1. Introduction

Welcome to the series of E-learning modules on Secondary Data. In this module we are going to cover the importance and sources of Secondary Data.

At the end of this session, you will be able to:

- Explain the meaning and importance of secondary data

- Explain the sources of Secondary data

- Explain the major sources of government publication

We face problems in various fields of our life, which force us to think and discover their solutions. When we are genuinely serious about the solution of a problem faced, a thinking process starts.

Statistical thinking or Statistical inquiry is one kind of thinking process which requires evidence in the form of some information, preferably quantitative, which is known as data or statistical information.

In a statistical inquiry, the first step is to procure or collect data. Every time the investigator may not start from the very beginning. He must try to use what others have already discovered, this will save in cost, efforts and time.

Statistics (that is in singular sense) is concerned with collection of data relevant to the solution of a particular problem.

According to Simpson and Kafka (Basic Statistics),"Data have no standing in themselves; they have a basis for existence only where there is a problem."

For example: If we want to analyze the sales pattern of an organization we need to gather information with respect to the number of sales happening in a day, sales done by various groups of salesmen, what is the lean period, effect of competitive products, income behavior and related matters which will help in gathering information called data set. This data set can be used for various other analyses like understanding the consumer's behavior, competitive products performance and so on.

Data can be classified as Quantitative and Qualitative Data, Sample and Census Data, Primary and Secondary data.

Quantitative and Qualitative Data:

Quantitative data are those set of information which are quantifiable and can be expressed in some standard units like rupees, kilograms, litres, etc.

Qualitative data on the other hand, are not quantifiable, that is, cannot be expressed in standard units of measurement they are 'features', 'qualities' or 'characteristics' like eye colours, skin complexion, honesty, good or bad, etc.

Sample and Census Data:
Information collected through sample inquiry is called sample data.

Whereas information collected through census inquiry is called census data. Example: Population census data are collected every ten years in India.

Primary and Secondary Data:
Primary data are collected by the investigator through field survey. Such data are in raw form and must be refined before use.

On the other hand, Secondary data are extracted from the existing published or unpublished sources that are from the data already collected by others.

Data are obtained through two types of investigations, namely, Direct Investigation which implies that the investigator collects information by observing the items of the problem under investigation. Investigation through Secondary Source which means obtaining data from the already collected data.

2. Definition ,Meaning and Examples of Secondary Data

Definition:
According to M.M. Blair, "Secondary data are those already in existence for some other purpose than the answering of the question in hand."

Meaning:
The statistical material which is not originated by the investigator himself but obtained from someone else's records, or for any other purpose at some subsequent enquiry it is termed as Secondary data.

This type of data is generally taken from newspapers, magazines, bulletins, reports, journals etc. Example: If the data published by Reserve Bank of India on currency, National Income, Exports or Imports, is used in some other statistical enquiry, it will be termed as Secondary data.

Secondary data review and analysis:
Secondary data review and analysis involves collecting information, statistics, and other relevant data at various levels of aggregation in order to conduct a situational analysis of the area.

Secondary data analysis can be literally defined as "second-hand" analysis.

It is the analysis of data or information that was either gathered by someone else (e.g., researchers, institutions, other NGOs, etc.) or for some other purpose than the one currently being considered, or often a combination of the two.

Example of secondary data are:
The following is a sampling of the types of secondary data and information commonly associated with poverty analysis:

- Demographic (population, population growth rate, rural or urban,gender, ethnic groups, migration trends, etc.)
- Discrimination (by gender, ethnicity, age, etc.)
- Gender equality (by age, ethnicity, etc.)
- Policy environment
- Economic environment (growth, debt ratio, terms of trade)
- Poverty levels (poverty and absolute)
- Employment and wages (formal and informal; access variables)

- Livelihood systems (rural, urban, on-farm, off-farm, informal, etc.)
- Agricultural variables and practices (rainfall, crops, soil types, and uses,irrigation, etc.)
- Health (malnutrition, infant mortality, immunization rate, fertility rate,contraceptive prevalence rate, etc.)
- Health services (level, services by level, facility-to-population ratio;access by gender, ethnicity. etc.)
- Environmental status and problems
- Education (adult literacy rate, school enrolment, drop-out rates, male-to-female ratio, ethnic ratio, etc.)
- Schools (level, school-to-population ratio, access by gender, etc.)
- Infrastructure (roads, electricity communication, water, sanitation, etc.)
- Harmful cultural practice

3. Characteristics and Sources of Secondary Data

Let us now move forward to the characteristics of Secondary Data:

Characteristics of secondary data are:

1. Reliability
2. Suitability
3. Adequacy

1. Reliability of data: Reliability is an essential pre-requisite for validity. Hence to check the reliability of the secondary data it is better to ask questions like who collects the data, what are the sources of the data, where they collected by using proper methods, time of collection of the data, any bias of the complier, what level of accuracy was desired, was it achieved. These points will help in realizing the reliability of the data.
2. Suitability of data: Data are generally collected with a specific scope, objective and nature. Hence, it is necessary to found whether the secondary data that is collected for one enquiry is suitable for another enquiry. Thus, the data should be checked for various terms, units of collection and the time of collection of the data from the primary source.
3. Adequacy of data: The available data should be checked for its level of accuracy, authenticity, as needed for the present level of enquiry, if found suitable then the data is called adequate data and can be used as it will be economical. If the data collected is narrower or wider than the area of the present enquiry then it is considered to be inadequate.

Importance of Secondary Data:
If secondary research and data analysis is undertaken with care and diligence, it can provide a cost-effective way of gaining a broad understanding of research questions.

Secondary data are also helpful in designing subsequent primary research and can provide a baseline to compare your primary data collection results. Therefore, it is always wise to begin any research activity with a review of the secondary data (Novak 1996).

There is a vast amount of published information from which statistical studies are made; also efforts are put in constantly to produce fresh statistics for research purposes. The sources of secondary data are broadly classified under two headings:

1. Published sources
2. Unpublished sources

Published sources are:

- Reports and official publication of the government
- Official publication of foreign countries
- Semi official publications
- Publications of autonomous and private institutions
- Financial and economic Journals
- Annual reports of Joint stock companies
- Publications bought by various autonomous research institutes
- Reports of Committee and Commissions
- Technical Reports
- Scholarly Journals
- Literature Review Articles
- Trade Journals
- Reference Books
- Newspapers and Magazines

Unpublished sources are:

- Unpublished findings of certain inquiry committees
- Diaries
- Letters
- Biographies
- Auto biographies

Published sources:
Generally, published sources are international, national, govt., semi-Govt, private corporate bodies, trade associations, expert committee and commission reports and research reports.

They collect the statistical data in different fields like national income, population, prices, employment, wages, export, import etc. These reports are published on regular basis i.e., annually, quarterly, monthly, fortnightly, weekly, daily and so on. These published sources of the secondary data are given below.

Reports & official publication of the government:

Central Statistical Organisation (C.S.O.): It publishes data on national income, savings, capital formation, etc. in a publication called National Accounts Statistics.

National Sample Survey Organisation (N.S.S.O.): Under Ministry of Statistics and Programme Implementation, this organisation provides us data on all aspects of national economy, such

as agriculture, industry, and labour & consumption expenditure.

Reserve Bank of India Publications (R.B.I): It publishes financial statistics. Its publications are Report on Currency and Finance, Reserve Bank of India Bulletin, Statistical Tables Relating to Banks in India, etc.

Labour Bureau: Its publications are Indian Labour Statistics, Indian Labour Year Book, Indian Labour Journal, etc.

Population Census: Undertaken by the office of the Registrar General India, Ministry of Home Affairs. It provides us different types of statistics about population.

Official publication of foreign countries:

All foreign governments and international agencies publish regular reports of international significance these contain reports which are regularly published by the agencies like; International bodies such as the world bank, International Labor Organization, Statistical office of the United Nations, International Monetary Fund, United Nations Organization, World Health Organization, Food and Agriculture Organization, International Bank for Reconstruction and Development, World Meteorological Organization.

Semi official publications:

Semi official publications include publication of information by various local bodies such as the Municipal Corporations, District bodies, etc.

Semi government, organizations municipalities, District Boards and others also publish reports in respect of birth, death and education, sanitation and many other related fields.

4. Sources of Secondary Data Continued

Publications of Private and Autonomous Institutions:

Trade and professional bodies like Federation of Indian Chambers of Commerce and Industry, Institute of Chartered Accountants, The Institute of Foreign Trade.

The prestigious journals of these institutes are 'Economic Trends', 'The Chartered Accountant' and the 'foreign Trade Review'.
Financial & Economic Journals

Financial and economic journals like:

- Financial analysts journal
- The journal of business
- The journal of political economy
- Applied economics
- Ecological Economics etc.

Annual reports of Joint stock companies

Annual reports of joints stock companies are also the published source of secondary data.

Publications bought by various Research Institutes:

Some commercial and research institutes publish reports regularly. They are like Institutes of Economic Growth, Stock Exchanges, National Council of Education Research and Training (NCERT), National Council of Applied Economic Research (NCAER) etc.

Reports of Committee and Commissions:

Central Government, or State Government, sometimes appoints committees and commissions on matters of great importance. Reports of such committees are of great significance as they provide invaluable data. These reports are like, Shah Commission Report, Sarkaria Commission Report and Finance Commission Reports etc.

Technical Reports:

Technical reports are accounts of work done on research projects.

They are written to provide research results to colleagues, research institutions, governments, and other interested researchers.

A report may emanate from completed research or on-going research projects.

Scholarly Journals:

Scholarly journals generally contain reports of original research or experimentation written by experts in specific fields.

Articles in scholarly journals usually undergo a peer review where other experts in the same field review the content of the article for accuracy, originality, and relevance.

Literature Review Articles:

Literature review articles assemble and review original research dealing with a specific topic.

Reviews are usually written by experts in the field and may be the first written overview of a topic area.

Review articles discuss and list all the relevant publications from which the information is derived.

Trade Journals:

Trade journals contain articles that discuss practical information concerning various fields. These journals provide people in these fields with information pertaining to that field or trade.

Reference Books:

Reference books provide secondary source material. In many cases, specific facts or a summary of a topic is all that is included. Handbooks, manuals, encyclopaedias, and dictionaries are considered reference books.

Newspapers and Magazines:

Various newspapers as well as magazines also do collect data in respect of many social and economic aspects. Some of them are as: Economic Times, Financial Express, Hindustan Times, Indian Express, Business Standard, Economic and Political Weekly, Main-stream, Business Outlook and The WEEK etc.

Problems faced in published sources:

The published sources are sources that have been collected by various agencies. These are useful to researchers as the information is easily obtainable, comprehensive and covers information of long time periods.

However, this information is often "characterized by unreliability, data gaps, over-aggregation, inaccuracies, mutual inconsistencies, and lack of timely reporting." It is important to critically analyze these data for accuracy and validity. There are several reasons why these problems exist:

1. The scale of official surveys generally requires large numbers of enumerators (interviewers) and, in order to reach those numbers enumerators contracted are often under-skilled;

2. The size of the survey area and research team usually prohibits adequate supervision of enumerators and the research process; and

3. Resource limitations (human and technical) often prevent timely and accurate reporting of results.

Unpublished Sources:

Diaries: A diaries are sources where information is gathered by individuals on time spent on professional or personal journals of thoughts. Diaries can record either quantitative or qualitative data, and provides information in management research about work patterns and activities.

Letters: Letters like diaries are also a rich source but should be checked for their reliability before using them.

For example A commercial business letter is a letter written in formal language, usually used when writing from one business organization to another, or for correspondence between such organizations and their customers, clients and other external parties.

There are many reasons to write a business letter. It could be to request direct information or action from another party, to order supplies from a supplier, to identify a mistake that was committed, to reply directly to a request, to apologize for a wrong or simply to convey goodwill all of these can be a source of information.

Biographies and Autobiographies: These are also important sources of information.
A biography is a detailed description or account of someone's life from birth to death whereas autobiography is a biography of a person written by that person.

5. Merits and Demerits of Secondary Data

Merits of secondary data are:

- Secondary data analysis can be carried out quickly when compared to formal primary data gathering and analysis exercises

- Where good secondary data is available, researchers save time and money by making good use of available data rather than collecting primary data, thus avoiding duplication of effort

- Using secondary data provides a relatively low-cost means of comparing the level of well-being. However, keep in mind that data collection methods vary (between researchers, countries, departments, etc, which may impair the comparability of the data

- Depending on the level of data disaggregation, secondary data analysis lends itself to trend analysis as it offers a relatively easy way to monitor change over time

- It informs and complements primary data collection, saving time and resources often associated with over-collecting primary data

Demerits of secondary data are:

- Secondary data helps us understand the condition or status of a group, but compared to primary data they are imperfect reflections of reality. Without proper interpretation and analysis they do not help us understand why something is happening

- The person reviewing the secondary data can easily become overwhelmed by the volume of secondary data available, if selectivity is not exercised

- It is often difficult to determine the quality of some of the data in question

- Sources may conflict with each other

- As the secondary data is usually not collected for the same purpose as the original researcher had, the goals and purposes of the original researcher can potentially bias the study

- As the data were collected by other researchers, they decide what to collect and what to omit, hence all of the information desired may not be available.

Effective presentation of secondary data:

- A well-documented secondary data review and analysis allows for easier use of the material by other interested parties

- Properly citing the publication date of the sources you used will allow subsequent researchers to use your work to make comparisons over time and between countries,

communities, towns, regions, etc

- Proper citation allows subsequent researchers to use your work, thus preventing unnecessary duplication of research efforts

Conclusion:

Secondary data review and analysis is a form of research and data compilation that is demanding and time-consuming; however, without proper citation (i.e., author, date, title) of materials that you used, your work will often be disregarded as it will only have limited use by those who wish to follow in your footsteps.

Here's a summary of our learning in this session:

- Meaning and importance of secondary data

- Sources of Secondary data

- Major sources of government publication