

# Techniques for data collection

Technical workshop on survey  
methodology:  
Enabling environment for sustainable  
enterprises in Indonesia

Hotel Ibis Tamarin, Jakarta  
4-6 May 2011

Presentation by  
Mohammed Mwamadzingo, ILO/ACTRAV Geneva

1

## The research process

- Topic
- Statement of the research problem
- Objectives
- Research questions
- Literature review (theoretical and empirical)
- **Data collection**
- Data analysis
- Report writing

2

## Data collection

### Organization of the presentation

- a) Data collection methods
- b) Sampling processes
- c) Data collection procedures
- d) Data collection forms (questionnaires)
- e) Exercise: Example of questionnaire to be used by trade unions in Indonesia (2011)

3

### (a) Data collection methods

- Once the research problem is identified and clearly defined, the research effort logically turns to data collection.
- Operating Rule:  
Consider a *survey* akin to *surgery*: to be used only after other possibilities have been exhausted.

4

## (a) Data collection methods

- First attempts at data collection should logically focus on secondary data.
  - **Secondary Data** = Statistics not gathered for the immediate study at hand, but for some other purpose.
  - **Primary Data** = Data originated by the researcher for the purpose of the investigation at hand.

5

## (a) Data collection methods: Using secondary data

### *Advantages of Secondary Data*

1. Cost and time economies

Secondary Data: Only one visit to the library, database.

Primary Data: Design forms, pre-testing forms, selection and training of staff, sampling, data collection, coding and tabulation, etc.

2. Help to better state the research problem.
3. Suggest improved methods or data for better understanding with the problem.
4. Provide comparative data by which primary data can be more insightfully interpreted.

6

## (a) Data collection methods: Using secondary data

### *Disadvantages of Secondary Data*

1. Problems of fit: Since secondary data are collected for other purposes, it will be rare that they fit the problem as defined perfectly

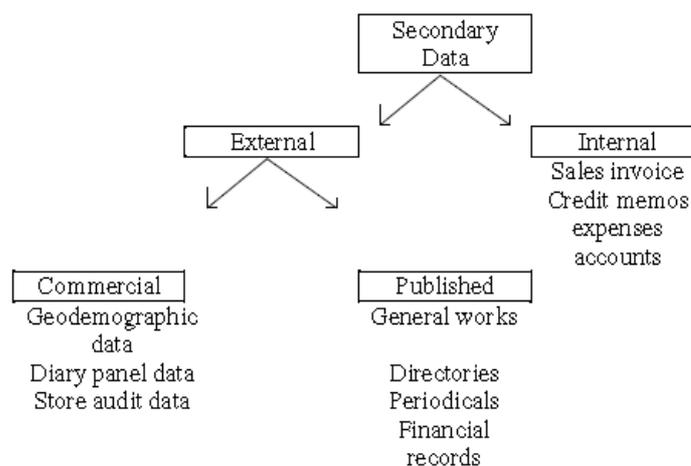
Problems relating to:

- (i) units of measurement
- (ii) class definitions
- (iii) publication currency

2. Problems of Accuracy

7

## (a) Data collection methods: Using secondary data



8

(a) Data collection methods:  
Using primary data

*Types of primary data*

**1. Demographic/socio-economic characteristics**

- Examples: age, education, occupation, marital status, sex, income.
- These variables are used to cross-classify the collected data and in some way make sense of them.
  - e.g., the workers' attitudes/opinions towards trade unionism in Kenya, vis-à-vis their education level.

9

(a) Data collection methods:  
Using primary data

*Types of primary data*

**2. Psychological/life-style characteristics**

- Examples: personality traits, activities, interests, and values.
  - e.g., personality (normal patterns of behaviour exhibited by an individual) influence choice of political party affiliation or interest in international, national or localised politics.

10

(a) Data collection methods:  
Using primary data

*Types of primary data*

**4. Awareness/knowledge**

- This refers to what respondents do and do not know about some object or phenomenon.
  - Workers' awareness on the effectiveness of political party ideologies on working conditions and welfare.
  - Which is most effective way of undertaking a recruitment campaign for trade union organizing? Door-to-door campaign, service to membership, political activism, use of print and electronic media?

11

(a) Data collection methods:  
Using primary data

*Types of primary data*

**3. Opinions/attitudes**

- What are attitudes? An individual's preference, inclination, views, or feelings towards some phenomenon.
- What are opinions? Verbal expressions of attitudes.

***But attitudes and opinions are normally used interchangeably***

12

(a) Data collection methods:  
Using primary data

*Types of primary data*

**5. Intentions**

- A person's intentions refer to the individual's anticipated or planned future behaviour.
  - What will workers do in case of changes in leadership at national trade union centre? In leadership of mainstream political parties?

13

(a) Data collection methods:  
Using primary data

*Types of primary data*

**6. Motivation**

- The concept of motivation seem to contain more semantic confusion than most terms in the behavioural sciences.
- For our purposes a 'motive' is a need, a want, a drive, an urge, a wish, a desire, an impulse, or any inner state that directs or channels behaviour towards goals.
- A trade unionists could be interested in finding out *why* there are normally more women voters than men.

14

## (b) Sampling process

The sampling process is about selecting those elements from which the information will be collected.

Two ways of selecting elements:

1. Census: from each member of the population  
 (Population is the totality of cases that conform to some designated specifications--thus requiring a clear definition of what constitutes the elements.  
 e.g., list of members in a given trade union, members of a political party, residents in a housing complex)
2. Sample: A portion of the population.

15

## (b) Sampling process

*Why take a sample?*

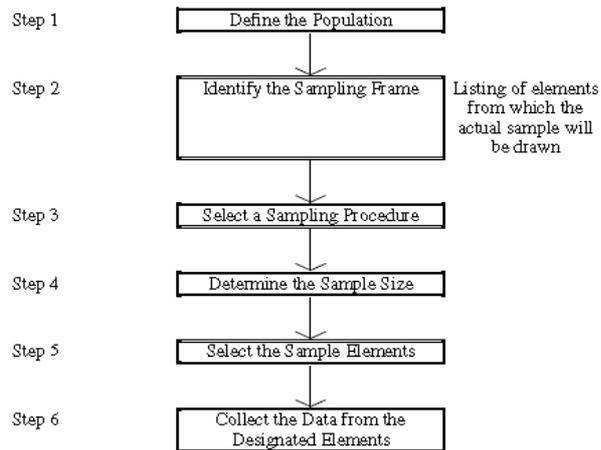
1. Complete count on populations of moderate size is very costly.
2. Census is time consuming.
3. Sometimes a census is impossible. e.g., Causing mental stress, anguish, and coercion to everybody could be undesirable!!
4. For purposes of accuracy: Census involves large field staff, thus introducing non-sampling errors.

16

## (b) Sampling process

### The How of Taking a Sample

#### *Required Steps*

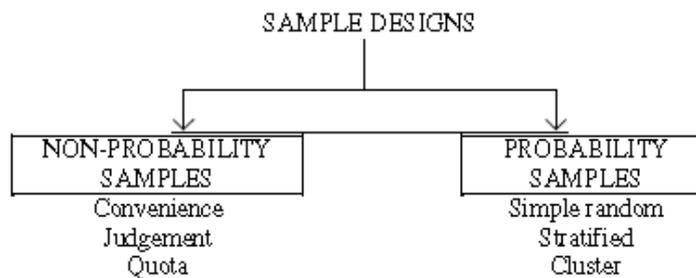


17

## (b) Sampling process

### Sampling Procedures

#### *Classification of Sampling Techniques*



18

## (b) Sampling process

### Probability samples:

each element has **known, non-zero** chance of being included in the sample.

#### (i) Simple Random Sampling:

each population element has not only a known but an equal chance of being selected.

$$\bar{x}_k = \sum_{i=1}^n X_i / n$$

19

## (b) Sampling process

### Probability samples:

(ii) Stratified Sample: is a probability sample that is distinguished by two-step procedure:

1. The parent population is divided into mutually exclusive and exhaustive subsets.
2. A simple random sample of elements is chosen independently from each group or subset

20

## (b) Sampling process

### **Probability samples:**

(iii) Cluster Sampling: Involves the following steps:

1. The parent population is divided into mutually exclusive and exhaustive subsets.
2. A random sample of the subsets is selected.
  - If the investigator then uses all of the population elements in the selected subsets for the sample, the procedure is one-stage cluster sampling. Otherwise, multi-stage cluster sampling.

21

## (b) Sampling process

### **Non-probability samples:**

we cannot estimate the probability that any population element will be included in the sample.

(i) Convenience samples: Accidental samples-the elements happen to be where the information is being collected.

22

## (b) Sampling process

### **Non-probability samples:**

(ii) Judgement Samples: Purposive samples-the sample elements are hand-picked because it is expected that they can serve the research purpose.

- The **snowball sample** is a judgement sample that is sometimes used to sample special populations. This relies on the researcher's ability to locate an initial set of respondents with the desired characteristics. These individuals are then used as informants to identify others with the desired characteristics.

23

## (b) Sampling process

### **Non-probability samples:**

(iii) Quota samples: Certain characteristic is approximately the same as the proportion of the elements with the characteristic in the population.

24

## (b) Sampling process

### Determination of sample size

The question of sample size is complex because it depends on (among other things):

- the type of sample
- the statistic in question
- the homogeneity of the population
- time, money, and personnel available for the study

25

## (b) Sampling process

### Determination of sample size

#### *Basic considerations*

1. The standard error of the estimate obtained from the known sampling distribution (which indicates how the sample estimate vary as a function of the particular sample selected) of the statistic.
2. Precision desired from the estimate. Precision is the size of the estimating interval when the problem is one of estimating a population parameter.
3. The desired degree of confidence associated with the estimate.

26

## (b) Sampling process:

### Determination of sample size

**A: Population variance known    B: Population variance unknown**

$$n = \frac{Z^2}{H^2} \sigma^2$$

n = sample size

Z = size of population

H = half length of the interval

$\sigma$  = population variance (known from previous studies)

$$n = \frac{Z^2}{H^2} \text{est } \sigma^2$$

27

## (c) Data collection procedures

- Interview
- Telephone
- Postal/email survey
- Focus group discussions

Discuss advantages and disadvantages of each.

28

### (d) Data collection forms (questionnaires)

A questionnaire is a means of eliciting the feelings, beliefs, experiences, perceptions, or attitudes of some sample of individuals.

A questionnaire could be structured or unstructured.

A questionnaire is a written or printed form.

29

### (d) Data collection forms (questionnaires)

#### **Advantages**

- Economy - Expense and time involved in training interviewers and sending them to interview are reduced by using questionnaires.
- Uniformity of questions - Each respondent receives the same set of questions phrased in exactly the same way. Questionnaires may, therefore, yield data more comparable than information obtained through an interview.
- Standardization - If the questions are highly structured and the conditions under which they are answered are controlled, then the questionnaire could become standardized.

30

## (d) Data collection forms (questionnaires)

### **Disadvantages**

- Respondent's motivation is difficult to assess, affecting the validity of response.
- Unless a random sampling of returns is obtained, those returned completed may represent biased samples.

31

## (d) Data collection forms (questionnaires)

### **Characteristics of a good questionnaire**

- Deals with a significant topic.
- Seeks only that information which cannot be obtained from other sources.
- As short as possible, only long enough to get the essential data.
- Attractive in appearance, neatly arranged, and clearly duplicated or printed.
- Directions are clear and complete.
- Questions are objective, with no leading suggestions to the desired response.
- Questions are presented in good psychological order, proceeding from general to more specific responses.
- Easy to tabulate and interpret.

32

## (e) Example

- Example of questionnaire to be used by trade unions in Indonesia (2011)

33

## **Techniques for data collection**

**Thank you**

34