South Africa - General Household Survey 2018

Statistics South Africa

Report generated on: June 16, 2026

Identification

SURVEY ID NUMBER
zaf-statssa-ghs-2018-v1

TITLE
General Household Survey 2018

ABBREVIATION OR ACRONYM
GHS 2018

COUNTRY

Name	Country code
South Africa	zaf

STUDY TYPE
Household Survey [hh]

ABSTRACT
The GHS is an annual household survey which measures the living circumstances of South African households. The GHS collects data on education, health, and social development, housing, access to services and facilities, food security, and agriculture.

KIND OF DATA
Sample survey data

UNIT OF ANALYSIS
Households and individuals

Version

VERSION DESCRIPTION
v1.0: Edited, anonymised dataset for public distribution

VERSION DATE
2019-05-28

VERSION NOTES
Version 1 was originally downloaded from Stats SA in June 2019

Scope

NOTES
The scope of the General Household Survey includes:

Household characteristics: Dwelling type, home ownership, access to water and sanitation, access to services, transport, household assets, land ownership, agricultural production
Individuals' characteristics: demographic characteristics, relationship to household head, marital status, language, education, employment, income, health, fertility, mortality, disability, access to social services

Coverage

GEOGRAPHIC COVERAGE
The General Household Survey has national coverage.

GEOGRAPHIC UNIT

The lowest level of geographic aggregation for the data is Province (and metropolitan municipality, where this applies)

UNIVERSE

The survey covers all de jure household members (usual residents) of households in the nine provinces of South Africa, and residents in workers' hostels. The survey does not cover collective living quarters such as student hostels, old age homes, hospitals, prisons, and military barracks.

Producers and sponsors

PRIMARY INVESTIGATORS

Name	Affiliation
Statistics South Africa	Government of South Africa

Sampling

SAMPLING PROCEDURE

From 2015 the General Household Survey (GHS) uses a Master Sample (MS) frame developed in 2013 as a general-purpose sampling frame to be used for all Stats SA household-based surveys. This MS has design requirements that are reasonably compatible with the GHS. The 2013 Master Sample is based on information collected during the 2011 Census conducted by Stats SA. In preparation for Census 2011, the country was divided into 103 576 enumeration areas (EAs). The census EAs, together with the auxiliary information for the EAs, were used as the frame units or building blocks for the formation of primary sampling units (PSUs) for the Master Sample, since they covered the entire country, and had other information that is crucial for stratification and creation of PSUs. There are 3 324 primary sampling units (PSUs) in the Master Sample, with an expected sample of approximately 33 000 dwelling units (DUs). The number of PSUs in the current Master Sample (3 324) reflect an 8,0% increase in the size of the Master Sample compared to the previous (2008) Master Sample (which had 3 080 PSUs). The larger Master Sample of PSUs was selected to improve the precision (smaller coefficients of variation, known as CVs) of the GHS estimates. The Master Sample is designed to be representative at provincial level and within provinces at metro/non-metro levels. Within the metros, the sample is further distributed by geographical type. The three geography types are Urban, Tribal and Farms. This implies, for example, that within a metropolitan area, the sample is representative of the different geography types that may exist within that metro.

The sample for the GHS is based on a stratified two-stage design with probability proportional to size (PPS) sampling of PSUs in the first stage, and sampling of dwelling units (DUs) with systematic sampling in the second stage.After allocating the sample to the provinces, the sample was further stratified by geography (primary stratification), and by population attributes using Census 2011 data (secondary stratification).

WEIGHTING

The sample weights were constructed in order to account for the following: the original selection probabilities (design weights), adjustments for PSUs that were sub-sampled or segmented, excluded population from the sampling frame, non-response, weight trimming, and benchmarking to known population estimates from the Demographic Analysis Division within Stats SA.

The sampling weights for the data collected from the sampled households were constructed so that the responses could be properly expanded to represent the entire civilian population of South Africa. The design weights, which are the inverse sampling rate (ISR) for the province, are assigned to each of the households in a province.

Mid-year population estimates produced by the Demographic Analysis Division were used for benchmarking. The final survey weights were constructed using regression estimation to calibrate to national level population estimates cross-classified by 5-year age groups, gender and race, and provincial population estimates by broad age groups. The 5-year age groups are: 0–4, 5–9, 10–14, 55–59, 60–64; and 65 and over. The provincial level age groups are 0–14, 15–34, 35–64; and 65 years and over. The calibrated weights were constructed such that all persons in a household would have the same final weight.

The Statistics Canada software StatMx was used for constructing calibration weights. The population controls at national and provincial level were used for the cells defined by cross-classification of Age by Gender by Race. Records for which the age, population group or sex had item non-response could not be weighted and were therefore excluded from the dataset. No additional imputation was done to retain these records.

Household estimates that were developed using the UN headship ratio methodology were used to weight household files. The databases of Census 1996, Census 2001, Community Survey 2007 Census 2011 were used to analyse trends and develop models to predict the number of households for each year. The weighting system was based on tables for the expected distribution of household heads for specific age categories, per population group and province.

Data Collection

DATES OF DATA COLLECTION

Start	End
2018-01	2018-12

DATA COLLECTION MODE
Face-to-face [f2f]

Questionnaires

QUESTIONNAIRES
Data was collected with a household questionnaire and a questionnaire administered to a household member to elicit information on household members.

Data Appraisal

DATA APPRAISAL
Please note that DataFirst provides versioning at dataset and file level. Revised files have new version numbers. Files that are not revised retain their original version numbers. Changes to any of the data files will result in the dataset having a new version number. Thus version numbers of files within a dataset may not match

Access policy

ACCESS CONDITIONS
Public access data, available to all

CITATION REQUIREMENTS
Statistics South Africa. General Household Survey 2018 [dataset]. Version 1. Pretoria: Statistics SA [producer], 2019. Cape Town: DataFirst [distributor], 2019. DOI: https://doi.org/10.25828/9tmn-fz97

ACCESS AUTHORITY

Name	Affiliation	Email	URL
DataFirst	University of Cape Town	support@data1st.org	Link

Metadata production

PRODUCERS

Name	Affiliation	Role
DataFirst	University of Cape Town	Metadata Producer

DATE OF METADATA PRODUCTION
2021-12-09

DDI DOCUMENT VERSION

Version 3

Data Description

Data file	Cases	Variables
ghs-2018-house-1	0	0
ghs-2018-person-1	0	0

Data file: ghs-2018-house-1

Cases:	0
Variables:	0

Variables

ID	Name	Label	Question

Total: 0

Data file: ghs-2018-person-1

Cases:	0
Variables:	0

Variables

ID	Name	Label	Question

Total: 0