South Africa - General Household Survey 2021

Government of South Africa, Statistics South Africa (Stats SA)

Report generated on: June 2, 2026

Visit our data catalog at: https://microdata.fao.org/index.php

Identification

SURVEY ID NUMBER
ZAF_2021_GHS_v01_EN_M_v01_A_ESS

TITLE
General Household Survey 2021

ABBREVIATION OR ACRONYM
GHS 2021

COUNTRY

Name	Country code
South Africa	ZAF

STUDY TYPE
Other Household Survey [hh/oth]

SERIES INFORMATION
The General Household Survey is one of Statistics South Africa's longest-running surveys. It has been conducted for more than twenty years, with its first round conducted in July 2002, and it was originally designed to meet user need of a survey conducted regularly to measure the level of development and the performance of government programs and projects.

ABSTRACT
The General Household Survey is an annual household survey which measures the living circumstances of South African households. The survey 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

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 (namely, the 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
Government of South Africa, Statistics South Africa (Stats SA)

Sampling

SAMPLING PROCEDURE

From 2015, the General Household Survey uses a Master Sample frame, which was developed in 2013 as a general-purpose sampling frame to be used for all Statistics South Africa's household-based surveys. This Master Sample has design requirements that are reasonably compatible with the General Household Survey. The 2013 Master Sample is based on information collected during the Census 2011 conducted by Statistics South Africa.

In preparation for the 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. Census EAs were used because they covered the entire country and provided crucial information for the stratification and creation of PSUs. There are 3324 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 (3324) reflects an 8.0 percent increase in the size of the Master Sample compared to the previous: the 2008 Master Sample had 3080 PSUs.

The larger Master Sample was selected to improve the precision (resulting in smaller coefficients of variation, known as CVs) of the General Household Survey estimates. The Master Sample is designed to provide survey estimates that are 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 General Household Survey is based on a stratified two-stage design with probability proportional to size (PPS) sampling of PSUs in the first stage, and sampling of DUs with systematic sampling in the second stage. After allocating the sample to the provinces, it was further stratified by geography (primary stratification), and by population attributes using Census 2011 data (secondary stratification).

WEIGHTING

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
- population excluded from the sampling frame
- non-response
- weight trimming
- benchmarking to known population estimates from the Demographic Analysis Division within Statistics South Africa

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

Mid-year population estimates, produced by the Demographic Analysis Division, were used for benchmarking. 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 were: 0-4, 5-9, 10-14, 55-59, 60-64, and 65 and over. The provincial level age groups were: 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 South Africa 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 South Africa ensures that all households in the house file are also represented in the person file.

Data collection

DATES OF DATA COLLECTION

Start	End
2021-09-01	2021-12-01

DATA COLLECTION MODE
Computer Assisted Telephone Interview [cati]

DATA COLLECTION NOTES
Due to the COVID-19 Pandemic, Statistics South Africa changed the mode of collecting General Household Survey 2020 data from Computer Assisted Personal Interviews (CAPI) to Computer-assisted Telephone Interviews (CATI).

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

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

CITATION REQUIREMENTS
Statistics South Africa. General Household Survey 2021 [dataset]. Version 1. Pretoria: Statistics SA [producer], 2021. Cape Town: DataFirst [distributor], 2022. DOI: https://doi.org/10.25828/7h7t-df42

ACCESS AUTHORITY

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

Metadata production

DDI DOCUMENT ID
DDI_ZAF_2021_GHS_v01_EN_M_v01_A_ESS_FAO

PRODUCERS

Name	Abbreviation	Affiliation	Role

DataFirst		University of Cape Town	Metadata Producer
Statistics Division	ESS	Food and Agriculture Organization	Metadata adapted for FAM

DDI DOCUMENT VERSION
ZAF_2021_GHS_v01_EN_M_v01_A_ESS_v01

Data Dictionary

Data file	Cases	Variables
ghs-2021-hhold-v1	1000	117
ghs-2021-person-v1	1000	90