decision making (World Bank, 2001).

To illustrate, certain functions or roles related to sex, such as childbearing by females, are predetermined. However, when a female is denied benefits such as access to education or is disadvantaged on account of being female, this constitutes a gender bias and creates a gender difference (Bekele, 1997). While sex roles are predetermined, gender roles are shaped through social interactions and the historical, political, economic and cultural climates of a particular society.

The inequality of men and women is an important factor contributing to their participation or non-participation in the development process. Population census data can provide an overview of the gender gaps that exist in South Africa. Highlighting these gender gaps is important for an understanding of the contribution of men and women to the development process in South Africa.

Gender gaps are the differences between men and women in relation to a particular socio-economic or demographic indicator.

To illustrate, census data can provide an overview of gender gaps in South Africa by displaying differential access of males and females to education, income, employment and other resources. The census data can also provide a synopsis of the differences between men and women in terms of age-specific mortality rates, life expectancy, the proportion of the population and the spatial and age distribution of the population as well as a synopsis of the number of households headed by men and by women and the differences in living conditions between these households. Census data can also provide a snapshot of the differential nature of disability prevalence between men and women and the extent to which men migrate compared to women.

Gender gaps displayed through census data come about as a result of unequal power relations between men and women. The absence of gender gaps is an

indication of a balance of power (Bekele, 1997). The existence of gender gaps is an indication of inequality of access to resources and an imbalance of power between men and women. In the words of Amartya Sen (UNFPA, 2000), "…the overarching objective of development is to maximize people's capabilities … to lead the kind of lives they value …" Persistent gender inequalities in South Africa mean that South African society carries a higher financial and social cost. For example, unequal power relations in the household and reduced reproductive health care may increase the incidence of maternal mortality from puerperal causes (causes related to pregnancy and childbirth).

11.2 Uses of census data

Regarded as one of the most important tools for obtaining data on the characteristics of a population, the population census has become an important source of information on the men and women in South Africa in order to understand their contribution to development and the impact of development on both.

The demand for gender-specific data and indicators incorporating a gender perspective is increasing rapidly. Such data include information on who does what, levels of access to resources for both men and women and the demographic implications for both men and women There is therefore a growing realisation that gender-specific statistics and the analysis of such statistics are needed to meet government priorities and to integrate gender into all South Africa's policies. As the user audience has expanded to include decision makers at every level and in every area of social and economic development policy formulation, there is an even greater need for pertinent statistics.

Analysis that takes gender into account is essential for policy formulation and programme planning to ensure equity in resource allocation. When gender gaps are identified, planners can develop appropriate strategies to correct imbalances, influence budgetary allocations and focus on improving human resources for sustainable development.

Statistics incorporating a gender perspective are essential for advocates of gender equity; planners for use in economic and social policy formulation, implementation and monitoring; development experts who want to review and analyse gender aspects and interactions; the international community, governments and NGOs for use in project and programme design, implementation and evaluation; and the general public for a better understanding of society.

The local and international context of gender frameworks is testimony to the fact that as awareness of policies and programmes concerning gender increases, even more needs to be done. The International Conference on Population and Development

(ICPD) Programme of Action, the Beijing Platform of Action and, closer to home, the SADC Platform of Action call for gender-disaggregated data in order to enhance knowledge on the position and role of gender in social and demographic processes. Reporting and analyses of census and other data sources should thus be designed to contribute to a better understanding of selected aspects of population, gender and development in South Africa.

In this context, the objective of a gendered analysis of population census data should be to:

1. identify gender disparities and gaps in the demographic and socio-economic indicators captured
2. provide useful information for future policy formulation and development planning
3. make recommendations for improvement of data collection in future censuses which will enable a more complete compilation of gender statistics.

11.3 Methodologies, indicators and checklists for a gendered analysis of population data

The need for sex-disaggregated data has been stressed in numerous international conventions and declarations, including the 1979 Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW), the 1985 Nairobi Forward-Looking Strategies for the Advancement of Women and the 1995 Platform for Action of the Fourth UN World Conference on Women in Beijing.

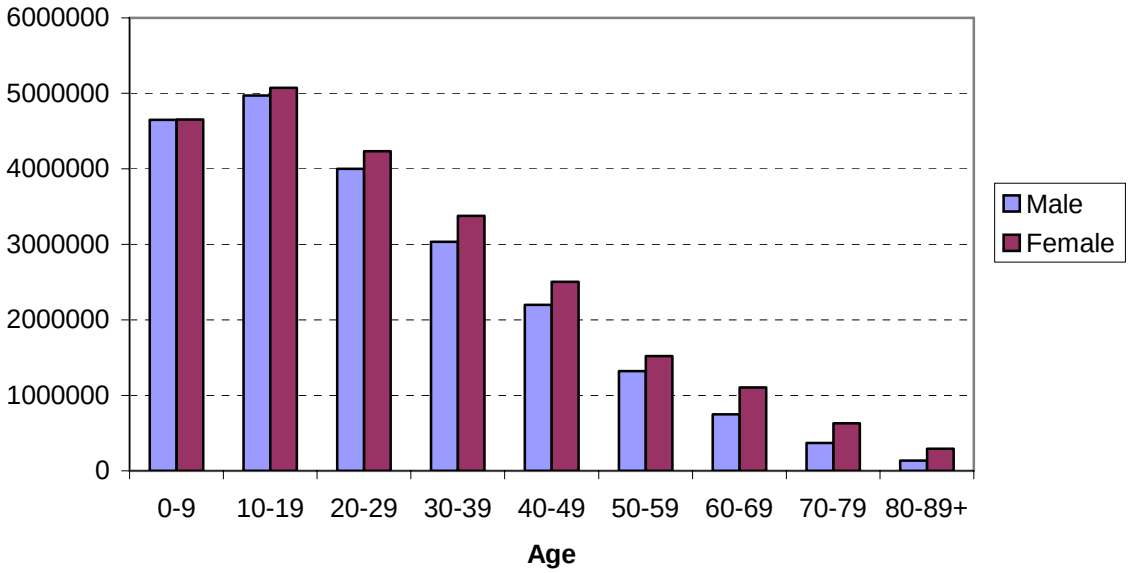
11.3.1 General methodology for a gendered analysis of population data

Census data can be used effectively to identify socio-economic and demographic gender gaps in the South Africa population. A number of indicators can be created, many of which are listed in the different gender-analytical frameworks that outline approaches to gender analysis. Such frameworks include the Moser Framework, which focuses on inadequacies in living conditions such as housing, or the ABC of Gender Analysis, which offers guidelines to identify gender gaps in education (see for example *Statistical Indicators of Female Participation in Education in Sub-Saharan Africa*).

A number of very simple computations can be undertaken to measure the magnitude of inequality and disparity between women and men, depending on the indicator used for measurement.

While population distribution does not provide a sense of inequality between males and females, Figure 1 is a simple illustration of the distribution of males and females in South Africa by age group.

Figure 1: Population distribution by age and sex



The graph indicates an excess of females over males. Sex ratios in South Africa also indicate conformity to the general global pattern of high sex ratios at birth and younger ages and low sex ratios at more advanced ages (Table 1).

Table 1: Sex distribution and sex ratios by age, 2001

Age	Percentages		Sex ratio
	Male	Female	
0–9	49,98	50,02	99,91
10–19	49,50	50,50	98,04
20–29	48,59	51,41	94,50
30–39	47,34	52,66	89,91
40–49	46,77	53,23	87,85
50–59	46,49	53,51	86,88
60–69	40,43	59,57	67,87
70–79	36,94	63,06	58,57
80+	31,93	68,07	46,90

Relative frequencies

The simplest and most commonly used indicators of gender inequality are **relative frequencies** (percentages) of males and females rather than absolute frequencies (total numbers). For example, out of the total population that has never been to school, what percentage is female, and what percentage is male? The inherent assumption under ideal conditions would be that female and male populations are close to 50%.

Table 2: Percentage distribution of persons aged 15 years or older with no education by sex and population group, 1996 and 2001

	African	Coloured	White	Indian
1996				
Male	21,5	12,2	4,7	6,1
Female	22,9	11,8	4,3	9,5
2001				
Male	14,8	8,7	3,2	4,5
Female	18,0	8,6	3,1	7,4

Source: Statistics South Africa, Population Censuses, 1996 and 2001

Table 2 shows the overall distribution of persons with no education by sex and population group for 1996 and 2001. It indicates that while there was some improvement in the overall access to education between 1996 and 2001, the racial disparity in education remained. Overall, the percentage of people with no education in 2001 was lower than in 1996. It also suggests that African males remained about 4,6 times as likely as white males to have had no education in 1996 and 2001.

In terms of access to education, African women were slightly worse off in 2001 than in 1996. African women were about 5,3 times as likely as white women not to have had any education in 1996 compared to nearly 6 times as likely in 2001. Moreover, the gender difference in access to education for Africans was greater in 2001 than in 1996. In 1996, nearly equal proportions of African men and women had no formal schooling; by 2001, African women were 1,2 times as likely as African males to have had no schooling.

Rates and ratios

Rates and ratios are also used to show general disparities in the selected indicators. Where appropriate, the magnitude of the disparity may be assessed in terms of gender ratios, gender gaps and relative gaps, especially where one group appears to be significantly disadvantaged.

The **gender ratio**, sometimes referred to as the **gender parity index (GPI)**, measures the extent of inequality. The GPI ranges from 0 to 1, where 0 stands for inequality and 1 for equality. When computed as a percentage (F/M x 100) it denotes the number of females per 100 males for any indicator under study.

The gender ratio is similar to the sex ratio computed in demographic analysis (M/F x 100). However, for gender analysis the female population is divided by the male population if the focus is to emphasise the position of females relative to males. The rationale for reversing the computation, especially in the analysis of socio-economic characteristics such as education and economic activity, is based on what needs to

be emphasised in the study. The method is recommended by Unesco (1997) for the compilation of gender-sensitive statistics**.**

The **gender gap** also measures the magnitude of disparities and is computed by the formula F-M. When computed from actual population figures (Fpop-Mpop), it is referred to as an **absolute gap.** The gap may be positive or negative, depending on whether the inequality is in favour of males or females.

Table 3: Gender disparities in water source for households

Source of water	% of female-headed households	% of male-headed households	Gender gap (F-M)
Regional / local water scheme	66,7	73,9	-7,3
Borehole	8,7	9,2	-0,4
Spring	3,9	2,5	1,4
Rainwater tank	1,4	13	0,1
Dam / pool / stagnant water	3,4	2,9	0,5
River / stream	13,2	7,7	5,5
Water vendor	2,0	1,8	0,2
Other	0,7	0,7	0,0

Table 3 indicates the gender disparities in terms of the main source of water for male-headed and female-headed households in South Africa. This information is important as water source impacts significantly on women's use of time and labour in particular. Table 3 indicates that significantly more female-headed households than male-headed households rely on rivers and streams for water for household use. Census data also suggest that regional and local water schemes (i.e. clean water) seem to benefit male-headed households disproportionately.

The **relative gap** measures the extent to which interventions are required or how much effort is required to achieve equality between men and women with respect to a particular indicator. The relative gap is computed as F-M/F x 100, denoting the percentage effort required to achieve gender equality. Again, the gaps can be in favour of men or women, depending on the indicator being measured.

11.3.2 Gender-sensitive indicators for monitoring progress

While Census 2001 provides valuable information on a cross-section of variables, the selection of indicators to measure gendered development is not prescribed by the census but rather may be defined by the analysis framework adopted. The extent to which the indicators may be generated is then determined by the information that can be gleaned from census data.

An ***indicator*** is an item of data that summarises a large amount of information in a single figure in such a way as to give an indication of change over time and in comparison to a norm. Indicators differ from statistics in that, rather than merely presenting facts, indicators involve comparison to a norm in their interpretation.

National-level gender-sensitive indicators are among the key means by which planners and policy makers measure gender inequality. These indicators also provide information on the basis of which gender specialists advocate policies likely to lead to greater gender equality. Gender-sensitive indicators support the Gender and Development (GAD) approach which focuses on changing the gendered nature of society through the promotion of gender equity rather than on women in isolation, which was the focus of the Women in Development (WID) model, and is reflected in an emphasis on gender statistics.

For example, according to South Africa's National Gender Framework adopted by the Office on the Status of Women country-specific indicators for South Africa would be as listed in Table 4.

Table 4: Country-level gender-sensitive indicators and checklists

Measure	Indicator (sex-disaggregated data)
Gender Empowerment Measure (GEM)[26]	• Seats in parliament • Legislators, senior officials and managers • Professional and technical workers
Gender-related Development Index (GDI)	• Life expectancy at birth • Adult literacy rate • Combined primary, secondary and tertiary gross enrolment ratio • Estimated earned income
Gender inequality in education	• Adult literacy rate • Youth literacy rate • Primary net enrolment ratio • Secondary net enrolment ratio • Tertiary net enrolment ratio
Gender-related Education for All Index (GEAI)	• Total participation in primary and secondary education • Adult literacy
Gender inequality in economic activity	• Economic activity rate • Employment in agriculture • Employment in industry

[26] The Economic and Social Commission for Asia and the Pacific (ESCAP), the Canadian International Development Agency (CIDA), the Bangladesh Rural Advancement Committee (BRAC) and Grameen Bank have developed other gender-empowerment measures.

	• Employment in services • Contributing family workers • Unemployment rate • Youth unemployment rate • Long-term unemployment rate
Women's political participation	• Women in government at ministerial level • Seats in cabinet, national assembly and provincial parliaments
Gender inequality in health	• People living with HIV/AIDS
Service Deprivation Index (SDI)	• Female-headed households: housing, energy (cooking, heating and lighting), water, toilet facilities and refuse removal
CEDAW[27]	• Development and advancement of women: legislation and admin procedures, National Gender Machinery (NGM), policy, budgeting, training and research • Equality between women and men: public service and private sector • Sex roles and stereotyping: custom, tradition, religion, language, education and media • Exploitation of women: prostitution and trafficking • Political and public life: parliamentary representation, provincial leadership, local government and the judiciary • International representation and participation: women in SA foreign missions and diplomatic service training • Nationality: citizenship, immigration and refugees • Education: schools, Early Childhood Development (ECD) and Adult Basic Education and Training (ABET) • Employment: rates, institutional arrangements, rights, employment opportunities, training, work-related social security, health and safety and child care • Health care: health profile, major health problems, access to health care and reproductive health • Social and economic benefits: child and family grants, old age pension, sport, arts and culture and credit • Rural women: land reform, women farmers, women farm workers, ICT, health services, training and education and development programmes • Equality before the law and in civil matters: legislation, access to legal services and freedom of movement and domicile

[27] The Convention on the Elimination of All forms of Discrimination Against Women was ratified by South Africa in 1995. CEDAW is meant to measure the extent to which women are discriminated against compared to men in different spheres of life.

	• Marriage and family law: forms of marriage, rights in marriage, divorce, rights in respect of children, inheritance and family court • Violence against women: incidence, legislation and victim support

It is clear that census data cannot provide input for the computation of all the gender indicators in Table 4. Data from more specific surveys and other gender-specific sources are required. This should be noted by national statistical offices and those in charge of national statistics systems for them to meet the challenge of gender-mainstreaming their statistical systems through the provision of relevant data inputs and to ensure the development of relevant indicators of measurement.

South Africa's commitment to the Millennium Development Goals (MDGs), together with that of 191 other nation states, has presented a new set of challenges in terms of measuring progress towards achieving gender equality. MDGs are the goals and targets from the Millennium Declaration to be achieved universally by the year 2015, with specific indicators to monitor progress.

Notwithstanding arguments by development feminist writers that only one MDG makes specific reference to gender equality, gender equality is essential for reaching all eight millennium development goals.

Table 5: Excerpt from Millennium Development Goals (Goal 3)

Millennium Development Goal 3: Promote Gender Equality and Empower Women	
Goal and Target *(From Millennium Declaration)*	**Indicators for Monitoring Progress**
Target 4: Eliminate gender disparity in primary and secondary education preferably by 2005 and on all levels of education no later than 2015.	• Ratio of girls to boys in primary, secondary and tertiary education • Ratio of illiterate females to males of 15–24 years old • Share of women in wage employment in the non-agriculture sector • Proportion of seats held by women in national parliament

Again, it becomes clear that the census is not a sufficient tool to generate all relevant indicators for this goal.

11.4 Gender indicators and gender-sensitive analysis of population census data

A number of GAD approach indicators allow for detailed analysis of existing gender gaps using census data. These include the United Nations Women's Indicators and Statistical Databases (WSTAT) of 1999, which outline a list of demographic and socio-economic indicators that can be used. The United Nations Scientific and Cultural Organisation (UNESCO) has also produced a guideline for gender-sensitive indicators.

The following classes of gender-sensitive indicators have been developed from Commonwealth government priorities, UN recommendations and priority areas determined by the Beijing Platform for Action:

1. Population composition and change
2. Human settlements and geographic distribution
3. Households and families, marital status and fertility
4. Learning in informal and non-formal education
5. Health, health services and nutrition
6. Economic activity and labour force participation
7. Access to land, equipment and credit
8. Legal rights and political power
9. Violence against women
10. Macroeconomic policy and gender

The three main data systems that can be used to produce gender-sensitive indicators are population censuses, sample surveys of the population and the National System of Accounts (SNA). Censuses and similar sample surveys provide data for indicators 1 to 6. Data for indicators 7 to 10 can be derived from other sources, such as those suggested by CEDAW, special surveys such as time-use studies and a revised System of National Accounts (SNA). Measures such as the SNA and Gross Domestic Product (GDP), with their concentration on measuring paid employment, are less useful. They have been strongly criticised for having a gender bias and in particular for ignoring women's contribution to the economy and to society as a whole.

State parties to CEDAW are required to submit reports every four years on all articles of CEDAW. Reporting on CEDAW offers governments an excellent opportunity to synthesise available gender-related data, to measure advancements in the status of women and to identify and fill gaps. While census and other national-level surveys deal mainly with demography, work, health and education, CEDAW offers governments the opportunity to synthesise and produce gender-sensitive indicators of empowerment, violence against women, cultural issues and women's rights.

A checklist of methodological points to bear in mind when using gender-sensitive indicators at the national level:

- **Comparability**: This entails comparison to a norm, for example the situation of men in the same country or the situation of women in another country. In this way the indicator can focus on questions of gender equality and equity rather than only on the status of women.
- **Sex-disaggregation**: Data need to be disaggregated by sex and where possible also by age, socio-economic grouping and national and/or regional origin.
- **Accessibility**: Sex-disaggregated data that can be used to generate indicators to measure gender equality need to be easily accessible in a usable format.
- **Scope**: This means scope of availability. (Data should be available for the whole country.)
- **Reliability**: No data is absolutely reliable, but reliability checks should be carried out.
- **Measurability**: Concepts such as women's empowerment or gender equity may be difficult to define and measure. In this case proxy indicators may be used, for example relating to greater choice for women in accessing health care or education.

- **Time-frames**: Gender-sensitive indicators should be reliable enough to use as a time series. The time span that the indicator covers should be clearly stated.
- **International comparability**: Gender-sensitive indicators should be collected using internationally accepted definitions.
- **Measurement of impact**: The indicator should, where feasible, measure the outcome or impact of a situation rather than the input. For example, women's literacy is often a better measure of women's educational status than female enrolment rates because literacy measures the impact of enrolment rates.
- **Participation**: Indicators should be used and developed in as participatory a process as possible. This involves the collection of qualitative data to supplement quantitative data.

11.5 Data quality

South Africa's Census 2001 provides information on all members of households who were present on the night of 9–10 October 2001. The census questionnaire was used to collect information on various demographic and socio-economic variables. However, the population census cannot be used independently of other data sources to explain social concepts and the situation in South Africa with respect to gender equality. Two observations can be made.

First, because the information on some household members obtained through the census questionnaire was not provided by the individuals themselves but by other persons present at the time of enumeration, there could be an element of inaccurate reporting on some of the socio-economic characteristics. It is difficult to assess the degree of inaccuracy for the various indicators used, but imputations done on the "Not Stated" category of responses for some socio-economic indicators such as education may indicate some inaccuracy in reporting.

Second, while the information collected in Census 2001 captured a number of demographic and socio-economic characteristics of the population, the data are limited in terms of the indicators that can be used for computation of summary indices such as the Gender Development Index (GDI). This limits comparability with other developing countries as well as with more developed countries. The GDI, for example, takes into consideration the literacy rates of the male and female population, the income levels in terms of Purchasing Power Parities (PPP) and differences in life expectancy. Data for the first two indicators must be obtained from alternative sources.

11.6 Conclusion

This chapter set out to raise awareness of the fact that the population census, although not deliberately intended to measure gender inequalities and disparities, can be used for this purpose. The chapter highlights how population census data can be engendered and what key indicators can draw attention to disparities that need to be addressed in policy planning and formulation.

The chapter also describes simple computations that can be carried out to measure gender gaps and disparities. On data quality, attention is drawn to the fact that there is a need in South Africa to improve the collection of quality gender-disaggregated data for efficient planning and development of programmes that are more gender-sensitive. This is important to help planners critically understand the relationship between population, gender and development. Whether more detailed questions can be incorporated into the census questionnaire in future in order to capture indicators necessary for as sessment of levels of gender development and for international comparison is an area for further debate.

Bibliography

PCAS (2003) *Towards a 10-Year Review*. Pretoria: Government Printers

Mason, D and E. King (2001). *Engendering Development*. Oxford University Press.

Bekele, F (1997). *The Gender Dimension in Development Projects: Concepts, Approaches and Framework for Gender Analysis*. Addis Ababa: Ossrea

Sen, A (2000) “Counting the Cost of Gender Inequality” in UNFPA (2000) The State of the World Population 2000

Teresa Hartnett and Ward Heneveld. 1993. **Statistical Indicators of Female Participation in Education in Sub-Saharan Africa**. AFTHR Technical Note No. 7. Africa Technical Department. Washington, D.C.: World Bank

UNESCO (1997) ***Gender-Sensitive Education Statistics and Indicators**. A Practical Guide*. Paris

Office on the Status of Women (undated). South Africa’s National Policy Framework for Women’s Empowerment and Gender Equality. Pretoria.

The Economic and Social Commission for Asia and the Pacific (ESCAP), the Canadian International Development Agency (CIDA), the Bangladesh Rural

Advancement Committee (BRAC) and Grameen Bank have developed other gender-empowerment measures.

The Convention on the Elimination of All forms of Discrimination Against Women was ratified by South Africa in 1995.

Chapter 12: Provision of basic services

12.1 Introduction

The South African government has formulated various policies aimed at alleviating poverty affecting households as well as individuals. These policies underscore the need to improve access to basic services, particularly clean water, electricity and sanitation. However, the service delivery process is slow due to the political policies of inequity before 1994 (The Presidency, 2003). The information collected in the 1996 and 2001 censuses provides an opportunity to assess what has been achieved since 1994. No comparable data exist for the period before 1994.

According to Goal 7 of the Millennium Development Goals (MDGs), which seeks to halve the number of people without sustainable access to safe drinking water by 2015, an adequate supply of clean water would reduce disease and improve the quality of life in developing countries. The government on the other hand is also committed to ensuring provision of electricity, housing and appropriate sanitation to the population as one of the pillars of the development agenda.

12.2 Potential uses of data on services

Data on service delivery are mainly used for planning. However, government departments providing these services also need to evaluate their progress. Statistics on household service delivery are used by international agencies for monitoring and evaluation of their programmes. The community at large is interested in progress made in improving living standards. Researchers need to inform policymakers about the progress with regard to poverty alleviation.

12.3 Delivery of basic services as an indicator of development

The delivery of basic services to households is among the key indicators with which a country's level of development can be measured. The following are regarded as the most important services because of their impact on health status and general living standard of households:

- clean piped water
- electricity for cooking, heating and lighting
- flush toilet facilities
- at least weekly refuse removal by local authorities
- communication facilities (telephone).

12.4 Objectives of the chapter

The main objective of this chapter is to highlight the importance of using data from the two censuses to analyse change in service delivery over time. The specific objectives are:

- to inform readers of the reliability of census data with regard to provision of basic services
- to highlight differences between 1996 and 2001 census questions with regard to household services
- to highlight trends in service delivery from 1996 to 2001

12.5 Questions on basic services

Collection of data on basic services by censuses is easier than collection of data on population variables. This is because most of these services can be seen by those mandated to collect the information.

Questions on basic services were mostly identical in 1996 and 2001. However, there were a few deviations with regard to categorisation, with the 2001 census categorising further than the 1996 census.

12.5.1 Water supply

In the 1996 census, the question on water services was as follows: “What is this household’s main water supply?” The following options were provided:

1 = Piped (tap) water in dwelling
2 = Piped (tap) water on the site or in the yard
3 = Public tap
4 = Water-carrier / tanker
5 = Borehole / rainwater tank / well
6 = Dam / river / stream / spring
7 = Other (e.g. from shops, hospitals, schools etc.)

In the 2001 census, the question was divided into two. The first part was posed as: “In which way does this household obtain piped water for domestic use?” Choices given were the following:

1 = No access to piped (tap) water
2 = Piped water on community stand: distance greater than 200 m from dwelling
3 = Piped water on community stand: distance less than 200 m from dwelling
4 = Piped water inside yard
5 = Piped water inside dwelling

The second part of the question on water supply in 2001 was the following: “What is this household’s main source of water for domestic use?” Choices were the following:
1 = Regional / local water scheme (operated by a water service authority or provider)
2 = Borehole
3 = Spring
4 = Rainwater tank
5 = Dam / pool / stagnant water
6 = River / stream
7 = Water vendor
8 = Other (specify)

The 2001 water supply questions were combined to derive a new variable comparable with that of 1996.

12.5.2 Electricity for cooking, heating and lighting

The 1996 and 2001 censuses used the same question and categories for the type of energy used for domestic purposes: “What type of energy/fuel is used by this household for cooking, heating and lighting?” Response categories were the following:
1 = Electricity
2 = Gas
3 = Paraffin
4 = Wood (cannot be used for lighting)
5 = Coal (cannot be used for lighting)
6 = Candles (cannot be used for cooking and heating)
7 = Animal dung (cannot be used for lighting)
8 = Solar
9 = Other (specify)

12.5.3 Sanitation facilities

In the 1996 and 2001 censuses, the question on sanitation facilities was asked in the same way: “What is the main type of toilet facility used by this household?” Response categories for the 2001 census were more than those for the 1996 census. In the 1996 census, there were four response categories:
1 = Flush toilet / chemical toilet
2 = Pit latrine
3 = Bucket latrine
4 = None of the above

In the 2001 census, there were seven response categories:
1 = Flush toilet (connected to a sewerage system)
2 = Flush toilet with septic tank
3 = Chemical toilet
4 = Pit latrine (with ventilation)
5 = Pit latrine (without ventilation)
6 = Bucket latrine
7 = None

For the sake of this comparison, the first three categories were combined into one, “flush or chemical toilet”, in order to match the categorisation of 1996. Categories 4 and 5 for the 2001 census were also combined into one, “pit latrine”, in order to match the categorisation of 1996.

12.5.4 Access to telephone facilities

In the 1996 census, the question on access to communication facilities was as follows: “Where do members of this household mainly use a telephone?” Response categories were the following:
1 = In this dwelling / cellular phone
2 = At a neighbour nearby
3 = At a public telephone nearby
4 = At another location nearby, e.g. work
5 = At another location not nearby
6 = No access to a telephone

However, in the 2001 census, the categories “telephone in the dwelling” and “cellular phone” were put under the question of household goods. Therefore, a new variable was derived with all six categories to match the categorisation of 1996.

12.5.5 Refuse removal

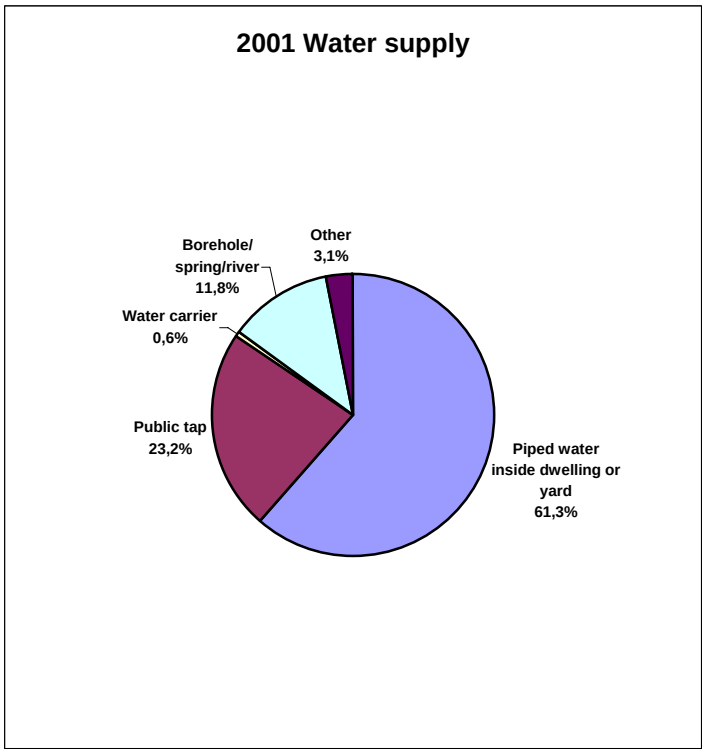
In both censuses, the question on household refuse removal was posed in the same way, with the same response categories:
1 = Removed by local authority at least once a week
2 = Removed by local authority less often
3 = Communal refuse dump
4 = Own refuse dump
5 = No rubbish disposal
6 = Other (specify)

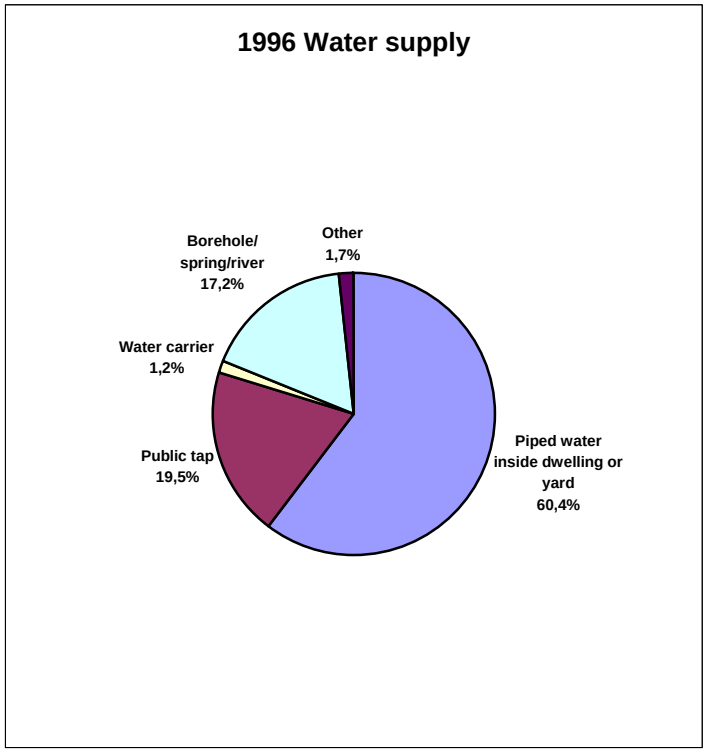
12.6 Trends in provision of basic services (1996–2001)

12.6.1 Water supply

The graphs below show that there was a substantial increase in the installation of public taps from 1996 to 2001. This is illustrated by the decrease in households relying on water-carriers / tankers, boreholes, springs and rivers as their source of water for domestic use and the increase in households using public taps in 2001. However, it should be noted that public taps in this context include those that are situated closer than 200m to or farther than 200m from the dwellings. Households with piped water inside the dwelling or in the yard also increased from 60,4% in 1996 to 61,3% in 2001. The above figures highlight the need for the government and other partners to intensify their efforts to provide all households in the country with access to clean piped water.

Figure 1: Comparison of the proportion of households' main source of water supply, 2001 and 1996





Total figures for different categories were divided by the overall column total for all the categories and multiplied by 100 to obtain the percentages as shown in Table 1.

Table 1: South African households categorised by main type of water supply in 2001 and 1996

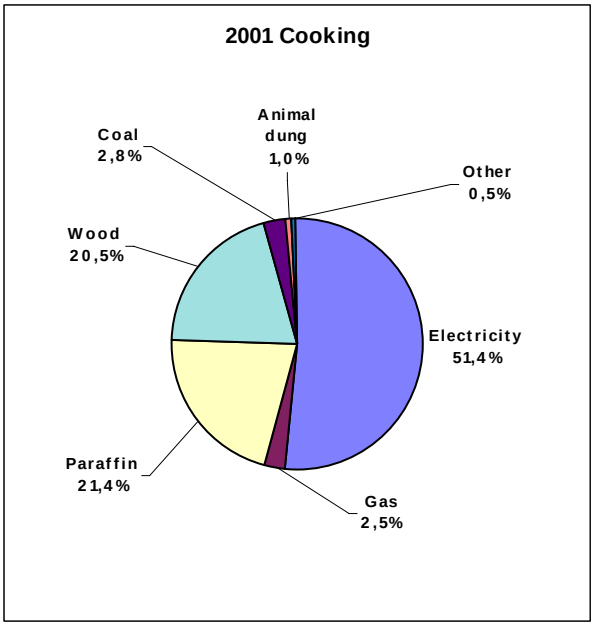
Type of water supply	2001		1996	
	Number	Per cent	Number	Per cent
Piped water inside dwelling or yard	6 871 464	61,3	5 468 083	60,4
Public tap	2 594 904	23,2	1 765 945	19,5
Water-carrier	67 680	0,6	111 204	1,2
Borehole / spring / river	1 320 937	11,8	1 558 368	17,2
Other	350 720	3,1	155 970	1,7
Total	**11 205 705**	**100,0**	**9 059 570**	**100,0**

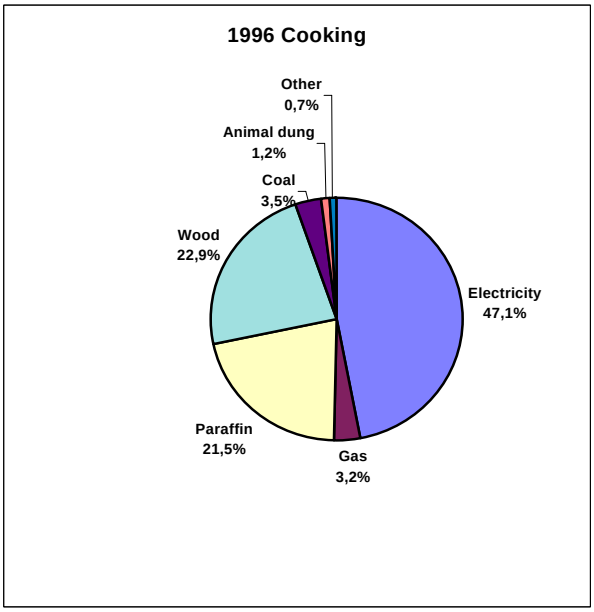
12.6.2 Energy for cooking

According to the graphs below, households that used electricity for cooking increased from 47,1% in 1996 to 51,4% in 2001. Households that used coal, animal dung and wood for cooking decreased from 28% in 1996 to 24% in 2001. However, those households that used paraffin for cooking constituted the same percentage in 2001 as in 1996: about 21%. This may be due to the formation of new households in informal settlements. Households that used gas for cooking also decreased from 3,2% in 1996 to 2,5% in 2001. The increase in the percentage of households using

electricity for cooking implies development and improvement in the living standards of South Africans.

Figure 2: Distribution of households by main source of energy for cooking, 2001 and 1996

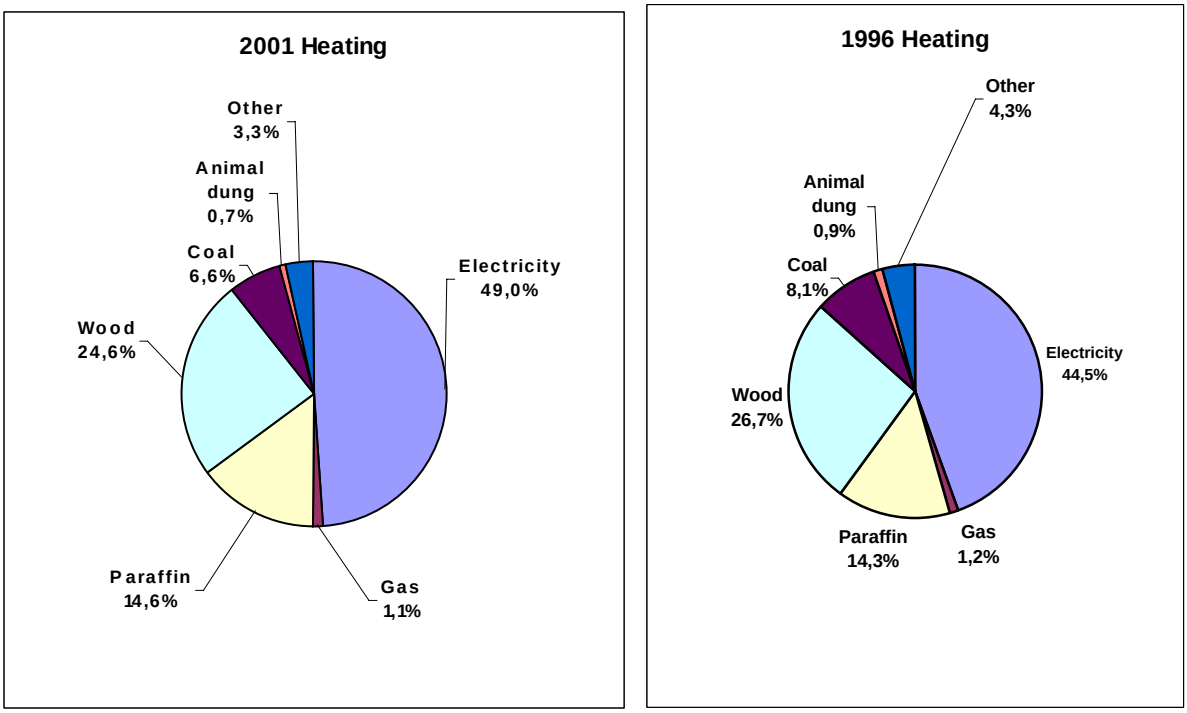




12.6.3 Energy source for heating

The graphs below show that households that used electricity for heating increased from 44,5% in 1996 to 49% in 2001. Households that used coal for heating decreased from 8% in 1996 to about 7% in 2001. Households that used wood for heating decreased from 26,7% in 1996 to 24,6% in 2001.

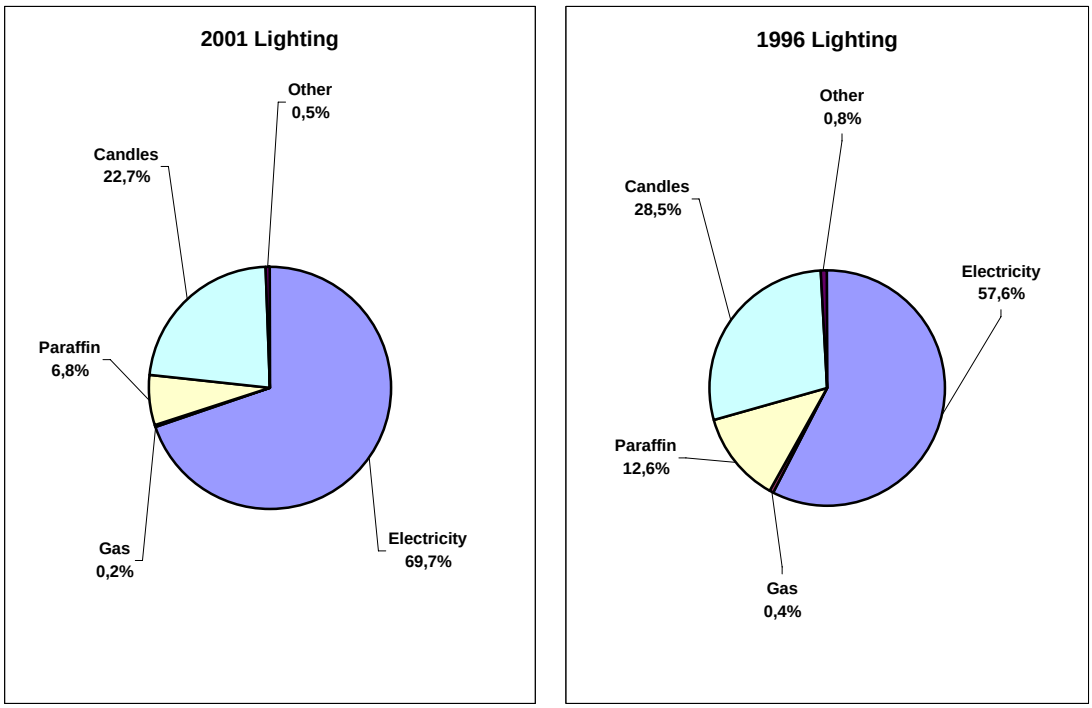
Figure 3: Distribution of households by main source of energy for heating, 2001 and 1996



12.6.4 Energy for lighting

Households using paraffin for lighting decreased from 12,6% in 1996 to 6,8% in 2001. Paraffin and gas usage for lighting did not change from 1996 to 2001. Households that used candles for lighting decreased substantially from 28,5% in 1996 to 22,7% in 2001. All in all, electricity was mainly used for lighting in a majority of households in 2001.

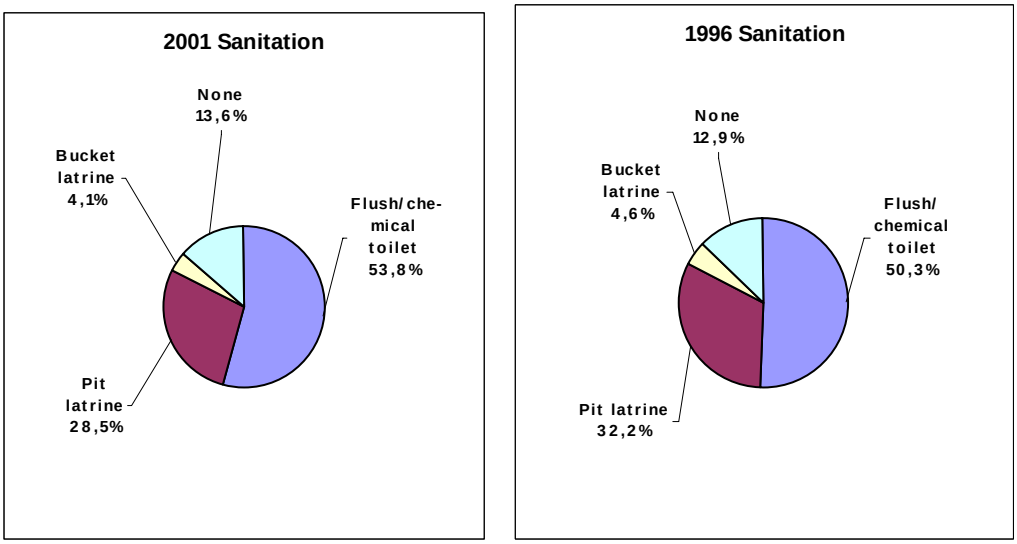
Figure 4: Distribution of households by main source of energy for heating, 2001 and 1996



12.6.5 Sanitation

The graphs below show that minimal success has been achieved as far as the improvement of sanitation is concerned from 1996 to 2001. However, The percentage of households that used flush or chemical toilets slightly increased from 50,3% in 1996 to 53,8% in 2001. The percentage of households that used pit latrines decreased from 32,2% in 1996 to 28,5% in 2001. However, the percentage of households that had no toilet facilities increased slightly from 12,9% in 1996 to 13,6% in 2001. This may be due to the increase in new households in informal settlements. The percentage of households that used bucket latrines declined slightly from 4,6% in 1996 to 4,1% in 2001.

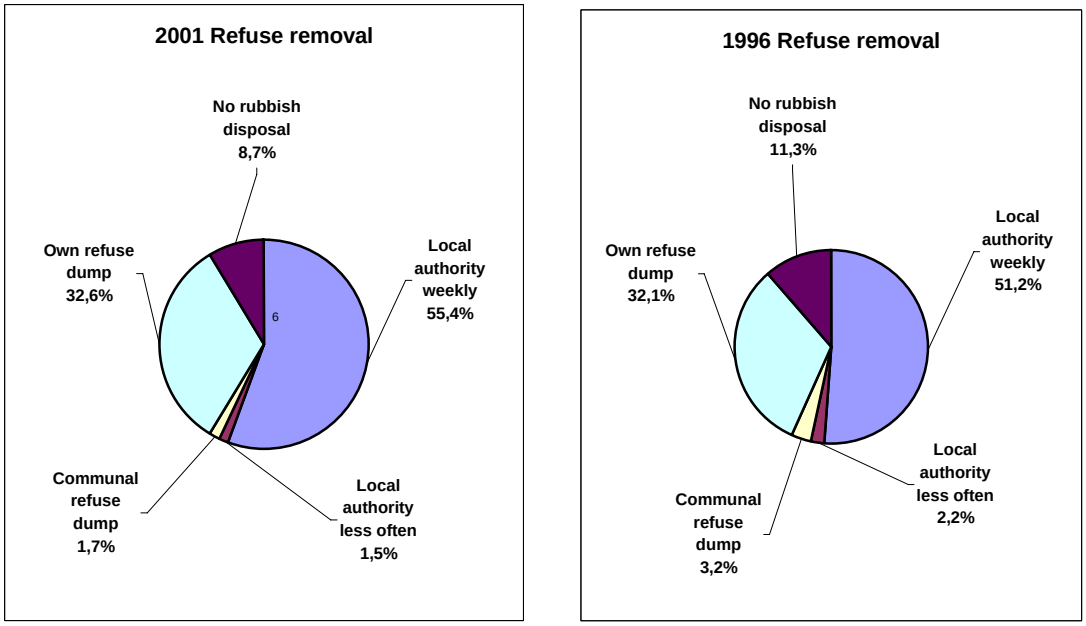
Figure 5: Distribution of households by type of sanitation used, 2001 and 1996



12.6.6 Refuse removal

Refuse removal poses some challenges, especially in rural or semi-rural areas. However, from the graphs below it is evident that there were substantial changes in the type of refuse removal from 1996 to 2001. The percentage of households that reported that their refuse was removed on a weekly basis by local authorities increased from 51,2% in 1996 to 55,4% in 2001.

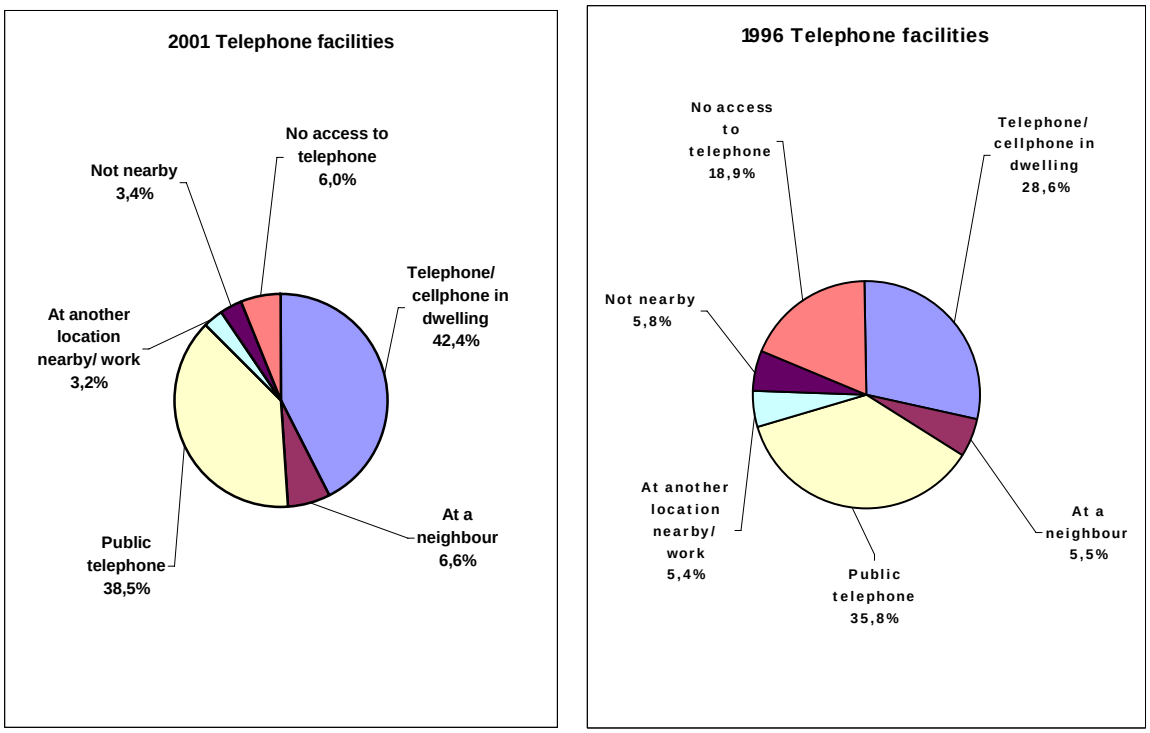
Figure 6: Distribution of households by type of refuse removal, 2001 and 1996



12.6.7 Telephone facilities

The graphs below provide a profile of the changes that have taken place with regard to availability of communication facilities between 1996 and 2001. The percentage of households that had no telephone facilities decreased significantly from 18,9% in 1996 to 6% in 2001. The percentage of households that had a cellular phone or a telephone in the dwelling increased from 28,6% in 1996 to 42,4% in 2001. The percentage of households that used public telephones increased slightly from 35,8% in 1996 to 38,5% in 2001.

Figure 7: Distribution of households' access to a telephone, 2001 and 1996



12.7 Conclusion

In the five years between 1996 and 2001, there was an increase in the proportion of households that had access to piped water, sanitation, electricity, refuse removal and communication facilities.

More than half of the households enumerated in the 2001 census used electricity for cooking. The proportion of households that used gas and paraffin for cooking remained more or less constant. There were fewer households that used wood, coal and animal dung for cooking in 2001 than in 1996.

Electricity was the most common energy source used for heating from 1996 to 2001. The proportion of households that used paraffin and gas for heating also remained

constant whereas the proportion that used wood, coal and other sources of energy decreased from 1996 to 2001.

Most households relied mainly on electricity for lighting, about 70% in 2001. The proportion of households that used paraffin and candles decreased from 1996 to 2001. The use of gas and solar energy was insignificant.

There was a slight increase in the proportion of households that used flush or chemical toilets from 1996 to 2001. The proportion of households that used pit latrines decreased whilst those that had no toilet facilities increased slightly from 1996 to 2001.

There was a marked increase in the proportion of households whose refuse was removed by local authorities weekly in 2001. The proportion of households that had their own refuse dump remained constant. The proportion of households that had no refuse removal decreased from 1996 to 2001.

There was a significant increase in the proportion of households that had a telephone or cellular phone inside their dwelling from 1996 to 2001.The proportion of households that accessed public telephones also increased from 1996 to 2001.

Bibliography

Statistics South Africa. (1998). *1996 Census in brief.* Pretoria: Statistics South Africa. http://www.statssa.gov.za

Statistics South Africa. (2003). *2001 Census in brief*. Pretoria: Statistics South Africa. http://www.statssa.gov.za

The Presidency, Policy Coordination and Advisory Services. (2003). *Towards a ten-year review: Synthesis report on implementation of government programmes.* Discussion document. Pretoria: Government Printer.

ANNEXURE 1: Community Profiles Training Document

Background

The SuperTABLE software package is an Australian product that was first used by Stats SA for the dissemination of Census 1996 data. Access to the software was limited to those individuals and organisations that could afford to pay the licence fees. This time around, no licence fee is required to be paid. The SuperTABLE software is available free of charge together with the Census 2001 *Community Profiles* databases.

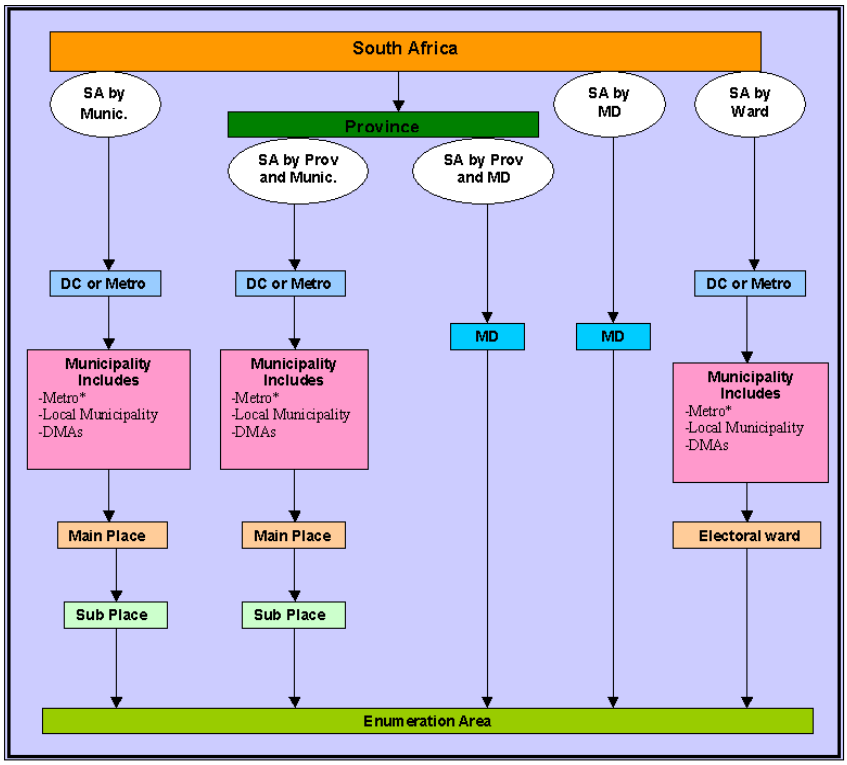
The *Community Profiles* databases package comprises a set of 12 CDs in a black CD wallet. All government departments, major public entities, constitutional institutions, research institutions and educational institutions will receive the first multi-user copy of the *Community Profiles* package free of charge. Private businesses, non-government organisations, international agencies and private individuals will be required to pay the cost of printing and packaging. Installation and initial training will be provided free of charge. (Please refer to the Pricing Policy and Tariffs on the Statistics SA website, http://www.statssa.gov.za, for further details.)

Census 1996 was the first population census conducted after the 1994 elections, and Census 2001 was the second. Comparisons between the data from the two censuses must be made with caution. First, the geographic hierarchies for the two censuses differ. For example, in 1996 reference was made to place names. There were 12 852 place names in the 1996 *Community Profiles* databases. In 2001 these place names were regrouped into 3 109 main places (towns and tribal authorities) and 21 243 sub-places (suburbs and villages). Place names may therefore differ between the two censuses and should be verified before comparisons are made. (For more information on the geographic hierarchies, please refer to the metadata contained in the *Community Profiles* databases or to the geography metadata on Statistics SA's website).

Information on geography

Geography hierarchy

The flow diagram below explains the hierarchical structure used for geographic areas in Census 2001.



* Will be expanded to metropolitan substructures when available

The structure consists of seven levels, as follows:

Level 1 – South Africa
Level 2 – Province
Level 3 – District council (Category C) or metropolitan area (Category A)
Level 4 – Local municipality (Category B) or district management area (DMA)
Level 5 – Main place
Level 6 – Sub-place
Level 7 – Enumeration area (EA)

While the structure is intended to be hierarchical, South Africa's geography has cross-boundary entities at all seven levels, which complicates the picture. For example, there are eight municipalities that lie across provincial boundary lines. These cross-boundary areas require special codes to indicate to which region each portion belongs.

ANNEXURE 2: BASIC FUNCTIONALITY OF SuperTABLE

1. Introduction

1.1 Background

This document outlines the basic functionality of SuperTABLE with respect to the Census 2001 *Community Profiles* databases. SuperTABLE is a user-friendly cross-tabulation software that enables a user to design tables, charts and basic maps from Census 2001 data.

By completing all the exercises in this document, you will learn the basic procedures involved in accessing data from Census 2001. The activity segment of this document has been divided into three main sections. The first section covers the steps involved in creating tables, the second outlines the steps in creating charts and the third covers the steps involved in designing maps. Start with Section 2 (creating a table) and work through the exercises in sequential order. Each step in every exercise has been numbered. Section 5 contains an additional exercise that will test what you have learnt.

1.2 Objectives

By the end of the training the trainee should be able to do the following:

- work with Statistics SA's Census 2001 databases and with census data in general
- extract Census 2001 data to create simple tables and cross-tabulations within the SuperTABLE environment
- group variables together and save these groups as recode files
- export census tables in Excel spreadsheet format
- express the census figures in the form of charts within the SuperTABLE environment and export these charts to other Windows-based software
- create simple maps from census data and export these maps to other software packages (e.g. Word or PowerPoint).

2. Creating a table

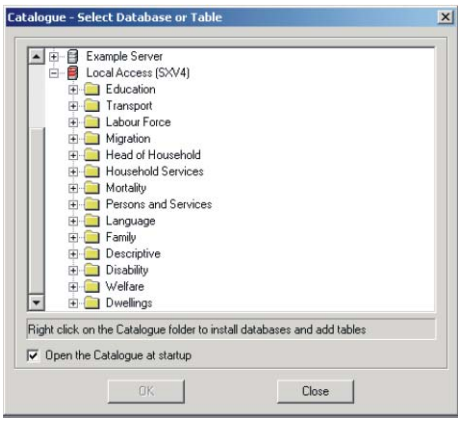
This section outlines the steps that are required to create a table. It basically involves a drop-and-drag process whereby variables are added and moved to various locations on a table. At the end of this section you will know how to perform the following tabular functions: adding and removing fields, adding total and percentage columns, grouping various values and creating recode files. The resulting table can also be saved in SuperCROSS format or exported as an Excel spreadsheet.

2.1 Opening SuperTABLE and selecting databases

1) Double click on the SuperTABLE icon on the desktop.

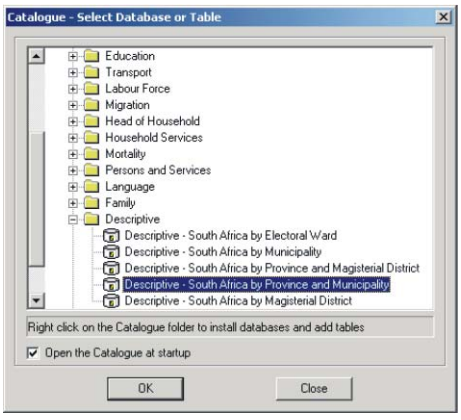
SuperTABLE will open to reveal what is called the Catalogue dialogue box.

2) Open the Local Access (SXV4) folder. This will open folders for each of the 14 census themes. Depending on the speed of your computer, this might take a couple of moments.



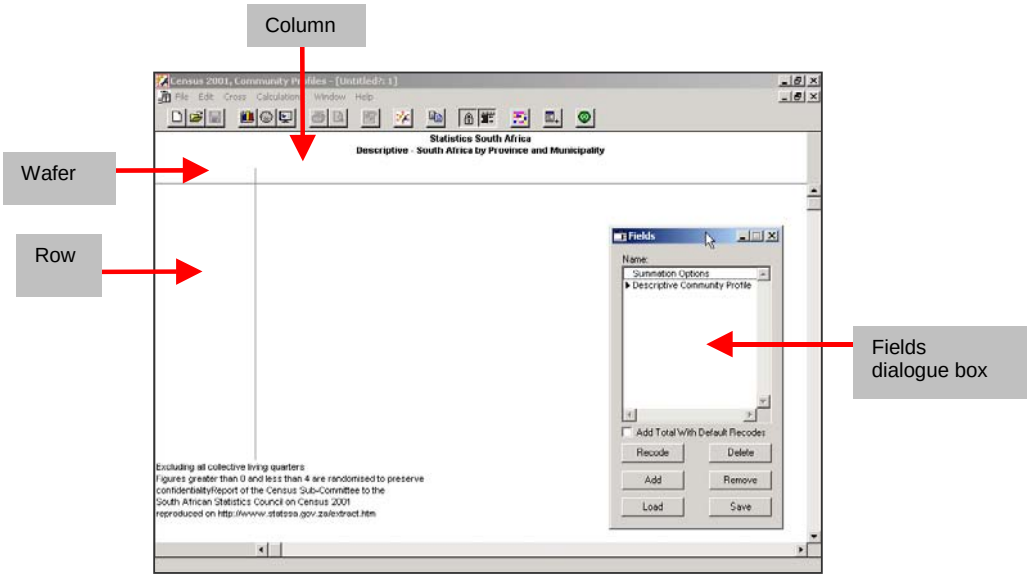
Each theme contains a set of variables that can be used to extract data for a specific topic. For example, if you are interested in obtaining employment data, select the Labour Force folder. If, however, you are looking for basic population demographics (such as gender, population group or age), select the Descriptive folder. Each folder contains five databases, each one representing a separate geographic hierarchy.

3) Open the Descriptive folder.
4) Within the Descriptive folder, double click on the database labelled Descriptive – South Africa by Province and Municipality.



2.2 *Adding fields*

Once the database is selected, the Catalogue dialogue box will be replaced by a blank table and the Fields dialogue box (see below). The Fields dialogue box contains the fields that can be added to the table. The blank table contains three areas in which the values can be placed, namely the Column area, the Wafer area and the Row area.



Designing a table involves dragging the names of the selected fields from the Fields dialogue box to any of these three areas. In this exercise, a table will be created showing the number of individuals by population group for all district councils in the Eastern Cape. The end result will be the following table:

Statistics South Africa
Descriptive - South Africa by Province and Municipality
Geography by Population group
for Person weighted

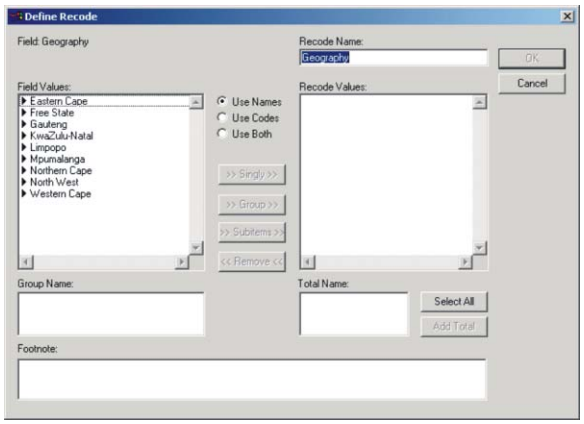
	Black African	Coloured	Indian or Asian	White
DC10: Cacadu District Municipality	202,541	140,851	730	44,082
DC12: Amatole	1,539,223	52,289	4,706	68,039
DC13: Chris Hani District Municipality	762,369	31,489	507	15,935
DC14: Ukhahlamba District Municipality	321,261	11,611	93	8,360
DC15: O.R.Tambo	1,668,007	5,448	880	2,148
DC44: Alfred Nzo District Municipality	549,271	881	128	126
Port Elizabeth: Nelson Mandela	592,568	236,058	11,152	165,996

Begin by adding the fields of Geography (district councils) and Population Group to create the above table.

1) In the Fields dialogue box, double click on the small black arrow next to the words Descriptive Community Profile.

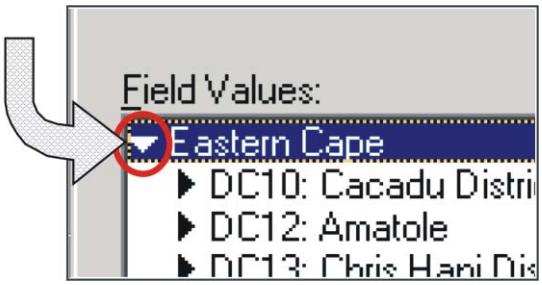
By double clicking on the black arrow, all the fields in the Descriptive database will be listed, starting with Age and ending with Present School Attendance.

2) Double click on the field Geography. The Define Recode dialogue box will appear.



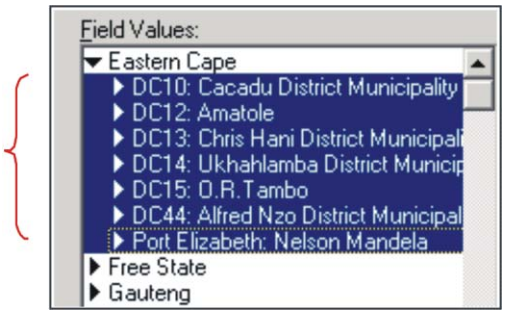
Use this dialogue box to extract figures for provinces, municipalities, towns and suburbs. In this example, the district councils in the Eastern Cape will be selected.

3) Double click on the small black arrow next to Eastern Cape. In SuperTABLE, these black arrows indicate folders. By double clicking on the arrow next to Eastern Cape, you will bring up all the district councils within this province.

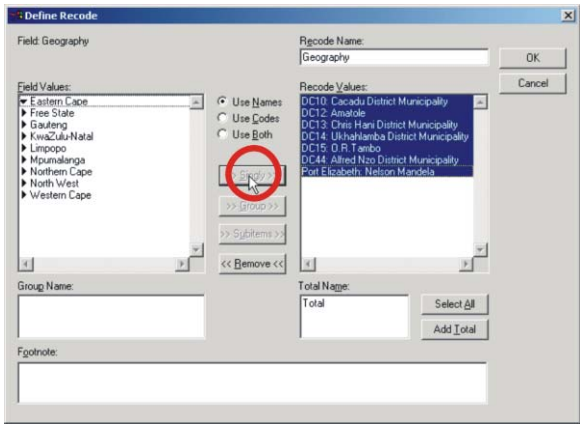


4) Select all the district councils in the Eastern Cape.

The selection is made by clicking on the first district council (DC10: Cacadu District Municipality) and holding the left mouse button down while dragging the cursor to the last district council (Port Elizabeth: Nelson Mandela). Alternatively, you can use the Shift key on the keyboard to select all the district councils.



5) Once the district councils have been selected, click on the option button labelled Singly to move the selected councils from the Fields Values list box to the Recode Values list box.

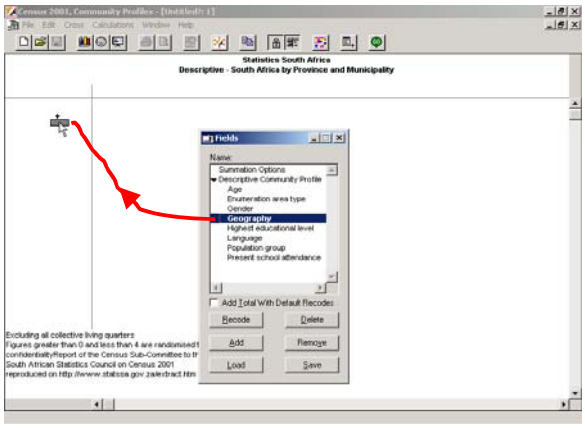


The required values (i.e. all the district councils in the Eastern Cape) have been selected.

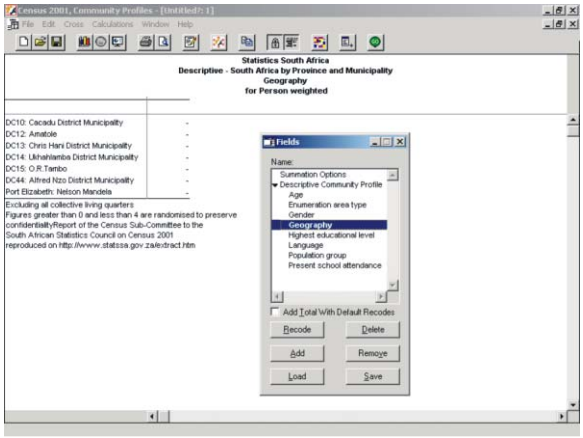
6) Click OK on the Define Recode dialogue box.

The Define Recode dialogue box will disappear. You will notice that Geography in the Fields dialogue box is now in bold font. This means that a selection has been made and that the selected values are ready to be placed into the table.

7) To place the district councils into your table, click on Geography and, while holding down the mouse button, drag it to the Row area of the empty table.



The district councils that were selected should now appear as rows on the table (see below).



Now the field Population Group will be added as columns to complete the table.

8) Double click on Population Group in the Fields dialogue box.

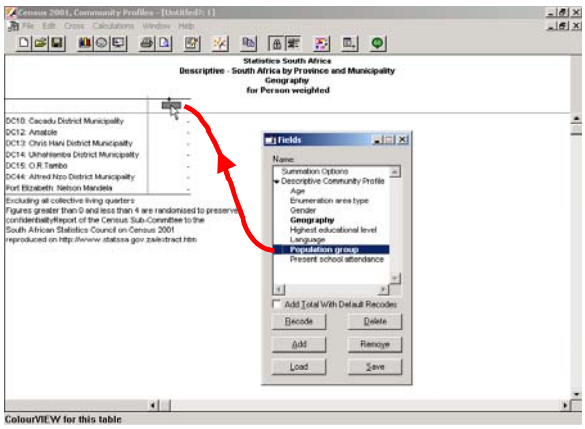
The Define Recode dialogue box for Population Group will appear. All population groups (i.e. Black African, Coloured, Indian or Asian and White) will appear in the Field Values list box.

9) Select all population groups in the same way that the district councils were selected.

10) Click on the Singly option button to transfer the selected values to the Recode Values list box.

11) Click OK on the Define Recode dialogue box. This box will disappear. Population Group in the Fields dialogue box is now in bold font.

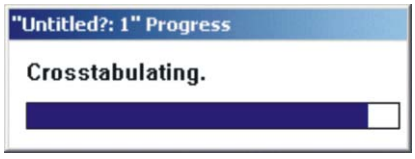
12) Click on Population Group in the Fields dialogue box and drag it to the Column area of the table (see below).



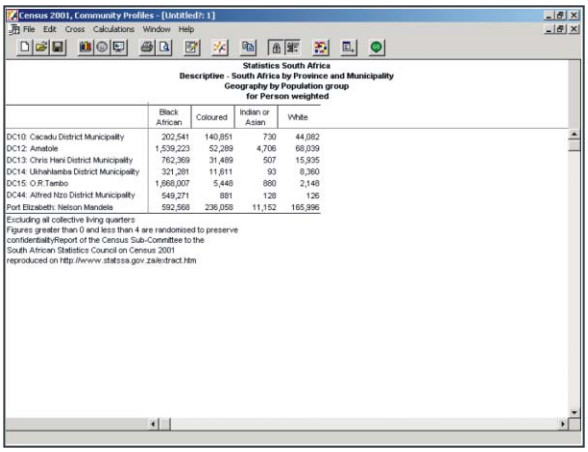
A complete table should now appear on the screen with the district councils as rows and the population groups as columns. However, as yet there are no figures in the table.

13) To populate the table with figures, press the green Go button on the top icon bar.

A progress bar will appear, showing the progress of the cross-tabulation.



After a few moments the progress bar will disappear and the table will be populated with figures (see below).



Well done! You have just created your first table in SuperTABLE. This table is a basic population table showing the number of individuals in each Eastern Cape district council by population group. The next part of the exercise shows how to save a table once it has been created.

2.3 Saving a table in SuperCROSS format

There are a number of ways of saving tables created in SuperTABLE. If you think that you may need to manipulate the table within SuperTABLE at a later stage, for example to add additional fields, you will need to save the table in SuperCROSS format.

1) Click on File > Save As. The Save As dialogue box will appear.

2) For the Save In option, select the Desktop.

3) Enter Population_E Cape_DC for the File Name option.

4) You will notice that the SuperCROSS (SCS) format has been chosen as the default Save as Type. Do not change this format.

5) Click Save

Your table has now been saved as a SuperCROSS file. You can open it at a later stage if you need to add or remove fields.

2.4 Saving a table in Excel format

You can also export your table in various formats, including Excel spreadsheet format. This makes it easier to send your tables to colleagues who do not have SuperTABLE.

1) Click on File > Save As. The Save As dialogue box will appear.

2) For the Save In option, select the Desktop.

3) Enter Population_E Cape_DC for the File Name option.

4) For the Save as Type, select Excel (XLS).

5) Click Save

2.5 Accessing metadata

All databases in the *Community Profiles* contain metadata files. Metadata files enable you to find information about the fields themselves. For example, the metadata file on Gender contains notes on how the gender question was phrased in the questionnaire, how the data were collected and how the results were coded. In the following steps, the metadata file for Gender will be opened.

6) In the Fields dialogue box, right click on Gender. The following option box will appear.

7) Click on Field Meta-Information. The metadata document for Gender, in the form of a Word document, will appear.

8) Close the Word document.

2.6 Adding additional fields

In the previous exercise, the two fields of Geography (district councils) and Population Group were cross-tabulated. In this exercise, Gender will be added as a third field.

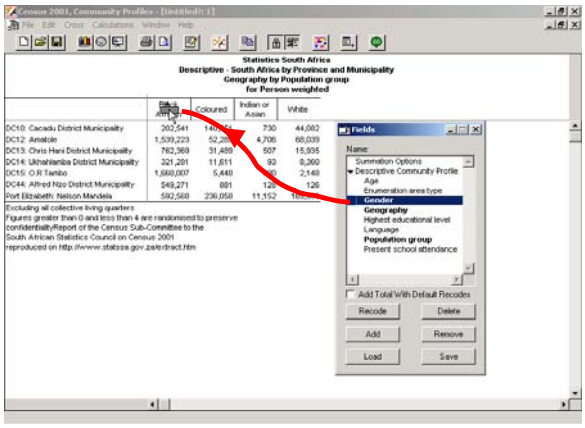
1) Double click on Gender in the Fields dialogue box. The Define Recode dialogue box for Gender will appear.

2) In the Field Values list box, select both Male and Female.

3) Click Singly to transfer Male and Female to the Recode Values list box.

4) Click OK. The Define Recode box will disappear, and Gender will become bold in the Fields dialogue box.

5) Click and drag Gender from the Fields dialogue box and drop it over the title of the Black African column in the table (see below).



The field of Gender will now appear in the table. Each population group should now be divided into both gender categories.

6) Press the Go button.

	Black African		Coloured		Indian or Asian		White	
	Male	Female	Male	Female	Male	Female	Male	Female
DC10: Cacadu District Municipality	96,810	105,731	67,645	73,207	364	365	21,087	22,995
DC12: Amatole	710,940	828,283	25,222	27,067	2,468	2,238	32,946	35,093
DC13: Chris Hani District Municipality	351,210	411,160	15,169	16,319	269	238	7,799	8,136
DC14: Ukhahlamba District Municipality	149,305	171,977	5,692	5,918	57	36	4,008	4,352
DC15: O.R.Tambo	753,672	914,334	2,706	2,742	467	412	1,157	990
DC44: Alfred Nzo District Municipality	246,297	302,974	483	398	80	48	78	48
Port Elizabeth: Nelson Mandela	281,208	311,360	113,004	123,054	5,525	5,627	80,166	85,830

You can add and cross-tabulate as many fields from the Fields dialogue box as you like. This ability is the greatest benefit of SuperTABLE as it enables you to design your own census tables with any fields that you require.

2.7 Removing fields from a table

Selected fields can also be removed from a table. In the following example, Gender will be removed from the table (using the same click-and-drag method).

1) Click on any of the Male / Female headings in the table and drag back to the Fields dialogue box. All gender headings should disappear.

   The Gender field has now been removed and the table should now only contain the population groups as columns and district councils as rows (as in the original example).

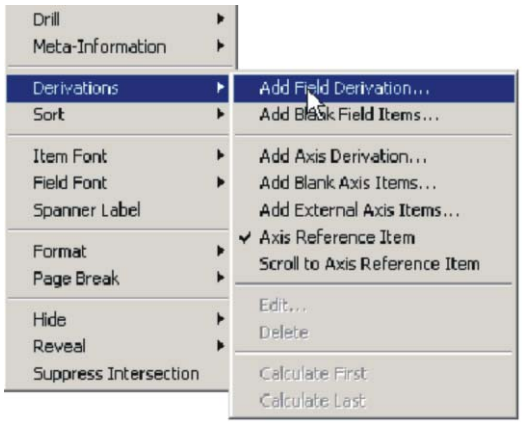
2) Press the Go button to populate the table with figures.

2.8 Adding a Total column

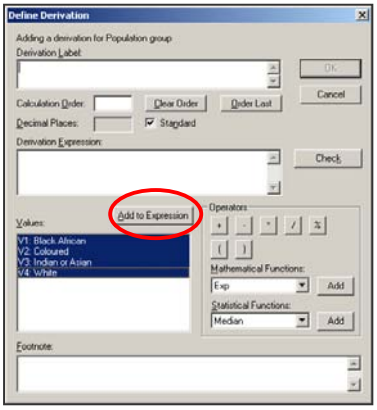
The table looks great but lacks something quite important. In all tables a Total column or Total row is important to provide a summary of what the table is portraying. Adding totals in a SuperTABLE table is quite easy and it makes use of the Define Derivation dialogue box.

1) Right click on the column heading White (this will create a Total column to the right of the White column). A menu will appear.

2) Select Derivations > Add Field Derivation.



The Define Derivation dialogue box will appear. It is within this box that the Total column will be created.



3) All the population groups will be highlighted. Click on the Add to Expression button.

The following expression will appear in the Derivation Expression box: sum(V1:V4). You will also notice that Total appears in the Derivation Label list box – this will be the heading of the Total column.

Click OK. The Define Derivation dialogue box will disappear, and the Total column will be added to the right of the White column.

2.9 Adding a Total row

A Total row can also be added at the bottom of the table to show the total number of individuals per population group for the entire Eastern Cape (i.e. for all district councils).

1) Right click on the bottom row (i.e. on Port Elizabeth: Nelson Mandela) and select Derivations > Add Field Derivations. The Define Derivation dialogue window should appear with all district councils listed in the Values list box.

2) Click the Add to Expression button.

The following expression will appear in the Derivation Expression list box: sum(V1:V7). Total will also appear in the Derivation Label list box.

3) Click OK. The Define Derivation dialogue box will disappear and the Total row will be added under Port Elizabeth: Nelson Mandela.

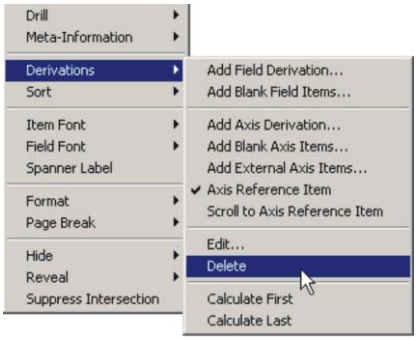
DC44: Alfred Nzo District Municipality	549,271	881	128	126	550,406
Port Elizabeth: Nelson Mandela	592,568	236,058	11,152	165,996	1,005,774
Total	5,635,261	478,627	18,194	304,686	6,436,769

2.10 Removing a Total column/row

The Total row under Port Elizabeth: Nelson Mandela will be removed.

1) Right click on the word Total (underneath Port Elizabeth: Nelson Mandela).

2) Select Derivations > Delete.



The Total row will disappear.

2.11 Adding a percentage column

The Define Derivations dialogue box can also be used to add percentages. A percentage column showing the percentage of Black African individuals for all district councils will be added to this table.

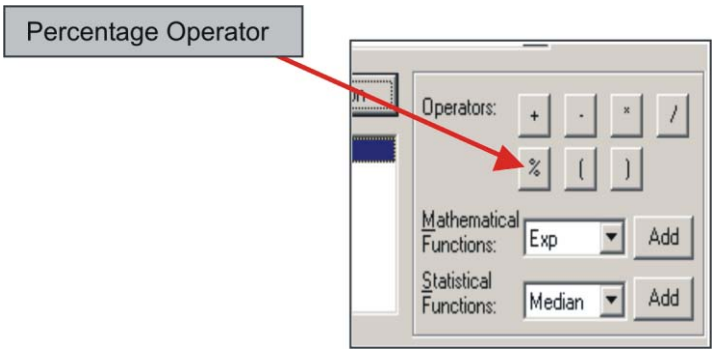
1) Right click on the title of the Black African column. Select Derivations > Add Field Derivations. The Derivations dialogue box should appear.

2) Enter % Black African in the Derivation Label list box. This will be the heading of the percentage column.

3) In the Values list box, click on Black African *only* to highlight this value on its own.

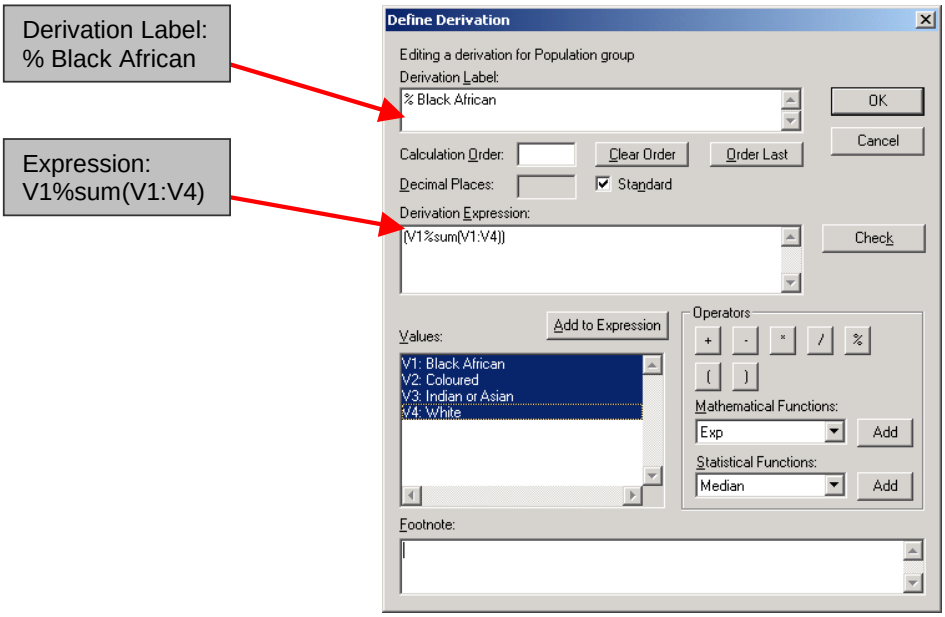
4) Click Add to Expression. V1 will appear in the Derivation Expression list box.

5) Click on the % operator. The percentage sign will now appear after V1 in the Derivation Expression list box.



6) Highlight all the population groups in the Values list box and then click Add to Expression. The expression V1%sum(V1:V4) should now appear in the Derivation Expression list box.

The Define Derivation dialogue box should now look like this:



7) Click OK. The Define Derivation dialogue box should disappear and the percentage column (with the heading of % Black African) should now appear in the table next to the Black African column.

	Black African	% Black African	Coloured
DC10: Cacadu District Municipality	202,541	52	140,851
DC12: Amatole	1,539,223	92	52,289
DC13: Chris Hani District Municipality	762,369	94	31,489
DC14: Ukhahlamba District Municipality	321,281	94	11,611
DC15: O.R.Tambo	1,668,007	99	5,448
DC44: Alfred Nzo District Municipality	549,271	100	881
Port Elizabeth: Nelson Mandela	592,568	59	236,058

8) Save the table by clicking on File > Save (the file should still have the file name Population_E Cape_DC).

2.12 Grouping values

In this exercise the grouping function of SuperTABLE will be explored. Different values can be grouped together for particular purposes. This is useful for variables with many values, such as age. A new table will be designed, showing the number of households in selected main places that fall within each income category. These 12 income categories will then be grouped into three broader categories, thus creating a simpler table.

1) Close SuperTABLE, then reopen it by double clicking on the SuperTABLE icon on the desktop.

2) In the Catalogue dialogue box, open the folder Local Access (SXV4). This will open folders for each of the 14 themes.

3) Open the folder Household Services.

4) Open the database Household Services – South Africa by Province and Municipality.

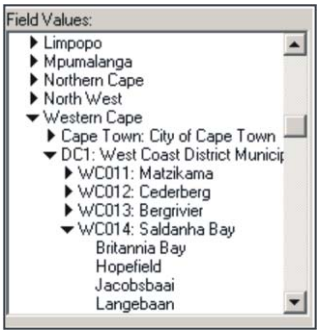
5) In the Fields dialogue box, double click on the small arrow next to Household Services Community Profile to view all the fields.

6) Double click on Geography. The Define Recode dialogue box for Geography will appear.

7) Under the Field Values list, double click on the small black arrow next to Western Cape to open the district councils in this province.

8) Double click on the small black arrow next to DC1: West Coast District Municipality to open the list of local municipalities.

9) Double click on the small black arrow next to WC014: Saldanha Bay to open the list of main places.



10) Select the following main places and click on Singly to bring each of them across to the Recode Values box: Hopefield, Langebaan, Saldanha Bay, St Helena Bay and Vredenburg.

11) Click on OK to close the Define Recode dialogue box. Note that Geography is now in bold font.

12) Click and drag Geography from the Fields dialogue box to the column area of the blank table. Your selected main places should now appear in the table (as columns).

13) Double click on Annual Household Income in the Fields dialogue box. The Define Recode dialogue box for Annual Household Income will appear.

14) Select all the income categories and click on Singly to bring them across to the Recode Values list box.

15) Click on OK to close the Define Recode dialogue box. You will notice that Annual Household Income in the Fields dialogue box is highlighted.

16) Select Annual Household Income and drag it to the Row area of the table.

17) Click the Go button. You should now have a table showing the number of households in all the main places according to each income group. For example, there are 37 households in Langebaan that have no income at all.

    This table is useful as you can see how many households fall within each of Statistics SA's defined income categories. However, this table might contain too much detail for certain purposes. In SuperTABLE, you can group values together into custom groupings. In the remaining part of this exercise, these income categories will be grouped into the three categories of Poor, Middle and Upper.

18) Remove the Annual Household Income field by clicking on any of the income categories and dragging it back to the Field dialogue box. All the income categories should disappear, leaving the main places on their own.

19) Double click on Annual Household Income in the Fields dialogue box to bring up the Define Recode dialogue box. You will notice that all the income categories are in the Recode Values list box.

20) Click on the Select All button to select all the categories and then click Remove. This will move the categories back to the Field Values list box.

21) Highlight the income categories from No Income to R9 601–R19 200 (the first four categories) in the Field Values list box.

22) In the Group Name list box, type Poor as the name for this group. You will notice that the Group button is highlighted.

23) Click on the Group button to move the first four income categories (as a single group) to the Recode Values list box.

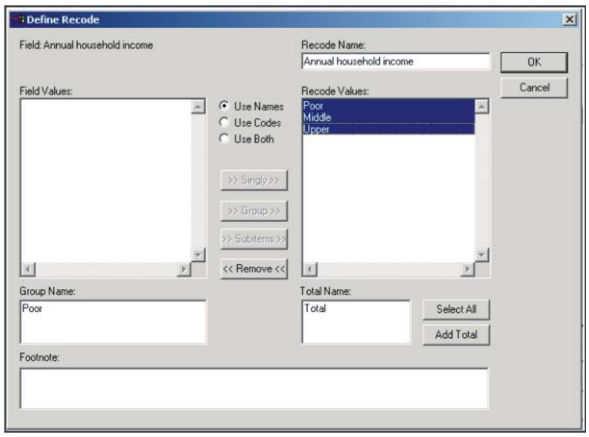
24) To add the second group, highlight the income categories from R19 201–R38 400 to R76 801–R153 600 in the Field Values list box.

25) In the Group Name list box, type Middle as the name for this group.

26) Click on the Group button to move the highlighted income categories (as a single group) to the Recode Values list box. It will be added as a second group in the Recode Values list box.

27) Similarly, group the income categories R153 601–R307 200 to R2 457 601 and More into a third group and name it Upper.

You should now have the following groupings in your Recode Values list box:



28) Click OK.

29) Click and drag the Annual Household Income category from the Fields dialogue box to the Row section of the table. The grouped categories should now appear in the table.

30) Click Go to populate the table with figures.

Your table should look like this:

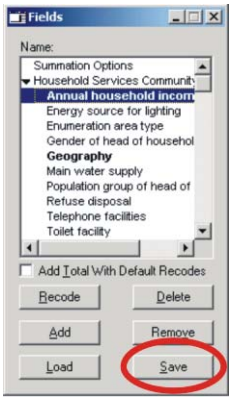
	Hopefield	Langebaan	Saldanha Bay	St Helena Bay	Vredenburg
Poor	122	201	433	706	2,971
Middle	298	686	554	1,248	3,843
Upper	21	215	72	61	596

31) Save the table on the Desktop in SuperCROSS format and give it the file name Income_Groups_W Cape

2.13 Creating recode files

In the previous exercise, the grouping function of SuperTABLE was explored. All income categories were grouped into the three groups of Poor, Middle and Upper. If you want to use these custom-made income groups for other tables, it will be wise to save the groupings as a recode file. Once saved, a recode file can be used to automatically group categories into custom-made groups whenever a new table is designed. This exercise will explore the steps involved in creating recode files.

1) Open the SuperTABLE file Income_Groups_W Cape if it is not already open. Click on the Save button on the Fields dialogue box (this save option is used to save groupings as recode files). A Save As dialogue box will appear as indicated below.



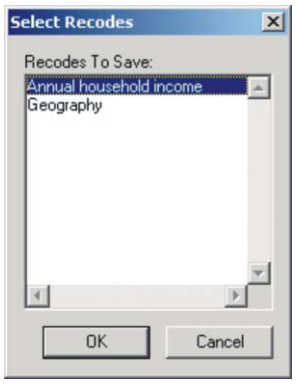
2) For the Save In option, select the Desktop.

3) Type the file name Income_W Cape_custom in the File Name list box and leave the Save as Type option as Recode (RCD).

4) Click Save. The Save As dialogue box will be replaced by a Select Recodes dialogue box.

   More than one grouping can be saved in a single recode file. The Select Recodes dialogue box allows you to choose which groupings to save. All groupings are highlighted by default. However, in this exercise, we are only interested in saving the income groupings for this recode file.

5) Select Annual Household Income only.



6) Click OK.

The custom-made income groupings have now been saved as a recode file. This recode file can now be loaded into any new SuperTABLE table where the custom income groupings are needed. The next exercise will explore the steps involved in loading recode files.

2.14 Loading recode files

In the last exercise, the recode file Income_W Cape_custom was created. This recode file, when loaded into any new SuperTABLE table, will automatically group all the single income categories into the custom-made groupings of Poor, Middle and Upper. This will avoid having to group individual income categories every time a new SuperTABLE table is created.

In this exercise, a new table will be designed showing income groupings for all the municipalities in South Africa. The Income_W Cape_custom recode will be used to automatically group the income categories into the custom-made groups.

1) If SuperTABLE is still open, close it.

2) Re-open SuperTABLE by double clicking on the SuperTABLE icon on the desktop.

3) In the Catalogue dialogue box, open the folder Local Access (SXV4).

4) Open the folder Household Services.

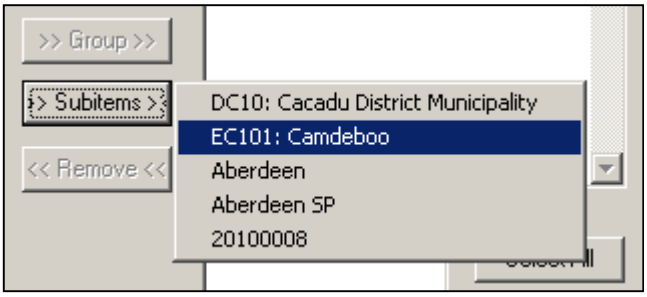
5) Open the database Household Services – South Africa by Province and Municipality.

6) In the Fields dialogue box, double click on the small arrow next to Household Services Community Profile to view all the fields.

7) Double click on the field Geography. The Define Recode dialogue box for Geography will appear.

8) In the Field Values list box, select all the provinces and click on the Sub Items button.

9) Select EC 101: Camdeboo by clicking on it (see below).



The Sub Items option allows you to select all lower values in a hierarchy. By clicking on EC 101: Camdeboo, all the municipalities from all provinces are selected and placed into the Recode Values list box.

10) Click on OK to close the Define Recode dialogue box. You will notice that Geography in the Fields dialogue box is in bold font.

11) Click and drag Geography from the Fields dialogue box to the Row area of the blank table. The municipalities should now appear in the table as rows.

The recode file that will automatically arrange all the income categories into the custom groupings will now be loaded into the table.

12) Open the Fields dialogue box, and click on the Load button. The Open dialogue box will appear.

13) Select the Income_W Cape_custom.rcd file and click Open.

14) The Select Recodes dialogue box will appear. Select Annual Household Income and click OK. Annual Household Income in the Fields dialogue box will now appear in bold font.

15) Select Annual Household Income and drag it to the Column section of the table. Your income groupings should now appear in the table as Poor, Middle and Upper.

16) Click Go to view the figures.

17) Save the table on the Desktop in SuperCROSS format and give it the file name Income_Groups_Municipality

Well done! You have learnt all the steps for creating tables within SuperTABLE. After a little practice you will soon be able to extract and manipulate any of the census variables from all the census databases.

3. Creating a chart

SuperTABLE can be used to design simple charts. These charts graphically display the figures presented in created tables. Once a chart has been designed, it can be saved as a Windows Bitmap (BMP) image file, which can then be imported into other Windows-based programmes (such as Microsoft Word or PowerPoint).

At the end of this section you will know how to perform the following charting functions: creating a chart, changing chart style, adding chart titles, adding chart legends and exporting charts in BMP image format.

3.1 Creating a simple chart

In this exercise, a simple chart will be created from the percentage column created in the SuperTABLE table Population_E Cape_DC.scs.

1) Open the SuperCROSS file Population_E Cape_DC if it is not already open. The percentage column % Black African should still be in the table.

2) Highlight the entire percentage column by clicking once on the heading % Black African.

Black African	% Black African	Coloured
202,541	52	140,851
1,539,223	92	52,289
762,369	94	31,489
321,281	94	11,611
1,668,007	99	5,448
549,271	100	881
592,568	59	236,058

3) Click on the Chart button on the tool bar.

A graph should appear, together with the Layer Control dialogue box. The Layer Control dialogue box can be used to change the properties of the chart.

4) Close the Fields dialogue box.

5) Close the Layer Control dialogue box.

6) Maximise the chart window

This is a basic chart in SuperTABLE. The following steps will show how to change the style of the chart (e.g. from bar chart to pie chart).

3.2 Changing chart style

SuperTABLE contains a variety of different chart styles that can be used to reflect data in different ways. In this exercise, the presently displayed chart will be changed to a pie chart.

1) Open the Layer Control dialogue box by clicking on View > Layer Control on the top menu.

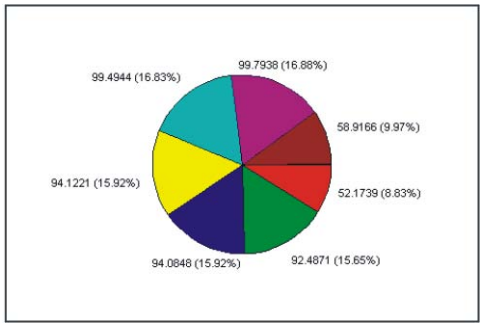
2) Click on the Style button on the Layer Control dialogue box. The Style dialogue box should appear.

3) Select the Chart – Pie option from the Style drop-down list.

4) Click OK.

5) Close the Layer Control dialogue box.

The chart should now appear as a pie chart (see below).



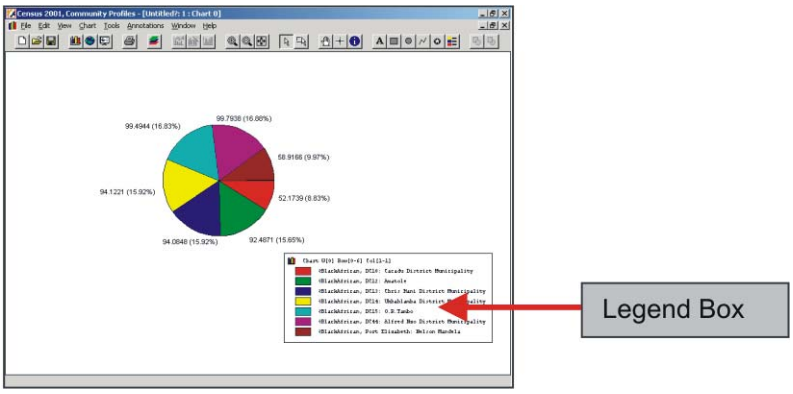
3.3 Adding a chart legend

The pie chart looks great, but it needs a legend to show what the different colours (or slices) of the chart represent.

1) Click on the Legend button 

2) Move your cursor across the screen. You will notice that it takes the form of a cross-hair.

3) Click and drag the cross-hair next to the pie chart to create a legend box. Use the box handles to size the legend so that all legend entries can be seen.

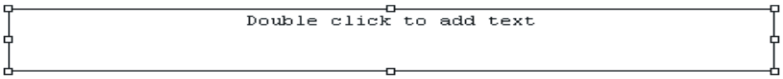


3.4 Adding a chart title

1) Click on the text box icon 

2) Move your cursor across the screen. You will notice that it takes the form of a cross-hair.

3) Click and drag the cross-hair above the pie chart to create a text box. It will contain the words Double Click to Add Text.



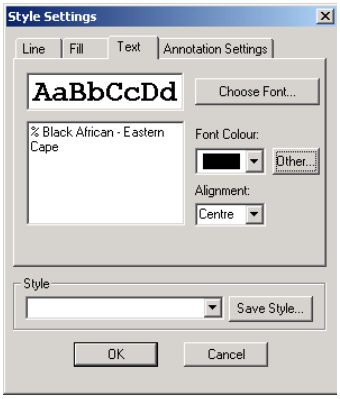
4) To enter text, double click on the text box to open the Style Settings dialogue box

5) Click on the Text tab on the Style Settings dialogue box.

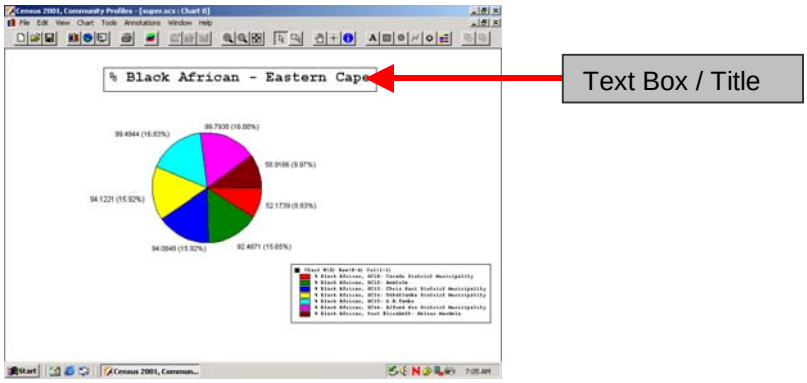
6) In the Style Settings dialogue box, delete Double Click to Add Text and enter % Black African – Eastern Cape in its place.

7) Click on the Choose Font button to open up the Font dialogue box and increase the font size to 20 and make it bold.

8) Click OK to close the Font dialogue box. The Style Settings dialogue box should now look like the following screenshot:



9) Click OK to close the Style Settings dialogue box. The title % Black African – Eastern Cape should appear in the text box.



3.5 Saving a chart in SuperTABLE format

A chart can be saved within the SuperTABLE file together with the table.

1) Click on File > Save

The chart will be saved as a separate window in the file Population_E Cape_DC.scs together with the associated table. By holding down the Ctrl key and pressing the Tab key at the same time, you can hop between the table and chart windows in Population_E Cape_DC.scs.

3.6 Exporting a chart in BMP format

A chart can also be exported in BMP image format, which can then be imported into Word documents and PowerPoint presentations.

1) Maximise the chart window and make sure that it is active.

2) Click on File > Export Chart. The Save As dialogue box will appear.

3) For the Save In option, select the Desktop.

4) Enter the name Population_Chart_E Cape in the File Name option. Notice that the File Type option is Windows Bitmap (BMP).

5) Click Save.

3.7 Importing a chart (BMP image) into Microsoft Word

If you have Microsoft Word or PowerPoint, you can import the BMP image of the chart into your documents. The following exercise will outline the steps for importing the BMP image into Microsoft Word; the same steps can be used for PowerPoint.

1) Close SuperTABLE.

2) Open Microsoft Word.

3) On the Microsoft Word menu, click on Insert > Picture > From File. The Insert Picture dialogue box should open.

4) Select the Bitmap image Population_Chart_E Cape and click Insert. The chart should now appear in your Word document.

5) Save the Word document on the Desktop with the file name of E Cape_Chart.

6) Close Microsoft Word.

These are the simple steps involved in creating a chart within the SuperTABLE environment. Although the charting function is not as comprehensive as in Excel, it does enable the user to represent data in basic graphic format.

The creation of tables and charts within SuperTABLE has now been covered. The next section will explore the steps involved in creating basic maps.

4. Creating a map

Not only can charts be created within the SuperTABLE environment, but simple maps can also be designed to represent the data (through another software package called ArcExplorer). Within the *Community Profiles* package, geographic boundaries have been included for provinces, district councils, municipalities, magisterial districts, wards, main places and sub-places.

This section of the manual explores the steps for the following mapping functions: creating maps, adding feature names and exporting maps in BMP image format.

4.1 Creating a simple map

In this exercise a simple map will be created from a percentage column created in the SuperTABLE table of Income_Groups_Municipality.scs

1) Close SuperTABLE and reopen it. Open the SuperCROSS file Income_Groups_Municipality.scs

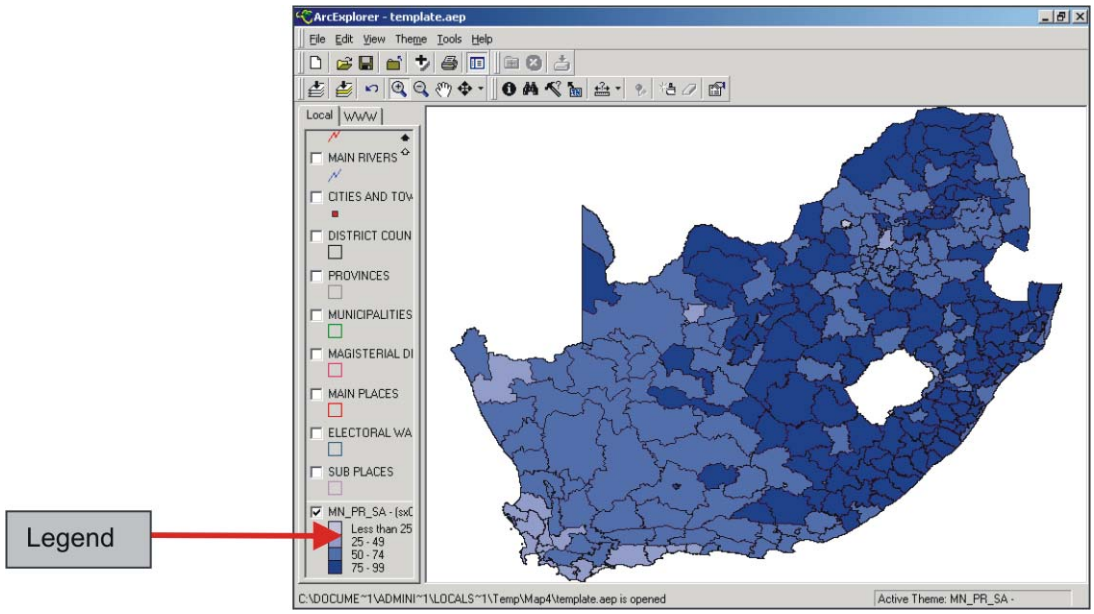
2) Create a percentage column for the Poor income category and name it % Poor (see page **Error! Bookmark not defined.** {check at final stage} for how to add a percentage column).

3) Highlight the Poor percentage column.

4) Click on the Map button.

A map of South Africa should appear in the software package called ArcExplorer.

Maximise the map



This is the basic structure of the map. Notice the legend on the left-hand side. The blue colour variations represent different percentages of households that fall within the Poor income group, and these are reflected for all municipalities on the map. The following exercises will show how to add feature names and how to export the map in BMP image format.

4.2 Adding feature names and zooming into the map

In this exercise, the names of the provinces will be added to the map.

In the legend box, check the box next to Provinces.

1) The province names will appear on the map for all the provinces.

2) In the legend box, check the following boxes: Main Roads, Cities and Towns and Municipalities.

3) Use the Zoom function tool  and Zoom into the Gauteng area.

4) Press the Zoom to Full Extent icon  to bring the entire map of South Africa back on screen.

4.3 Printing a map

A map can be printed with a scale bar, title and north arrow.

1) On the map window, click on File > Print. The Print Map dialogue box will appear.

2) For the Map Title option, type Percentage of Poor Households in Each Municipality – Census 2001

3) Click Print.

4.4 Exporting a map in BMP format

A map can also be exported in BMP image format, which can then be imported into Word documents and PowerPoint presentations.

1) Click on Edit > Copy to File (BMP). The Save Export File As dialogue box will appear.

2) For the Save In option, select the Desktop.

3) Enter the name Poor_Households in the File Name option. Notice that the File Type option is Windows Bitmap (BMP).

4) Click Save.

4.5 Importing a map (BMP image) into Microsoft Word

If you have Microsoft Word or PowerPoint, you can import the BMP image of the map into your documents. The following exercise will outline the steps for importing the BMP image into Microsoft Word. The same steps can be used for PowerPoint.

1) Close SuperTABLE.

2) Open Microsoft Word.

3) On the Microsoft Word menu, click on Insert > Picture > From File. The Insert Picture dialogue box should open.

4) Select the image Poor_Households and click Insert. The map should now appear in your Word document.

5) Save the Word document on the desktop with the file name of Poverty_Map.

6) Close Microsoft Word.

5. Evaluation / self-assessment

Now that you are familiar with SuperTABLE, you can try the next exercise on your own.

5.1 Exercise: education levels

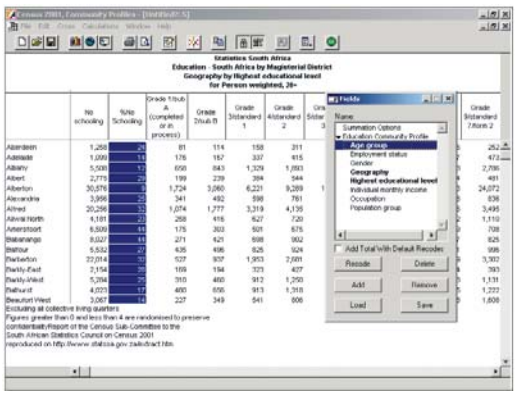
For this exercise, a table and map will be created, showing the percentage of those aged 20 and above with no schooling for all magisterial districts in South Africa.

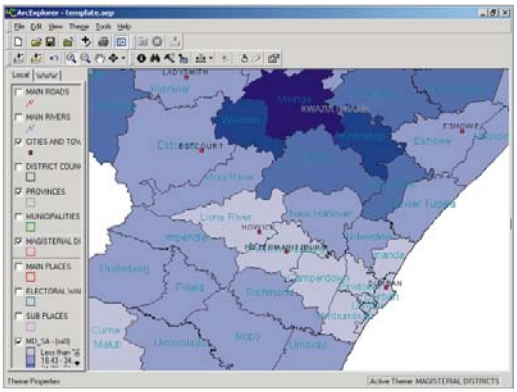
1) Create a table in the Education – South Africa by Magisterial District database with the Highest Education Level recode as columns and all the magisterial districts as rows. Add Age Group to the table as well (group all ages from 20 and above into a single group and add this to the table). Create a percentage column for those with no schooling, and map this column. On the map, add the province and magisterial district names, and then zoom into KwaZulu-Natal.

2) Save the table with the file name of Exercise 1 in Excel and SuperCROSS format.

3) Export the map of the Gauteng area into a PowerPoint slide.

4) Use the table to answer the following questions:
   (a) According to the map, does Durban in general have a high or low percentage of people aged 20 or above with no schooling?
   (b) What is the percentage of people aged 20 and above who have no schooling in Msinga magisterial district?





Answers: a. low; b. 68%

6. Conclusion

Thank you for working through this manual. If you have managed to complete all the exercises in this manual, you should now know enough to create simple tables, charts and maps from Census 2001 data. If you interested in learning more about SuperTABLE, you can take a look at the Acrobat Reader user guides, visit the website at http://www.statssa.gov.za or make contact using the details listed below.

- Statistics South Africa Marketing Department
  Telephone: (012) 310 2112 / 2107 /
  2111/ 2108 / 2109 / 2110 / 2105
  Fax: (012) 310 8966

- Manual compiled by Kevin Parry
  Tel: (012) 310 2111
  Fax: (012) 310 8966
  E-mail: kevinp@statssa.gov.za

This entire document is a summary of three very large Acrobat Reader user guides that are included in the SuperTABLE package. To access these guides, click on Help > User Guides. Alternatively, click on Start > Programmes > STR > SuperTABLE. You can then select between three different PDF documents: SuperTABLE User Guide, SuperCHART User Guide and SuperMAP User Guide. These user guides contain more extensive information on SuperTABLE functions that are not covered in this manual.

7. Appendix 2 – *PX-Web Databases Training Manual*

7.1 Introduction

This document outlines the basic functionality of PX-Web as a dissemination tool for economic, social and population databases.

General information is covered in the first part of the document, and in the second part a description is given on how to use PX-Web to get to the information that you require.

Why use PX-Web?

PX-Web and its software counterpart PC-Axis are currently in use in 23 different statistical offices across the globe. Although initially designed for dissemination of census data PX-Web has also proved equally valuable in the dissemination of economic and social statistics.

Statistics SA has been using this technology since October 2000, and it has proved its reliability, speed and overall cost effectiveness through the dissemination of time series data on the Statistics SA website.

The PX-Web system is a zero-footprint client. This means that you can use this software directly on our website without having to do any client installation.

The PC-Axis software suite (if required) has been available for download from the Statistics SA website since July 2003. This software allows the user to locally build reports and to do cross-tabulations of data with information downloaded via the PX-Web system on the Statistics SA website.

7.2 *Tip on how to mark in a list box*

At least one indication in every list box has to be made to show a result for each variable.

When you wish to see one or a few variables, mark these only. NB: To undo a mark, press the Ctrl button and click the mouse button on the text.

To select all lines in a list box

- Mark the first line in the list box, then press the Shift and End buttons simultaneously.

To select/undo a marked line in a list box

- Press the Ctrl button and click the left mouse button on the selected line simultaneously.

To select more than one line in a list box

- Mark lines in a row by pressing the left mouse button and slipping it over the requested lines. Or mark the first line on your list and then press the Shift button at the same time as you press the up or down arrows on the keyboard.
- Mark lines not in a row by pressing the Ctrl button at the same time as you click the left mouse button on the requested lines.

To search for a specific value in the list box

- Mark the first line in the list box, then type the first letter in the line requested. If you fail to get the right line, try typing again until you find the right value, or use the up or down arrows or the mouse button to look further.

7.3 *Explanation of the buttons used*

 Display the table in new layout

Save as PC-Axis file

 Save as Excel file

 Save as text file. You get to select text attributes.

 Save as delimited file. You get to select the delimiter.

 Save in any of the file formats, Html, dBase, Lotus

 Change order for the variables, clockwise

 Change order for the variables, counter clockwise

 Change order for the variables, in any order you select

 Remove lines with only zeros in the table cells

 Sum a variable

 Detailed information for tables. The button is only visible when such information is available.

 Save information about the selected table in a cookie for later use

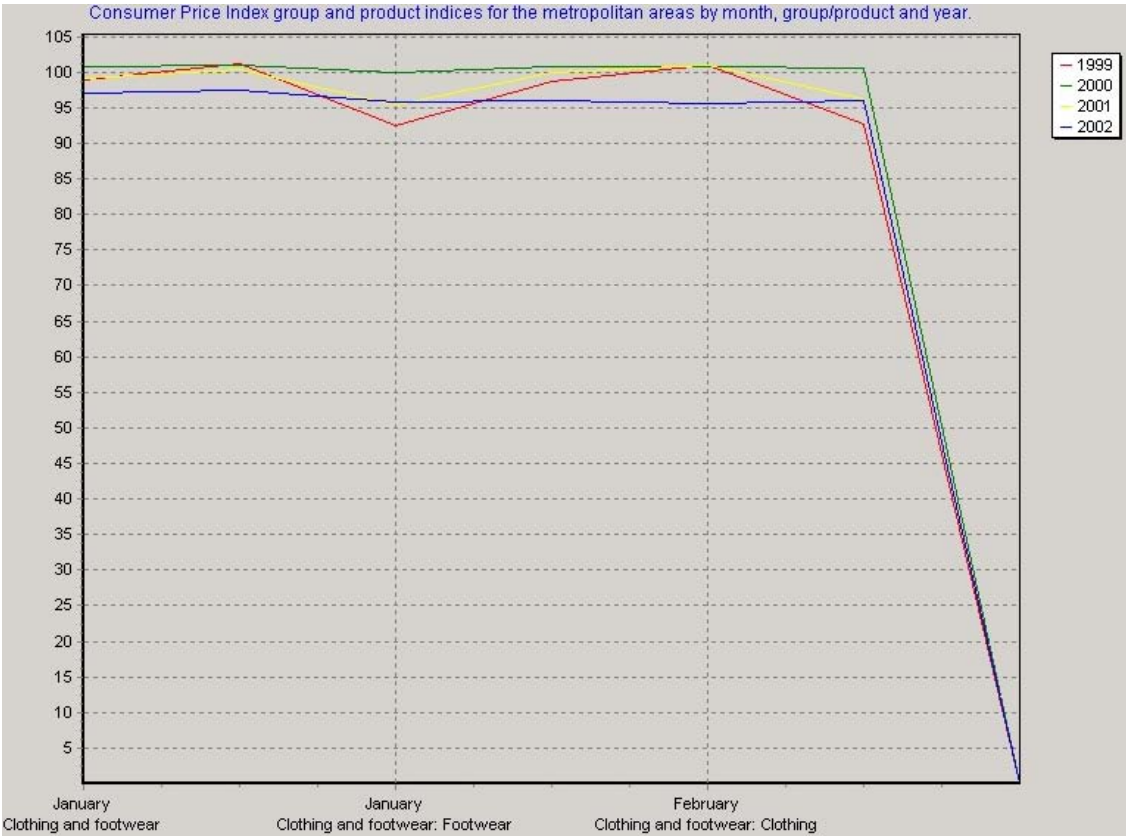
 A page explaining the buttons used.

7.4 Formatting a graph

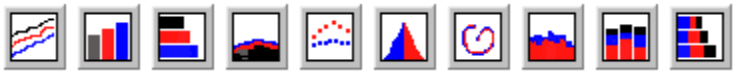
When a table is displayed on screen, you can generate a bitmap format graph from the information.

Click the graph button above the table to do this. The button looks like this:

A graph will be generated and displayed on screen. The default type of graph is a line graph as shown below.



By using the Graph toolbar, different types of graphs can be generated. This toolbar is displayed above the graph generated and looks like this:



It is important to note that although all options work, care should be taken when generating the Population Pyramid as this graph only works with very specific population variables.

7.5 Help functions for presentation formats and other functions

Presentation formats accessible from the selection page and the table page:

Presentation form	File type	Icon on table page	Software
Show table on screen	-	-	Presents selected material in a table on screen
Save PC-Axis file	.px		PC-Axis
Save Excel file	.xls		Excel
Save text file	.txt		Notepad, MS Word or any other compliant word processing software
Save delimited file	.prn		Spreadsheet software. See description below.
Save relational table as text file	.scb		Relational table that can be used in a local database system. See description below.
Show map	.px		The map software in SSD. A map can only be made if the geographic variable is the first variable in the table.

Additional presentation formats accessible from the table page:

Internet, HTML	.htm		File format for web browsers can be produced using the Print option and saving the table that is shown using the Save As option or View Source.
Save dBase file	.dbf		dBase and other systems using dBase files.
Show table with alternative layout	-		Toggle between alternative table layouts. If the option to remove zero lines is used the result will be shown in Table Layout 1.

Text files:

Save file. Select line length, page length and margin. Press the button.	
Line length: Page length: Margin:	80 (One line on an A4 page contains approximately 80 characters.) 70 (One A4 page contains approximately 70 lines.) 4
If the indicated line length is too small to contain the stub and one column, the file will be created without a page break. The table is limited to a maximum of 30 columns and 500 rows.	

Delimited files:

Save file. Select file type and delimiter. Press the button.	
Delimited file with heading Delimited file without heading	Tabulator Comma Space Semi colon
If you select the file format option PRN with heading or without heading, the table will be converted with and without table heading accordingly. All textual fields are quoted and delimited from next field using optional delimiters – tabulation sign, comma, blank or semicolon. The table cells are not quoted but are delimited using the selected delimiter option.	

Relational table as text file:

The file with the relational table contains column headings but not a table title. It is tabulator delimited.

Base files:

Save file. Select file format and press the button.
dBase(as relational table) dBase(with layout as table on screen)
If the file is to be used in a map program it is best to select table layout and place the geographic variable first in the table.

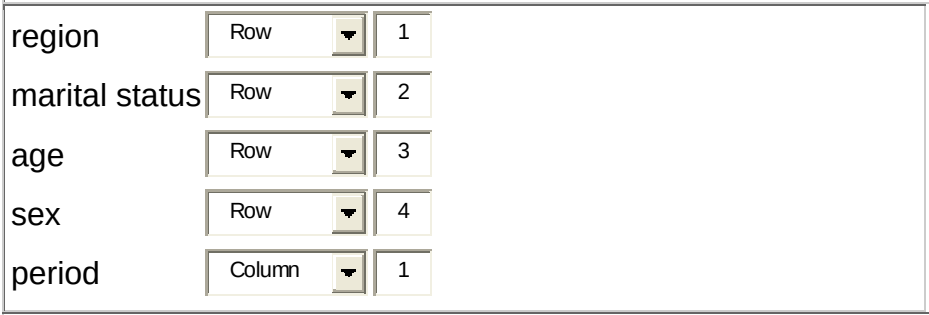
Other functions:

Pivot clockwise		Option to turn the table around clockwise
Pivot anti-clockwise		Option to turn the table around anti-clockwise
Pivot manually		Manual pivot option
Remove zero rows		Suppress lines containing only zeros. This is possible for tables that contain a maximum of three variables in the stub. The result is shown in Table Layout 1.
Sum variable		Summing all values for a selected variable
Print		Print the table.
Help		Help information on this page

 ***Manual pivot options:***

Change layout:

Select row or column for each variable and order number within row/column. Then press the OK button.



The user must enter a consecutive numbering for row and column starting from one.

Summing variables:

Sum variable. Select one variable and press the OK button.	
◉ region ○ marital status ○ age ○ sex ○ period	
New Value Name:	Total

Notice that it is not suitable to make aggregations of indices and averages.

7.6 Exercises

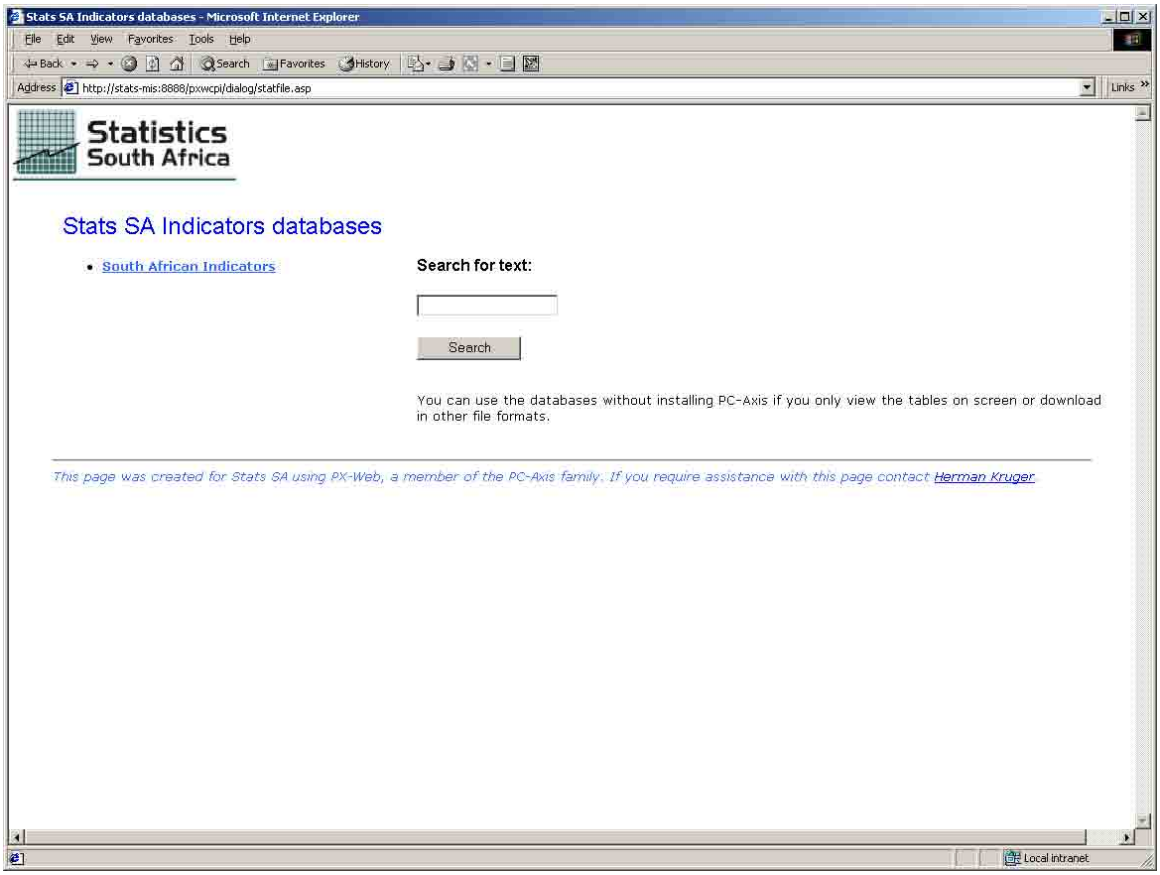
* Please note that this exercise will be on the Statistics SA time series database. Census data will be made available through the Census 2001 part of the website on the release date of Census 2001.

Exercise 1

Step 1: Accessing the system

For training purposes, we will access the Statistics SA Management Information system.

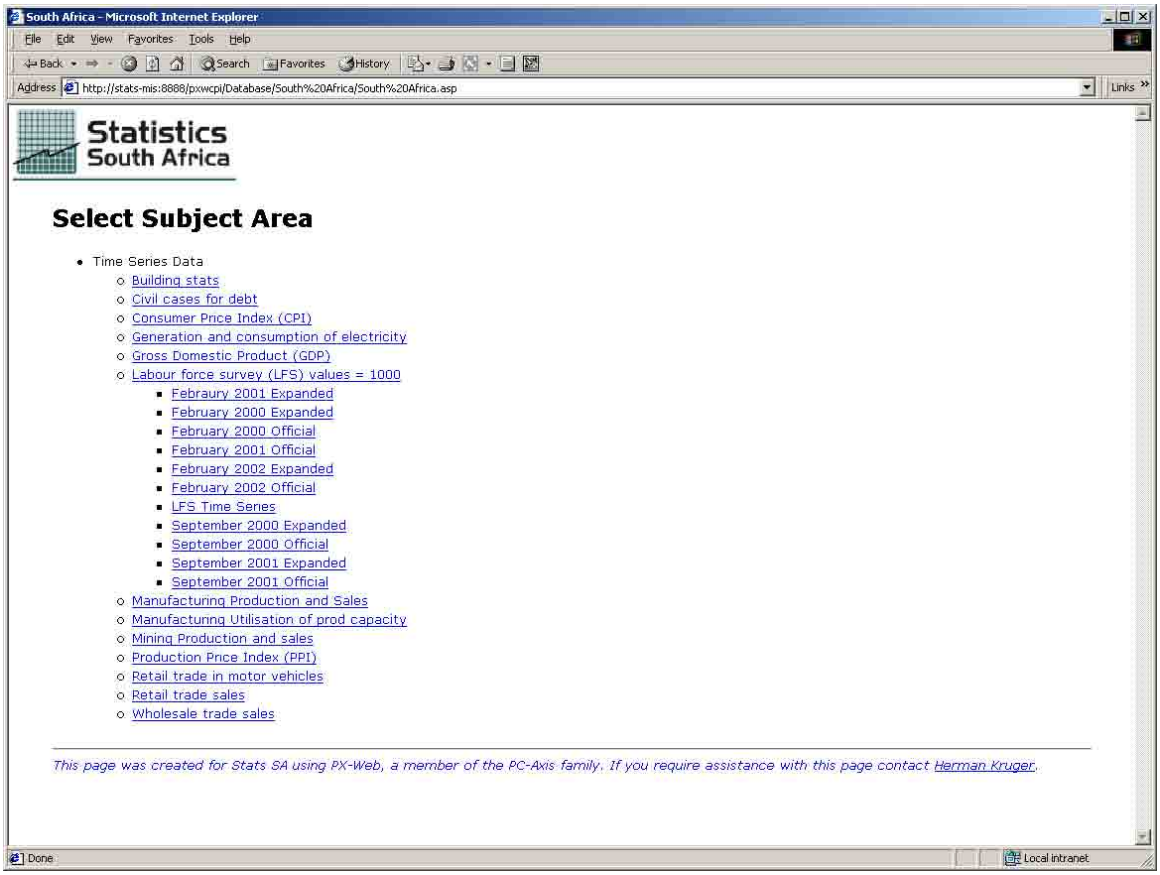
* All economic databases are available from the Statistics SA website (www.statssa.gov.za). Registration, although compulsory, is free. Census 2001 data will be available from the Census links on the Statistics SA website.



This homepage will be displayed for training purposes. The only database listed for this exercise is the South African Indicators. Select this link to progress to the next screen.

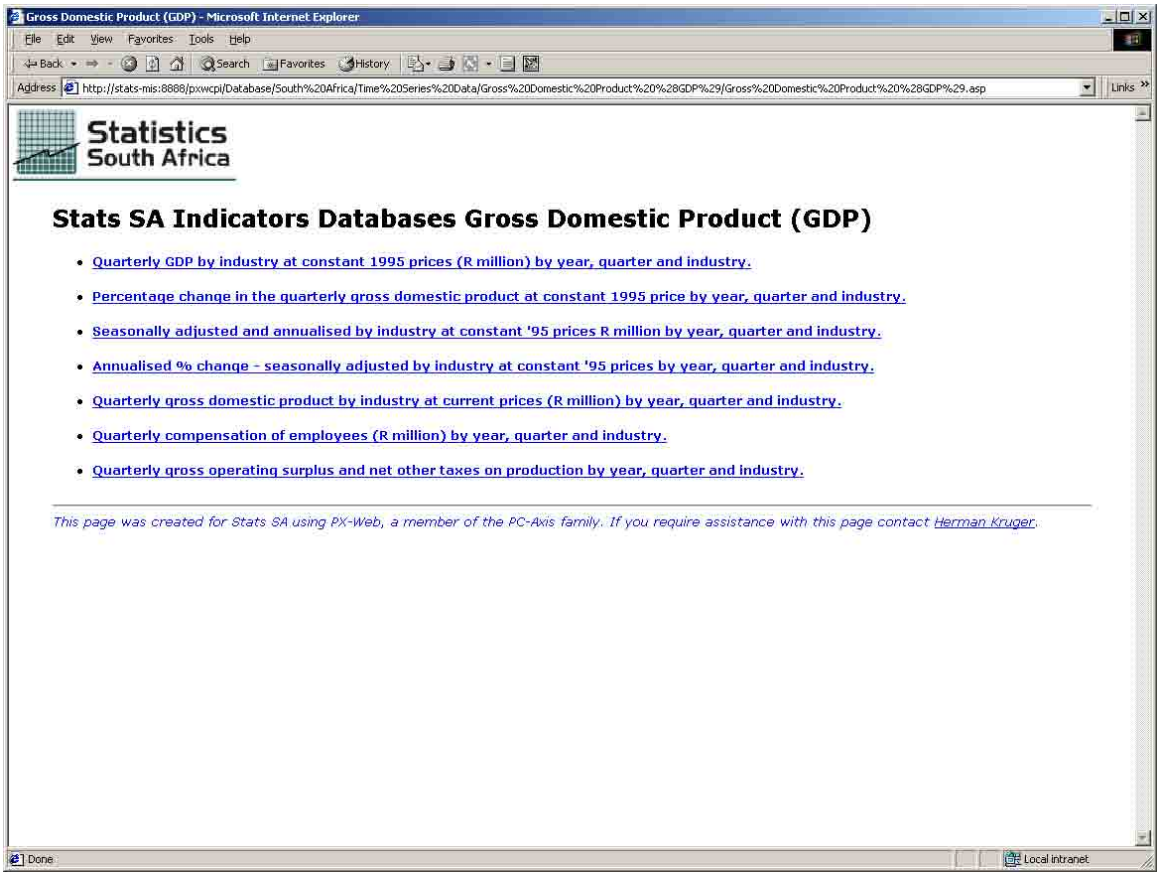
Step 2: Choosing your database to query

Now you can choose the database that you would like to query.
Choose the database for Gross Domestic Product (GDP).


South Africa - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://stats-mis:8888/pxwcpi/Database/South%20Africa/South%20Africa.asp
Links
Statistics South Africa
Select Subject Area
Time Series Data
Building stats
Civil cases for debt
Consumer Price Index (CPI)
Generation and consumption of electricity
Gross Domestic Product (GDP)
Labour force survey (LFS) values = 1000
Febraury 2001 Expanded
February 2000 Expanded
February 2000 Official
February 2001 Official
February 2002 Expanded
February 2002 Official
LFS Time Series
September 2000 Expanded
September 2000 Official
September 2001 Expanded
September 2001 Official
Manufacturing Production and Sales
Manufacturing Utilisation of prod capacity
Mining Production and sales
Production Price Index (PPI)
Retail trade in motor vehicles
Retail trade sales
Wholesale trade sales
This page was created for Stats SA using PX-Web, a member of the PC-Axis family. If you require assistance with this page contact Herman Kruger.
Done
Local intranet

Step 3: Choosing your table to query

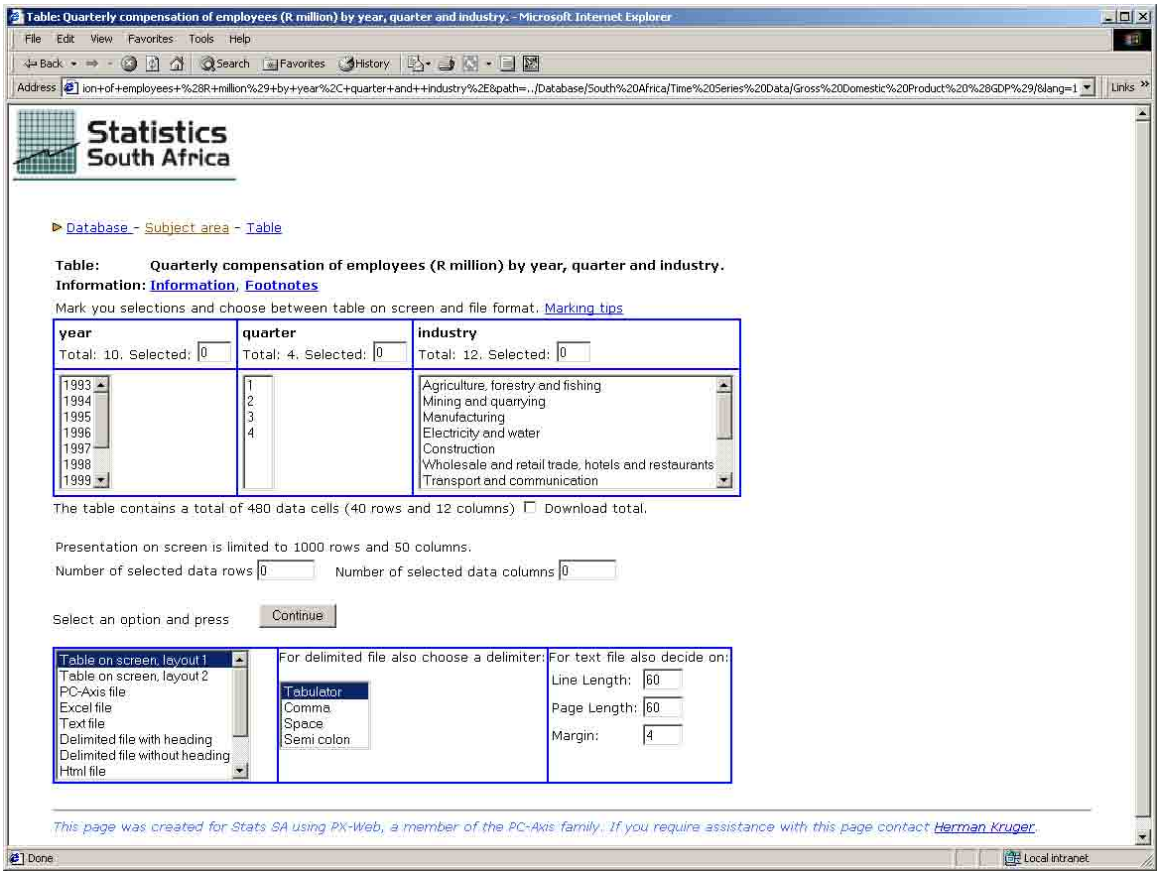
Every database consists of one or more tables. Choose the Quarterly Compensation of Employees table for further analysis.



Step 4: Choosing your variables

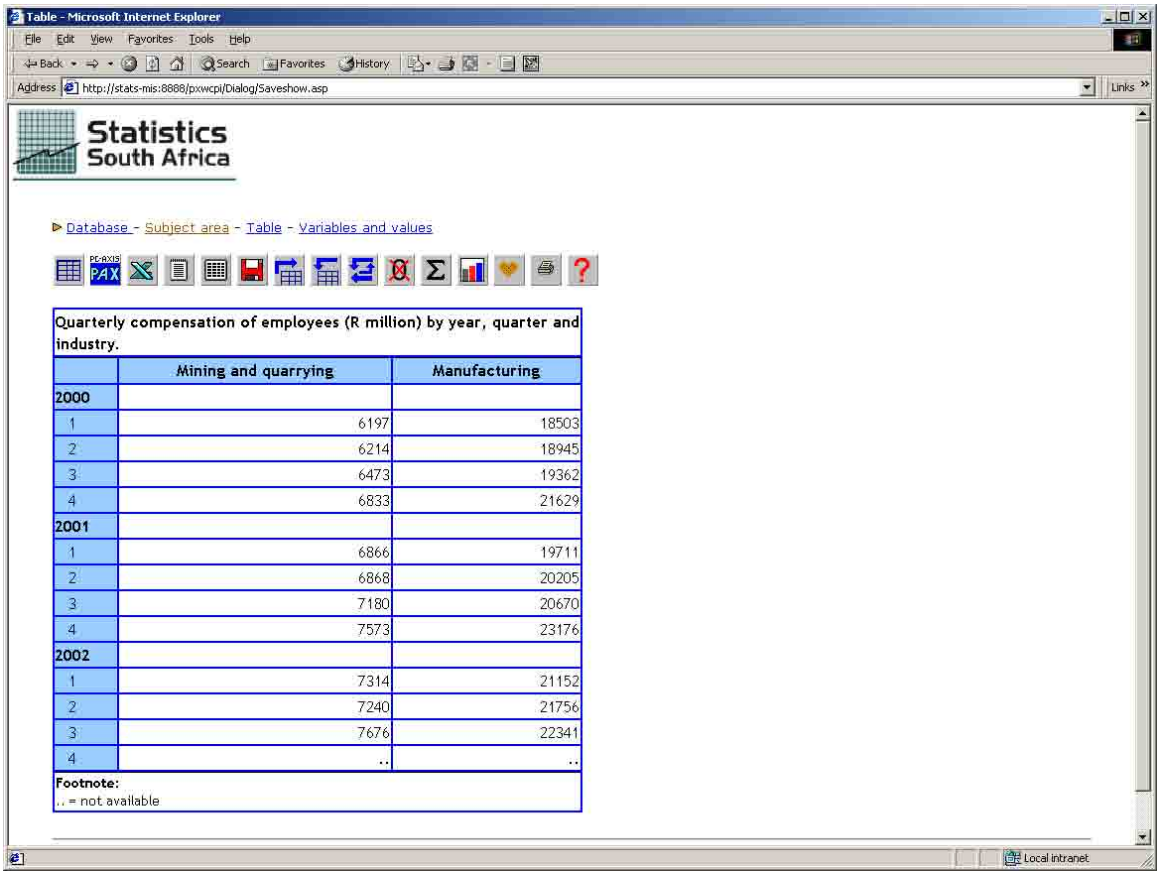
Let us choose the following variables:
From the Year column – 1999, 2000, 2001, 2002. (Multiple selection can be made by using Ctrl + click.)
From the Quarter column – 1, 2, 3, 4. (Multiple selection can be made by using Ctrl + click.)
From the Industry column – Mining and Quarrying, Manufacturing. (Multiple selection can be made by using Ctrl + click.)

Click the Continue button.



Step 5: Building your table

Your table is now being displayed and can be printed. Graphs can also be generated from this table as illustrated in Step 6. Click the graph icon to generate your first graph.



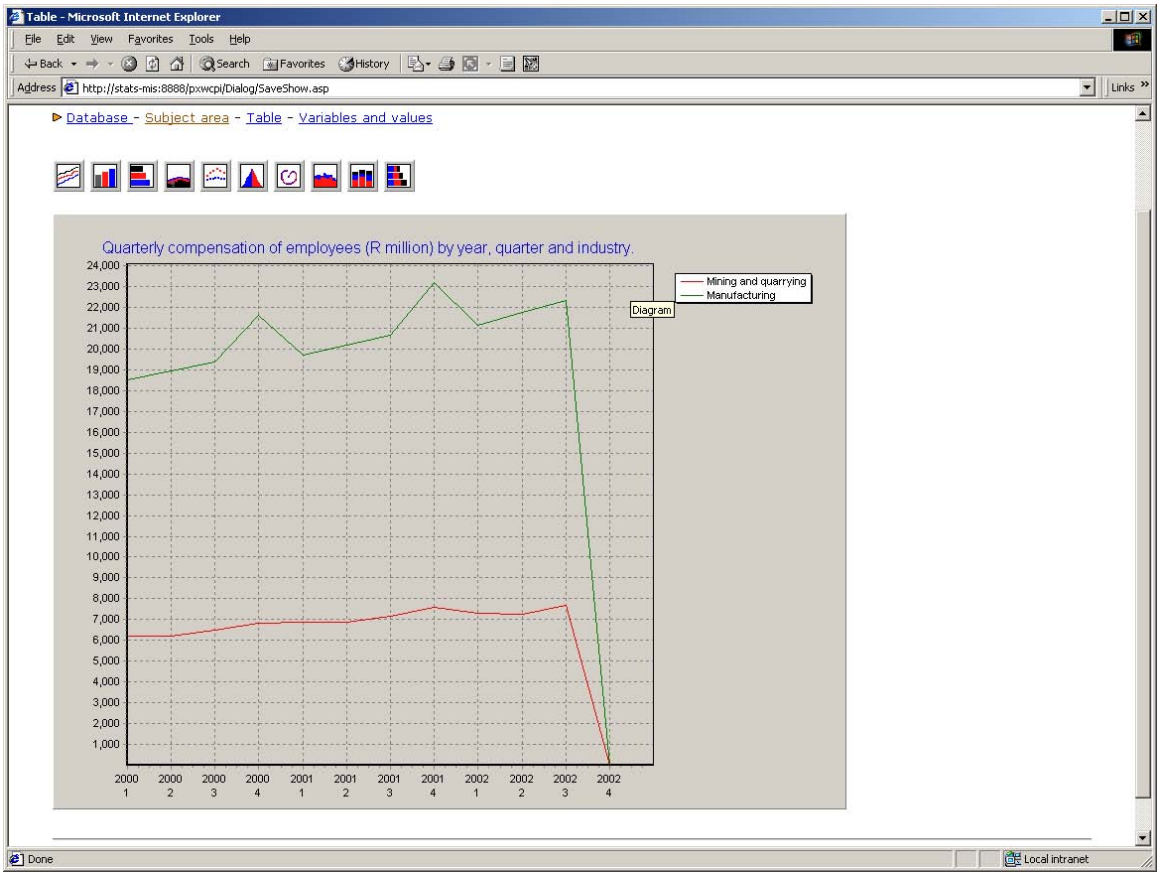
Quarterly compensation of employees (R million) by year, quarter and industry.

	Mining and quarrying	Manufacturing
2000		
1	6197	18503
2	6214	18945
3	6473	19362
4	6833	21629
2001		
1	6866	19711
2	6868	20205
3	7180	20670
4	7573	23176
2002		
1	7314	21152
2	7240	21756
3	7676	22341
4	..	..

Footnote:
.. = not available

Step 6: Generating a graph

The default graph generated is a line graph. This graph is in a picture format and can be printed or downloaded to your local machine.



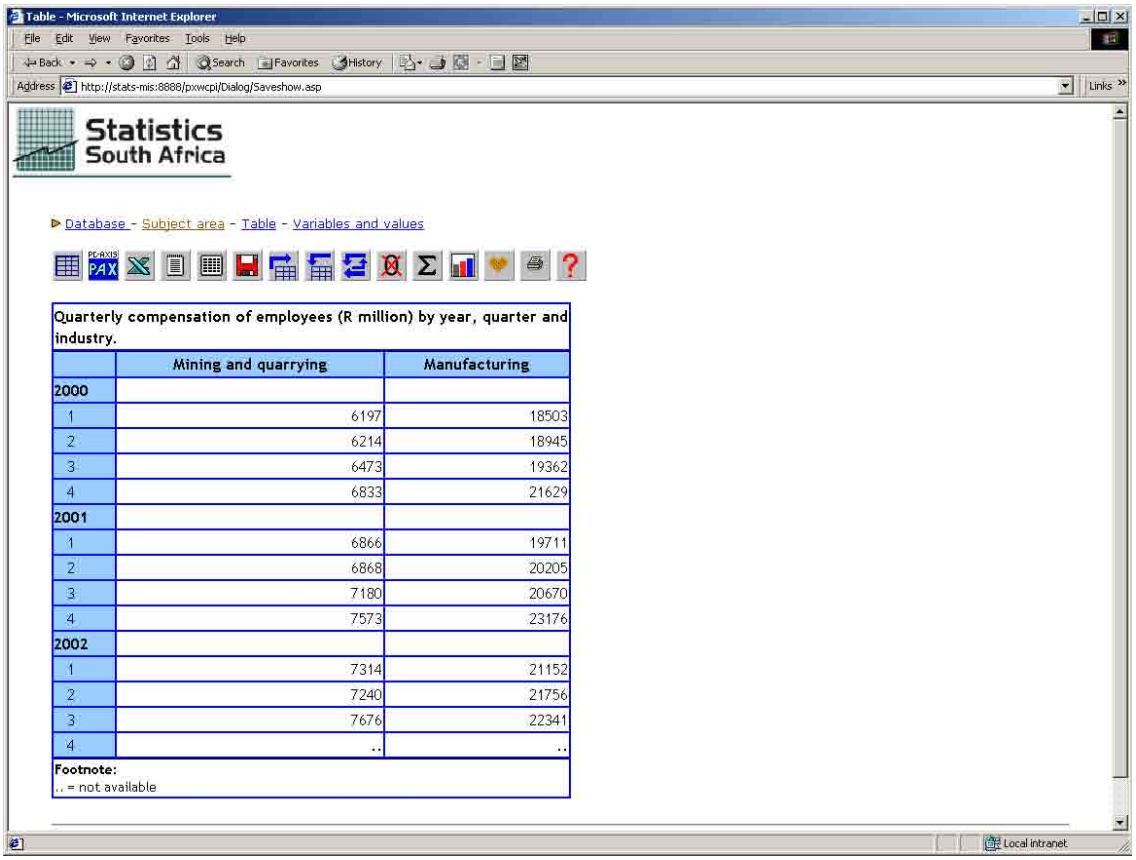
Exercise 2

Step 1: Accessing the system

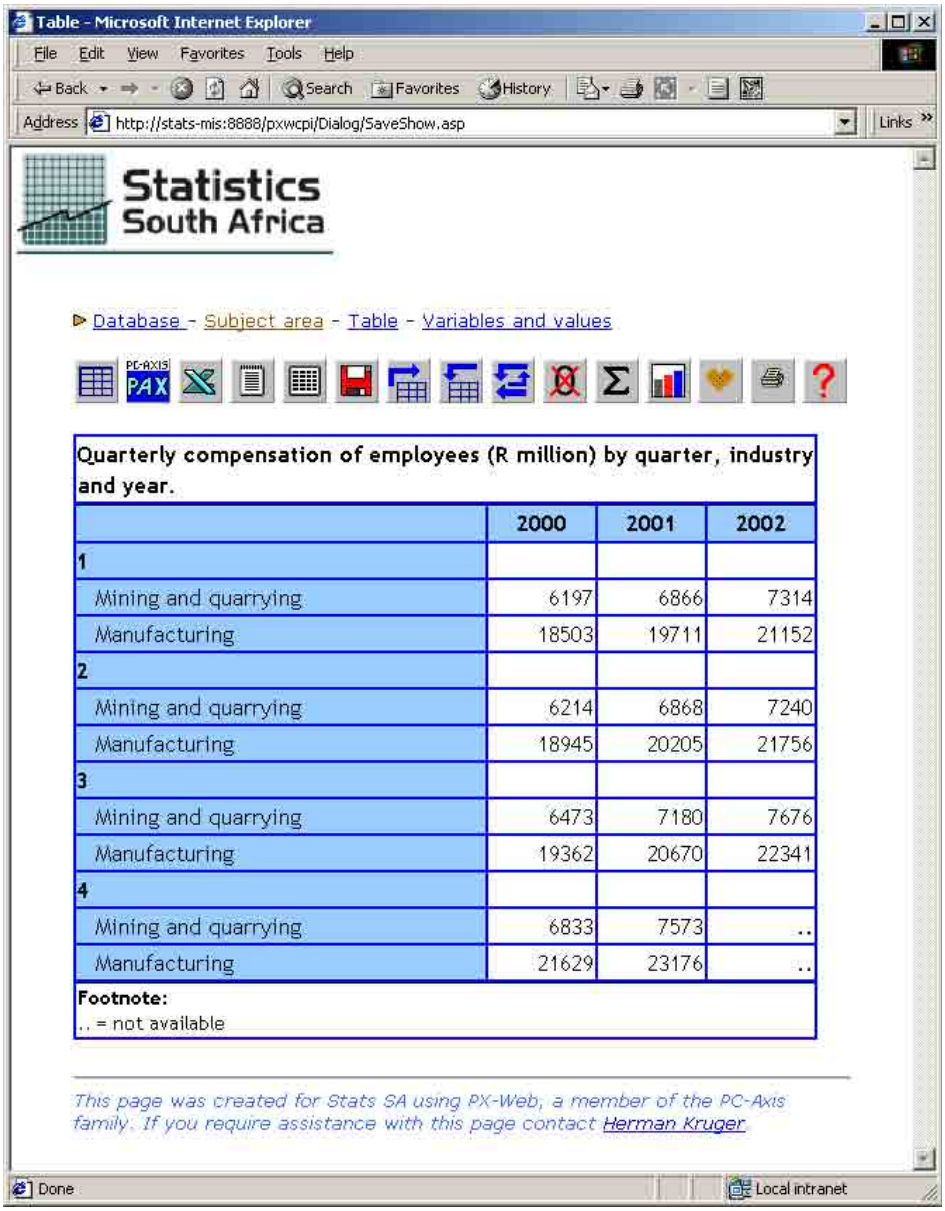
Repeat steps one to four as above.

Step 2: Building your table

Your table is now being displayed.

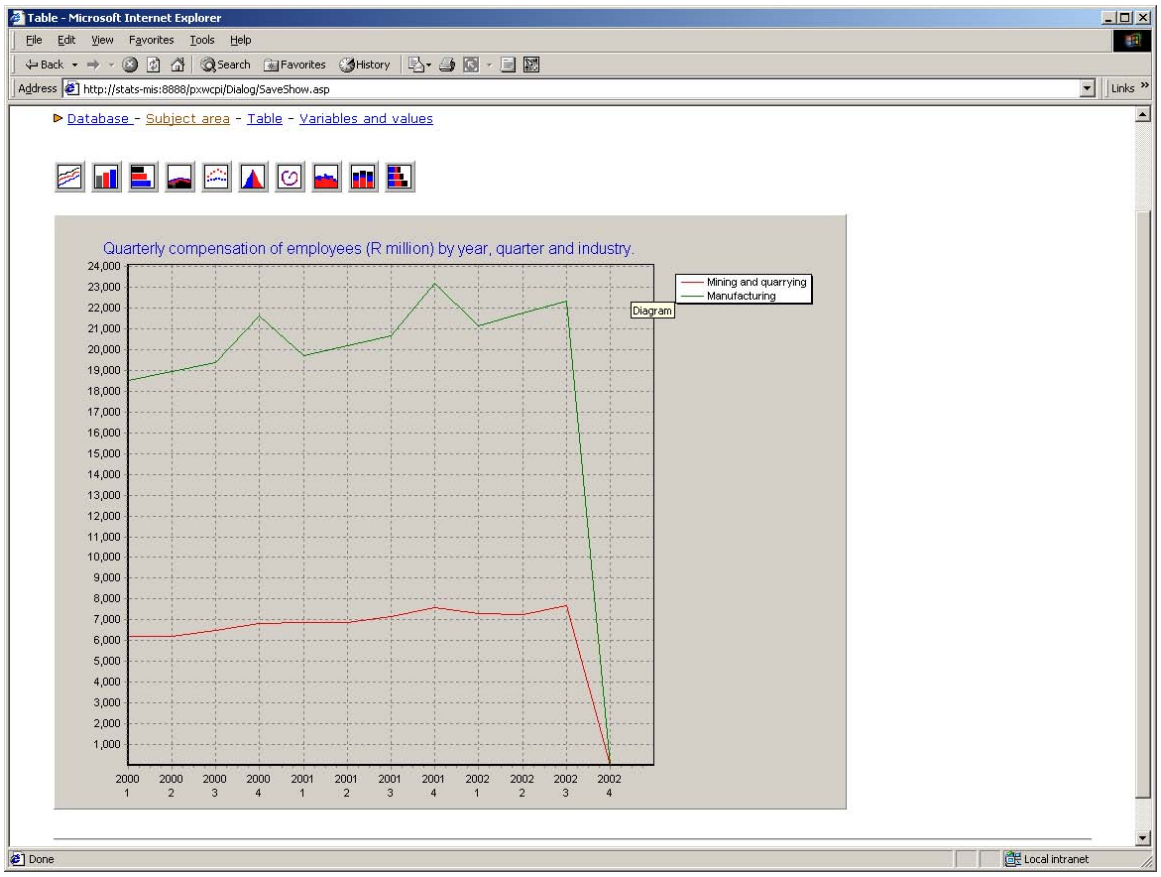


You can use the pivot buttons to pivot your table as shown below.

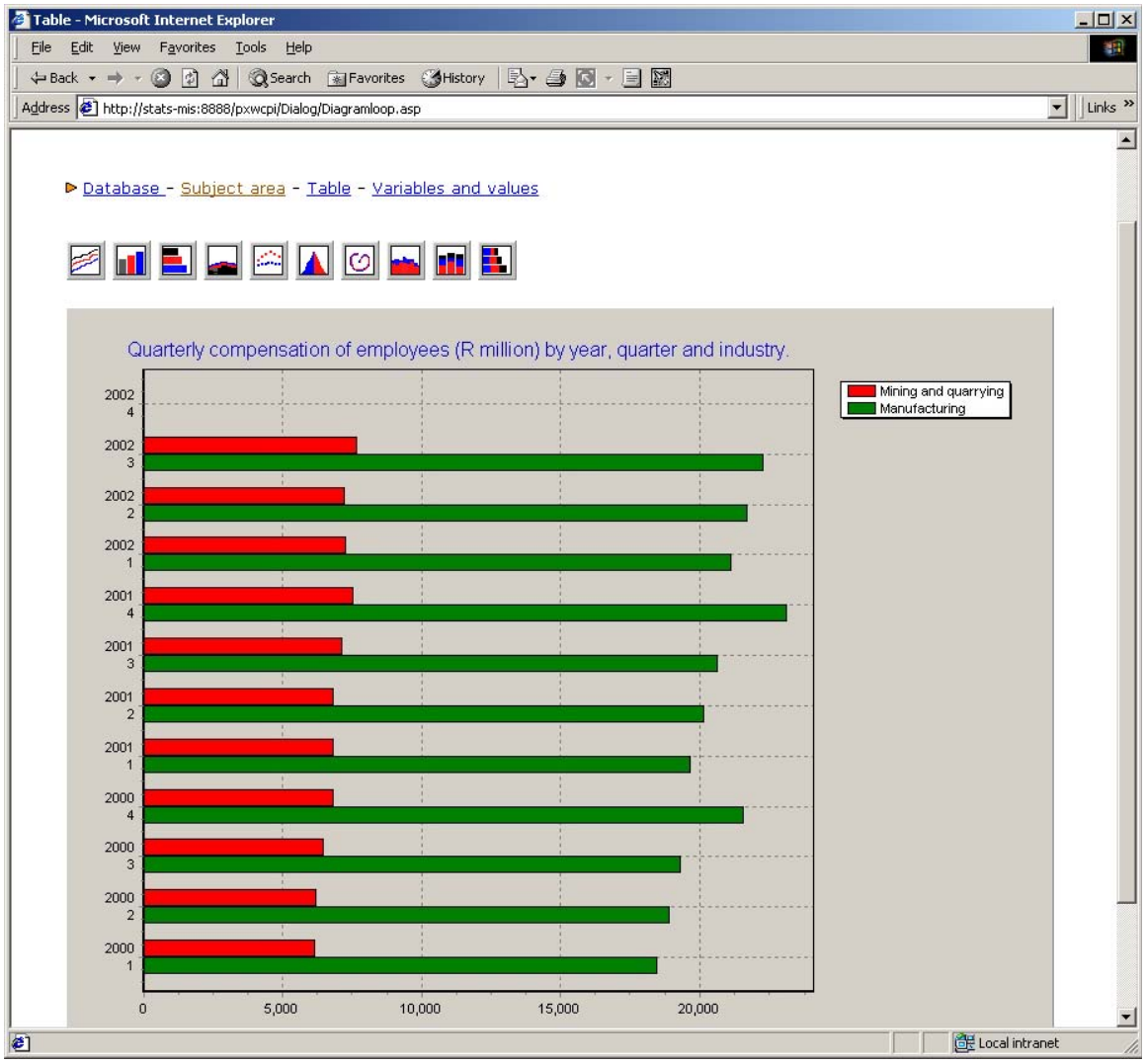


Step 3: Generating a graph

The default graph can be formatted to ten different graph types.



Formatting of graphs is done by using the different graph type buttons. A horizontal bar chart is shown below from the same data as the line chart.



Exercise 3

This is a free exercise; try building tables from any of the available databases.

Some things to try:

Generate an Excel file for download.
Save a graph to your local PC.
Save your query as a cookie.

7.7 Contact us

If you have any queries, questions or comments regarding PX-Web, you are very welcome to contact the Statistics SA Marketing Department at the following numbers:

- Statistics SA Marketing Department
  Telephone: (012) 310 2112 / 2107 /
  2111/ 2108 / 2109 / 2110 / 2105
  Fax: (012) 310 8966

- Manual compiled by Herman Kruger
  E-mail: hermank@statssa.gov.