



SADA

South African Data Archive

**GENERAL HOUSEHOLD SURVEY, 2010
(South Africa)**

Statistics South Africa (Stats SA)

CODEBOOK

SADA 0178

BIBLIOGRAPHIC CITATION

Any publication or other presentation based in whole or part on the data and documentation supplied by SADA must prominently use the following citation:

Statistics South Africa. *General Household Survey, 2010* [Computer file]. S0178. Pretoria: Statistics South Africa [producer], 2010. Pretoria: South African Data Archive, National Research Foundation [distributor], 2011.

DEPOSIT OF PUBLICATIONS

At least one copy of any published work or report based in whole or part on the dataset will be deposited with the South African Data Archive. Please indicate the title and number of the study utilised.

LIMITATION OF LIABILITY

The Archive and the depositor of the dataset supplied to the user bear no legal responsibility for their accuracy or comprehensiveness.

Postal Address:	SOUTH AFRICAN DATA ARCHIVE NATIONAL RESEARCH FOUNDATION P.O. BOX 2600 Pretoria 0001
Telephone:	+27 12 481 4158/ 4016/ 4120
Telefax:	+27 12 349 1179
Electronic mail:	sada@nrf.ac.za
World Wide Web:	http://www.nrf.ac.za/sada

STUDY DESCRIPTION

SADA 0178

TITLE: General Household Survey, 2010 (GHS 2010).

PRINCIPAL INVESTIGATOR: Statistics South Africa (Stats SA)

DEPOSITOR: Statistics South Africa

ORGANISATION HOUSING THE DATA:

Statistics South Africa

Private Bag x44

Pretoria, 0001

South Africa

Tel: (012) 310 8600

Fax: (012) 310 8500

Email: info@statssa.gov.za

Web: <http://www.statssa.gov.za>

KEYWORDS:

Employment; Unemployment; Household surveys; Housing; Education and Health.

SERIES DATA:

Series name:	Principal Investigator:	Currently available:
General Household Survey	Statistics South Africa	2002 - 2010

PURPOSE:

The GHS is an annual household survey since 2002, specifically designed to measure various aspects of the living circumstance of South Africa households.

GEOGRAPHIC LOCATION: South Africa

IMPORTANT VARIABLES: Employment, unemployment, education, health, access to services and facilities.

DEMOGRAPHIC VARIABLES: Age, gender, level of education, marital status, use of health services, economic activity and employment.

UNIVERSE: Households in the nine provinces of South Africa

METHOD OF DATA COLLECTION: Survey Questionnaire

SAMPLING:

The sample design for the GHS 2010 was based on a master sample (MS) that was originally designed for the QLFS and was used for the first time for the GHS in 2008. This master sample is shared by the Quarterly Labour Force Surveys (QLFS), General Household Survey (GHS), Living Conditions Survey (LCS), Domestic Tourism Survey and the Income and Expenditure Surveys (IES).

The master sample used a two-stage, stratified design with probability-proportional-to-size (PPS) sampling of PSUs from within strata, and systematic sampling of dwelling units (DUs) from the sampled primary sampling units (PSUs). A self-weighting design at provincial level was used and MS stratification was divided into two levels. Primary stratification was defined by metropolitan and non-metropolitan geographic area type. During secondary stratification, the Census 2001 data were summarised at PSU level. The following variables were used for secondary stratification; household size, education, occupancy status, gender, industry and income.

Census enumeration areas (EAs) as delineated for Census 2001 formed the basis of the PSUs. The following additional rules were used:

- Where possible, PSU sizes were kept between 100 and 500 dwelling units (DUs);
- EAs with fewer than 25 DUs were excluded;
- EAs with between 26 and 99 DUs were pooled to form larger PSUs and the criteria used was same settlement type;
- Virtual splits were applied to large PSUs: 500 to 999 split into two; 1 000 to 1 499 split into three; and 1 500 plus split into four PSUs; and
- Informal PSUs were segmented.

A Randomised Probability Proportional to Size (RPPS) systematic sample of PSUs was drawn in each stratum, with the measure of size being the number of households in the PSU. Altogether approximately 3 080 PSUs were selected. In each selected PSU a systematic sample of dwelling units was drawn. The number of DUs selected per PSU varies from PSU to PSU and depends on the Inverse Sampling Ratios (ISR) of each PSU.

UNITS OF OBSERVATION: Households

DATE OF DATA COLLECTION: July 2010

EXTENT OF DATA COLLECTION: 2 data files in ASCII and machine-readable documentation.

FILE SPECIFICATIONS:

Part 1: House Data

Number of cases:	25 548
Number of records:	25 548
Number of records per case:	1
Logical Record Length:	381
Number of variables:	224
Number of kilobytes:	9 556 KB

Part 2: Person Data

Number of cases:	95 918
Number of records:	95 918
Number of records per case:	1
Logical Record Length:	226
Number of variables:	153
Number of kilobytes:	18 114 KB

PUBLICATIONS:

Statistics South Africa. 2010. General Household Survey, 2010. Pretoria: Statistics South Africa.

NOTES:

Most questions in the General Household Survey questionnaire are pre-coded, i.e. there is a set number of choices from which one or more must be selected. For open-ended 'write-in' questions, the description will note that post-coding occurred and explain how this was done. Most variables have been pre-coded from the questionnaire and is not repeated in the variable description. Where the coding is not apparent, the description either provides the codes or indicates where code lists are to be found.

RESTRICTIONS:

Only available to bona fide researchers.

DATE STUDY ADDED:

November, 2011