

# South Africa - General Household Survey 2022

**Statistics South Africa**

Report generated on: August 25, 2023

Visit our data catalog at: <https://www.datafirst.uct.ac.za/dataportal/index.php>

## Identification

### SURVEY ID NUMBER

zaf-statssa-ghs-2022-v1

### TITLE

General Household Survey 2022

### COUNTRY

Name	Country code
South Africa	zaf

### STUDY TYPE

Household Survey

### 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: Edited, anonymised dataset for public distribution

### VERSION DATE

2022

### VERSION NOTES

Version 1 was originally downloaded from Stats SA in August 2023

## 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.

**Note on Independently Calibrated Weights for the Person and Household Data Files**

Until 2010 Statistics SA used an integrating weighting methodology. 'Integrated' weights allocated the same weight to all household members. The household head's weight was carried over the house file. This model allowed the replication of the population size if household sizes were multiplied with the household weight. However, this method provided variable household totals from year to year. Therefore from 2010 the Person and House files across the whole GHS series are calibrated independently from each other. The person data is calibrated using the mid-year population estimates from the 2017 series, while the house data is weighted using household estimates that are also based on the 2017 mid-year population series. However, this method means that the totals will not be aligned. For weights that are better aligned users can transfer the weight allocated to the household head to the household file. Statistics SA ensures that all households in the

house file are also represented in the person file.

## Data Collection

### DATES OF DATA COLLECTION

Start	End
2022-05-01	2022-12-01

### DATA COLLECTION MODE

Computer Assisted Personal Interview

### DATA COLLECTION NOTES

Due to the COVID-19 Pandemic, StatsSA changed the mode of collecting GHS 2020 data from Computer Assisted Personal Interviews (CAPI) to Computer-assisted Telephone Interviews (CATI). After the pandemic, in 2022, the GHS reintroduced the CAPI interviews.

## 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

Since 2019, the questionnaire for the GHS series changed and the variables were also renamed. For correspondence between old names (GHS pre-2019) and new name (GHS post-2019), see the document [ghs-2019-variables-renamed](#).

## Access policy

### CONTACTS

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

### ACCESS CONDITIONS

Public access data for use under a Creative Commons CC-BY (Attribution-only) License

### CITATION REQUIREMENTS

Statistics South Africa. General Household Survey 2022 [dataset]. Version 1. Pretoria: Statistics SA [producer], 2022. Cape Town: DataFirst [distributor], 2023. DOI:

### ACCESS AUTHORITY

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

## Metadata production

### PRODUCERS

Name	Affiliation	Role
------	-------------	------

DataFirst	University of Cape Town	Metadata producer
-----------	-------------------------	-------------------

DATE OF METADATA PRODUCTION  
2023-08-25

DDI DOCUMENT VERSION  
Version 1

**Data Description**

<b>Data file</b>	<b>Cases</b>	<b>Variables</b>
<b>ghs-2022-hhold-v1</b>	1000	193
<b>ghs-2022-person-v1</b>	1000	112