



PO Box 81443-80100 Mombasa
+254-707-628-499
grassrootseconomics.org

Data Analysis/Survey and Reporting Standards - 2017

Grassroots Economics Foundation (GE) adheres to the following standards as a research service provider to ensure proper documentation and comparability of reporting across all surveys. *These standards are adapted from the University of Cape Town & Canadian Government Standards and Guidelines*

Data Analysis Verification Standards

1 Analysis records

- GE shall keep accurate and descriptive records of the analysis process, to ensure that any analysis undertaken can be replicated at a later date.

2 Data analysis verification

- GE shall have in place procedures to ensure the tabulations and other outputs have been checked.
- As a minimum, these checks shall verify:
 - Completeness, i.e., that all tables are present as specified, including the results of all reported significance tests
 - That abbreviations for headings or open-ended responses accurately reflect the full content
 - That the base for each table is correct against other tables or frequency counts
 - That the standard breaks/banner points are checked against source questions
 - That all derived data items are checked against their source
 - That the figures for sub-groups and nets are correct
 - That there are no blank tables (i.e., with no data)
 - Weighting (e.g., by test tables)
 - Frequency counts prior to running tables, in order both to ensure the accuracy of data and to determine base sizes for sub-groups
 - Spelling and legibility

- That any statistical analysis used is appropriate and correct, both in its descriptive and inferential aspects
- For any subsequent outputs, appropriate checks shall be applied.

Delivery of Data Tables Standards

3 Delivery of stand-alone hard or soft copy of data tables

- When data are reported to the client, such as in a stand-alone hard or soft copy of data tables, the following shall be taken into account, as appropriate:
 - Reference to the actual source question to which the data pertains
 - Inclusion of a description of any weighting method applied to the data
 - Clear identification of any sub-groups used
 - Availability of the bases for each question, so that the number of respondents who have actually answered the question is identifiable
 - The number or proportion of respondents who replied "don't know" or gave "no answer"
 - Availability of both weighted and unweighted bases
 - Clear and complete definition and explanation of all variables used in the analysis of the data, including any significance testing, indexing, scoring, scaling and calculations of means, median, modes and standard deviations
 - The types of statistical tests being used and their level of precision
 - Information on cell suppression and other measures to assure confidentiality
 - Warnings on results which are unreliable due to very small sample sizes

4 Electronic data delivery

- GE shall provide the project authority with a data file.
- For data delivered to the project authority in electronic format, the following shall be checked prior to data release:
 - Compatibility of the file format with the software specification agreed with the client (we generally use excel for small surveys or .csv)
 - Completeness (i.e., the correct number of files and records are in each file)
 - Inclusion of all appropriate documentation to allow for replication of the data analysis and additional analyses, including where applicable
 - A structural description of the file

- Labelling of the contents of the file, i.e., fully labelled variables and value labels
- Identification and description of any computed or recoded variables, and instructions on limitations of use
- Labelled weighting variables and a description of how these were applied
- All personal identifiers per have been removed from the files
- Encryption of files upon request

Inferences and Comparisons

STANDARD

- GE must base statements of comparisons and other statistical conclusions derived from survey data on acceptable statistical practice.

GUIDELINES

- Before including statements in information products that two characteristics being estimated differ in the actual population, make comparison tests between the two estimates, if either is constructed from a sample. Use methods for comparisons appropriate for the nature of the estimates. In most cases, this requires estimates of the standard error of the estimates and, if the estimates are not independent, an estimate of the covariance between the two estimates.
- Given a comparison that does not have a statistically significant difference, conclude that the data do not support a statement that they are different. If the estimates have apparent differences, but have large standard errors making the difference statistically insignificant, note this in the text or as a note with tables or graphs.
- Support statements about monotonic trends (strictly increasing or decreasing) in time series using appropriate tests. If extensive seasonality, irregularities, known special causes or variation in trends are present in the data, take those into account in the trend analysis.
- When performing comparison tests, report only the differences that are substantively meaningful, even if other differences are also statistically significant.

Back-up, Retention/Security of Data

Documents related either to the method of interviewing or specific documents arising from the project have fundamental security/retention policies by which members are governed.

The method used in a research project should be collectively administered to allow the study to be replicated. The required information is usually found in the Technical Appendix, along with any specific documents related to the study. A member company retains such files for a minimum of three years.

Specific documents, like completed questionnaires, the actual data in electronic format and tabulations produced from the same data also have similar retention policies. Questionnaires can be disposed of after 12 months from the end of the data collection.

Unless prior agreement has been reached with the client, these documents are the property of the member. Should the client wish to have such documents, the client is expected to pay reasonable costs of ensuring confidentiality (i.e., removing names and addresses from questionnaires).

5 Back-up, security/retention of data

- All data shall be held securely to prevent unwarranted access, damage or accidental loss.
- The research service provider shall ensure that all data are stored and retained in accordance with professional codes, the applicable law and legislative regulations, and any negotiated agreement with the project authority. (Note: It was suggested this point be amended to state the specific requirements, once this is explored with .)
- Data pertaining to data processing and analysis may include, but is not limited to:
 - Raw data files
 - Other electronic files
 - Hard copy questionnaires and any materials/visual aids used by interviewers in the execution of interviews (e.g., cards, lists, s)
 - Code frames
 - Project files including project management information
 - Emails and other correspondence
- The research service provider shall establish and maintain procedures to ensure that DP (data processing) computer files are clearly identified.
- Where data has been edited, cleaned, recoded or changed in any other way from the format, content and layout of its original format, the original data, final data and programme files, including all documentation related to changes to the data (as a minimum) shall be kept so that the final data set can be easily reconstructed.
- Extra analyses may be requested after the main processing and such analysis shall conform to the DP (data processing) requirements of this standard.
- For any proposed release of tables or data, procedures shall be in place to minimize the risk of disclosure of respondent identifiable information.

6. Survey Documentation

In quantitative research the following minimum details shall be documented in the project report. These allow the reader to understand the way the research project was conducted and the implications of its results:

- The name of the client
- The name of the research service provider
- An executive summary of key results and conclusions, linked to the survey objectives, research questions
- Detailed description of background including at minimum:
 - Purpose, how the research will be used
 - Objectives, research questions
- Detailed description of methodology including:
 - The target group for the research project
 - The achieved sample size against projected sample size and reasons, if relevant, for not obtaining the projected sample
 - The date of fieldwork
 - The average interview length and the range
 - The sampling method, including the procedure for selecting respondents
 - The data collection method, and if applicable:
 - The type and amount of incentives
 - The number of interviewers
 - The interviewer validation methods
 - The call dispositions and response rate (in the case of probability samples) using the formula recommended by
 - State the level of precision, including the margin of error and confidence interval for the total sample and any key sub-groups (when probability samples are used)
 - Overview of the survey analytical plan
 - The weighting procedures, if applicable
 - The estimating and imputation procedures, if applicable
 - The results that are based on sub-groups and the number of cases used in sub-group analysis

- A brief summary of other quality controls and procedures used, including the results of each, which are to be detailed in the Technical Appendix
- Two separate appendices:
 - Study Materials, containing the questionnaires, any visual aids, and other relevant data collection documents, in all languages in which the research was conducted
 - Technical Appendix, containing:
 - Detailed call disposition record
 - A detailed description of the quality control procedures used and the results of each, measures/sources of sampling and non-sampling errors and, as appropriate any other information related to the quality of the survey