Community Survey 2007

METHODOLOGY, PROCESSES AND HIGHLIGHTS OF KEY RESULTS

Statistics South Africa
Preferred supplier of quality statistics

General information

Statistics South Africa
Private Bag X44
Pretoria 0001
South Africa

170 Andries Street
Pretoria

User information services
Tel: (012) 310 8600
Fax: (012) 310 8500

Main switchboard
Tel: (012) 310 8911

www.statssa.gov.za
Tel: (012) 310-8600

Community Survey 2007

Methodology, Processes and Highlights of Key Results

Statistics South Africa
2007

Pali Lehohla
Statistician-General

Report No. 03-01-20 (2007)

Community Survey 2007: Methodology, Processes and Highlights of Key Results

Published by Statistics South Africa, Private Bag X44, Pretoria 0001

Community Survey 2007: Methodology, Processes and Highlights of Key Results / Statistics South Africa. Pretoria: Statistics South Africa, 2007

[Report No. 03-01-20 (2007)]
ISBN 978-0-621-37411-7

A complete set of Stats SA publications is available at Stats SA Library and the following libraries:

National Library of South Africa, Pretoria Division
National Library of South Africa, Cape Town Division
Library of Parliament, Cape Town
Bloemfontein Public Library
Natal Society Library, Pietermaritzburg
Johannesburg Public Library
Eastern Cape Library Services, King William's Town
Central Regional Library, Polokwane
Central Reference Library, Nelspruit
Central Reference Collection, Kimberley
Central Reference Library, Mmabatho

This report is available on the Stats SA website: www.statssa.gov.za

Copies are obtainable from: Printing and Distribution, Statistics South Africa
Tel: (012) 310 8044
Fax: (012) 321 7381
(012) 310 2973
Email: publications@statssa.gov.za

Table of contents

Acronyms

Acronym	Term
CS	Community Survey
CSAS	Census and Surveys Administrative System
DDG	Deputy Director-General
DO	District Office
DPLG	Department of Local Government
DSC	District Survey Coordinator
DU	Dwelling Unit
EA	Enumeration Area
FWC	Fieldwork Coordinator
FWS	Fieldwork Supervisor
GIS	Geographic Information System
HH	Household
HO	Head Office
HR	Human Resource
HRM	Human Resource Management
MM	Mapping Monitor
PMO	Project Management Office
PO	Publicity Officer
PSC	Project Steering Committee
PSCs	Provincial Survey Coordinators
QA	Quality Assuror
SG	Statistician-General
Stats SA	Statistics South Africa
VPN	Virtual Private Network

1. Background and purpose

1.1 Objectives of the document

This document provides an overview of the methodologies and processes that were used during the undertaking of the Community Survey (CS) and also presents the key results of the survey.

1.2 Background

The Community Survey is a nationally representative large-scale survey that was carried out by Statistics South Africa (Stats SA) in February 2007. It is designed to provide demographic and socio-economic data up to municipal level. This survey differs from any other survey conducted by the organisation. The decision to undertake such a large survey in 2007 stemmed from Cabinet's decision to postpone Census from 2006 to 2011.

1.3 Primary objectives of the survey

The main objectives of the survey were to:

- provide data at lower geographical levels than existing household-based surveys;
- build human, management and logistical capacities for Census 2011; and
- establish a primary base for a mid-year population projection.

The project strove for maximum coverage and an acceptable degree of quality data. Quality was determined by the extent to which the listing was done correctly, information was correctly recorded on the questionnaires by enumerators, the extent of coverage, and how the data were captured, and how all inconsistencies were eliminated through editing.

1.4 Project planning and approach

Planning for the survey started in August 2004 with the aim of conducting the pilot in February 2006; undertaking the main survey in February 2007; finalising data processing and analysis by October 2007; and releasing the final results to the public by November 2007.

1.5 Management structure

The project was managed using the following management structures:

1.5.1 Executive Committee (EXCO)

EXCO was the primary decision-making body with a policy-setting focus and guidance role. The Statistician-General (SG) is the Chairperson with participation from other Deputy Directors-General (DDGs). EXCO had the overall mandate on policy direction and implementation of the project.

1.5.2 The Statistics Council

The Statistics Council advises the Minister of Finance and Stats SA on all matters relating to the collection, processing, analysis and dissemination of statistics – including the undertaking of population censuses. The Council played a key role in all the project processes.

1.5.3 Project Steering Committee (PSC)

The PSC is an advisory body to the project and is chaired by the Deputy Director-General (DDG) of Population and Social Statistics, with participation from other DDGs in the organisation and nominated representatives from other departments and stakeholders, namely the Department of Local Government (DPLG), Treasury and Agri SA (a farmers' union).

1.5.4 Work streams

The survey was executed by four main operational work streams, namely geography; content development and products; data collection; and data processing and dissemination. These were assisted by five support work streams, namely the Project Management Office (PMO); Public Relations; Provinces; Resource Management; and Information and Communication Systems (ICT).

1.5.5 Project Nerve Centre (NC)

The NC was the nucleus of the project, chaired on a daily or weekly basis (depending on the project tasks at hand) by the Project Director. All work-stream leaders and other support structures attended these sessions. All aspects of the project planning, implementation, monitoring and evaluation were cross-examined against the baseline plans and intervention strategies put in place to address possible delays.

1.5.6 Technical integration meetings

These meetings entailed technical sessions where interaction of work streams and technical tasks as required by the work streams were tabled. These technical sessions discussed and agreed on methodological or technical processes and steps to foster integration among work streams, and were attended by work-stream

leaders, their members and other members in the project structure who were associated with the topic of discussion.

1.5.7 CS group management model

A challenge facing all projects, including the CS, is to execute tasks to meet the required quality standards, within the minimum possible time, cost and resources. Human resources constitute an important aspect in the execution of a project and it is always a challenge to have sufficient human resources with the desired skills.

The Community Survey was no exception. During the resource-planning phase of the CS, it became apparent that there were not sufficient human resources or sufficient skills to undertake the project. The CS management model adopted was therefore based on the team cells approach. Although the project governance model was work-stream based, the day-to-day methodological and operations processes were driven by teams who were headed by an expert in a specific field. This was a deliberate strategy adopted mainly because the project did not have sufficient and experienced personnel and also became a conduit for on-the-job training of new and existing personnel on all aspects of the statistical processes.

Another alternative adopted was to second or delegate individuals to leadership positions, not based on their skills and experience, but solely based on potential. At project level, individuals were seconded to positions as work-stream leaders, where their duties, responsibilities and authority were just below that of the project director. At work-stream level, particularly Data Collection, it became apparent that finding suitable Provincial Survey Coordinators (PSCs) would be a challenge, and as a result, individuals involved in the project at a lower level were seconded to PSCs in the provinces.

At project management level, clear communication lines with provinces and other internal stakeholders were established. An integrated planning process with the provinces and other internal stakeholders was found essential to establish clear communication lines and continuous feedback on updates.

1.6 Questionnaire development and design

1.6.1 User consultation, data items and the questionnaire

The overall process of design and development of the questionnaire was done in different phases, namely:

- Users' needs assessment phase;
- Stakeholders' consultation phase;
- Design and development phase;
- Testing/evaluation of developed tools and methods;
- Approval process of developed tools and methods;
- Final approval of the questionnaire; and
- Monitoring of quality for implementation/production/printing/training of developed tools or methods.

1.6.2 Consultative process

Ten stakeholder workshops were held across the country during August and September 2004. Approximately 367 stakeholders, predominantly from national, provincial and local government departments, as well as from research and educational institutions, attended. The workshops aimed to achieve two objectives, namely to better understand the type of information stakeholders need to meet their objectives, and to consider the proposed data items to be included in future household surveys. The output from this process was a set of data items relating to a specific, defined focus area and outcomes that culminated with the data collection instrument (see Annexure B for all the data items).

1.6.3 Design of the questionnaire

The design of the CS questionnaire was household-based and intended to collect information on 10 people. It was developed in line with the household-based survey questionnaires conducted by Stats SA. The questions were based on the data items generated out of the consultation process described above. Both the design and questionnaire layout were pre-tested in October 2005 and adjustments were made for the pilot in February 2006. Further adjustments were done after the pilot results had been finalised.

2. Sample design and estimation

2.1 Introduction

The sample design that was adopted took into consideration the sampling frame, methodology to be used, the objectives of the survey, and the data items that were to be collected. The sample design also took cognisance of the challenges that were experienced during Census 2001.

2.2 The sampling approach

The sampling approach consisted of two stages, namely the selection of enumeration areas, and the selection of dwelling units.

Each municipality was considered a unique stratum. The stratification is done for those municipalities classified as category B municipalities (local municipalities) and category A municipalities (metropolitan areas) as proclaimed at the time of Census 2001. However, the newly proclaimed boundaries as well as any other higher level of geography such as province or district municipality, were considered as any other domain variable based on their link to the smallest geographic unit – the enumeration area.

The weights were minimally adjusted to ensure consistency between the estimates of the survey and those of the census as well as expected numbers of surviving children and men relative to women, using estimates of demographic parameters from the survey. These adjustments were kept to a minimum.

2.2.1 The frame

The Census 2001 enumeration areas were used because they give a full geographic coverage of the country without any overlap. Although changes in settlement type, growth or movement of people have occurred, the enumeration areas assisted in getting a spatial comparison over time. Out of 80 787 enumeration areas countrywide, 79 466 were considered in the frame. A total of 1 321 enumeration areas were excluded (919 covering institutions and 402 recreational areas).

On the second level, the listing exercise yielded the dwelling frame which facilitated the selection of dwellings to be visited. The dwelling unit is a structure or part of a structure or group of structures occupied or meant to be occupied by one or more households. Some of these structures may be vacant and/or under construction, but can be lived in at the time of the survey. A dwelling unit may also be within collective

living quarters where applicable (examples of each are a house, a group of huts, a flat, hostels, etc.).

The Community Survey universe at the second-level frame is dependent on whether the different structures are classified as dwelling units (DUs) or not. Structures where people stay/live were listed and classified as dwelling units. However, there are special cases of collective living quarters that were also included in the CS frame. These are religious institutions such as convents or monasteries, and guesthouses where people stay for an extended period (more than a month). Student residences – based on how long people have stayed (more than a month) – and old-age homes not similar to hospitals (where people are living in a communal set-up) were treated the same as hostels, thereby listing either the bed or room. In addition, any other family staying in separate quarters within the premises of an institution (like wardens' quarters, military family quarters, teachers' quarters and medical staff quarters) were considered as part of the CS frame. The inclusion of such group quarters in the frame is based on the living circumstances within these structures. Members are independent of each other with the exception that they sleep under one roof.
The remaining group quarters were excluded from the CS frame because they are difficult to access and have no stable composition. Excluded dwelling types were prisons, hotels, hospitals, military barracks, etc. This is in addition to the exclusion on first level of the enumeration areas (EAs) classified as institutions (military bases) or recreational areas (national parks).

2.2.2 The selection of enumeration areas

The EAs within each municipality were ordered by geographic type and EA type. The selection was done by using systematic random sampling. The criteria used were as follows:

- In municipalities with fewer than 30 EAs, all EAs were automatically selected.
- In municipalities with 30 or more EAs, the sample selection used a fixed proportion of 19% of all sampled EAs. However, if the selected EAs in a municipality were less than 30 EAs, the sample in the municipality was increased to 30 EAs.

Table 1 gives the distribution of municipalities by the number of EAs selected.

Table 1: Number of municipalities by minimum number of EAs selected and province

Province	<=30 EAs	>30 EAs	Total
Eastern Cape	18	23	**41**
Free State	11	9	**20**
Gauteng	6	7	**13**
KwaZulu-Natal	31	21	**52**
Limpopo	7	19	**26**
Mpumalanga	14	12	**26**
Northern Cape	30	1	**31**
North West	11	14	**25**
Western Cape	26	4	**30**
Total	**154**	**110**	**264**

Table 2 shows the total number of EAs sampled and the distribution of these EAs across the provinces.

Table 2: Distribution of EAs in the sample and frame by province

Province	Number of EAs in sample	Sample % distribution	Number of EAs in sample frame	Sample % distribution	% EAs sampled
Eastern Cape	3 699	22%	18 160	23%	20,4
Free State	1 087	6%	5 115	6%	21,3
Gauteng	2 493	15%	12 949	16%	19,3
KwaZulu-Natal	2 767	16%	12 547	16%	22,1
Limpopo	2 045	12%	10 342	13%	19,8
Mpumalanga	1 166	7%	5 629	7%	20,7
Northern Cape	846	5%	1 485	2%	57,0
North West	1 369	8%	6 400	8%	21,4
Western Cape	1 626	10%	6 839	9%	23,8
Total	**17 098**	**100%**	**79 466**	**100%**	**21,5**

2.2.3 The selection of dwelling units

The second level of the frame required a full re-listing of dwelling units. The listing exercise was undertaken before the selection of DUs. The adopted listing methodology ensured that the listing route was determined by the lister. This

approach facilitated the serpentine selection of dwelling units. The listing exercise provided a complete list of dwelling units in the selected EAs.

Only those structures that were classified as dwelling units were considered for selection, whether vacant or occupied. This exercise yielded a total of 2 511 314 dwelling units.

The selection of the dwelling units was also based on a fixed proportion of 10% of the total listed dwellings in an EA. A constraint was imposed on small-size EAs where, if the listed dwelling units were less than 10 dwellings, the selection was increased to 10 dwelling units. All households within the selected dwelling units were covered. There was no replacement of refusals, vacant dwellings or non-contacts owing to their impact on the probability of selection. Concerted efforts were made to improve the response rates through multiple visits.

Table 3 gives the number of dwelling units that were selected after the listing exercise.

Table 3: Sampled EAs and dwelling units

Province	Number of EAs in the sampling frame	Number of EAs sampled	Sampled EAs with identifiable dwelling units	Number of dwelling units
Western Cape	6 839	1 626	1 589	31 532
Eastern Cape	18 160	3 699	3 231	40 404
Northern Cape	1 485	846	820	13 139
Free State	5 115	1 087	1 060	16 726
KwaZulu-Natal	12 547	2 767	2 695	45 988
North West	6 400	1 369	1 323	23 154
Gauteng	12 949	2 493	2 448	58 997
Mpumalanga	5 629	1 166	1 115	17 849
Limpopo	10 342	2 045	1 973	26 559
Total	**79 466**	**17 098**	**16 255**	**274 348**

2.3 The weights calculation

The CS sample has equal probabilities for all elements in the cluster which make it a self-weighting systematic random sample. Since the sample is stratified by municipalities as demarcated at the time of Census 2001, the inclusion probability of selection of an EA at the first level of selection, and the dwelling unit at the second level of selection, is the product of first and second-level probabilities. Also, since all

households within the dwelling unit are considered, their probability of being in the dwelling unit is always one.

Since the probability of selection of an EA and dwelling unit is: $\prod_{hi} = \frac{n_h}{N_h} \times \frac{m_i}{M_i}$ the design weight is a straightforward inverse of the probability represented as $W_{hi} = \frac{1}{\prod_{hi}}$ where n_h represents the numbers of selected EAs in the municipality (explicit stratum h), N_h the total number of EAs in the municipality, m_i represents the numbers of selected dwelling units in a given EA and M_i the total numbers of listed dwelling units.

It is important to note that non-responses in the CS can occur at EA level, at dwelling-unit level and at household level. For instance, there were two EAs in Western Cape where fieldworkers were unable to gain access because of political unrests. On dwelling-unit level, only 238 067 out of 274 348 sampled dwelling units returned completed questionnaires. This means that out of 16 255 EAs with listed dwelling units that were visited for enumeration, the completed questionnaires came from only 16 173 EAs.

There are also non-responses at household level which occur inside the dwelling unit. However, the undercount of households in the dwelling unit as well as the undercount of some persons in the households was not easy to account for without any dual estimation approach in place, such as the post-enumeration survey. In general, non-responses can be dealt with as either non-coverage, undercount, or proper non-responses, depending on the situation.

It is imperative that the lack of responses be adjusted in order to take care of the unknown condition at the time of the design. The adjustment of non-responses is based on the classification of dwelling units or households, based on their enumeration status (enumeration completed, partially completed, non-contact, refusal, unusable information, listing error, unoccupied dwelling, demolished dwelling, vacant dwelling, and others). These categories are grouped into three (respondent, non-respondent and invalid/out of scope). The ‘respondent’ is the household within a dwelling unit where the survey information is available. The ‘non-respondent’ is the household within a dwelling unit that did not provide any positive response; and ‘invalid cases’ are dwelling units that were included in the frame at the time of listing,

but that would have been ignored if the information of their status was known (e.g. listing error, unoccupied, demolished and vacant).

Table 4 gives a summary of the CS response rates, with the percentage of responses totalling 93,9% as opposed to a non-response rate of 6,1%.

Table 4: CS response rates

Response category	Number of dwelling units	Percentage of responses
Completed cases responding	238 067	**93,9%**
Non-response cases	15 393	**6,1%**
Invalid or out-of-scope cases	20 888	
Total number of dwelling	274 348	

In view of the above response rate, the non-responses are adjusted from the design weight using the following formula applicable to each municipality *h* and EA *i*:

$$W_{hi} = \left[\frac{N_h}{n_h} \times \frac{M_i}{m_i}\right] \times \left(\frac{t_h}{x_h}\right) \times \left(\frac{(m_i - o_i)}{y_i}\right)$$

where n_h represents the number of selected EAs in the municipality (explicit stratum *h*), N_h the total number of EAs in the municipality, m_i represents the number of selected dwelling units in a given EA, M_i the total number of listed dwelling units, x_h is the realised EA in the sample, y_i is the realised DUs (dwelling units), t_h is the number of listed EAs, and O_i represents cases out of scope.

2.4 The benchmarking of CS results

The calibration process forces the weighted estimates of certain variables to match known or alternatively estimated population totals called benchmarks. In the case of the CS, the calibration will be used to correct non-responses or coverage errors resulting from undercoverage of the frame, or from any possible overcount. Also, the response errors may occur due to a lack of accurate answers to the CS questions. This may be due to poor understanding of the questions, lack of reliable information, or lack of effort to remember or retrieve information from the past. Also, this is due to the fact that one respondent provides information on behalf of other members of the

household. In the South African context, enumerators might be translating the questions inaccurately into the language of the respondent.

It is well known that every sample is more or less skewed because it gives a reflection of a population within certain confidence intervals. Also, fieldworkers may have introduced some erroneous information in addition to household refusals, non-contacts and other issues. This may result in the data obtained from members of a household being incomplete by overlooking some persons or omitting information. This is why the need to reduce the skewedness by either benchmarking or using an adjustment on a reliable post-enumeration survey (PES).

In the context of the Community Survey, the available alternative for reducing possible skewedness in the sample was to prepare fresh population estimates using the demographic parameters as measured from the survey. Then, after the target population had been determined, the adjustment process of calibration was applied to yield adjusted CS results.

The determination of the new population estimates was done using Census 2001 as the base reference population, which was projected by using the cohort-component method from 2001 to 2007 with some interpolation to February 2007. The population projection took into account the inter-provincial migration as measured from the survey. Mortality and fertility over the period were estimated to be consistent with the rates from the Community Survey. There was also an adjustment in respect of men mostly between 20 and 39 years old, by using sex ratios to compensate for men who were undercounted relative to women.

After obtaining the best reliable population estimate, the CS data were calibrated by an adjustment factor in each cell of the sub-class of adjustment composed by province, population group, sex, and each single year. The adjustment factor was derived from the difference between the target projected population in each sub-class and the CS population adjusted (taking into account the number of non-responses).

In order to have a complete population estimate for South Africa, those cases not in scope such as collective living quarters (institutions) and some normal households in EAs classified as recreational areas or institutions, needed to be added to the CS estimates. However, no recent account of these cases out of the CS scope was available, except the Census 2001 data. Hence, the adopted approach was to consider each individual record as counted in 2001 without any change over time,

since they represent 1,3% of the overall population. For easy manipulation of the data, a sub-sample of 17 000 of those people outside the CS scope was extracted and reweighted to give the overall expected count of 651 769 persons.

The adjustment of households was based on the general adjustment of the people within households across all ages, which was applied to the household weight adjusted of non-responses. This gave a solution of dealing with general undercounts (and occasional overcounts) that may have occurred at household level.

Finally the household weights were adjusted for any general adjustments made to population numbers in any sex-population group-province combination in order to ensure that the average number of people per household remained relatively unchanged.

3. Pilot survey

The pilot survey was conducted in February 2006. The purpose of the pilot was to test all the developed strategies, methodologies, systems, and the questionnaire. A total of 782 EAs were covered using a staff complement of 15 District Survey Coordinators, 15 District Logistics Officers, 18 Fieldwork Coordinators, 51 Fieldwork Supervisors and 208 Enumerators.

Table 5: Sampled EAs for the pilot

Province	No. of sampled EAs	No. of sampled DUs
Western Cape	100	2 516
Eastern Cape	100	1 686
Northern Cape	60	1 726
Free State	60	1 229
KwaZulu-Natal	100	2 521
North West	60	1 351
Gauteng	100	3 230
Mpumalanga	101	2 594
Limpopo	101	2 146
Total	**782**	**18 999**

During the pilot survey the effectiveness of instruments, processes and methods used within the scope of the CS were tested. A range of lessons were learnt which led to the refinement of processes, methods and systems towards the main survey.

4. Pre-enumeration activities

4.1 Publicity

Publicity plays a crucial role in the survey process since it is responsible for laying the foundation for enumeration. Successful enumeration and a high response rate is closely dependent on an effective publicity campaign.

4.1.1 Publicity methodology

The publicity and advocacy campaign consisted of various phases; namely:

- Promotion and advocacy;
- Education and call-to-action campaign through mass media; and
- Thank you campaign.

4.1.2 Publicity at an EA level

The 'EA Level' methodology of publicity was used in EAs where there were no refusals during listing publicity. At an EA level, the Publicity Officer made an appointment with gatekeeper(s) with the aim of explaining the importance of the Community Survey. Only when permission had been granted by the relevant gatekeeper(s) was the Publicity Officer able to start publicity in the selected EA. The Publicity Officer then placed publicity posters at the most suitable and visible location that was accessible to the general public. Pamphlets were also distributed in public areas, e.g. bus stations, taxi ranks, municipal offices, traffic departments, clinics, and police stations. On completion of the publicity campaign in the EA, the Publicity Officer announced his/her departure to the gatekeeper(s), specifically thanking them for granting access and permission to work in the area.

4.1.3 Publicity at a DU level

The 'DU Level' methodology of publicity was used in EAs where refusals were encountered during listing publicity, as well as in areas that were known to be problematic. The Publicity Officer visited all sampled dwelling units, knocked on the door and made a detailed explanation about the Community Survey to the respondent. The aims and objectives of the CS 2007 were then explained to ensure that the respondent understood all the details that were required. Upon leaving the premises, the Publicity Officer left publicity materials with the respondents to inform them of the enumeration period (7–28 February, 2007).

4.1.4 Promotion and advocacy

Four hundred Publicity Officers were recruited across the country to promote CS 2007 and inform communities about the upcoming data collection. According to the quality assurance plan, Publicity Coordinators and Supervisors were responsible for

quality checks within their respective districts. The aforesaid supervisors had to be familiar with the districts where they operated. They also conducted spot-checks on publicity booklets submitted by each team.

In the promotion and education campaign, Publicity Officers visited selected areas to inform respondents about the survey. Pamphlets and posters were also used as part of a promotional campaign to create awareness around the organisation. As part of the campaign, Publicity Officers also embarked on a wider consultation phase with gatekeepers such as councillors, body corporates, indunas, farmers' unions, etc. in order to facilitate access to specific communities.

4.1.5 Education and call-to-action campaign through the mass media

Stats SA also made use of the mass media to intensify its call-to-action campaign. Key messages aimed at the general public for support were designed in the form of print, radio and television. Interviews with Subject-matter Experts (SMEs) were carried out on different radio stations across the country. Publicity and advocacy were also conducted at dwelling-unit level across the country.

4.1.6 Thank you campaign

This campaign came in the form of a press conference and billboards, which were put up across the country at the end of the enumeration process.

4.2 Listing of structures

Listing is the compilation of a register of all dwelling units and other structures in all selected Enumeration Areas (EAs). Listing fieldwork commenced on 7 August 2006 and was completed on 15 November 2006 with the target sample consisting of 17 098 EAs. A three-phase listing approach was adopted to have a manageable team of listers, to ensure that mistakes made during listing were corrected while still in the field, and to ensure that EAs that change rapidly (e.g. squatter camps) were listed towards the beginning of data collection.

4.2.1 Development of listing methodology, quality assurance and map-reading guidelines

The methodology for listing was developed in line with that adopted by other surveys within the organisation. A few key adaptations were made in order to suit the purposes of the Community Survey. Some changes were made to the design of listing forms and elements that were to be captured on the listing forms. Some forms were collapsed and consolidated into a single A3 size (CS-HH3). These forms were then bound together with orientation and EA maps in an A3 size book format.

4.2.2 Technical consultation

During the process of developing the methodology and other documents, several technical consultation meetings were held between CS Geographical Operations and the Social Statistics division. Presentations were also made to a technical committee consisting of representatives from all the work streams within the Community Survey. Documents were reviewed and technical advice was given where necessary. The documents were only signed off when the members of the technical committee were satisfied with the structure and content of the documents. The listing and quality assurance methodologies as well as listing tools were developed simultaneously. The development of the listing and quality assurance methodology, EA Summary Book and other listing tools started in May 2005. These manuals and tools were tested during the pilot phase in October 2005. Thereafter, necessary changes were effected based on pilot listing debriefing/reviews and recommendations submitted from the provinces on the usability and quality of the manuals. After the listing pilot phase in September 2005, CS Geographical Operations held numerous methodological updating and refinement meetings with the provinces. Several meetings were held with Social Statistics personnel in order to ensure that the Community Survey outputs would continue to be useful to both parties.

4.2.3 Listing methodology

During listing fieldwork, listing field-teams were provided with EA Summary Books and writing materials. Field staff listed all dwelling units and other structures within the EA according to the listing rules and guidelines. More specifically, the following were listed: all dwelling units, all units or rooms within collective living quarters, non-residential buildings, vacant stands, sports fields, parks, parking lots, cemeteries, demolished and semi-demolished structures etc.

4.2.4 Listing quality assurance

Quality assurance was an integral part of the listing field operations. This process involved field and office quality assurance. All personnel involved in listing fieldwork were trained on quality assurance, how to identify common errors as well as how to avoid or correct them, but mostly the process involved checking for errors on all elements on the EA Summary Book and beyond.

At the end, all EAs except two in Western Cape Province were listed. The two in Western Cape could not be listed because fieldworkers were denied access to these EAs.

4.2.5 Map-reading

Map-reading was an important tool for EA orientation; locating the sampled EA and identifying the boundary limits. A lot of time was assigned to the development of fieldworker map reading skills through both theory and practical training.

4.2.6 EAs listed

Table 6 below depicts the number of EAs successfully listed and the status of the others that were never listed.

Table 6: Number of EAs with sampled dwelling units (from listing)

Province	Total sampled EAs	Number of EAs with sampled DUs	EAs that were vacant	EAs where there were refusals
Eastern Cape	3 699	3 231	468	0
Free State	1 087	1 060	27	0
Gauteng	2 493	2 448	45	0
KwaZulu-Natal	2 767	2 695	72	0
Limpopo	2 045	1 973	72	0
Mpumalanga	1 166	1 116	50	0
Northern Cape	846	820	26	0
North West	1 369	1 323	46	0
Western Cape	1 626	1 589	35	2
Total	**17 098**	**16 255**	**841**	**2**

4.3 Recruitment

4.3.1 Method

Recruitment was decentralised to the provincial and district offices, and technically supported by a centralised control structure at Head Office. The strategy adopted for recruitment was to consider the internal candidates who were involved in other processes within the project first, and thereafter source for external candidates to fill the remainder of the positions through the normal advertisements. The profiling of such candidates and placement was done through the Census and Surveys Administration System (CSAS) that had links down to the regional offices. The system used geographic locations to help with the administration of automatic placements. All recruitment activities were supported by HR Officers and progress was monitored on a daily basis from a control point at Head Office. The EAs within which field activities were conducted for the CS 2007, were linked to either a district

municipality or metro and thus the recruitment process followed the same hierarchy approach.

4.4 Training

4.4.1 Training manuals

Two types of manuals were developed, i.e. field manuals and training guides. The field manuals that were developed were the Enumerator's Manual, Supervisor's Manual and the Fieldwork Coordinator's Manual. The training guides developed were the Trainers' Guide and Trainees' Guide. The field manuals were intended to be used as both training manuals and reference manuals. The Enumerator's Manuals were intended for all field staff and everyone involved in the data collection process, as they contained details on methodologies and procedures for enumeration. The Supervisor's and Fieldwork Coordinator's Manuals focused mainly on quality assurance and the management of field processes since these were their main responsibilities.

4.4.2 Training plan

Training was planned and executed at national, provincial and district levels. The trainees at national level did the training at provincial level, and those that were trained at provincial level did the training at district level. The cascade method of training was at three levels, with the exception of publicity and listing training. Publicity training was done only at national and provincial levels for Publicity Coordinators and Publicity Officers respectively. Likewise, listing training was done at national and provincial levels for supervisors and fieldworkers respectively. Data collection training involved the largest number of trainees. Training was started by subject matter specialists training the trainers at Head Office. The national trainers trained Provincial Survey Coordinators (PSCs), District Survey Coordinators (DSCs), Mapping Monitors (MMs) and GIS Officers at national level. The DSCs trained Fieldwork Coordinators (FWCs) at provincial level with the supervision and monitoring of the PSCs and Head Office monitors. Finally, FWCs trained Fieldwork Supervisors (FWSs) and Enumerators at district level.

During the training of fieldworkers, video training technology was used in addition to the instructor-led training approach. Although video training can never replace the trainer completely, it offered an ideal opportunity to access large numbers of trainees, in different training venues, at different times, with customised training solutions, quickly and cost effectively. After facilitation by the trainer during training sessions, a training video was used to consolidate knowledge learnt and to clarify issues that were not clear to trainees during training.

4.4.3 Evaluation of training

Every training session was evaluated by both the trainees and trainers. Trainers completed a daily evaluation form in order to identify problems that trainees had experienced during that particular day's training. Areas that needed remedial training were revisited the following day.

4.5 Field logistics

All logistics pertaining to the distribution of materials were coordinated from the data processing warehouse in Pretoria. Activities included receiving materials from suppliers, packaging materials according to province or district requirements, dispatching materials to the province or district, and receiving materials from the province or district.

4.5.1 Quality assurance of packaged materials and goods

Using the inventory list for each PO/DO, quality assurors (logistics officers) checked and double-checked against the inventory list items that were assembled for a particular PO/DO. Five teams working in pairs were assigned to carry out this duty.

4.5.2 Reverse logistics

Reverse logistics involved receiving questionnaires and materials from the districts and provinces back to the central warehouse in Pretoria. The reverse logistics for questionnaires were prioritised over other field materials to ensure data processing preparation started as planned.

5. Enumeration

The main objective of enumeration is to collect and document particulars of all individuals and housing units with the selected respondent(s).

5.1 Methodology

The adopted enumeration method for CS 2007 was canvassing, whereby the enumerator conducts a face-to-face interview with the respondent while simultaneously completing the questionnaire. The Community Survey adopted both the de jure and de facto approach in order to compare with other Stats SA social statistics definitions as well as to give a comparison over time between the censuses with the ultimate objective of having two estimates of the population – the de jure population estimates are mostly useful for long-term planning, and the de facto population estimates are mostly used for demographic estimations.

Enumerators visited the selected sampled dwelling units to interview households and ensure that the information required from them was captured on the questionnaires.

Self-enumeration was not allowed. The enumeration was carried out over a three-week period with a non-response follow-up period of one week as planned, that is on 7 February 2007. The mop-up exercise was carried out from 1 to 7 March. This included follow-up on non-contacts, vacant dwellings, and unoccupied dwellings. However, due to the high number of dwelling units that were being mapped for the non- response follow-up period, the contracts of enumerators were extended beyond 28 February to assist the supervisors during that period.

5.2 Field organisation

The data collection approach revolved around the use of a mobile team of four enumerators and a supervisor. The team was assigned a fixed number of EAs to enumerate. The team worked together in each sampled EA and moved to the next one once the targeted EA had been completed. The advantage of this method was that the supervisor was in daily contact with the team, which improved the quality of the data collected during fieldwork.

During enumeration, supervisors (who doubled up as drivers) and their teams of four enumerators each, identified the selected EA. They then dropped off each enumerator at a selected dwelling unit, and ascertained that each enumerator had been accepted to conduct the interview. They picked up the enumerator who had completed the interview and immediately checked the questionnaire for errors, consistency and completeness. Where errors were found, the enumerator was sent back to the household to correct the information that had been recorded. If the supervisors were satisfied, they signed off the questionnaire and stored it in a safe place. Supervisors did the same for all the members of their teams until the EA had been completed. The team then moved to another selected EA.

Based on the number of sampled EAs (17 098), 1 182 teams comprising 1 182 Fieldwork Supervisors and 4 728 enumerators were formed. Each team was expected to enumerate about 16–20 EAs in three weeks, with an additional one week assigned for non-contacts and refusals. The supervisors were supervised by 236 Fieldwork Coordinators (FWCs), resulting in a Fieldwork Coordinator–to-Supervisor ratio of 1:5. Fieldwork Coordinators were supervised by 55 District Survey Coordinators (DSCs), resulting in a DSC–to-FWC ratio of 1:4. The DSCs were supervised by nine Provincial Survey Coordinators (PSCs), resulting in a PSC-to-DSC ratio of approximately 1:6. The PSCs were based in their respective provincial offices. They coordinated data collection for their assigned province.

The 55 DSCs were based in 55 district offices (DOs) that were temporarily created specifically for CS 2007. Each of the 236 FWCs had a temporary local office or fieldwork station that was used as a base during the fieldwork phase of the project, and also for the training of enumerators and supervisors attached to them.

5.3 Non-response procedures

5.3.1 Non-contact and unoccupied dwellings

Non-contact describes the situation where an enumerator failed to make contact with a household at an address for the survey period. An unoccupied dwelling is a dwelling whose inhabitants were absent at the time of the visit(s). During enumeration, supervisors and their teams handled non-contacts as follows:

- They made at least three attempts to contact a household at an address or dwelling.
- If no one was at home, they revisited the dwelling at least three times at different times of day, on different days; that is, they came in the morning, and tried again in the afternoon or in the evening. This could be on the same day or another day, but they tried at a time when someone was more likely to be in. They also asked from the neighbours when someone was likely to be at home again.
- At both the first and second unsuccessful visits, the enumerators left the non-contact form to say when they will be coming back. The form also contained the contact details of the enumerator.
- In all cases, non-responses resulted in the work being checked and verified by the supervisor, before the non-contact result code could be recorded.
- After three unsuccessful visits, the supervisor reported the matter to the Fieldwork Coordinator for further action.

5.3.2 Refusals, unusable information and partly completed questionnaires

At a number of DUs that were sampled, respondents refused to take part in the survey. The policy of Stats SA is to do everything possible to persuade people to cooperate and this was emphasised to Community Survey personnel. Enumerators had to try to convince respondents of the importance of the CS, and give them the introduction letter. However, if these efforts failed, a refusal form was completed detailing the refusal and reported to the supervisor. The supervisor would then visit the household and try to persuade the respondent to cooperate. Thereafter, it was reported to the FWC who was authorised to inform respondents about penalties for refusing to cooperate.

In some cases, the questionnaire was only partially completed – the respondent refused to answer some sections of the questionnaire, or became agitated or irritated during the interview process and refused to complete the questionnaire. The code pertaining to the category 'partly completed' was selected and was only considered after the supervisor confirmed that there was no expectation of completion. In some instances the respondent did not make any sense when answering the questionnaire. This may be in the case of mentally disabled people, or perhaps those under the influence of alcohol. If the enumerator was able to ascertain that the information received was unlikely to be true, they classified the questionnaire as 'unusable information'.

5.4 Quality assurance

The FWS and FWC conducted 100% quality checks for accuracy and completeness on all completed questionnaires. In addition, the DSCs, PSCs and Monitors also did quality checks on randomly selected questionnaires and DUs and addressed problematic questions as they came up. In addition, the FWC did 2% spot checks of selected dwelling units within their assigned fieldwork coordination unit to minimise bogus enumeration.

Training played a big role in ensuring good quality data from the field. At district level, retraining was done in areas where fieldwork monitors felt that the work was not of the expected quality. A close watch was also kept on individual enumerators, and Fieldwork Supervisors and Fieldwork Coordinators who had problems performing according to expectations were retrained where necessary. Their work was also checked more frequently.

5.4.1 Handling of completed questionnaires and boxes

FWS were required to package questionnaires in their EA boxes and hand them over to the FWCs soon after the completion of the EA. The FWCs were required to sign for the receipt of the boxes after verifying the contents of the boxes. They were in turn required to hand over the completed boxes to the DLOs for reverse logistics. DLOs were also required to sign for the receipt of the boxes after verifying the contents of the boxes. The boxes were then stored in designed storage areas awaiting shipping back to the data processing centre in Pretoria.

5.4.2 Progress reporting

Progress reporting for data collection was done on a daily basis. Provinces were provided with procedures and timelines for progress reporting and were able to report progress on a daily basis though at the initial stages, there were problems as outlined below.

5.5 Census and Surveys Administration System

The core objective was to design, develop, and implement a Census and Surveys Administration System (CSAS) that will assist in the management and monitoring of the overall CS 2007 project from start to completion. The core preparatory work was done based on the previous CAS (Census Administration System) that was utilised in Census 2001. The CSAS module had the following core modules:

- Recruitment modules
- Logistics management
- Item tracking
- Vehicle monitoring

All regional offices were connected with Head Office through to the Stats SA Virtual Private Network (VPN).

CSAS was successfully used for the capturing and profiling of over 80 000 candidates in their respective work areas for proper selection of fieldworkers. Furthermore, detailed questionnaire tracking, with key linkages between questionnaires, enumerators and EAs was accomplished for the ±300 000 questionnaires and ±20 000 boxes that were in circulation.

6. Data processing

Data processing refers to a class of programmes that organise and manipulate usually large amounts of numeric data. Data processing involved the processing of completed questionnaires. Information received from questionnaires collected during fieldwork was converted into data represented by numbers or characters. The two methods used for this conversion were manual capturing (key-entry) and scanning. The scanning method was used as the main process and the key-entry application was used for questionnaires that were damaged and not scannable.

6.1 High-level processes

In general, the high-level processes covered the following activities:
Boxes were received and questionnaires were checked to ensure that–

1) they belonged to the box; and
2) were not damaged.

Data were then captured and converted into electronic format through scanning or Key-from-Paper (KFP). Thereafter, an account of all sampled dwelling units was prepared and data were balanced to verify whether the data collected for each household contained the four sections – General, Persons, Mortality, and Household. Data were then checked for consistency and prepared for final output based on the tabulation plan.

6.2 Data capturing

Two methods were used for capturing the data, namely scanning and manual capturing (key-entry).

6.2.1 Scanning

The scanning process proceeded as follows:

The data processor scanned the box number, and then entered the estimated number of pages in each batch. At this stage, the batches were ready to be scanned. One box at a time was given to each of the six Scanning Operators to avoid scanning the questionnaires twice. The batches were then taken out of the box and placed next to the tray on the scanner. The box number was then scanned using the small hand-held scanner and the number of pages per batch was entered into the Input Station.

A visual check was performed on the scanning to ensure that the images were clear of any noise and that the data were clear and readable. The barcode as well as the actual data on the questionnaire was checked. In the case where the image was either too light or too dark, parameters were adjusted and the batch was rescanned.

Validations were automatically executed to confirm scanning parameters and image quality.

Questionnaires that could not be scanned were de-activated from their boxes and assigned to a new box. Images were transferred to the server and their barcodes were tracked. These questionnaires were then sent to Key-from-Paper.

6.2.2 Manual capturing (Key-from-Paper)

Key-from-Paper (KFP) is an application for manual data capturing. The application was developed to capture questionnaires that were not suitable for scanning. Such questionnaires included those which were torn or where pencil entries were not bold enough for the interpretation of the scanner, or those that were in a bad condition. Duplicate application was created for quality assurance purposes. The same questionnaires that were captured in application one, were also captured on application two. Each questionnaire captured in both applications, was compared to one another using corresponding fields. Validation checks were not implemented in the applications.

The application was used by data processors to capture information as was reflected on the questionnaires. EA and DU numbers were placed into the look-up table to validate the sampled frame. In cases where an EA or DU was found to be invalid, the EA Summary Book was then used for corrections.

6.2.3 Coding of open-ended questions

Coding is the process of assigning numerical values to responses to facilitate data capturing and processing in general. The code lists for occupation and industry were based on the International Standard Classifications done to the five-digit level. The variables covered were occupation, industry, and place names.

6.2.4 Quality assurance (QA)

Quality Assurance (QA) was applied at each process on a sample basis. At the end of the capturing process, via scanning recognition, the captured questionnaires were grouped into sets of 100 questionnaires called QA groups; from which a 10% sample was selected. Subsequently, a second level sample of 10% of fields were selected and presented to the data capturers who keyed information as it appeared on the image.

The quality of capturing was determined by accepting all questionnaires in the group where the scanned information agreed with the keyed information above a 95% threshold. Any group below the threshold was sent for verification or recapturing to determine the cause of failure.

6.3 Post-capture processes

6.3.1 Automated and manual editing

The automated cleaning was implemented based on editing rules specification defined with reference to the approved questionnaire. Most of the editing rules were

categorised into structural edits looking into the relationship between different record type, the minimum processability rules that removed false positive reading or noise, the logical editing that determine the inconsistency between fields of the same statistical unit, and the inferential editing that search similarities across the domain. The edit specifications document for the structural, population, mortality and housing edits was developed by a team of Stats SA subject-matter specialists, demographers, and programmers. The process was successfully carried out during the months of July/August 2007.

6.4 Product development

The Community Survey 2007 gathered information on individual and household characteristics in sampled enumeration areas. Therefore, all products developed will focus mainly on the data items contained in the questionnaire and will be provided in different formats as required by users. The products earmarked are the Basic Results pamphlet, Key Municipal Data, Province at a Glance, and Methodology, Processes and Highlights of Key Results.

The Community Survey data will be released in two stages. The first release is planned for 24 October 2007 and will comprise the following:

- Methodology, processes and highlights of key results
- Basic results pamphlet
- Statistical Release

The second set of releases is scheduled for release after November 2007 and will comprise the following:

- Key Municipal Data
- Interactive Internet web products
- Community profile database
- Thematic reports Atlas
- Unit records on CD for analysts

7. Challenges

'A good plan violently executed today is better than a perfect plan tomorrow' (General George S. Patton[1]), and 'anything that is measured and monitored improves'[2].

The qualities embedded in the CS project show a project that had plans, monitoring and quality assurance systems in place. However, a project of this magnitude cannot run smoothly without newly emerging challenges – even after carefully taking into account lessons learnt from previous surveys. The challenges that were observed during the CS are summarised as follows:

- The CS project established a good rapport with the communities by reaching households and community leaders to expedite listing and data collection operations. The publicity campaign worked well to minimise refusal rate. However, it did not work well in areas where there was political unrest. In future, publicity needs to work around improving consultations and relations with hard-to-count areas like farms.
- Data processing included applying weights according to the sample design to estimate the total population and number of households. As a result, this process involved a lot of work, including evaluation of data and interrogation of weights by using other parameters. This process was time consuming.

The lessons learnt from the Community Survey and the above challenges will greatly assist the organisation in planning and implementing the activities scheduled to be carried out to have a successful census in 2011.

[1] Source: WorldofQuotes.com – Historic Quotes and Proverbs Archive: http://www.worldofquotes.com/topic/Military/1/index.html

[2] Source: Robert's rules of survival @ American Digest

8. Conclusion

Statistics South Africa is pleased to continue informing social debates and decision-making processes amongst other things, by providing relevant, accurate, reliable and timeous data as mandated through the Statistics Act.

For the first time as an organisation, we have undertaken a large-scale household-based survey with the main objective being to deliver data at a geographic level lower (below the provinces) than any other existing household survey. The data will be used for a variety of purposes in the policy-making arena, population projections and the generation of some indicators for the Millennium Development Goals (MDGs).

The method of data collection involving a team of supervisors checking and verifying questionnaires on the spot assures Stats SA of high-quality data from the respondents. Stats SA would like to express its appreciation for the overwhelming support from all the households that spared their time to supply the CS with this valuable information.

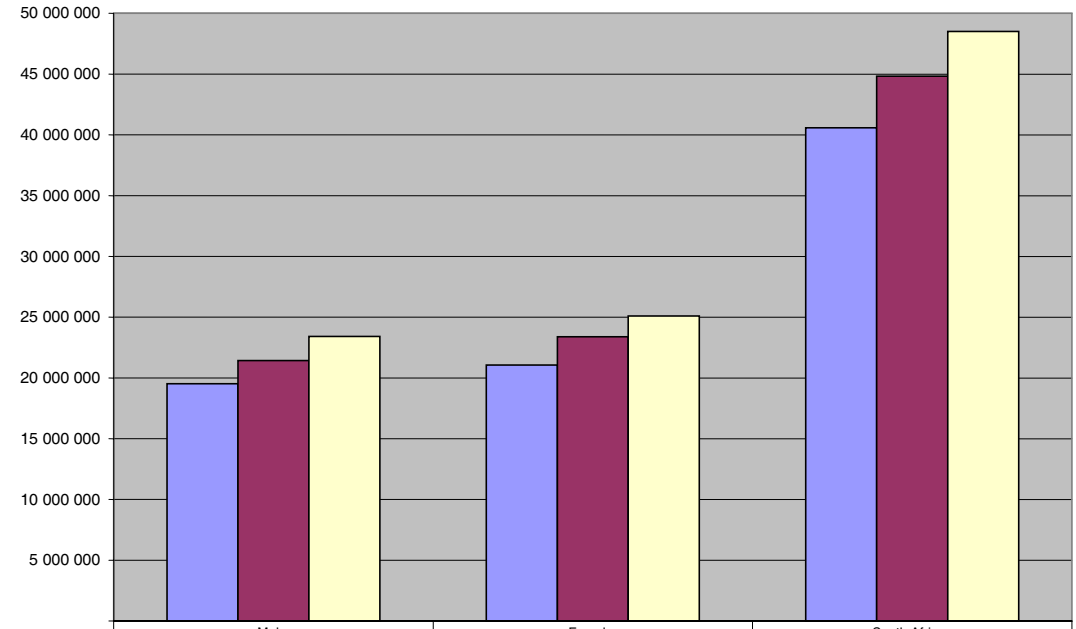
This report affirms our continued efforts and commitment to provide users with data that meet their requirements through the application of internationally acclaimed practices.

From the assessment of project objectives and the evaluation processes undertaken to date, it is clear that the project is a success. However, there are still challenges that sample surveys will always carry with them. We will continue interrogating this rich dataset with the aim of improving our methods and processes.

Highlights of key results

Key figures

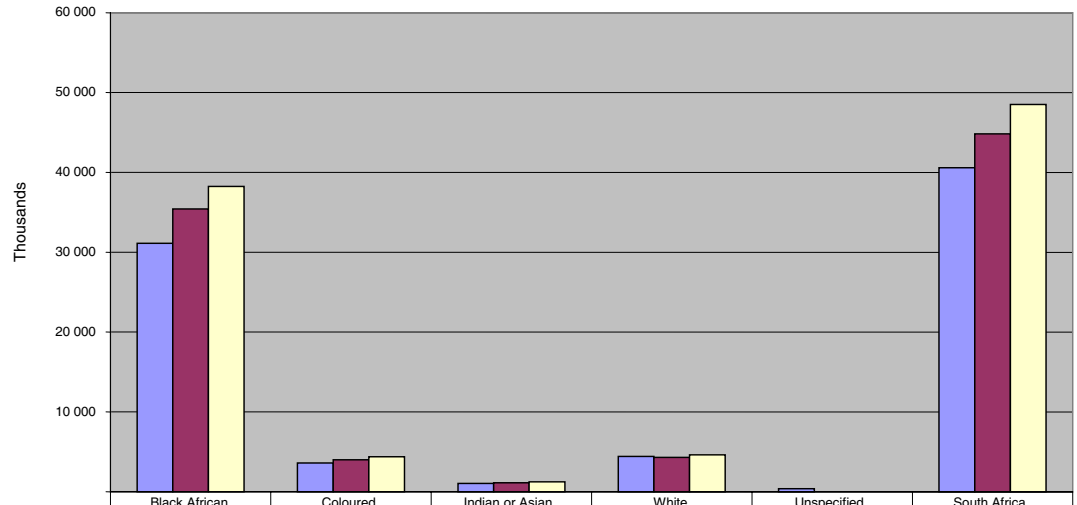
Figure 1: Population of South Africa by sex



	Male	Female	South Africa
Census 1996	19 520 887	21 062 685	40 583 573
Census 2001	21 434 040	23 385 737	44 819 778
CS 2007	23 412 064	25 089 999	48 502 063

The census results showed that the population of South Africa increased from 40,5 million in 1996 to 44,8 million in 2001. The Community Survey has returned an estimated population of 48,5 million, showing an overall increase of 8,2% since 2001

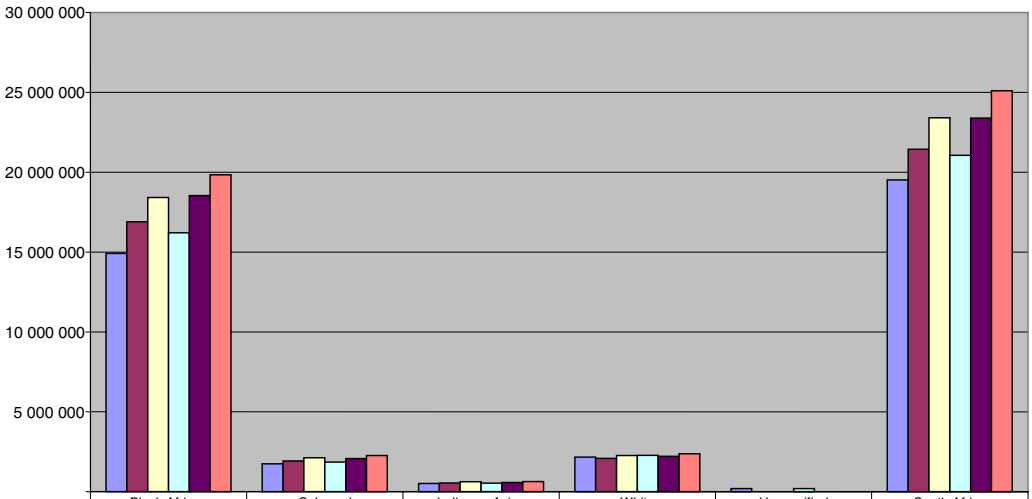
Figure 2: Population distribution by population group



	Black African	Coloured	Indian or Asian	White	Unspecified	South Africa
Census 1996	31 127 631	3 600 446	1 045 596	4 434 697	375 204	40 583 573
Census 2001	35 416 166	3 994 505	1 115 467	4 293 640	-	44 819 778
CS 2007	38 255 165	4 375 527	1 244 634	4 626 738	-	48 502 063

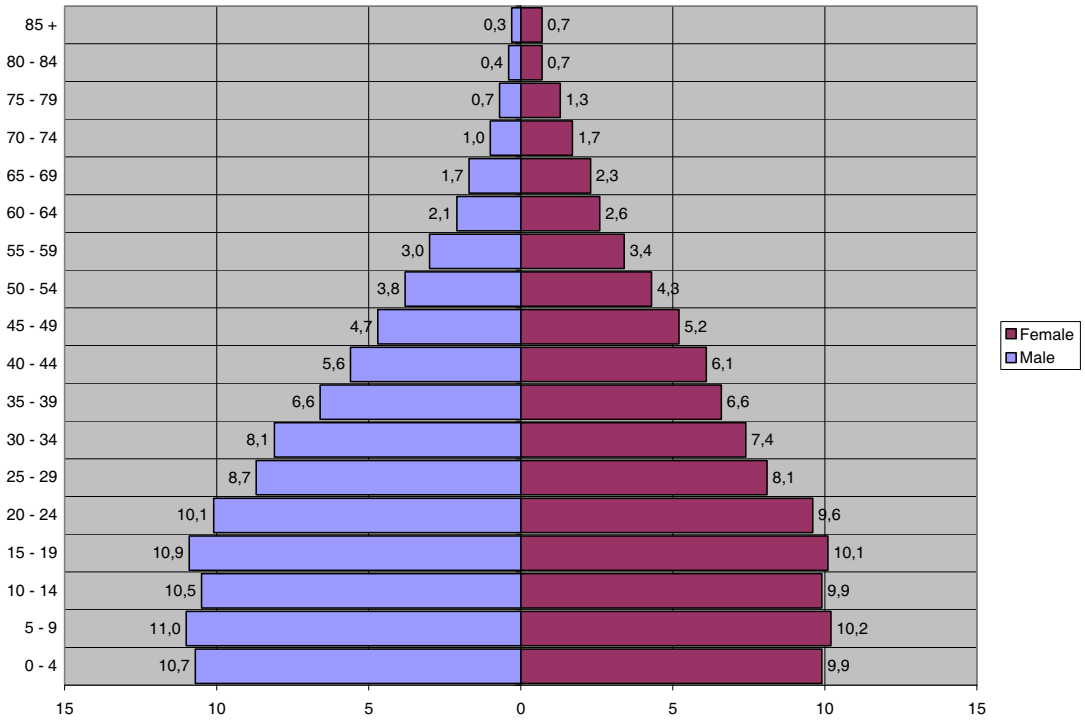
Overall, more than three quarters of the population is black African and it is increasing. It was 76,7 % in 1996 and stayed at 79% in 2001 and CS 2007. The percentage of the Coloured population has remained constant at 8,9%. The percentage of the Indian or Asian population has gone down from 2,6% in 1996 to 2,5% in 2001 and CS 2007. The percentage of the white population, on the other hand, has declined slightly from 10,9% in 1996 to 9,6% in 2001 and CS 2007.

Figure 3: Population distribution by population group and sex



	Black African	Coloured	Indian or Asian	White	Unspecified	South Africa
Male Census 96	14 916 712	1 744 920	512 231	2 162 699	184 326	19 520 887
Male Census 01	16 887 830	1 920 426	545 050	2 080 734	0	21 434 040
Male CS 07	18 417 433	2 117 608	615 970	2 261 053	0	23 412 064
Female Census 96	16 210 919	1 855 526	533 365	2 271 998	190 878	21 062 685
Female Census 01	18 528 336	2 074 079	570 417	2 212 905	0	23 385 737
Female CS 07	19 837 332	2 257 919	628 663	2 365 684	0	25 089 998

Figure 4: Distribution of total population by five-year age groups and sex



The population pyramids for the years 1996, 2001 and CS 2007 show smaller percentage of populations in the age-group 0–4 years. Data also show that population in the age group 0–9 has been underestimated in previous censuses and picked up in subsequent censuses when the cohort is older.

Figure 5: Percentage distribution of persons aged 20 years and older by level of education

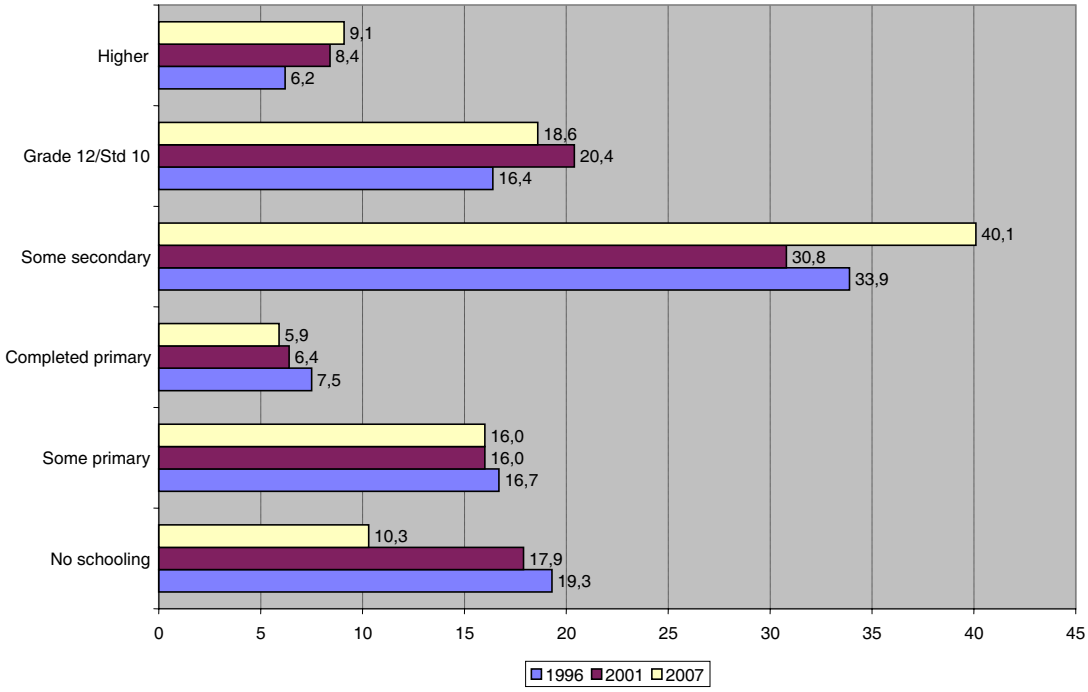
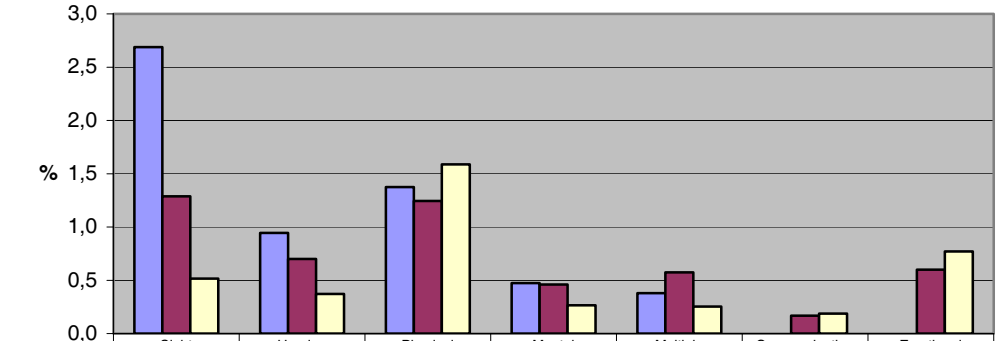


Figure 5 shows that in 2007, a large percentage of the population aged 20 years and older (40,1%), completed some secondary school. An overwhelming 10,3% of the population in this age group did not have any formal schooling. However, this percentage decreased from 2001 and 1996, where no schooling accounted for 17,9% and 19,1% respectively. The percentage of the population who completed a higher education (includes certificates, diplomas above Grade 12, degrees and postgraduate qualifications) increased from 6,2% in 1996 to 8,4% in 2001 and 9,1% in 2007, showing a small but vital gain in educational attainment. Primary school educational attainment showed no significant changes between the years.

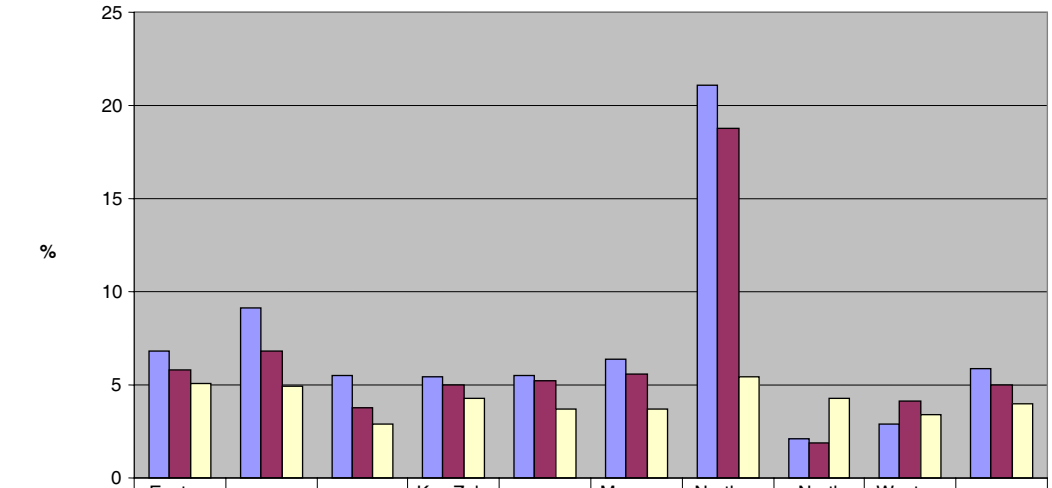
Figure 6: Percentage of population with disability by type of disability



	Sight	Hearing	Physical	Mental	Multiple	Communication	Emotional
Census 1996	2,7	0,9	1,4	0,5	0,4	0	0
Census 2001	1,3	0,7	1,2	0,5	0,6	0,2	0,6
CS 2007	0,5	0,4	1,6	0,3	0,3	0,2	0,8

The variables communication and emotional disability were not asked during Census 1996. Therefore comparisons have been made for persons who were affected by various disabilities that cut across the three years 1996, 2001 and 2007. Sight, hearing, physical, mental and multiple disabilities were compared. The prevalence of sight, hearing, and mental disability show a downward trend through the three years 1996, 2001 and 2007. Physical disability on the contrary, shows a significant increase of 0,4% from 2001 to year 2007. This could be a reflection of high accidents on the South African roads due to the high volumes of vehicles that are currently experienced in the country.

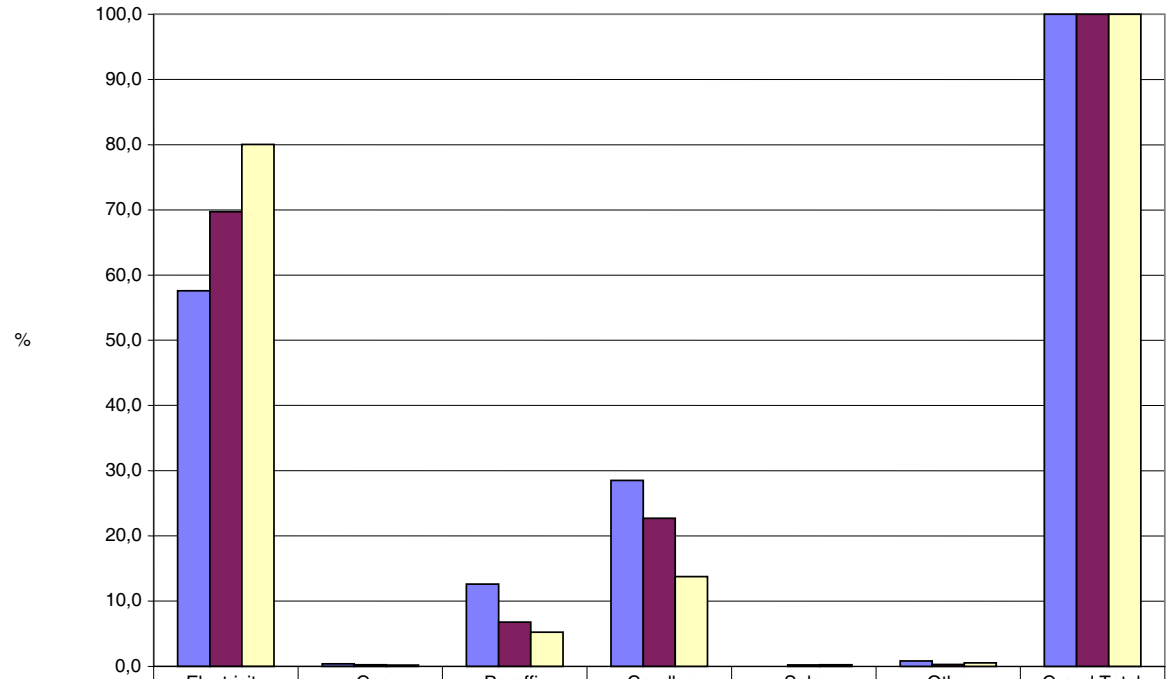
Figure 7: Percentage of population with disability by province



	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpuma-langa	Northern Cape	North West	Western Cape	RSA
Census 1996	6,8	9,1	5,5	5,4	5,5	6,4	21,1	2,1	2,9	5,9
Census 2001	5,8	6,8	3,8	5,0	5,2	5,6	18,8	1,9	4,1	5,0
CS 2007	5,1	4,9	2,9	4,3	3,7	3,7	5,4	4,3	3,4	4,0

The table above shows that in the three years 1996, 2001 and 2007, Northern Cape had the highest prevalence (21,1%, 18,8% and 5,4% respectively) compared to all other provinces. During 1996 and 2001, North West province recorded the lowest prevalence (2,1% and 1,9%) of disability and made a significant increase (4,3%) during 2007. Gauteng province had the lowest prevalence (2,9%) of disability in 2007.

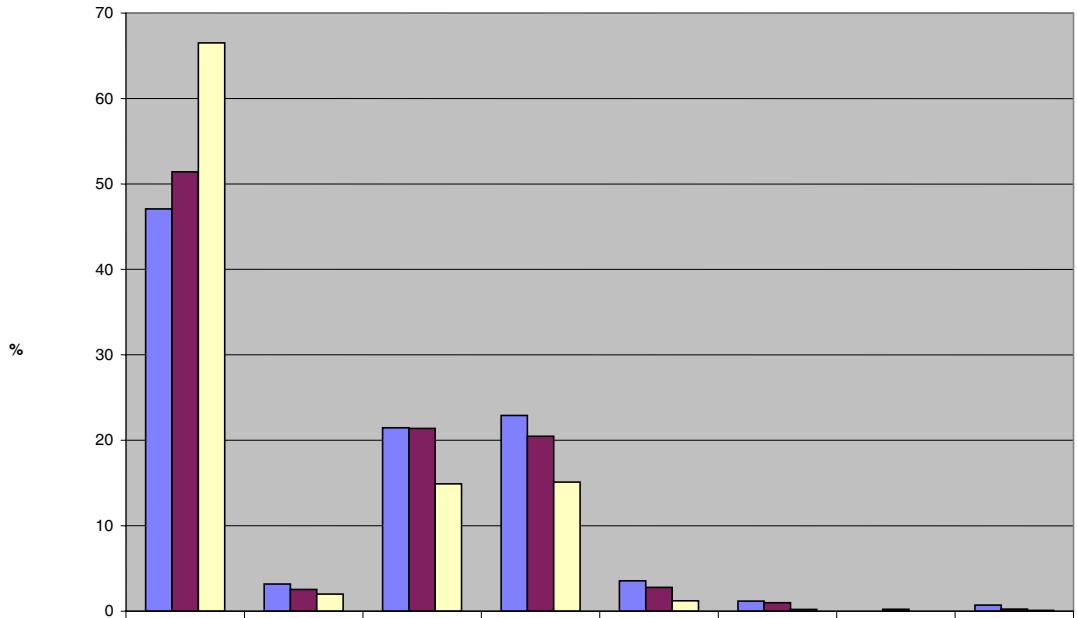
Figure 8: Percentage distribution of households by type of energy used for lighting



	Electricity	Gas	Paraffin	Candles	Solar	Other	Grand Total
Census 1996	57,6	0,4	12,6	28,5	0,0	0,8	100,0
Census 2001	69,7	0,2	6,8	22,7	0,2	0,3	100,0
CS 2007	80,0	0,2	5,3	13,8	0,2	0,5	100,0

Figure 8 shows that in 1996, 2001 and 2007, electricity was the main energy source for lighting. The use of electricity by households as the main source of lighting increased significantly in 2007. There was a corresponding decrease of households using paraffin and candles for lighting over the period (paraffin: 12,6% in 1996, to 6,8% in 2001 to 5,3% in 2007, and candles: 28,5% in 1996, to 22,7% in 2001 to 13,8% in 2007).

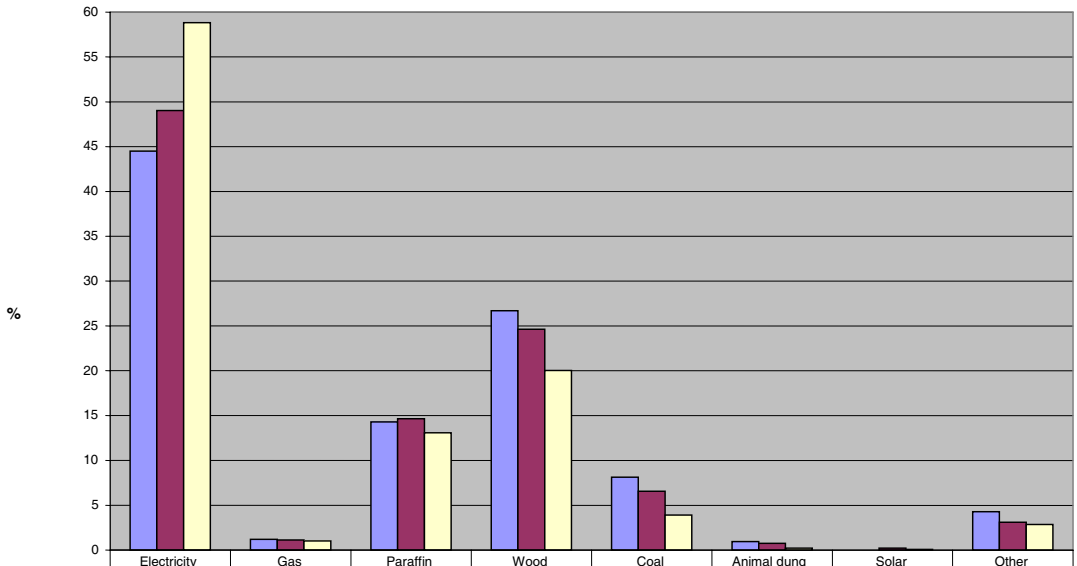
Figure 9: Percentage distribution of households by type of energy used for cooking



	Electricity	Gas	Paraffin	Wood	Coal	Animal dung	Solar	Other
Census 1996	47,1	3,2	21,5	22,9	3,5	1,2	0,0	0,7
Census 2001	51,4	2,5	21,4	20,5	2,8	1,0	0,2	0,2
CS 2007	66,5	2,0	14,9	15,1	1,2	0,2	0,0	0,1

Figure 9 shows that most of the households in the country used electricity for cooking. The use of electricity increased from 47,1% in 1996, to 51,4% in 2001, to 66,5% in 2007. Wood and paraffin followed with 15,1% and 14,9% of households using these for cooking. The use of gas, paraffin, wood, coal and animal dung as the source of energy for cooking has been declining from 1996. The graph clearly shows that households are moving away from using other sources of energy for cooking, and are using electricity to a larger extent.

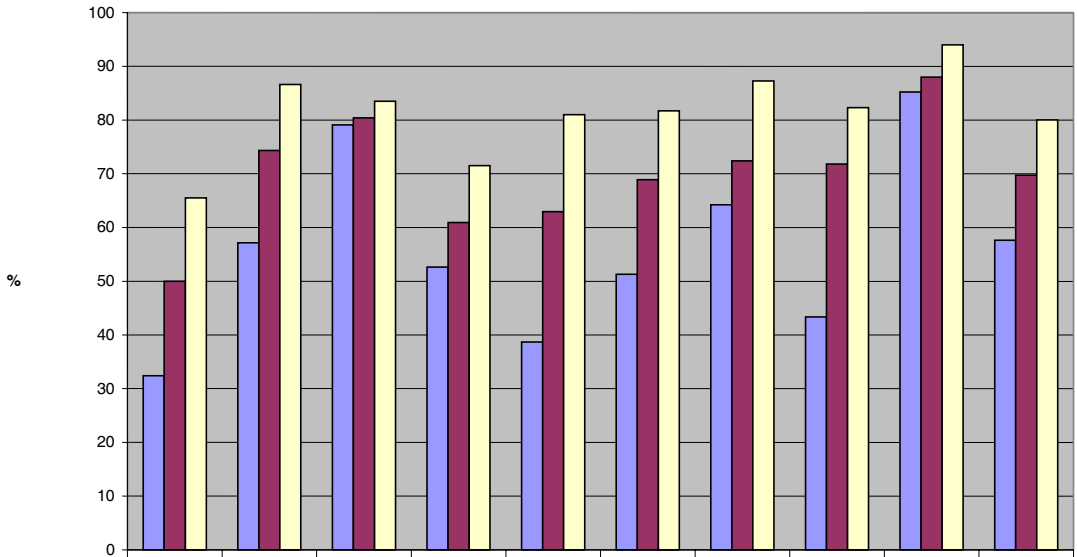
Figure 10: Percentage distribution of households by type of energy used for heating



	Electricity	Gas	Paraffin	Wood	Coal	Animal dung	Solar	Other
Census 1996	44,5	1,2	14,3	26,7	8,1	0,9	0,0	4,3
Census 2001	49,0	1,1	14,6	24,6	6,6	0,7	0,2	3,1
CS 2007	58,8	1,0	13,1	20,0	3,9	0,2	0,1	2,9

Figure 10 shows that more than half of the households in the country used electricity for heating. The use of electricity by households as the main energy source for heating increased from 44,5% in 1996, to 49,0% in 2001, to 58,8% in 2007. All the other sources of energy for heating with exception of solar has been decreasing from 1996 to 2007. The graph clearly shows that households were moving away from using other sources of energy for heating, and relying more on electricity.

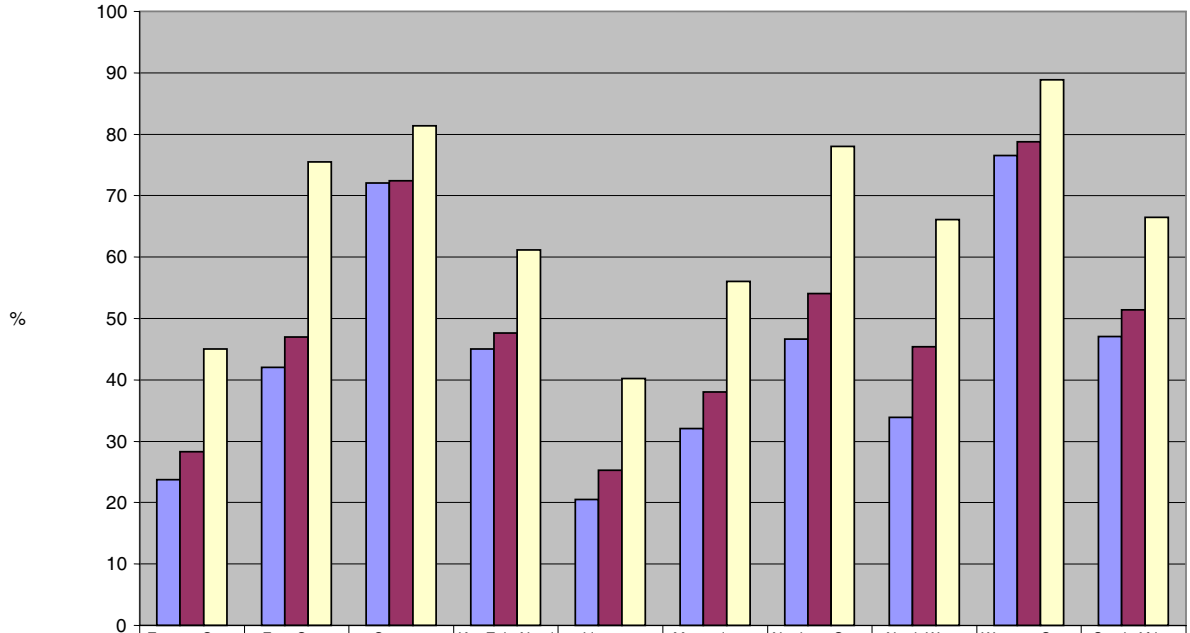
Figure 11: Percentage of households using electricity for lighting by province



	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	South Africa
Census 1996	32,4	57,1	79,1	52,6	38,7	51,3	64,2	43,3	85,2	57,6
Census 2001	50,0	74,4	80,4	60,9	62,9	68,9	72,4	71,8	88,0	69,7
CS 2007	65,5	86,6	83,5	71,5	81,0	81,7	87,3	82,3	94,0	80,0

Figure 11 shows that electricity used for lighting has increased in all the provinces, with 80,0% of households in South Africa using electricity for lighting. Western Cape was the province with the highest proportion (94,0%) of households using electricity for lighting, followed by Northern Cape with 87,3% and Free State with 86,6%. Eastern Cape was still the province with the lowest proportion (65,5%) using electricity for lighting in all the years (1996, 2001 and 2007). The proportion of households in Eastern Cape and KwaZulu-Natal using electricity for lighting is below the national average.

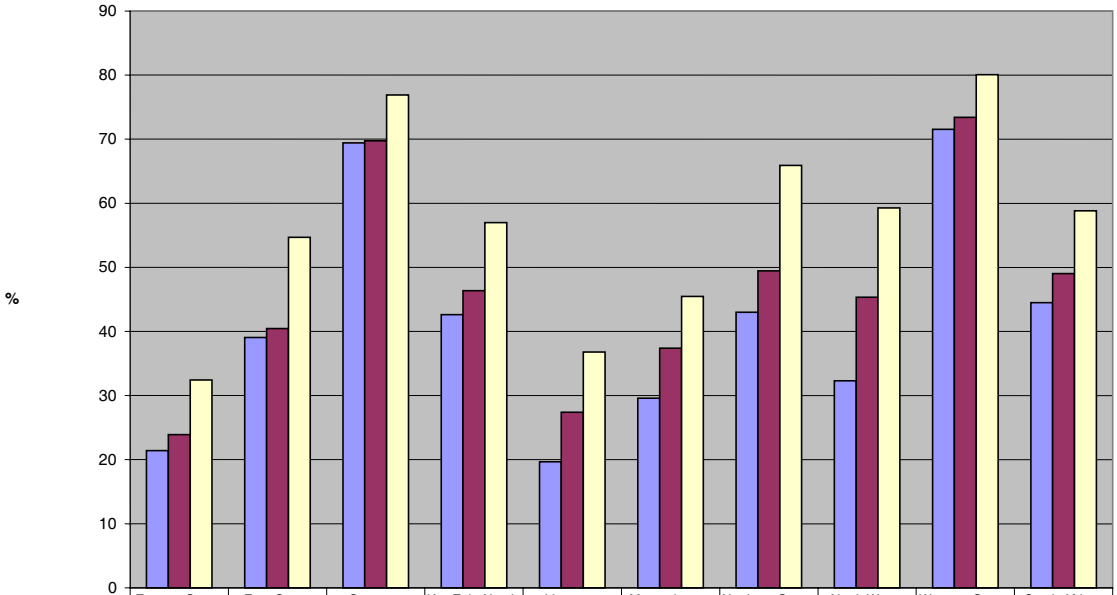
Figure 12: Percentage of households using electricity for cooking by province



	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	South Africa
Census 1996	23,7	42,0	72,1	45,0	20,5	32,1	46,7	33,9	76,5	47,1
Census 2001	28,3	47,0	72,4	47,6	25,3	38,0	54,1	45,4	78,8	51,4
CS 2007	45,0	75,5	81,4	61,1	40,2	56,0	78,0	66,1	88,9	66,5

Figure 12 shows households using electricity for cooking. It shows that the use of electricity for cooking by households increased from 47,1% in 1996 to 51,4% in 2001 to 66,5% in 2007. This increase in the proportion of households using electricity for cooking applied to all the provinces. Western Cape was still the province with the highest percentage of households (88,9%) using electricity for cooking. Northern Cape and Free State followed with 78,0% and 75,5% of households respectively using electricity for cooking. Limpopo had the lowest percentage (40,2%) of households using electricity for cooking.

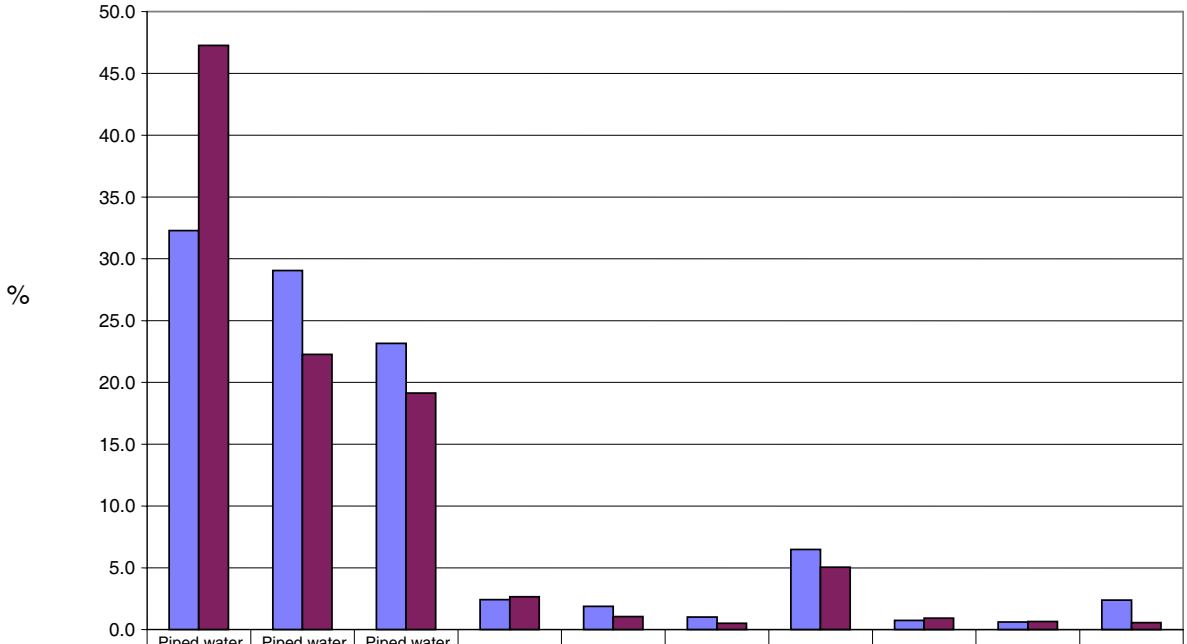
Figure 13: Percentage of households using electricity for heating by province



	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	South Africa
Census 1996	21,4	39,0	69,4	42,6	19,7	29,6	43,0	32,3	71,5	44,5
Census 2001	23,9	40,5	69,8	46,4	27,4	37,4	49,4	45,3	73,4	49,0
CS 2007	32,4	54,7	76,9	57,0	36,8	45,5	65,9	59,3	80,1	58,8

Figure 13 shows that 58,8% of households in the country were currently using electricity for heating. This was an increase since 1996. This increase in the proportion of households using electricity for heating applied to all the provinces. Western Cape was still the province with the highest percentage of households (80,1%) using electricity for heating, followed by Northern Cape and Gauteng with 76,9% and 65,9% respectively. Eastern Cape had the lowest proportion (32,4%) of households using electricity as a source of energy for heating.

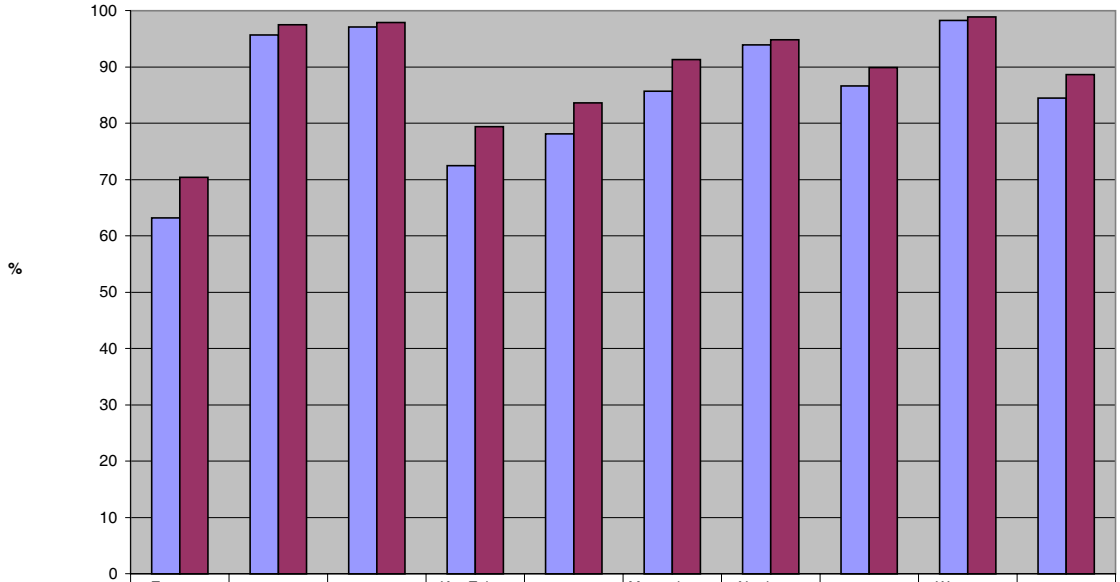
Figure 14: Percentage distribution of households by type of source of water



	Piped water inside the dwelling	Piped water inside the yard	Piped water from access point outside	Borehole	Spring	Dam/pool	River/stream	Water vendor	Rain water tank	Other
Census 2001	32.3	29.0	23.2	2.4	1.9	1.0	6.5	0.7	0.6	2.4
CS 2007	47.3	22.2	19.1	2.6	1.0	0.5	5.1	0.9	0.6	0.6

Figure 14 shows an increase in the proportion of households obtaining water from piped water inside the dwelling from 32,3% in 2001 to 47,3% in 2007. It also shows a corresponding decrease of the proportion of households obtaining water from piped water inside the yard from 29,0% in 2001 to 22,2% in 2007, and obtaining piped water from an access point outside the yard from 23,2% in 2001 to 19,1% in 2007. The proportion of households obtaining water from a dam, river/stream or spring declined, whereas water obtained from a borehole or water vendor increased.

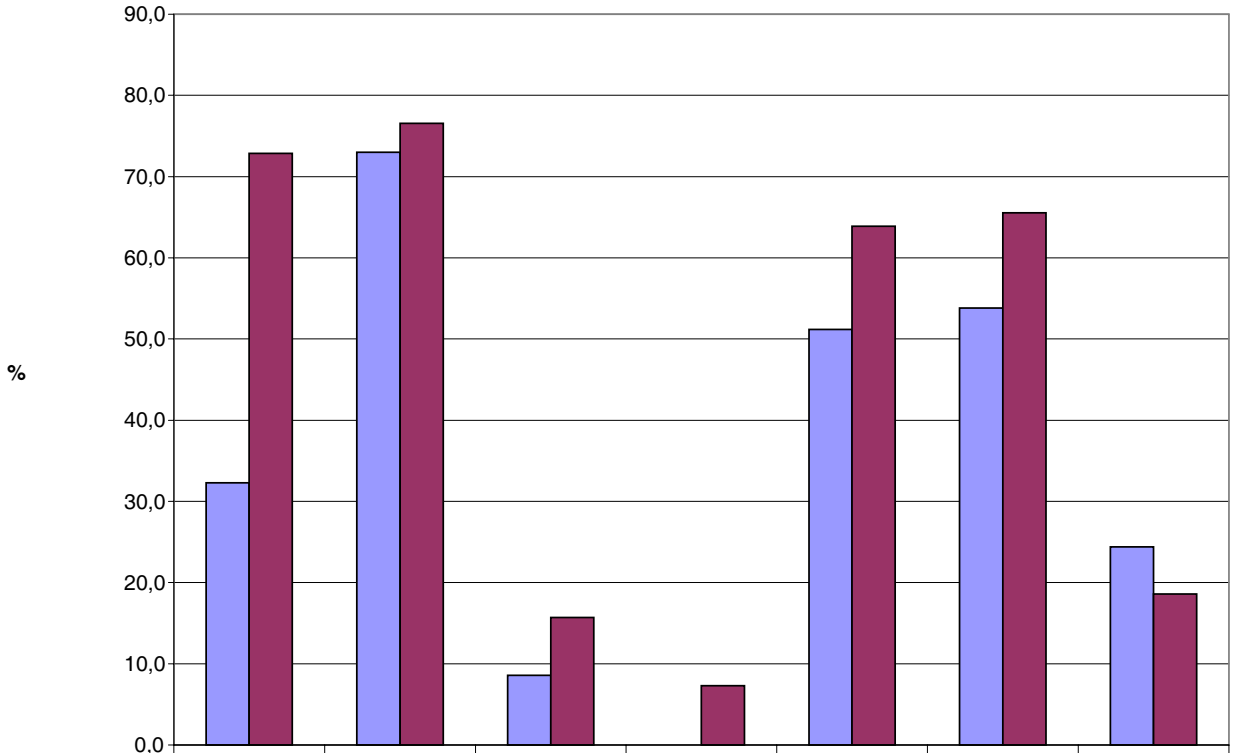
Figure 15: Percentage of households that have access to piped water by province



	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	South Africa
Census 2001	63,2	95,7	97,1	72,5	78,1	85,7	93,9	86,6	98,3	84,5
CS 2007	70,4	97,5	97,9	79,4	83,6	91,3	94,8	89,9	98,9	88,6

Figure 15 shows that the percentage of households with access to piped water has increased from 84,5% in 2001 to 88,6% in 2007. In all the nine provinces there was an increase in the percentage of households that had access to piped water. Western Cape was the province with the highest proportion (98,9%) of households with access to piped water, followed by Gauteng and Free State with 97,9% and 97,5% respectively. The percentage of households that had access to piped water in Eastern Cape, KwaZulu-Natal and Limpopo is below the national average.

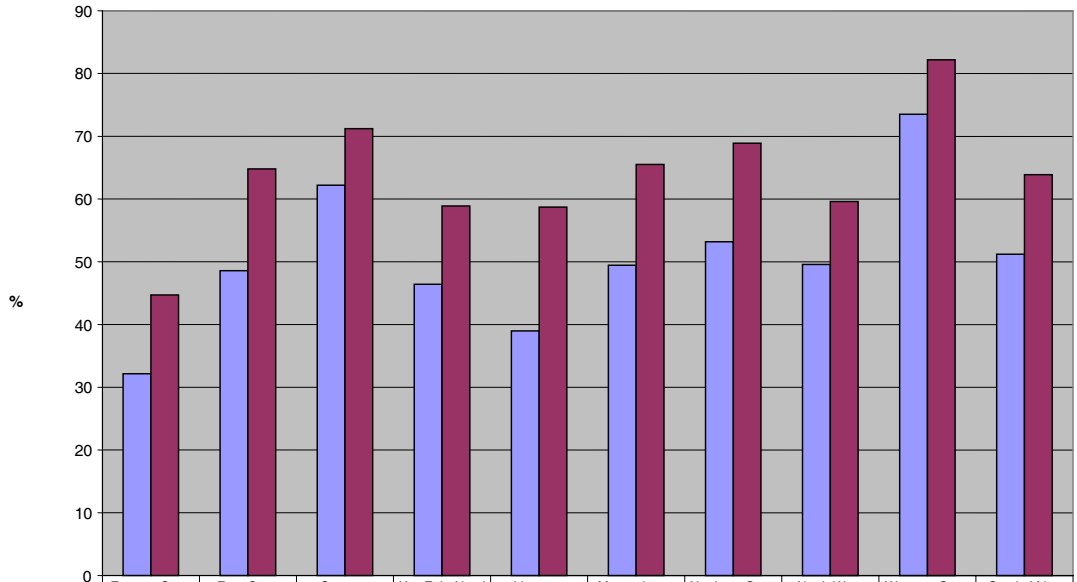
Figure 16: Percentage of households with household goods in working order



	Cellphone	Radio	Computer	Internet facilities	Fridge	Television	Landline telephone
Census 2001	32,3	73,0	8,6		51,2	53,8	24,4
CS 2007	72,9	76,6	15,7	7,3	63,9	65,6	18,6

Figure 16 shows the percentage of households with household goods in working order. Ownership of a radio, television, computer, refrigerator and cellphone has increased considerably between 2001 and 2007. However, the demand for landline telephones has decreased owing to a rise in the popularity of cellphones. Census 2001 did not ask a question pertaining to Internet facilities.

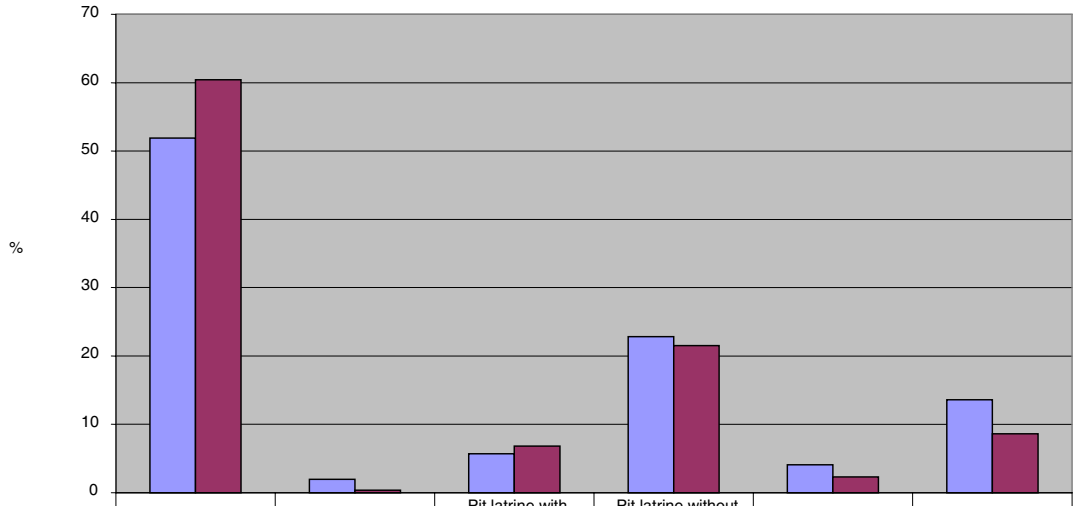
Figure 17: Percentage of households that have a refrigerator by province



	Eastern Cape	Free State	Gauteng	KwaZulu-Natal	Limpopo	Mpumalanga	Northern Cape	North West	Western Cape	South Africa
Census 2001	32,1	48,6	62,2	46,4	39,0	49,4	53,2	49,6	73,5	51,2
CS 2007	44,7	64,8	71,2	58,9	58,7	65,5	68,9	59,6	82,2	63,9

Figure 17 shows that the percentage of households that have a refrigerator has increased in all provinces, from 51,2% in 2001 to 63,9% in 2007. Western Cape was the province with the highest percentage of households having a refrigerator with 73,5% in 2001 and increasing to 82,2% in 2007. Eastern Cape had the lowest percentage of households that had a refrigerator with 44,7% in 2007.

Figure 18: Percentage distribution of households by type of toilet facility

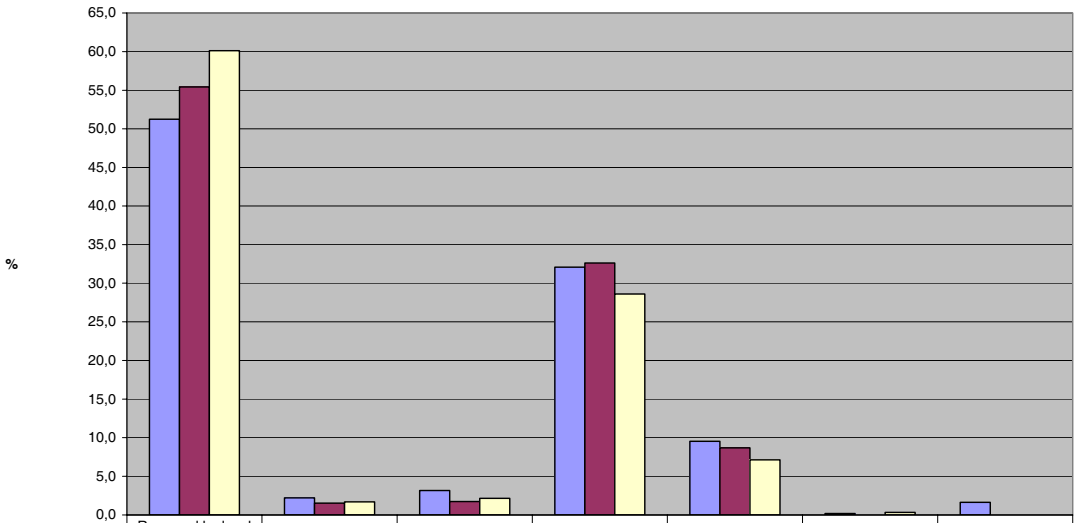


	Flush Toilet	Chemical toilet	Pit latrine with ventilation (VIP)	Pit latrine without ventilation	Bucket toilet system	None
Census 2001	51,9	1,9	5,7	22,8	4,1	13,6
CS 2007	60,4	0,4	6,8	21,5	2,3	8,6

Note: Data for a dry toilet facility have been excluded.

Figure 18 indicates that the percentage of households with access to a flush toilet increased from 51,9% in 2001 to 60,4% in 2007, while the use of pit latrines (without ventilation) decreased from 22,8% in 2001 to 21,5% in 2007. The proportion of households without access to any toilet facility also decreased from 13,6% in 2001 to 8,6% in 2007.

Figure 19: Percentage of households by type of refuse disposal



	Removed by local authority at least once a week	Removed by local authority less often	Communal refuse dump	Own refuse dump	No rubbish disposal	Other	Unspecified
Census 1996	51,2	2,2	3,2	32,1	9,5	0,2	1,6
Census 2001	55,4	1,5	1,7	32,6	8,7	0,0	0,0
CS 2007	60,1	1,7	2,1	28,6	7,1	0,3	0,0

Figure 19 shows that the percentage of households of which the refuse was removed by local authority at least once a week increased from 51,2% in 1996 to 55,4 % in 2001, and to 60,1% in 2007. The graph also shows that the percentage of households that relied on their own refuse dump has declined by 4,0% in 2007.

Annexure A – Key Concepts and Definitions

Main terms used in this report are:

- **Data editing**

The process of detecting and correcting errors (logical inconsistencies) in the data.

- **Data imputation**

The substitution of estimated values for missing or inconsistent data items (fields).

- **De facto**

A census in which people are enumerated according to where they were on census night.

- **De jure**

A census in which people are enumerated according to where they usually live (where a person has his/her goods, pays his/her rates and taxes).

- **Dwelling unit**

A unit of accommodation for a household, which may consist of one structure, more than one structure, or part of a structure. (Examples of each are a house, a group of rondavels, and a flat.) It may be vacant, or occupied by one or more than one household. A dwelling unit usually has a separate entrance from outside or from a common space, as in a block of flats.

- **EA Summary Book**

The Enumerator Area Summary Book (formerly known as the 09 Book) is a register of mapping and listing information pertaining to a particular EA.

- **Enumeration area**

The smallest geographical unit (piece of land) into which the country is divided for census or survey purposes, or a geographical unit of a size able to be enumerated by one fieldworker/enumerator within a specified period. The EA is the building block of any geography.

- **Enumerator**

A person who visits each household and other individuals in a specific EA and administers the questionnaires or arranges for self-enumeration.

- **Household**

A household is a group of persons who live together and provide themselves jointly with food and/or other essentials for living, or a single person who lives alone.

- **Key-from-Image (KFI)**

Capturing of data by looking at the image.

- **Key-from-Paper (KFP)**

Capturing of data by looking at the physical questionnaire.

- **Listing**

Compiling a register of all dwellings, possible dwellings and landmarks in a given EA, including all housing units, all units or rooms within collective living quarters, all non-residential buildings and all vacant stands.

- **Listing error**

For the purpose of the CS, a listing error is a mistake committed during listing, e.g. an office given a Dwelling Unit (DU) number.

- **Multiple households**

Two or more separate households living in the same dwelling unit.
Multiple households occur when:
There is more than one household at one address, or
There is more than one household at one dwelling unit.
Multiple households can be found, for example, in polygamous or extended family situations.
During enumeration, multiple households must always be given separate interviews and separate questionnaires, and a household number.

- **Non-contact form**

The form on which the data collector records failure to make contact, and the reasons why.

- **Non-contacts**

A situation where data collection is incomplete as the enumerator fails to make contact with a household at an address, or an individual in collective living quarters because no one was at home at the time of the visit.

- **Partly completed**

For the purpose of the CS questionnaire, partly completed is when the questionnaire is incomplete.

- **Record number**

A unique number – usually from 001–600 pre-printed in the Enumerator Area Summary Book. During listing, each dwelling unit or other place to be visited for purposes of enumeration is associated with a separate record number in the EA Summary Book. Each record number is on a separate line or row.

- **Refusals**

A situation where a household or individual refuses to answer the questions or complete the questionnaire.

- **Respondents**

The person (or persons) responding in this interview should be a member (members) of the household and be in a position to answer the questions. This will preferably be any responsible adult for the flap and section 1. For the rest of the questionnaire the respondents should answer these questions for themselves, if possible.

- **Unoccupied dwelling**

A dwelling whose inhabitants are absent at the time of the visit or during the reference period during a census or survey.

- **Unusable information**

For the purpose of the CS questionnaire, unusable information is information given by the respondents during the interview that does not make sense.

- **Vacant dwelling**

A dwelling that is uninhabited, i.e. no one lives there.

Annexure B – Community Survey data items

Demographic	**Economic activity**
Age and sex	Any work in the last 7 days
Fertility	Reason for not working
Children ever born (life-time fertility)	Active steps to seek work
Children alive	Availability
Last child born (date of birth, sex, alive/dead)	Status in employment
Mortality	Informal/formal
Did anybody die in the past 12 months prior to survey (how many)	Occupation
In which year and month	Industry
Sex of deceased	Income
Age of deceased	**Geography**
Cause of death (accident, pregnancy, violence, etc.)	Province
Father still alive	District Council
Mother still alive	Municipality
Migration	
Main and sub-place of previous residence (if moved in the past 4 years)	
Month and year of last move	
Main and sub-place of usual residence	
Social	**Services**
Marital status	Type of housing
Relationship	Number of rooms
Population group	Water
Disability	Energy
Social grant	Sanitation
Present school attendance	Communication
Level of education	Refuse removal
Household goods	