USER GUIDE

SOUTH AFRICAN SOCIAL ATTITUDES SURVEY (SASAS) 2008

HSRC Study ID

SASAS 2008 Q1

SASAS 2008 Q2

Human Sciences Research Council

2014

Data Curation

HSRC Human Sciences Research Council

Table of Contents

Human Sciences Research Council (2014) *South African Social Attitudes Survey (SASAS) 2008 User guide. Version 2.* Pretoria: Human Sciences Research Council.

South African Social Attitudes Survey (SASAS) 2008

1. Introduction

The primary objective of SASAS is to design, develop and implement a conceptually and methodologically robust study of changing social attitudes and values in South Africa. In meeting this objective, the HSRC intends to position itself so as to be able to carefully and consistently monitor and explain changes in attitudes amongst various socio-demographic groupings.

It is intended that SASAS will explore a wide range of value changes, including the distribution and shape of racial attitudes and aspirations, attitudes towards democratic and constitutional issues, and the redistribution of resources and power.

Moreover, there is also an explicit interest in mapping changing attitudes towards some of the great moral issues that confront and are fiercely debated in South Africa, for instance abortion, gender issues, AIDS, crime and punishment, governance, service delivery, etc.

In sum, then, SASAS is intended to provide a unique long-term account of the social fabric of modern South Africa, and of how its changing political and institutional structures interact over time with changing social attitudes and values.

To accommodate the wide variety of topics that was included in the 2008 survey, two questionnaires were administered simultaneously. This will continue to occur in future phases of the SASAS survey. The core module will remain constant from round to round, with the aim of monitoring change and continuity in a variety of socio-economic and socio-political variables. (Note that in 2008, not all the questions were included in the core module for questionnaire two.)

In addition, a number of themes will be accommodated on a rotational basis. This rotating element of the survey consists of two or more topic-specific modules in each round of interviewing and is directed at measuring a range of policy and academic concerns and issues that require more detailed examination at a specific point in time than the multi-topic core module would permit. The two different versions of the questionnaire were administered concurrently in each of the chosen sampling areas.

Fieldworkers were required to complete a paper-based instrument while interviews were conducted face-to-face.

International Social Survey Programme. (ISSP web page: www.issp.org/)

The International Social Survey Programme (ISSP) is run by a group of research organisations, each of which undertakes to field annually an agreed module of questions on a chosen topic area. SASAS 2003 represents the formalisation of South Africa's inclusion in the ISSP, the intention being to include the module in one of the SASAS questionnaires in each round of interviewing. Each module is chosen for repetition at intervals to allow comparisons both between countries (membership currently stands at 40) and over time. In 2008, the chosen subject was religion, and the module was carried in version 2 of the questionnaire (Questions1-74). This data can be accessed through the ISSP data portal – see link above.

Thematic content:

Democracy, intergroup relations, education, moral issues, crime and safety, voting, water, demographics and other classificatory variables.

Rotating modules:

Questionnaire 1: National identity and pride, gender attitudes, poverty, household expenditure, service delivery, hope scale, soccer world cup, democracy (part 2).
Questionnaire 2: ISSP module – religion, water and sanitation, hunger scale, voting.

Please refer to the Study information document on the web site for a detailed description of the study.

2. Available data sets

SPSS data sets:

SASAS2008_Q1.sav (n=3321); 366 variables, 'benchwgt' (weight variable)

SASAS2008_Q2.sav (n=3292); 287 variables, 'benchwgt' (weight variable)

STATA data sets:

SASAS2008_Q1.dta (n=3321); 366 variables, 'benchwgt' (weight variable);

SASAS2008_Q2.dta (n=3292); 287 variables, 'benchwgt' (weight variable);

SAS data sets and catalog files:

SASAS2008_Q1.sas7bdat (n=3321); 366 variables, 'benchwgt' (weight variable);
SASAS2008_Q1.sas7bcat

SASAS2008_Q2.sas7bdat (n=3292); 287 variables, 'benchwgt' (weight variable);
SASAS2008_Q2.sas7bcat

ASCII data sets:

Fixed format

SASAS2008_Q1.dat (n=3321)

SASAS2008_Q2.dat (n=3292)

Comma Separated Values data file (csv)
SASAS2008_Q1.csv (n=3321); 366 variables

SASAS2008_Q2.csv (n=3292); 287 variables

2.1. Program code to generate the different statistical data sets above:

SPSS:

SASAS2008_Q1.sps

SASAS2008_Q2.sps

SAS:

SASAS2008_Q1.sas

SASAS2008_Q2.sas

STATA:

SASAS2008_Q1.do; SASAS2008_Q1.dct (dictionary)

SASAS2008_Q2.do; SASAS2008_Q2.dct (dictionary)

3. Conversion of data

Ingest format(s) of the data: SPSS (.sav)

Additional comment about each data format is given below, please read the wording corresponding to the data format you have downloaded:

SPSS:

If SPSS has not been the ingest format this format has been created according to StatTransfer standards. Three files are created:
ASCII data file (.dat), SPSS syntax file (.sps) and a SPSS data set (.sav).
Use the SPSS syntax file to convert data to an SPSS data set (if not provided). Be sure to change the path in the DATA LIST FILE= statement to point to where you stored the ASCII data file.

SAS:

If SAS is not the ingest format, SAS files are created according to StatTransfer standards. Three files are created:

ASCII data file (.dat), SAS program file (.sas) and SAS data set (.sas7bdat). The formats.sas7bcat file contains the formats of the variables and must be used with the data set as illustrated under the heading: 'Additional notes on SAS'.

In the SAS program file, change the infile statement to point to where you stored the ASCII data file. To create a SAS data set (if not provided) open the SAS program in the SAS program editor and submit. If you do not want the value labels permanently associated with the variable values in the data set, comment out the format statement in the data step and rather use it in the PROCS.

STATA:

If STATA is not the ingest format, STATA files are created according to StatTransfer standards. Four files are created:
ASCII data file (.dat), STATA program file (.do), dictionary file (.dct) and a STATA data file (.dta).
To create a STATA data set (if not provided) open the .do program in (e.g. STATA editor) and change the path name to point to where you stored the ASCII data and the .dct (dictionary) files. Also indicate the path of the .dta output data file. When the .do file is opened in STATA, the STATA data set is created.

The data sets contain a mix of ordinal, nominal and scale (continuous) variables. (See Appendix B for an explanation.)

3.1. Additional notes on SPSS

When using SPSS please note the following: if statistical analysis, other than descriptive statistics is done, the measurement level for variables should be set to the appropriate level as per the example below.

Example: Measurement Level syntax (SPSS):

```
VARIABLE LEVEL
uniqueid ea hhnum pers16 hhpers hhpers16 benchwgt q136a q142a q226
(SCALE)
/ prov q59 q64 q69 geotype metro
(NOMINAL)
/ q1 to q33 q103
(ORDINAL).
```

3.2. Additional notes on SAS

In the downloadable SAS program the formats were stored in a SAS catalog, SASAS2008_Q1.sas7bcat. This was the result of using the library option in proc format and putting the format statement to associate the value labels with the variables, in the data step. Whenever the SAS data set is used, a libname statement must indicate where the permanent SAS data set and the format catalog are located. Store them in the same folder.

Example: Referencing the permanent SAS data set and the format catalog

Libname library '*full path where the data set and catalog are stored e.g. c:\myfolder*';
PROC FREQ DATA=library.SASAS2008_Q1;
Run;

Please note that there is no need to specify the formats catalog if it is ***formats.sas7bcat*** *(default) as long as it is in the same folder as the data set.*

If however the format has a name such as ***SASAS2008_Q1.sas7bcat,*** *you need to put in additional statements (see red below).*

libname library 'c:/myfolder';

OPTIONS FMTSEARCH=(library.SASAS2008_Q1); (This is to locate the formats)

PROC FREQ DATA=library.SASAS2008_Q1; (This is for the SAS data set SASAS2008_Q1.sas7bdat)

Note:

When searching for a format or informat, SAS always searches in WORK.FORMATS first, and then LIBRARY.FORMATS, unless one of them appears in the FMTSEARCH= list. SAS searches the catalogs in the FMTSEARCH= list in the order that they are listed until the format or informat is found.

If you want the values and not the formats in the data set, then use the following statements to disassociate the formats from the variables in the data set. You can then run the proc format separately and associate the variables with the formats in a PROC.

WARNING: If the value labels are in the data set, SAS will do the analysis on the value labels which are text instead of the numeric values.

Example: Getting rid of the formats in the data set

The following code will remove all formats associated with the variables in the SAS data set.

Libname library '*full path where the data set and catalog are stored e.g. c:\myfolder*';

DATA *library.newdsn*;
SET *library.dsn*;
FORMAT _all_;
RUN;

- *library* refers to name you used in the libname statement.
- *dsn* is the name of the SAS data set, whose formats need to be removed.
- *newdsn* is the name of the new SAS data set with the formats removed.
- format _all_ will delete all the formats in the SAS data set.

The following code will remove only the formats of specific variables in the SAS data set. Specify the variables without their formats in the FORMAT statement. This is also applicable to PROCS.

DATA *library.newdsn*;
SET *library.dsn*;
FORMAT q1a q1b q2 q3;
RUN;

3.3. Additional notes on STATA

Q1:

In the analysis of the data, the multilevel complex design of the survey should be taken into account. The following design variables should be specified when analysing these data:

- the weight variable is **benchwgt**
- the strata are **prov, geotype, race (q225)**
- the primary sampling unit is **ea**

Q2:

In the analysis of the data, the multilevel complex design of the survey should be taken into account. The following design variables should be specified when analysing these data:

- the weight variable is **benchwgt**
- the strata are **prov, geotype, race (Q156)**
- the primary sampling unit is **ea**

3.4. Other formats:

If you do not have any of the above mentioned statistical software you can use the Comma Separated Values data file (.csv) as it is compatible with most software. The variable and value labels are provided in any of the program files (.sas, .sps or .do) which can be opened in any text editor. The program files can also be used to locate the positions of the variables if you want to use the .dat fixed format file.

4. Data and documentation notes

4.1. Deleted / recoded variables

Certain variables that were not relevant to research were deleted from the data sets.

4.2. Missing values within the data sets

Missing values in the data are indicated as 'System missing' denoted by a dot (.).

Inconsistencies and missing values in the data could be attributed to various aspects during fieldwork:

In certain cases questions had to be skipped depending on the answer provided at a previous question and follow-up question(s) were left blank.

Missing values could also be attributed to respondents refusing to answer particular questions or due to fieldworker inaccuracy.

4.3. Variables and labels within the data sets

The following code books may be consulted for the variables, variable labels and frequencies within the respective data sets:

SASAS 2008 Q1 Code book.pdf
SASAS 2008 Q2 Code book.pdf

4.4. Weighting of the data

The metadata record on the web site provides a detailed description of the weighting procedure that was followed for the SASAS 2008 survey.

The weighting variable to be used in the analysis: 'benchwgt' for both Q1 and Q2.

The sample is weighted to represent the South African population older than 16 years. With certain exceptions analysis should be performed on weighted data.

4.5. Coding of occupation

The occupation was coded to two digits on the basis of the International Standard Classification of Occupations (ISCO 88), URL: www.ilo.org/public/english/bureau/stat/isco/isco88/major.htm

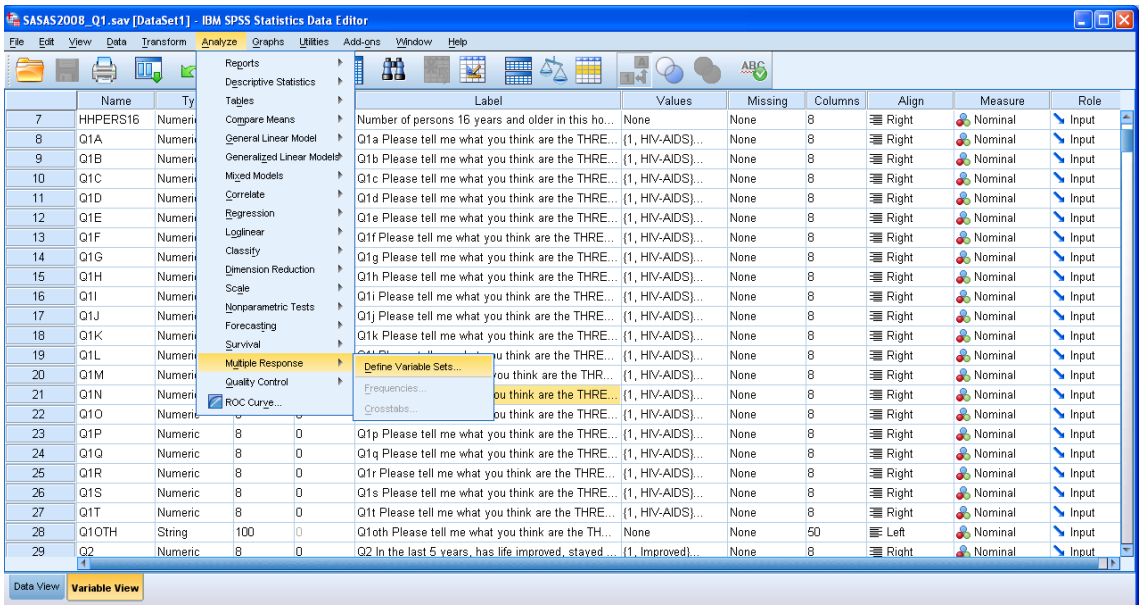
APPENDIX A

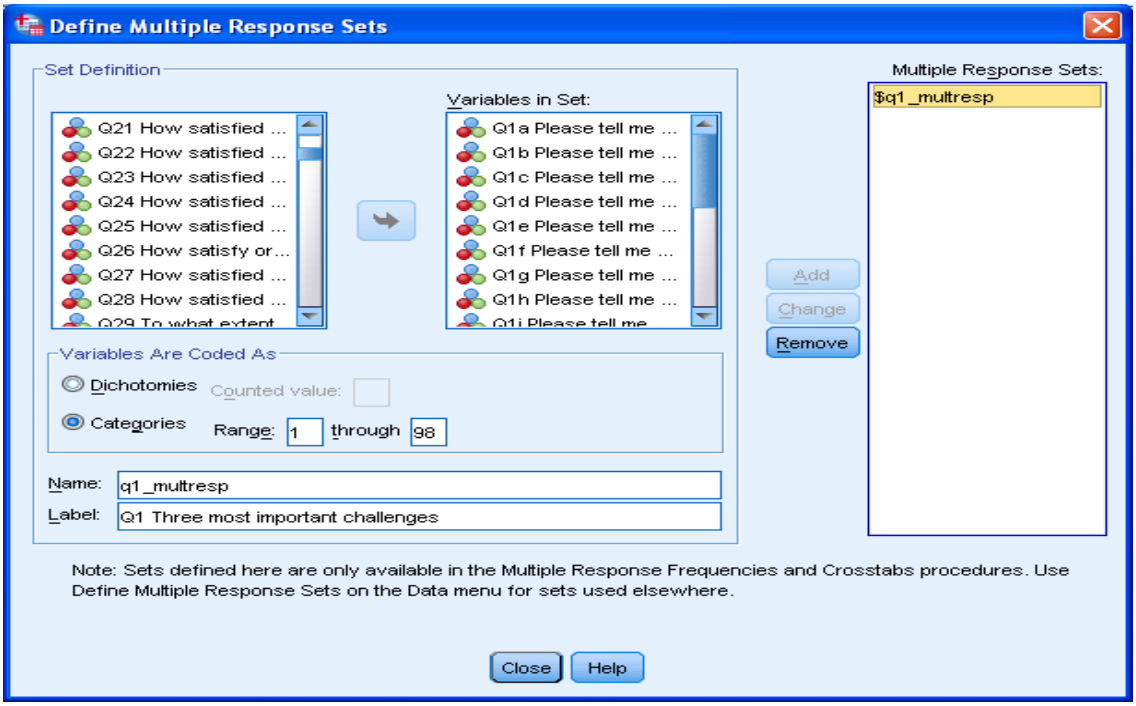
Analysing Multiple Response Questions:

Multiple Response Questions:

Multiple response questions: Questionnaire 1: q1a -q1t, q60, q173

If you want to analyse multiple response questions to get an overview of similar responses, you can use SPSS multiple response set functionality under *Analyze, Multiple response, Define Variable Sets:*





Select the variables you will need for the multiple response set: e.g. Q1a-Q1t and move to *Variables in set:* box by ticking the ► arrow.

In this illustration *Categories* will be selected, as answer options in Q1a to Q1t vary from code 1 to code 19 and code 98. (In the case where variables are coded as *'Yes/No'* , *Dichotomies* will be selected and *Counted value* will be assigned in the box.)

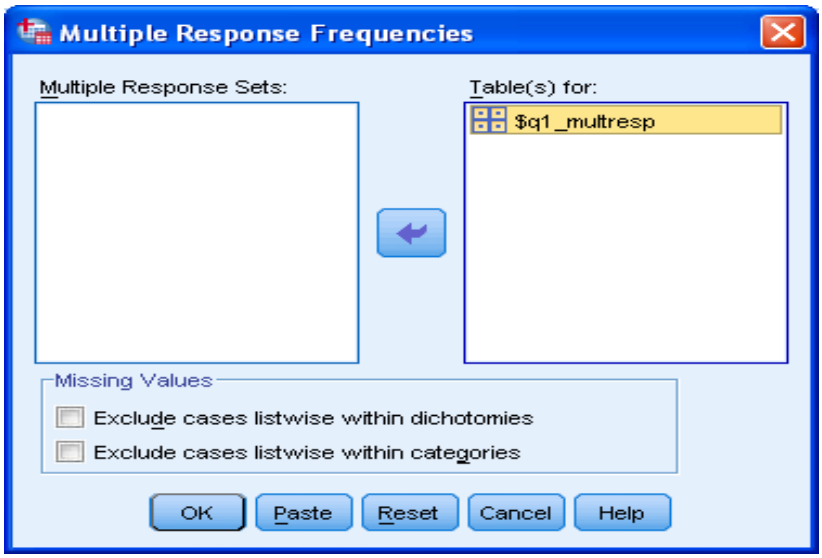
Assign a variable name in the *Name* box (q1_multresp) and a description in the *Label* box (Q1 Three most important challenges).

Select *Add*, *Close*.

To run a frequency or cross tabulation:

Analyze, Multiple response, select *Frequencies….*or *Crosstabs…:*

The created variable **$Q1_multresp** can be used in a frequency table or crosstab.



Example of output for multiple response questions:

$Q1_multresp Frequencies

		Responses		Percent of Cases
		N	Percent	
$q1_multresp Q1 Three most important challenges[a]	1 HIV-AIDS	1528	15.5%	46.0%
	2 Unemployment	2226	22.6%	67.0%
	3 Racism	165	1.7%	5.0%
	4 Xenophobia	173	1.8%	5.2%
	5 Crime and safety	1609	16.4%	48.5%
	6 Service provision or delivery	412	4.2%	12.4%
	7 Affordable housing	356	3.6%	10.7%
	8 Land reform issues	78	.8%	2.3%
	9 Human rights	142	1.4%	4.3%
	10 Education	297	3.0%	8.9%
	11 Economic and financial issues	304	3.1%	9.2%
	12 Work-related issues	92	.9%	2.8%
	13 Family and youth issues	102	1.0%	3.1%
	14 Religion and culture issues	36	.4%	1.1%
	15 Environmental issues	91	.9%	2.7%
	16 Political issues	254	2.6%	7.7%
	17 Corruption	522	5.3%	15.7%
	18 Poverty	1295	13.2%	39.0%
	19 Other	116	1.2%	3.5%
	98 Do not know	37	.4%	1.1%
Total		9835	100.0%	296.2%

a. Group

Note that percentages do not add up to 100.0% in the 'Percent of Cases' column as one person could select more than one answer option. (The above is an example and calculations were not performed on weighted data.)

APPENDIX B

Nominal

A variable can be treated as nominal when its values represent categories with no intrinsic ranking; for example, the department of the company in which an employee works. Examples of nominal variables include region, zip code, or religious affiliation.

Ordinal

A variable can be treated as ordinal when its values represent categories with some intrinsic ranking; for example, levels of service satisfaction from highly dissatisfied to highly satisfied. Examples of ordinal variables include attitude scores representing degree of satisfaction or confidence and preference rating scores. For ordinal string variables, the alphabetic order of string values is assumed to reflect the true order of the categories. For example, for a string variable with the values of low, medium, high, the order of the categories is interpreted as high, low, medium which is not the correct order. In general, it is more reliable to use numeric codes to represent ordinal data.

Scale

A variable can be treated as scale when its values represent ordered categories with a meaningful metric, so that distance comparisons between values are appropriate. Examples of scale variables include age in years and income in thousands of dollars.