**Data collection** is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes. The data collection component of research is common to all fields of study including physical and social sciences, humanities, business, etc. While methods vary by discipline, the emphasis on ensuring accurate and honest collection remains the same.

**The factors influencing choice of Data Collection Method**

- Time Factor
- Cost Factor
- Quality of Decision-Making
- Type of Research Work
- Nature of Researcher
- Objectives of Research
- Availability of Respondents
- Availability of Research Staff

**There are 2 types of data, namely:**

Primary data is data that is collected by a researcher from first-hand sources, using methods like surveys, interviews, or experiments. It is collected with the research project in mind, directly from primary sources.

Secondary data is public information that has been collected by others. It is typically free or inexpensive to obtain and can act as a strong foundation to any research project — provided you know where to find it and how to judge its worth and relevance.
Advantages and limitations of primary data

<table>
<thead>
<tr>
<th>PRO’S</th>
<th>CON’S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provides first-hand information</td>
<td>Paper work</td>
</tr>
<tr>
<td>In depth information</td>
<td>Expensive</td>
</tr>
<tr>
<td>In depth information</td>
<td>Sampling errors</td>
</tr>
<tr>
<td>Reliable information</td>
<td>Time consuming</td>
</tr>
<tr>
<td>Accurate data</td>
<td>Respondent bias</td>
</tr>
<tr>
<td>Specific data</td>
<td>Interviewer bias</td>
</tr>
<tr>
<td>Supplement secondary data</td>
<td>Processing of data</td>
</tr>
<tr>
<td>Enhances the quality of research work</td>
<td>Problem of quick decision making</td>
</tr>
<tr>
<td>Helps in formulation of hypothesis</td>
<td></td>
</tr>
<tr>
<td>Flexibility</td>
<td></td>
</tr>
<tr>
<td>Helps to overcome resistance</td>
<td></td>
</tr>
</tbody>
</table>
# Methods of Primary Data Collection

## Observation Method
- a) Structured and Unstructured Observation
- b) Disguised and Undisguised Observation
- c) Mechanical Observation

## Experimental Method
- First Hand information
- Reliable and relevant information
- Develop new techniques

## Interview Method
- face to face interaction
- Reliability
- Detailed Information

## Survey Method

### Schedules

**PURPOSE**
- To provide a standardized tool
- To act as memory tickler
- To facilitate the work of tabulation and analysis

**TYPES**
- Rating Schedule
- Documents Schedule
- Survey Schedule
- Observation Schedule
- Structured or unstructured

**FEATURES**
- Personal contact
- Nature of Respondents
- Response Rate
- Area Coverage
- Use of computers for data collection
**Questionnaire**

A Questionnaire is a set of questions, which act as an instrument to collect data from the respondents to a survey or interview.

**SIGNIFICANCE**
- Relevant Data
- Convenience to the Respondents
- Structured and Unstructured Responses
- Qualitative and Quantitative Data
- Large Coverage
- Proper Processing
- Easy to Alter
- Sensitive Information

**ESSENTIALS OF GOOD QUESTIONNAIRE**
- Relevant questions
- Clarity
- No. of questions
- Types of questions
- Decision on wording
- Sequence on questions
- Physical appearance:
  - a) Paper quality & colour
  - b) size
- Pilot study

**Sources of Secondary data**

- **Internal sources**
  - Purchase and sales record
  - Operational records or unit wise production
  - Debtor’s and creditor’s record
  - Financial statements

- **External sources**
  - Employee’s record including performance claims etc.
  - General publications
  - Syndicated sources
  - Government publications

---

LSRC/2019-20/tutoriallesson  Page 4
**Advantages and limitations of secondary data.**

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less expensive</td>
<td>Problem of accuracy</td>
</tr>
<tr>
<td>Less time consuming</td>
<td>Problem of reliability</td>
</tr>
<tr>
<td>Less processing of data</td>
<td>Problem of adequacy</td>
</tr>
<tr>
<td>Quick decisions</td>
<td>Lack of in-depth option</td>
</tr>
<tr>
<td>Supplements primary data</td>
<td>Problem in quality decision making</td>
</tr>
<tr>
<td>Less paper work</td>
<td>Problem of specific data</td>
</tr>
<tr>
<td>Large volume of data</td>
<td>Unsuitability</td>
</tr>
<tr>
<td>No sampling errors</td>
<td>Problem of biased information</td>
</tr>
</tbody>
</table>

**Sampling**

Sampling design is a plan designed to select the appropriate sample in order to collect the right data as to achieve research objectives.

*Donald Tull and Dell Hawkins defines sample as “those individuals chosen from the population of interest as subjects in an experiment or to be the respondents to a survey. “*
**Methods of sampling:**

![Sampling Method/Sampling Techniques]

**Sample Size.**

There is a need to select the right sample size. Over-sized sample of respondents may lead to waste of time, efforts and money. Under-sized sample of respondents may lead to inaccuracy of data, and therefore, poor research results. Therefore, there is a need for appropriate size of sample of respondents for research activity.

Factors such as:

- Area of Research
- Availability of funds
- Availability of manpower
- Time frame
- Nature of Research
- Method of Sampling
- Method of Data collection
- Judgement of the researcher
- Precision / Accuracy
Stages in Data Processing.

Processing of data is the process of editing, coding, classification, tabulation, and graphic presentation of data. Data processing is required for the purpose of analysis and interpretation.

The various stages in data processing are as follows:

PIE CHART:

A pie chart is a circular chart used to compare parts of the whole. It is divided into sectors that are equal in size to the quantity represented. It is diagrammatically represented as:
**BAR GRAPHS:**

A Bar chart or a bar graph is a chart with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. Following are the diagrammatical representation of bar chart/bar graph.

![Bar Graph Diagram]

**CHARACTERISTICS OF A GOOD MEASURE OF CENTRAL TENDANCE:**

- It should be simple to calculate and easy to understand
- It should be rigidly defined
- It should be based on all the observations
- It should not be affected by extreme items
- It should be capable of further algebraic treatment
- It should have sampling stability
- It can be easily calculated in the case of distribution containing open end class-intervals
- It should be in the form of mathematical formula
### MERITS AND DEMERITS OF ARITHMETIC MEAN

<table>
<thead>
<tr>
<th>MERITS</th>
<th>DEMERITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to calculate</td>
<td>Absurd results</td>
</tr>
<tr>
<td>Rigidly defined</td>
<td>Not suitable for incomplete data</td>
</tr>
<tr>
<td>Based on all items</td>
<td>Problem of extreme values</td>
</tr>
<tr>
<td>Further mathematical treatment</td>
<td>Mean value may not be in the data</td>
</tr>
<tr>
<td>Stability</td>
<td>Problem of open-end class intervals</td>
</tr>
<tr>
<td>Comparison</td>
<td>Limited application</td>
</tr>
<tr>
<td>Arrangement of order</td>
<td>-</td>
</tr>
<tr>
<td>Determining the values</td>
<td>-</td>
</tr>
</tbody>
</table>

### MERITS AND DEMERITS OF GEOMETRIC MEAN

<table>
<thead>
<tr>
<th>MERITS</th>
<th>DEMERITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rigidly defined</td>
<td>Difficult to understand and compute</td>
</tr>
<tr>
<td>Less affected by extreme values</td>
<td>Value may not be in the data</td>
</tr>
<tr>
<td>Useful to obtain average</td>
<td>Difference in the value of ratio change</td>
</tr>
<tr>
<td>Capable of further algebraic treatment</td>
<td>More weightage to smaller items</td>
</tr>
<tr>
<td>Least affected by fluctuations of sampling</td>
<td>Cannot be valued if negative numbers</td>
</tr>
<tr>
<td>Can average and construct index number</td>
<td>-</td>
</tr>
<tr>
<td>Suitable in social and economic areas</td>
<td>-</td>
</tr>
</tbody>
</table>

### MERITS AND DEMERITS OF MEDIAN

<table>
<thead>
<tr>
<th>MERITS</th>
<th>DEMERITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to calculate</td>
<td>Based on certain items</td>
</tr>
<tr>
<td>Found by mere inspection</td>
<td>Affected by sampling fluctuations</td>
</tr>
<tr>
<td>Not affected by extreme values</td>
<td>Not capable for further calculation</td>
</tr>
<tr>
<td>Value exists in the data</td>
<td>Prior arrangement of data required</td>
</tr>
<tr>
<td>Appropriate average</td>
<td>Problem of vast data</td>
</tr>
<tr>
<td>Can be calculated from incomplete data</td>
<td>Ignores extreme values</td>
</tr>
</tbody>
</table>
**MERITS AND DEMERITS OF MODE:**

<table>
<thead>
<tr>
<th>MERITS</th>
<th>DEMERITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to calculate</td>
<td>Not defined rigidly</td>
</tr>
<tr>
<td>Representative</td>
<td>Not based on all items</td>
</tr>
<tr>
<td>Not affected by extreme values</td>
<td>Sampling fluctuations</td>
</tr>
<tr>
<td>Calculated in open-end class intervals</td>
<td>No further mathematical treatment</td>
</tr>
<tr>
<td>No need to know all values</td>
<td>Limited scope</td>
</tr>
<tr>
<td>Inspection</td>
<td>Poor measure</td>
</tr>
<tr>
<td>Determination through graph</td>
<td>-</td>
</tr>
<tr>
<td>Most descriptive average</td>
<td>-</td>
</tr>
</tbody>
</table>

**USE OF COMPUTER AND INTERNET IN DATA COLLECTION AND PROCESSING:**

- Conception phase
- Design and planning phase
- Empirical phase
- Analysis phase
- Dissemination phase

**GOOD REPORT WRITING:**

- A report is a statement of facts and figures, prepared for the purpose of information and action. The ‘oxford dictionary’ defines a report as “a record of ascertained facts”.

**ESSENTIALS:**

- INFORMATIVE
- CLARITY
- CONCISE
- ACCURACY
- RELIABILITY
- OBJECTIVITY
- LOGICAL ARRANGEMENT
- SECRECY
- TIMELY SUBMISSION OF REPORTS
- REFERENCES
- IMPERSONAL STYLE
- PROPER FORMAT
- SIGNATURE AND DATE
 STRUCTURE/LAYOUT OF RESEARCH REPORT:

There are 3 types of report:

- TECHNICAL
- INTERIM
- SUMMARY

Use these points and solve all the questions given in question bank