

Unit -2 DATA COLLECTION

2.1 Difference between term Data & Information

Basis	Data	Information
Meaning	Raw, unorganized facts and figures without context.	Processed, organized, and meaningful data.
Form	Numbers, symbols, characters, or raw facts.	Summarized, analyzed, and structured facts.
Context	Lacks meaning without processing.	Has context and relevance for decision-making.
Example	100, 200, 300 (just numbers).	"Sales increased from 100 to 300 units in 3 months."
Purpose	To be processed or analyzed to gain meaning.	To provide insights, knowledge, or support decisions.
Processing	Needs processing to become useful.	Already processed and useful.
Dependency	Cannot be used directly for decisions.	Directly useful for decision-making.

2.2 Differentiate between Primary & Secondary data

Meaning Primary Data collected first-hand by the researcher for a specific purpose.

Meaning Secondary Data, which is collected by **someone else** and already available for use.

Basis	Primary Data	Secondary Data
Meaning	Data collected first-hand by the researcher for a specific purpose.	Data collected by someone else and already available for use.
Source	Direct sources like surveys, interviews, observations, experiments.	Indirect sources like books, reports, journals, websites, government records.
Specificity	Specific to the researcher's objective .	May or may not be directly related to the current research objective.
Cost	Usually expensive and time-consuming to collect.	Generally cheaper and quicker to obtain.
Time Required	More time-consuming to collect and process.	Requires less time as data is readily available.
Accuracy & Reliability	High accuracy if collected properly; directly under researcher's control.	May have questionable accuracy depending on the source.
Up-to-date	Data is usually current and recent .	Data may be outdated or obsolete.

2.2.1 Primary Data collected first-hand by the researcher for a specific purpose.

Merits of Primary Data

1. **Specific and Relevant:** Directly collected for the particular problem.
2. **High Accuracy:** Controlled by the researcher, ensuring reliability.
3. **Recent Information:** Data is fresh and up to date.

4. **Flexibility:** The researcher can choose the methods best suited to the study.

Demerits of Primary Data

1. **Costly:** Requires a high budget for surveys, experiments, or fieldwork.
2. **Time-Consuming:** Data collection and processing take a longer time.
3. **Requires Expertise:** Need trained personnel and careful planning.
4. **Limited Scope:** May not be feasible for very large populations or wide coverage.

2.2.2 Secondary Data, which is collected by **someone else** and already available for use.

Merits of Secondary Data

1. **Economical:** Less costly as data is already available.
2. **Time-Saving:** Can be collected and analyzed quickly.
3. **Easy Accessibility:** Easily available from books, journals, websites, etc.
4. **Helpful for Comparison:** Useful for historical analysis and comparisons over time.

Demerits of Secondary Data

1. **Less Reliable:** May lack authenticity if the source is not credible.
2. **Irrelevance:** May not perfectly match the current research needs.
3. **Outdated Information:** Data might be old or no longer applicable.
4. **Lack of Control:** Researcher cannot control data quality or methodology used in data collection.

2.3 Source of secondary data

1. Internal sources of secondary data:

This is one of the basic sources of secondary data, where information is available from inside the organization, from the following sources.

1. Accounting records:

Accounting records contain all the information regarding different accounting information such as source of income, expenses, comparative accounting statements, profit and loss statement, etc.

2. Sales force report:

Sales force reports provide information about different sales force activities such as sales expenses, sales performance, area-wise sales report, salesman-wise sales report, product-wise sales report, etc.

3. Internal expert:

Internal experts mean employees working within the company, who have highly experience, highly qualified, can be a good source of information.

4. Miscellaneous report:

Which include different reports which are routine reports and generated regularly such as daily attendance report, daily stock report, maintenance report, purchase & sales report, etc. will be handy in secondary data collection.

External Sources of Secondary data:

Apart from internal sources of secondary data there is a large amount of information which is available outside of the organization from following sources,

1. Government publication
2. Non Government publication.

1. Government Publication:

There are certain information and report which are generated by government, which are valuable source of information for any type of research, government publication can be of following type,

1.Regular publication:

Government generate certain report on a regular basis RBI bulletin, report of different ministry, census report (every ten year), etc. are generated by government on a regular basis.

2. Ad-hoc publication:

Such report are not generated by government on a regular basis, which may be available on demand such as inquiry report, certain ministry report, information through Right to information act

2.Non Government Publication

There are other information which are available from non government publication, from following sources,

1. Industry & Trade association:

Industry & trade association such as textile , diamond, and almost every product category in industry has their trade association, who publish their annual report, which include amount of production in industry, total export, import and other information regarding performance of their industry.

2. Universities

University which offers different courses where research work is conducted by faculty members and student related to industrial problems apart of their curriculum activity, which can be useful in research work.

3. Syndicated services

Syndicate means a group of individual or organization combines to promote common interest, which is somewhat similar to trade association. Which also produce report about their industry or common interest group.

4. Consumer research services.

These are independent research organization, which conduct survey independently and publish their report for general public, as well as they provide paid research services as well.

5. Publication of International organization

There are number of international organization who produce their report regularly, which can be useful in many research, such as report by WHO, UNICEF, WORLD BANK, FORBES list of different category.

3.5 General idea of survey

What is a survey

A survey is a research method used for collecting data from a predefined group of respondents to gain information and insights into various topics of interest.

Process of conducting survey

1. Preliminary step
2. Preparation of Question

3. Preparation of questionnaire
4. Training to the data collector
5. Pilot survey
6. Actual field survey
7. Post survey work

1. Preliminary step:

At this stage general management & business problem are derived, which also involve target population, sampling method, budget of survey, data collection method etc. are decided.

2. Preparation of Question:

After deciding basic methodology, different questions are designed to collect the response from the respondent, which include following types of questions

1. Open ended questions
2. Close ended questions

1. Open ended question

1. Dichotomous question
2. Single choice question
3. Multiple choice question
4. Declarative question
5. Rating question
6. Ranking question

2. Close ended question

1. Feedback question

3. Preparation of questionnaire

Questionnaire is a set of question designed to get answer from respondent to answer the research question, while preparing questionnaire certain things need to be consider such as, formatting , avoid technical words , logical sequence, multiple language etc.

4. Training to the data collector

When survey is of large scale and multiple data collector are appointed than they are provided with training about how to approach respondent, how to collect data, how to handle queries. While collecting data generally experience data collector are mostly preferred.

5. Pilot survey

It is also known as testing survey where few actual respondents are contacted and they are asked to fill up the questionnaire. From this testing data collector come to know about errors in questionnaire or problem in questions while collecting data. So that proper modification can be made questionnaire before conducting actual survey

6. Actual field survey

After examining the questionnaire by pilot testing next task is to conduct actual fieldwork, where actual respondents are contacted to get information about research, in this step two activities are mainly focused,

1. Supervision
2. Guidance

This stage is most time consuming step as compare to any other stage in survey process.

7. Post survey work

After completing the task of data collection the next activity is to conduct post survey activity such as editing, coding, tabulation, analysis, interpretation, of collected data to get the answer of research question.

Survey Methods

There are basically 4 types of survey methods

1. Personal interview:

A personal interview survey, also called as a face-to-face survey, is a survey method that is utilized when a specific target population is involved. The purpose of conducting a personal interview survey is to explore the responses of the people to gather more and deeper information.

1. Door to Door interview
2. Executive Interview
3. Mall intercept survey

Types of Personal interview

1. Structured Interview
2. Unstructured interview
3. Focused interview
4. Clinical interview
5. Non Directive interview
6. C A P I

Telephonic Survey

A telephone survey is one of the survey methods used in collecting data either from the general population or from a specific target population over telephone. Telephone numbers are utilized by trained interviewers to contact and gather information from possible respondents via telephone.

Step to perform telephonic survey

1. Select respondent and collect their contact detail.
2. Make telephone call.
3. Introduction and gain rapport with respondent.
4. To conduct interview or clarify convenient time of interview.
5. Schedule a call report.

E- Mail survey

An email survey is one that sends the survey instrument (e.g. questionnaire) to a respondent via email and most often samples respondents via email. These electronic mail surveys first came into use in the late 1980s, and many scholars at the time thought that they represented the future of survey research

This method has highest possibility to cover wide geographical area who all has internet access. It is one of the latest techniques of conducting survey

Post mail survey

A mail survey is one in which the postal service, or another mail delivery service, is used to mail the survey materials to sampled survey addresses.

In this method responses are recorded by respondent themselves without influence of interviewer.

1. Deliver by a post or courier

2. Attaching questionnaire to the product.
3. Advertising questionnaire in a news paper.

Process of post mail survey

1. Preparing a list of respondent with mailing address.
2. Preparation of cover letter along with questionnaire, which explain purpose and importance of study as well as introduction.
3. Advance information.
4. Sending questionnaire.
5. Reminder.
6. Follow up.

How to improve response rate in Post mail survey

1. Quality printing.
2. Cover letter.
3. Advance information.
4. Incentive (Free gift, discount coupon).
5. Large sample size.
6. Proper follow up.

2.6 Compare four survey methods on following parameter.

Meaning of Parameters

1. **Social Desirability:** Tendency of respondents to answer in a way that is viewed favorably by others.
2. **Perceived Anonymity:** How anonymous or private respondents feel their answers are.
3. **Response Rate:** Percentage of people who complete and return the survey.
4. **Control on Data Collection Environment:** How much the researcher controls the setting where responses are collected.
5. **Diversity of Questions:** Ability to use different types of questions (open-ended, close-ended, visuals, etc.)
6. **Interviewer's Bias:** Influence of the interviewer's behavior or presence on the respondent's answers.
7. **Use of Physical Stimuli:** Ability to show materials like products, samples, charts during the survey.
8. **Sample Control:** Ability to ensure the selected sample actually participates.
9. **Quantum of Data:** Amount and depth of information collected.

Parameter	Personal Survey	Telephonic Survey	Electronic Survey	Postal Mail Survey
1. Social Desirability	High — Respondents may give socially acceptable answers due to face-to-face interaction.	Moderate — Less pressure than face-to-face but still some influence.	Low — Less direct interaction reduces social desirability bias.	Low — No interaction; answers are usually honest.
2. Perceived Anonymity	Low — Identity is known to the interviewer.	Moderate — Identity may be partly disclosed.	High — Often anonymous depending on design.	High — Usually anonymous unless return address is required.
3. Response Rate	High — Face-to-face ensures higher responses.	Moderate to High — Especially if pre-screened participants.	Low to Moderate — Can be low unless well incentivized.	Low — Response rates are typically poor without follow-up.
4. Control on Data Collection Environment	High — Interviewer controls setting and clarifies doubts.	Moderate — Less control, but immediate clarification possible.	Low — No control over where or how the respondent completes it.	Very Low — No control over environment or respondent's conditions.
5. Diversity of Questions	Very High — Can ask complex and adaptive questions with explanations.	Moderate — Limited complexity due to time and medium.	High — Allows varied formats, multimedia, logic branching.	Low to Moderate — Limited by space and respondent's patience.
6. Interviewer's Bias	High — Personal presence can unintentionally influence answers.	Moderate — Tone and manner of interviewer can bias responses.	None — No interviewer involved.	None — Entirely self-administered.
7. Use of Physical Stimuli	High — Can show products, images, prototypes directly.	Low — Difficult to show physical items; verbal only.	High — Images, videos, interactive media can be used.	Low — Only printed materials can be shared.

Parameter	Personal Survey	Telephonic Survey	Electronic Survey	Postal Mail Survey
8. Sample Control	High — Can select respondents precisely (location, demographics).	Moderate — Depends on phone list accuracy and screening.	Low to Moderate — Open to anyone unless access is restricted.	Low — Cannot control who fills and returns the survey.
9. Quantum of Data Collected	High — Can collect large amounts of detailed data.	Moderate — Limited by respondent time and attention span.	High — Easily collect large datasets and complex responses.	Low to Moderate — Space constraints limit extensive data collection.

2.7 Definition of Observation

Observation is a data collection method in which researchers watch, record, and analyze the behavior, actions, or events of subjects in their natural or controlled environment without direct questioning.

When is Observation Method Suitable?

- When **direct communication is not possible** (e.g., infants, animals, or illiterate respondents).
- When studying **natural behaviors and real-life situations**.
- When **accurate and unbiased information** about actual actions is needed rather than self-reported data.
- In **market research** (e.g., tracking customer behavior in a store).
- In **psychology or social research** to understand group dynamics or non-verbal communication.
- For **process analysis** in production or service industries.

Advantages of Observation Method

1. **Direct and Real Data:** Captures actual behavior rather than verbal or biased reports.
2. **Non-Verbal Information:** Useful to study gestures, expressions, and non-verbal cues.
3. **No Recall Error:** Does not depend on respondent's memory or willingness to answer.
4. **Natural Environment:** Observations can be made without influencing participants if done unobtrusively.
5. **Versatile:** Applicable in varied fields like marketing, psychology, education, and anthropology.

Limitations of Observation Method

1. **Time-Consuming:** Observing behaviors over time can take long periods.
2. **Costly:** Requires trained observers and sometimes specialized equipment.
3. **Limited Scope:** Only visible behavior can be recorded, not thoughts or feelings.
4. **Observer Bias:** Personal biases of the observer may affect accuracy.
5. **Ethical Concerns:** Observing without consent can raise privacy and ethical issues.
6. **Non-repeatability:** Once an event is missed, it cannot always be re-observed.