

Introduction to Survey Design

- **Defining the survey objectives, variables and concepts**
- **Defining the population to be surveyed**
- **Frames and Coverage**

Cenker Burak METİN

Head of Survey and Questionnaire Design Group

Defining the survey objectives, variables and concepts

Poorly defined objectives will often lead to statistical activities that deliver data which do not meet the desired uses of the stakeholders and can lead to inefficiencies in the design of the activity.

For example, collecting information on income

- as a broad, cross classificatory variable expressed in a small number of broad ranges; or
- as a detailed dollar figure collecting only wage and salary income; or
- as detailed variables indicating dollar figures for all forms of income.

These different forms of income are collected to meet different intended uses of the data and have considerably different costs associated with their collection.



Balance the requirements of the user/users with the constraints of budget, resources and time available

Defining the survey objectives, variables and concepts

Suppose that a survey on poverty is to be conducted. It is not enough to indicate that the purpose of the survey is to provide information on, for example, ‘housing conditions of the poor’. Such a vague statement may serve as a broad description of the survey’s general theme, but ultimately it must be expanded into more specific language.

- What is meant by ‘housing conditions’?
- What precisely is meant by ‘poor’? Is poverty to be measured in terms of income, expenditures, debts, or all of these?



Concepts and Operational Definitions

For example, in a survey of businesses, respondents may need to be classified according to their industrial activity. If this is done using a standard format, the resulting statistics can be compared with existing published statistics or with a later repeat of the survey.

Likewise, in a household survey, if income is collected using the standard definition, it can be used on a comparable basis. By making use of standard concepts and data items it may also be possible to integrate data from different organisations. This integration improves the comparability and relevance of the data and reduces duplication of data collection.

- Concepts and variables should be clearly defined and linked to the research.
- Standard concepts, variables, classifications should be used.
- International definitions and classifications should be considered.
- Concepts, variables and classifications used in the research should be documented and differences from standards should be recorded.

Defining the population to be surveyed

✓ Target and Survey Populations

The target population or scope of the survey is the group about which inferences would like to be made from the survey data.

The following factors are essential in defining the target population;

- ❖ the type of units that comprise the population and the defining characteristics of those units (Who or What?)
- ❖ the geographical location of the units (Where?)
- ❖ the reference (time) period under consideration (When?)

The survey population is the population that is covered by the survey.

Frames and coverage

✓ Survey Frame

Once the client and statistical agency are satisfied with the definition of the target population, some means of accessing the units of the population is required. The survey frame (also called the sampling frame when applied to sample surveys) provides the means of identifying and contacting the units of the survey population.(Survey Methods and Practices, 2010)

The frame chosen determines the definition of the survey population and can affect the methods of data collection, sample selection and estimation, as well as the cost of the survey and the quality of its outputs.

✓ Selection of a Survey Frame

A statistical survey require the use of a correct sampling frame

The coverage, completeness, timeliness, information content and accuracy of the frame are critical factors

✓ Types of Frames

A list frame can be defined as a conceptual list or a physical list of all units in the survey population.

Examples of list frames are:

- Statistics register (e.g., a list of all births and or deaths in the population);
- Business register (e.g., a list of all businesses in operation);
- Address register (e.g., a list of households with civic addresses);
- Telephone directory (i.e., a list of all households with published telephone numbers);

An area frame is a special kind of list frame where the units on the frame are geographical areas. The survey population is located within these geographic areas.

Area frames may be used either when the survey is geographic in nature or when an adequate list frame is unavailable, in which case the area frame can be used as a vehicle for creating a list frame.

A multiple frame is a combination of two or more frames, (a combination of list and area frames or of two or more list frames).

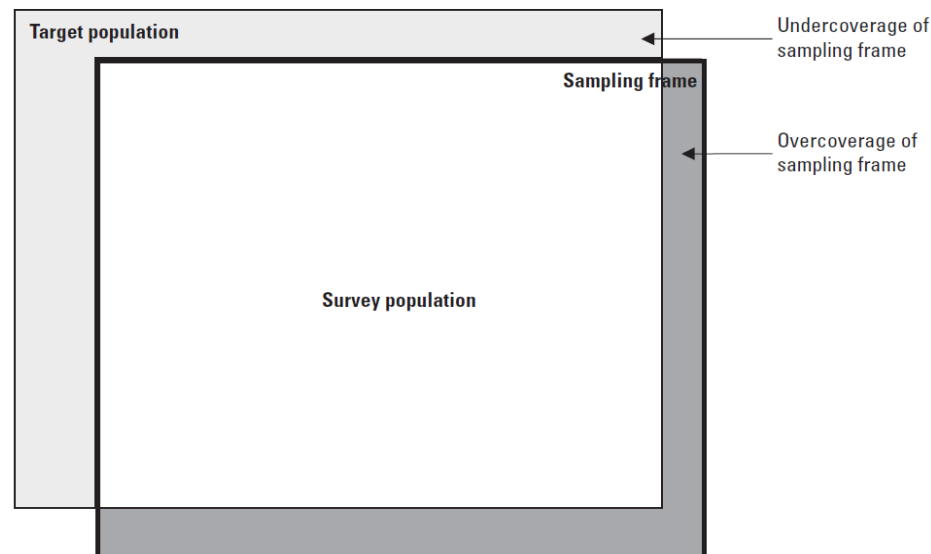
✓ Frame Defects

Undercoverage: exclusions from the frame of some units that are part of the target population.

Overcoverage: inclusions on the frame of some units that are not part of the target population

Duplication: the same unit appears on the frame more than once

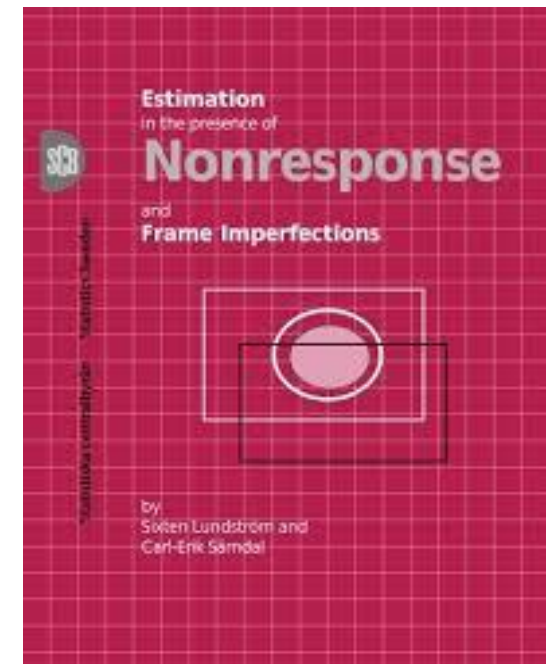
Misclarification: incorrect values for variables on the frame



Source: Quality Guidelines for Official Statistics (2002)

In order to choose and make the best use of the frame;

- ✓ assess different possible frames
- ✓ avoid using multiple frames
- ✓ use the same frame with the same target population
- ✓ incorporate procedures to eliminate duplication and to update frame
- ✓ emphasise the importance of coverage
- ✓ monitor the quality of the frame coverage periodically
- ✓ include descriptions in the survey documentation



References

Franklin, S., & Walker, C. (2010). Survey methods and practices. Statistics Canada. *Social Survey Methods Division, Ottawa*. (Originally published in October 2003)

Laiho, J., & Hietaniemi, L. (2002). Quality Guidelines for Official Statistics. Statistics Finland.