Research Methodology

Table of Content

Introduction to Research

Meaning

The word research is composed of two syllables, re and search former as a prefix meaning again, a new or over again and the latter as a verb meaning to examine closely and carefully, to test and try, or to probe. Together they mean a careful, systematic, patient study and investigation in some field of knowledge, undertaken to establish facts or principles. (Grinnell)

Grinnell further adds: 'research is a structured inquiry that utilises acceptable scientific methodology to solve problems and creates new knowledge that is generally applicable'

Definition

Burns defines research as 'a systematic investigation to find answers to a problem'.

According to Kerlinger , 'scientific research is a systematic, controlled empirical and critical investigation of propositions about the presumed relationships about various phenomena'.

Bulmer states: 'Nevertheless sociological research, as research, is primarily committed to establishing systematic, reliable and valid knowledge about the social world'

Features

From these definitions it is clear that research is a process for collecting, analysing and interpreting information to answer research questions. But to qualify to be called 'research', the process must have certain characteristics and fulfil some requirements: it must, as far as possible, be controlled, rigorous, systematic, valid and verifiable, empirical and critical.

- Controlled
- Rigorous
- Systematic
- Valid and verifiable
- Empirical
- Critical

Objectives

- To gain FAMILIARITY with a phenomenon or to achieve new insights into it (studies with this object in view are termed as exploratory or formulate research studies)
- To portray accurately the CHARACTERISTICS of a particular individual, situation or a group(studies with this object in view are known as descriptive research studies)
- To determine the FREQUENCY with which something occurs or with which it is associated with something else (studies with this object in view are known as diagnostic research studies)

- To test a HYPOTHESIS of a causal relationship between variables (such studies are known as hypothesis-testing research studies)

Importance of Research

- To remove doubts
- To solve operational & planning problems
- To enable government in making policies
- To predict& suggest remedial measures
- To keep business updated
- To test new products
- To avoid future failures
- Better decision making

Types of Research

Descriptive vs. Analytical Research

Descriptive: The surveys and fact- finding investigations of distinct types come under descriptive research.

Descriptive research focuses on expanding knowledge on current issues through a process of data collection. Descriptive studies are used to describe the behaviour of a sample population. For example, a study conducted to know if top-level management leaders in the 21st century possess the moral right to receive a huge sum of money from the company profit

In analytical research, a researcher has to avail data which already exists and has to make an accurate evaluation

Applied Vs Fundamental Research

Applied Research

The objectives of the applied research are to search an answer for the difficulty which is facing by a firm or a society. Applied research focuses on analysing and solving real-life problems. Studies play an important role in solving issues that impact the overall well-being of humans. For example: finding a specific cure for a disease.

Fundamental Research:

Theory's formulation and interference of general principles with particulars are related to fundamental research. The feature of collecting knowledge for the sake of knowledge is best suited to fundamental research. A basic research definition is data collected to enhance knowledge. The main motivation is knowledge expansion. It is non-commercial research that doesn't facilitate in creating or inventing anything. For example: an experiment to determine a simple fact. Mathematics Research- Pure research

Quantitative vs. Qualitative Research

Quantitative: Quantitative measurement of few characteristics which are explained in quantities is nothing but the quantitative research.

Quantitative research involves a larger population — more people mean more data. With more data to analyse, you can obtain more accurate results. This method uses close-ended questions because the researchers are typically looking to gather statistical data. For example: Online survey on consumer satisfaction

Qualitative research is related to the parameter of quality.

Qualitative research is heavily dependent on the experience of the researchers and the questions used to probe the sample. The sample size is usually restricted to 6-10 people. Open-ended questions are asked in a manner that encourages answers that lead to another question or group of questions. The purpose of asking open-ended questions is to gather as much information as possible from the sample.

Conceptual vs. Empirical Research

Conceptual : The conceptual research is concerned with ideas and theories. Generally used by philosophers and thinkers to develop new ideas Example: Sir Isaac Newton

Empirical research is perfect when we have the proof which explains that few variables generate a change on other variables in some way or the other. They rely on experiment or observation alone.

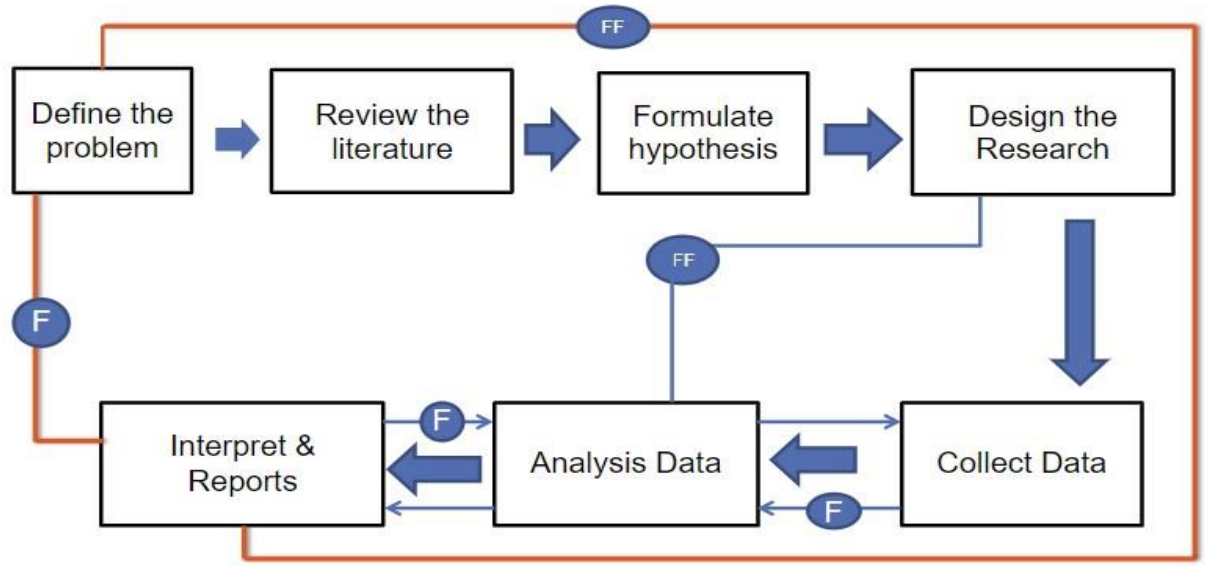
Example: Pharmaceutical companies use it to try out a specific drug on controlled groups to study the effect and cause.

Research Process

Also called as Research Methodology

A research process includes series of steps which need to be planned and performed in a specific order for orderly completion of research activity

RESEARCH PROCESS



Where, F = feed back (Helps in controlling the sub-system to which it is transmitted)
FF = feed forward (Serves the vital function of providing criteria for evaluation)

I.Define the Problem

The first step of research process is to define the problem

There are two types of research problem:

i. Those which relates to state of nature
ii. Those which relates to relationship between variables.

Essentially two steps are involved in define research problem:

i. Understanding the problem thoroughly.
ii. Rephrasing the same into meaningful terms from a point of view

II.Review OF Literature

Extensive literature survey connected with the problem need to be done

Academic journals, conference proceedings, government reports, books etc., must be tapped according to the nature of problem. Earlier studies must be carefully studied.

Search engines like ‘google’ or ‘google scholars’ can be used to search scholarly articles.

III.Formulate Hypothesis

The next steps are to formulate hypothesis. It is tentative assumption made in order to draw out test its logical or empirical consequences. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested.

The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track.

IV. Design the Research

The research problem having been formulated in clear cut terms, the researcher will be required to prepare a research design, i.e., he will have to state the conceptual structure within which research would be conducted.

The function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money.

Research purpose may be grouped into four categories,

i.e. (a). Exploration; (b). Description; (c). Diagnosis; (d).Experimentation.

V. Collect the Data

The next step is to collect the data.

There are several ways to collect the data are :

1.By Observation

2.Through personal interview

3.Through telephone interview

4. By mailing of questionnaires

5. Through schedules

The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, financial resources, available time and the desired degree of accuracy.

VI.Analysis Data

Requires a number of closely related operations

Coding, tabulation and drawing statistical inferences

Generally done by computers, which saves time and make it possible to study large number of variables.

Various statistical measures are used for analysis.

VII. Interest & Profit

Research has to prepare the report of what has been done by him.
Writing of report includes.:

1. The preliminary pages
2. The main text
3. The end matters

Characteristics of Good Research

- Systematic
- Logical
- Clearly defined purpose
- Outline research process

- Maintain high ethical standards
- Unambiguous presentation of findings
- Highlight limitations
- Empirical
- Free from personal bias

Factors Affecting Business Research

1. Time Constraints
2. Availability of Resources
3. Very Expensive
4. Availability of Data
5. Nature of Information Sought

Research Problem & Objectives

First step in research is to perceive a problem - either practical or theoretical

Existence of a problem motivates research

Finding answers to the problems is what is endeavoured in research

Formulation of research objectives is equally important as problem

Research goals need to be identified

The purpose of research is to find a solution to the difficulty

It is desirable that the researcher should propose a set of suggested solutions or explanations of the difficulty which the research proposes to solve.

Hypothesis

The hypothesis (plural hypotheses) is a tentative solution of a problem

The researcher always plans or formulate a hypothesis in the beginning of the problem.

The word hypothesis consists of two words: Hypo + thesis = Hypothesis.

'Hypo' means tentative or subject to the verification and 'Thesis' means statement about solution of a problem

A research hypothesis is a prediction of the outcome of a study. The prediction may be based on uneducated guess or a formal theory

Hypothesis –some examples

Example-1

Research question

What are the health benefits of eating an apple a day?

Hypothesis

Increasing apple consumption in over-60s will result in decreasing frequency of doctor's visits.

Example-2

Research question

Which airlines have the most delays?

Hypothesis

Low-cost airlines are more likely to have delays than premium airlines.

Hypothesis –some examples

Example-3

Research question

What effect does daily use of social media have on the attention span of under16s?

Hypothesis

There is a negative correlation between time spent on social media and attention span in under-16s.

Example-4

Can flexible work arrangements improve job satisfaction?

Hypothesis

Employees who have flexible working hours will report greater job satisfaction than employees who work fixed hours.

Nature & Significance

The hypothesis is a clear statement of what is intended to be investigated

It should be specified before research is conducted and openly stated in reporting the results

This allows to Identify…

▪ the research objectives;

▪ the key abstract concepts involved in the research; and

▪ its relationship to both the problem statement and the literature review

Functions of Hypothesis

1. It is a temporary solution of a problem concerning with some truth which enables an investigator to start his/her research works.

2. It offers a basis in establishing the specifics what to study for and may provide possible solutions to the problem.

3. Each hypothesis may lead to formulate another hypothesis.

4. A preliminary hypothesis may take the shape of final hypothesis.

5. Each hypothesis provides the investigator with definite statement which may be objectively tested and accepted or rejected and leads for interpreting results and drawing conclusions

Role Of Hypothesis

A hypothesis, which is a provisional formulation, plays significant role in empirical or socio-legal research. It not only navigates research in a proper direction but also contributes in testing or suggesting theories and describing a social or legal phenomenon.

A hypothesis, regardless of its source, states what a researcher is looking for. It also suggests some plausible explanations about the probable relationships between the concepts or variables indicated therein. In fact, it navigates the research. Without it, no further step is possible in empirical research

Role Of Hypothesis

A hypothesis helps the researcher in drawing 'meaningful conclusions' supported by 'relevant' empirical data.

A hypothesis serves as a sound guide to:

(I) the kind of data that must be collected in order to answer the research problem

(ii) the way in which the data should be organized most efficiently and meaningfully, and

(iii) the type of methods that can be used for making analysis of the data.

Types of Hypotheses

There are six forms of hypothesis and they are:

Simple hypothesis

Complex hypothesis

Directional hypothesis

Non-directional hypothesis

Null hypothesis

Associative and casual hypothesis

Simple Hypothesis

It shows a relationship between one dependent variable and a single independent variable. For example – If you eat more vegetables, you will lose weight faster. Here, eating more vegetables is an independent variable, while losing weight is the dependent variable.

Complex Hypothesis

It shows the relationship between two or more dependent variables and two or more independent variables. Eating more vegetables and fruits leads to weight loss, glowing skin, reduces the risk of many diseases such as heart disease, high blood pressure, and some cancers.

Directional Hypothesis

It shows how a researcher is intellectual and committed to a particular outcome. The relationship between the variables can also predict its nature. For example- children aged four years eating proper food over a five-year period are having higher IQ level than children not having a proper meal. This shows the effect and the direction of effect.

Non-directional Hypothesis

It is used when there is no theory involved. It is a statement that a relationship exists between two variables, without predicting the exact nature (direction) of the relationship.

Null Hypothesis

It provides the statement which is contrary to hypothesis. It's a negative statement, and there is no relationship between independent and dependent variable. The symbol is denoted by "HO".

Associative and Causal Hypothesis

Associative hypothesis occurs, when there is a change in one variable resulting a change in the other variable. Whereas, Causal hypothesis propose a cause-and-effect interaction between two or more variables.

Sources of Hypothesis

Analogies: Resemblance between the phenomenon.

Observations: from past studies, present day experiences, and from the competitors.

Imagination and thinking: General patterns that influence the thinking process of people.

Previous research

Personal experience

Academic literature

General culture

Research Design

Research design is a logical and systematic outline of research project prepared for directing, controlling and guiding research work

It acts as a broad outline of the research work

Definition:

" Research design is the blue-print for the collection and measurement and analysis of data"- Bernard Phylips

Need for Research Design

It provides appropriate guidelines-when to start and end, where to collect information, how the data collected etc.

Facilitates organising resources: fund, manpower, materials

Gives suitable directions: sources, data collection

Selection of appropriate techniques: statistical techniques to be used

Needed for the collection of relevant data

Motivation to staff

Helps in taking decisions

Importance of Research Design

Blueprint

Right track

Avoid errors

Proper planning

Higher efficiency

Time schedule

Removes uncertainty

Steps in Research Design

1.Define the problem:

Problem must be properly selected for study and explained

2. Objective of the study:

Objective of the study must be explained clearly as it gives direction and guide the researcher

3.Sources of data:

Primary Data: provides first-hand information, but costly, time consuming

Secondary Data: published data, economical

Researcher can combine both

4.Techniques of data collection:

Researcher has to adopt techniques to collect information through observation, face to face interviews, questionnaires

5.Decision on Population:

Called as universe. Researcher need to define the population, which needs to study as a 'unit'

6.Determine sample size:

Group of people carefully selected to represent population. It can be statistically estimated

7.Fixing geographical limit:

Data collected from specific group

8.Decision of resources

9.Preparing research design:

Objectives of the study, sources of data, data analysis, time to complete work, cost involved, sample size, techniques to collect data

10.Analysis of Data

Editing and tabulation

11. Submission of research report

Scope of research design

Sources of data collection

Methods of data collection

Methods of data analysis

Data collection specific

Availability of physical resource

Need for qualified staff

Time allotment

Population for research

Sample size

Essentials of good research design

Simplicity

Flexibility

Focus on objectives

Cost effective

Element of suitability

Pilot studying version of full-scale study

Smooth administration

Systematic approach

Training to research staff

Adopt right methods

Sampling

Sampling is the process of selecting or choosing elements (respondents) from the population under study to collect primary information.

Sampling elements

Population of the Study

The entire group of people of interest from whom the researcher needs to obtain information

Element (sampling unit) One unit from a population

Sampling

The selection of a subset of the population through various sampling techniques

Sampling Frame

Listing of population from which a sample is chosen. The sampling frame for any probability sample is a complete list of all the cases in the population from which your sample will be drawn

Example

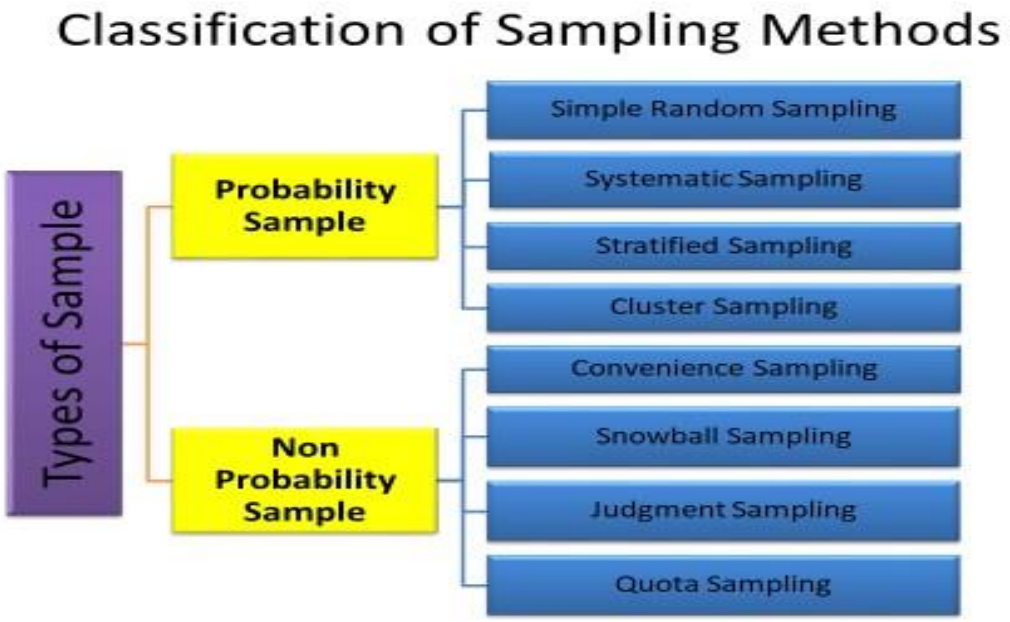
Topic of Research: To compare the online buying behaviour of Students in Mumbai and Students in Pune

Population: Total Students in Mumbai and Pune

Sampling Unit : 2 (1. Students in Mumbai 2. Students in Pune)

Sample size: Students in Mumbai – 400 and Students in Pune – 400)

Sampling Frame : All students studying in Commerce & Science degree colleges



Probability Sampling

Each and every unit of the population has the equal chance for selection as a sampling unit

Each unit in the population has unique identity

Population under study is identifiable or definite

Population is relatively small

Population is accessible or approachable

Types of Probability Sampling

Simple Random Sampling:

The purest form of probability sampling. Assures each element in the population has an equal chance of being included in the sample. Random number generators

Stratified sampling:

In stratified sampling entire population is bifurcated / divided into various mutually exclusive, homogeneous and non-overlapping subgroups known as strata. Then the sample is drawn randomly from each stratum. The procedure of drawing simple random samples from each stratum is called as stratified random sampling.

Cluster Sampling

The population is divided into subgroups (clusters) which are heterogeneous in nature. Then a sample of the cluster is selected randomly from the population. The term cluster refers to a natural, but heterogeneous, intact grouping of the members of the population. The most common variables used in the clustering population are the geographical area, buildings, school, etc. Heterogeneity of the cluster is an important feature of an ideal cluster sample design.

Systematic Sampling

Order all units in the sampling frame based on some variable and then every nth number on the list is selected**.** Gaps between elements are equal and Constant□□There is periodicity. N= Sampling Interval

Multistage Sampling

Multistage sampling refers to sampling plans where the sampling is carried out in stages using smaller and smaller sampling units at each stage. Not all secondary units sampled normally used to overcome problems associated with a geographically dispersed population

Non-Probability Sampling

The probability of each case being selected from the total population is not known.

Units of the sample are chosen on the basis of personal judgment or convenience

There are no statistical techniques for measuring random sampling error in a non-probability sample therefore, generalizability is never statistically appropriate

Non-Probability Sampling

Involves non-random methods in selection of sample

All have not equal chance of being selected

Selection depends upon situation

Considerably less expensive

Convenient

Sample chosen in many ways

Types of Non-Probability Sampling

Purposive Sampling /Judgement Sampling

Quota sampling (larger populations)

Snowball sampling

Self-selection sampling

Convenience sampling

Purposive Sampling /Judgment Sampling

Also called judgment Sampling. The sampling procedure in which experienced research selects the sample based on some appropriate characteristic of sample members to serve a purpose. Researcher relies on their own judgements. When taking sample reject, people who do not fit for a particular profile. Start with a purpose in mind

Example: Study the factors responsible for poor attendance of students in UG colleges, Perception of lawyers towards career in law.

Quota sampling (larger populations)

The population is divided into cells on the basis of relevant control characteristics. A quota of sample units is established for each cell. A convenience sample is drawn for each cell until the quota is met. It is entirely non-random and it is normally used for interview surveys

Example: Satisfaction survey of Metro Rail Commuters Mumbai

Central line – 500

Western Line - 500

Snowball sampling

The research starts with a key person and introduce the next one to become a chain. Make contact with one or two cases in the population. Ask these cases to identify further cases. Stop when either no new cases are given or the sample is as large as manageable

Example: Study of behaviour of Transgender, Mother of Special Child, Rape Victims etc.

Sequential Sampling

Sequential sampling is a non-probability sampling technique wherein the researcher picks a single or a group of subjects in a given time interval, conducts his study, analyses the results then picks another group of subjects if needed and so on. Mostly used in exploratory study or experimental study

Which is good sampling technique?

There is no answer to this question.

Decision on which technique to use one as to look into few factors such as:

1.Quantity and quality of Population under Study

2.Is the population is definite or indefinite

3.Can we obtain or allocate unique identifiable code or number to each element in population

4.Population is heterogeneous or homogeneous

Conditions for using Probability Sampling Technique

1.Population under study is definite

2.Each element of the population under study is identifiable and approachable

3.Unique identify of each element under population is available

4.Access to the identifiable population

Data Collection & Processing

Data Collection

Primary Data & Secondary Data

Primary Data:

Constitute of first-hand information which is collected for the first time in order to solve research problem. They are in the pure form; no statistical tests are done. It gives reliable factual information. Secondary Data:

Secondary data are readily available in published form. Collected by someone else and which already been passed through the statistical process.

Characteristics of Primary Data

Provides basic, first-hand information

Collect exactly as per the need

Provides reliable, valuable resources

Requires properly drafted questionnaire

Collected by using different methods such as surveys, interviews, observation & experimentation

Advantages of primary data

Resolve specific research issues-

Performing your own research allows you to address and resolve issues specific to your own business situation. The collected information is the exact information that the researcher wants to know and he reports it in a way that benefits the specific situation in an organization

Better accuracy

Primary data is much more accurate because it is directly collected from a given population.

Up-to-date information

The primary market research is a great source of latest and up-to-date information as you collect it directly from the field in real time. Usually, secondary data is not so up-to-date and recent data

You are the owner of the information

Information collected by the researcher is their own and is typically not shared with others. Thus, the information can remain hidden from other current and potential competitors

Disadvantages of primary data

More expensive: It could be very expensive to obtain primary data collection because the marketer or the research team has to start from the beginning. It means they have to follow the whole study procedure, organizing materials, process .

Time consuming: It is a matter of a lot of time to conduct the research from the beginning to the end. Often it is much longer in comparison with the time needed to collect secondary data.

Can have a lot of limits: Primary data is limited to the specific time, place or number of participants and etc. To compare, secondary data can come from a variety of sources to give more details.°

Not always possible: For example, many researches can be just too large to be performed by your company.

Secondary data

Secondary data is the data that have been already collected by and readily available from other sources. Such data are cheaper and more quickly obtainable than the primary data and also may be available when primary data cannot be obtained at all.

Advantages of Secondary data

It is economical. It saves efforts and expenses.

It is time saving.

It helps to make primary data collection more specific since with the help of secondary data, we are able to make out what are the gaps and deficiencies and what additional information needs to be collected.

It helps to improve the understanding of the problem.

It provides a basis for comparison for the data that is collected by the researcher.

Disadvantages of Secondary Data

It is difficult to collect specific secondary data related to the research problem be hand. For example, you want information on the disposable income of people, and you get data on the gross income of people.

Secondary data which is obsolete is not useful for the research.

The accuracy of the data cannot be ensured. Inaccurate data might impact the outcome of the research.

Sometimes attributes and values collected in secondary data don’t match with the attributes and values that you want to calculate. In such scenarios, secondary data is not useful for research.

It is difficult to obtain official data.

Insufficient details about secondary data make it difficult to use for research purpose

Sources of Secondary Data

Research institutes

Research institutes research various Problems and make information collected by the public. This type of data is usually reliable and can be used for research purpose.

Research institutes like C.S.O, Indian Statistical Institute, and N.S.S.O publish data collected by them for public use.

Data collected by scholars

Scholar's research various research problems and publish information collected by them. This source of information is also considered a useful source of secondary data.

One most important benefit of data collected from scholar research and paper is that the data is valid and can be relied on.

Government Publications

Government Publication is one of the most reliable sources of secondary data.

Countries conduct research on the population for different problems and make the information collected by the public for two reasons one is to make public aware of the prevailing problems in the nation and second to let researchers use this information for further research.

Examples of government Publications in India are reports on currency and finance, India trade journal, statistical abstract not India, Indian customs and central excise tariff, labour gazette, reserve bank of India bulletin, bulletin of agricultural prices, agricultural statistics of India, economic survey, and Indian foreign statistics, etc

International publications

Data published by international publications is collected by researching on a wider population. There are various international organizations such as World Bank, I.M.F, world health organization (W.H.O), world trade organization (W.T.O), United Nations organization, international labour organization, world meteorological organization, food, and agriculture organization, international bank and reconstruction and development, etc.

These organizations publish data related to their organization. For example, World Bank publish information about the growth rate of different countries or the currency values and W.H.O shares information about the health status of the population of different countries and information related to the different diseases.

Such information is very useful for the industries and businesses for import-export business.

Semi government Publications

Semi government organizations of a nation share information that they have collected through their work such as state boards and municipalities share information related to fields like education, death, and births in the state and also the information on sanitation, etc.

Newspapers and magazines

There are many newspapers and magazines which are a useful, important, and reliable source of secondary data. The organizations which publish newspaper and magazines daily and are required to research different fields to provide information on them.

Popular examples of international magazines which can be used as a secondary source of information are the economist, money, frontline, Bloomberg business week, entrepreneur, the New Yorkers, Forbes, fast company, the wall street journal, and business world, etc.

Commercial services:

There are many organizations which charge for providing information such as published market research reports and other publications produced by them. The information provided by commercial services is accurate, latest, and sorted.

Many market companies seek information like consumer information, media statistics. This information is used for business purposes and future investment decisions. These organizations work to collect a wide range of content and make money by selling that data to interested parties.

Libraries

Libraries are one of the best, cheapest, and reliable source of secondary data. There is at least one community library in every town. Otherwise, there are libraries in school or colleges which provides books for reading either free of cost or at very minimal charges.

From the books available in the library, you can collect good quality of secondary data.

Moreover, the information provided in the books are authentic and gone through much analysis.

Internet

Internet is a platform which provides you an abundance of data within a second. On the internet, you will find both published and unpublished secondary data on any topic.

However, the secondary data gathered from the internet is not reliable as there is nobody which regulates or check the validity of data being uploaded on the internet

Observation

Observation, as the name implies, is a way of collecting data through observing

Observation data collection method is classified as a participatory study

Observation method in data collection can be

Structured observation method – This is a systematic observation method where data is collected as per a pre-defined schedule. The specific variable is used in this method for data collection.

Unstructured observation method – The unstructured observation method is conducted in a free and open manner without using any pre-determined objectives, schedules or variables.

Advantages of Observation

Provides direct access to research phenomena

By observing first hand, the researcher can collect, check and record accurate data

Greater flexibility in terms of application

Generate a permanent record of phenomena and the researcher or others can refer with it later

The organization method is one of the simplest methods of data collection. It does not require too much technical knowledge

The observation method is one of the best ways to formulate a hypothesis. The researcher can observe and come to know about the activities, perceptions, likes and dislikes to form a theory on his subject

Observation method is one of the most common methods used in all sciences and is very easy to follow and accept

In some instances, observation is the only available tool to collect essential data and information

The observation method does not require the willingness of the participant to record. The researcher can observe from a distance and record his findings

Disadvantages of Observation

Faces a severe disadvantage because it takes a longer time frame compared to other data collection methods

There is a chance of higher observer bias in the observation method

Several personal behaviours are not open for observation and this proves a limitation in case of observation method

There is a higher chance of the observer influencing the behaviour of a sample group elements

Uncertainties of the event cannot determine the actual time when the event will take place, and this is why every occurrence that is open to observation cannot be observed

Many of the incidents are abstract like love, affection and the researcher can’t gain an exact and correct account of those factors.

The social phenomena generalization made by observation are not considered reliable as it cannot be used for lab experiments

In some cases, it is seen that two persons observing the same phenomena come at different results and this can lead to faulty perceptions

Observation method is considered an expensive affair as it requires hard effort, plenty of time and high cost

Methods of Observation

1.Simple Direct Observation

The behaviour of a person/customer observed as it occurs. For example, the observer acts as a shopper and studies the behaviour of the customers.

2. Indirect Observation

Past observation studied with the help of past records,films,photographers,etc.Behaviour itself is not observed, but effects are observed.

3.Structured Observation

Observation is clearly defined and observer knows whom to observe and what to observe.

4.Unstructured Observation

In this method approach of the observer is flexible.

5.Disguised Observation

The subjects do not know that they are observed.

6. Undisguised Observation

Purpose of the observation is brought to the notice of the respondents.

7.Mechanical Observation

Observation is made with the help of electro-mechanical devices.

Mechanical devices record the phenomenon of interest typically used for continuous recording of ongoing behaviour

e.g., Audiometer, people meters, traffic counters, on-site cameras, UPC system may also be used in situations that require respondent cooperation

e.g., eye-tracking monitors, pupilometers, psychogalvanometers, response latency devices assumption: physiological reactions are associated with cognitive and affective responses

Experimental Method

Experimentation may be defined as “a process where events occur in a setting at the discretion/option of the researcher and controls are used to identify the sources of variation in respondent’s response”

It tries to explain cause & effect relationship

There are two types of setting

A. Field Setting

B. Laboratory Setting

Types of Experimentation Methods

A. Field Setting:

Conducted at the market place

Purposes are not known to the participants

Normally conducted in test marketing in order to find out the acceptability of the new product from the customer

This method is extensively used under following areas

Product Design, Package design, Pricing policies, Distribution policies and Promotion policies

B. Laboratory Setting

The laboratory methods are accurate

But experiments are artificial due to controlled conditions

For example, it is in form of calling limited persons and offering them the product and asking them about reactions.

The purpose of experiment is known to the participants.

Laboratory experiments are less costly and less time consuming

Merits of Experimentation

The biggest advantage of the experimental method is its unique ability to isolate casual factors, since an experiment is highly controlled.

This method promises more accuracy in the study.

Reliable data can be collected.

This is more suitable to the problem with heterogeneous (varied) influencing factors.

Demerits of Experimentation

The disadvantage is that exactly this control may distort the validity of the obtained results, and especially the ecological validity.

This is very costly method.

This is suitable to simple problems with limited scope.

This is a time-consuming method.

Interview method of primary data collection

Interviews can be defined as a qualitative research technique which involves "conducting intensive individual interviews with a small number of respondents to explore their perspectives on a particular idea, program or situation."

Interviewing is a process of social interaction

It is method of data collection through verbal interaction between the respondent and the interviewers.

It is face to face contact with others for collecting required information

Advantages of Interview Method

It provides flexibility to the interviewers

The interview has a better response rate than mailed questions, and the people who cannot read and write can also answer the questions.

The interviewer can judge the non-verbal behaviour of the respondent.

The interviewer can decide the place for an interview in a private and silent place, unlike the ones conducted through emails which can have a completely different environment.

The interviewer can control over the order of the question, as in the questionnaire, and can judge the spontaneity of the respondent as well.

Disadvantages of Interview Method

Conducting interview studies can be very costly as well as very time-consuming.

An interview can cause biases. For example, the respondent's answers can be affected by his reaction to the interviewer's race, class, age or physical appearance.

Interview studies provide less anonymity, which is a big concern for many respondents.

There is a lack of accessibility to respondents (unlike conducting mailed questionnaire study) since the respondents can be in around any corner of the world or country.

Personal Interview

A face-to-face contact is made with the informants (persons from whom the information is to be obtained) under this method of collecting data.

The interviewer asks them questions pertaining to the survey and collects the desired information. Thus, if a person wants to collect data about the working conditions of the workers of the Company, he would go to the factory, contact the workers and obtain the desired information.

The information collected in this manner is first hand and also original in character.

Merits Of Personal Interview

1. Better cooperation from respondent:

Most often respondents are happy to pass on the information required from them when contacted personally and thus response is encouraging

2.Accurate Information: The information collected through this method is normally more accurate because interviewer can clear doubts of the informants about certain questions and thus obtain correct information. In case the interviewer apprehends that the informant is not giving accurate information, he may cross-examine him and thereby try to obtain the information

3.Availability of quality information: This method also provides the scope for getting supplementary information from the informant, because while interviewing it is possible to ask some supplementary questions which may be of greater use later.

4.Can collect desired information: There might be some questions which the interviewer would find difficult to ask directly, but with some tactfulness, he can mingle such questions with others and get the desired information. He can twist the questions keeping in mind the informant's reaction. Precisely, a delicate situation can usually he handled more effectively by a personal interview than by other survey techniques

5. Convenient Language can be used: The interviewer can adjust the language according to the status and educational level of the person interviewed and thereby can avoid inconvenience and misinterpretation on the part of the informant

6.Flexible

7.Improves quality of research work

8.Nonverbal responses are observed

Demerits Of Personal Interview

1.Expensive: This method can prove to be expensive if the number of informants is large and the area is widely spread.

2.Chance of personal bias: There is a greater chance of personal bias and prejudice under this method as compared to other methods.

3.Need for qualified interviewer: The interviewers have to be thoroughly trained and experienced; otherwise, they may not be able to obtain the desired information. Untrained or poorly trained interviewers may spoil the entire work.

4.Time Consuming: This method is more time taking as compared to others. This is because interviews can be held only at the convenience of the informants. Thus, if the information is to be obtained from the working members of households, interviews will have to be held in the evening or on weekend. Even during evening only an hour or two can be used for interviews and hence, the work may have to be continued for a long time, or a large number of people may have to be employed which may involve huge expenses.

Guidelines For Effective Personal Interview

Proper Procedure

Prior communication with respondent

Proper atmosphere for interview

Proper recording of replies

Appointment of trained , efficient and honest interviewers

Focused Group Interview

Type of interview in which a greater control is applied to the course of the interview

Definite framework of topics is decided in advance

Interviewer puts a number of related questions before the respondent

Attention of the respondent is focused on limited number of inter-related questions.

Unnecessary and unwanted discussion during interview is avoided

Focused interview is more precise, clear cut and to the point

It makes the interview result oriented

In-Depth Interview

It is qualitative research technique

This method allows the interviewer to collect different views and opinions of respondent through personal interview.

They are useful for the study of human behaviour

It relates to behavioural sciences such as sociology and psychology

Such interviews are undirected and require experience on the part of the interviewer.

Respondent is allowed to talk freely and the interviewer takes a position of good listener.

Survey Method of Primary Data Collection

Definition:

According to Toll & Hawkins, "survey is the systematic gathering of information from respondent for the purpose of understanding and for predicting some aspect of the behaviour of the population of interest"

Features:

Provides primary or original data

Data are collected through questionnaire and interview

Reliable , costly and time consuming

Different methods are used as per need of research project.

Types of Survey Data collection

1.Telephonic Survey

2.Mail Survey

3.E-mail Survey

4.Internet Survey

5.Social Media survey

6.Media Listening

Telephonic Survey Method

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

A telephone survey (CATI Computer Assisted Telephone Interviewing) involves phone agents that contact survey respondents by phone asking questions to collect information. The phone number is automatically displayed to the agents in the CATI interface and they can make calls using click-to-dial or automated or predictive dialling. The interviewer follows a script on the screen that includes an introduction to the survey, a questionnaire with different types of questions, and a conclusion.

Advantages of Telephonic Survey

Research can be gathered quickly because phone interviews are immediate and skilled interviewers can complete a lot of surveys in a day of work.

Most people have telephones, so you have an ample audience for gathering a representative sample to complete the survey.

A telephone interview has a personal touch, so it can lead to valuable brand-building benefits if the interviewer surveys in a professional and skilled way.

Telephone interviews can be cost-effective as you can have a higher response rate than web surveys, for example.

Disadvantages of Telephonic Survey

Sometimes telephone calls are perceived as telemarketing and thus negatively received by potential respondents. This might influence your response rate.

It can be challenging to design an effective phone survey because the questions need to be short and precise for easy comprehension.

Timing must be carefully considered. The administrators and supervisors should monitor both the time of the call and the length of the actual interview.

Mail Survey Method

A quantitative data collection method, mail survey is the pioneer of self-administered questionnaires. In this approach, the researcher sends the questionnaires enclosed with

postage-paid envelope through the postal system. Meanwhile, the participants will be asked to answer questions that are written on a paper. After completing the questionnaire, the respondents will send it back through mail.

straightforward, comprehensible, and having few open-end questions

mail surveys should be used when the researcher wishes to know if there are changes on the product or service that the consumer would want, the consumers' opinion regarding the company's eventual plans, or time-sensitive matters.

Advantages of a Mail Survey

Cost– Mail surveys need not much of manpower. A man alone can administer the entire survey process. Compared with telephone surveys and face-to-face interviews, the cost in conducting a mail survey is relatively cheaper. This type of survey is optimal of there are large sample size involved. Let us say that the participants are around 40,000. Mailing them is cost-effective than calling them one by one. On estimate, a typical medium-scale mail survey can cost at least $5,000. On the contrary, a telephone survey or a face-to-face interview requires double or triple of your budget for a mail survey.

Geographical stratification– A mail survey can specifically target different segments of the population.

Honesty– Research shows that participants of a survey give more honest answers compared with other data collection methods. This is because respondents are more comfortable giving their views or opinions through writing.

Convenience– Mail surveys provide convenience to respondents for they can answer the questionnaires at their own pace. Survey participants have the liberty to use as much time needed when answering the survey, which will result to more comprehensive and thorough responses. They can also answer the questionnaire anywhere they want to, as long as they have survey instrument.

Administration– For those who will administer and supervise the mail survey, not much of an experience are needed. This type of survey does not oblige the authority to make decisions during high-pressure scenarios. For researchers, they are permitted to curtail sampling errors. They also have the jurisdiction of what the respondents can see on the questionnaire, unlike online surveys where software compatibilities and technical issues are factors on how the survey will be displayed.

Other strengths of a mail survey include being able to contain lengthy and complex questions and visuals that will aid respondents in answering certain questions can be incorporated, and can ensure confidentiality and anonymity.

Disadvantages of a Mail Survey

Coverage errors and Response Rates– A mail survey usually generates 3-15% response rate. Having said that, it is not the primary drawback of engaging in this type of survey. The real problem is how to obtain a reliable and complete list of participants from the target population. When failed to do so, this will result to coverage errors. Examples are incomplete

mailing lists e.g., excluding members of the family that are temporarily away like college students. Biased results and outdated information are also included in coverage errors.

Questionnaire design– Since mail surveys do not offer the opportunity for follow-ups, the questionnaire design can make or break the survey. Questions must be brief, straightforward and accurate.

Respondents– Mail surveys are unseemly ineffectual for very young children, disabled or sick persons, to those with language barriers, and marginally literate or illiterate.

Administration– Researchers have no control as to whether or not the survey has been completely answered or what will happen to the questionnaire after being mailed.

Steps in Ensuring a Successful Survey

Establish your goal and the survey's objectives.

As much as possible, make sure that the survey is designed to be biased-proof.

Inform the respondents that they will be a part of the mail survey and don't forget the time estimate when they will receive the survey.

Fasten a signed cover letter and enclose a stamped return envelope.

Mail the survey. Remind those who did not respond while mail a thank you note to those who responded.

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.

Since then, Web (Internet) surveys have become the predominant model for electronic surveying, because of the relatively poor performance of email surveys in terms of ease of use and response rates.

Advantages & disadvantages of email survey

E- mail survey is quick and economical.

It is easily acceptable to computer friendly younger generation

Limitations of the email survey are:

Quality of the data may be reduced, if the respondent does not understand the questions

Post survey cleaning becomes necessary

Response under email may not be reliable

Internet survey Method

Internet/Web surveys use HTML language(Hypertext Markup Language) and posted in the Website.

Respondents may be recruited over the internet or from potential respondent data bases maintained by the marketing research firm or from an internet panel.

Respondents are asked to go to a particular web location to complete the survey.

This method of survey is reasonably convenient.

Information is available easily and quickly.

It can cover a larger sample at a lesser cost.

Internet research can be as representative and effective as other traditional methods, especially because the internet population continues to grow

Problems in conducting research over the internet need to be addressed and solved effectively.

However, these methods are not suitable for surveying the general population in India

The response to internet survey in India is not very encouraging.

Information may not be collected from those who do not have access to internet.

Advantages of Online Survey

Increased Response Rate

The low cost and overall convenience of online surveys bring in a high response. Respondents get to answer questions on their own schedule at a pace they choose.

Low Cost

Collecting data doesn't have to break the bank anymore. There are plenty of websites and platforms that make creating your survey fast and affordable.

Real-time Access

Respondents' answers store automatically so you get results at your fingertips in no time. These turns analysing your results into effortless and immediate action.

Convenience

Respondents answer questions on their own schedule and can even have flexibility with completion time.

Rapid deployment and return times are possible with online surveys that don't use traditional methods. And, if you have bad contact information for some respondents, you'll know it almost immediately.

Design Flexibility

Surveys can be programmed even if they're very complex. Intricate skip patterns and logic can be employed seamlessly. You can create the layout, questions, and answer choices with no hassle.

No Interviewer

Since respondents are not disclosing their answers directly to another person, it is easier for them to open up. Interviewers can also influence responses in some cases.

Disadvantages Online Survey

Survey Fraud

This is the biggest challenge. If your survey is long and/or confusing you might get fake answers. Since there is less accountability, the chances for people just hitting buttons to finish are high. Check the questions you use carefully. People often take surveys because they're promised a reward at the end, resulting in them not accurately contributing to your study.

Limited Sampling and Respondent Availability:

Certain populations are less likely to have internet access and to respond to online questionnaires. Drawing samples is harder based on email addresses or website visitations.

Possible Cooperation Problems:

Online surveys could be deleted and ignored. People hate feeling poked and if they get annoyed, they just have to click delete.

No Interviewer:

The lack of a trained interviewer to clarify and probe can lead to less reliable data.

Social Media Survey Method

Social media refers to " the many relatively inexpensive and widely accessible electronic tools that enable anyone to publish and access information, collaborate on a common effort or build relationships"

Social media enables create and engage communities online.

Connect greater number of likeminded people who share opinions ideas and information of interest with one another.

The network that evolves on social media platform are Facebook, Instagram, Twitter, LinkedIn, Google Plus .

Social media provide opportunity to meet other businessmen, create contacts and promote business.

Social media offers regular feedback from customers

Benefits of monitoring social media

Keep track of customers

Convenient to speak

Feasible to plan the next blog post

Possible to make new customers and retain the old customers

Can come out with improved product

Possible to know about brand image of the company

Media Listening Survey Method

Media listening is a process of using social media channels to track, gather and store the information and data of individuals, groups and organisations.

Media listening is also known as social media listening and social media measurement.

Business houses use this to reach out to customers information gathering.

The process has become simple with readily available outlets such as news sites, message boards, video/photo sharing sites, wikis, forums etc

There are number of media listening tools, customised data mining software and web crawlers.

Hence it is process of identifying and assessing what is being said about the product, performance by the consumers so that a strategy can be devised to better influence the customers.

Information collected can be utilised for analysing consumer behaviour, determining the public's perception of a company and comparing a company's reputation against competitors.

More effective Ad campaign can be made on the basis of viewers likes.

Media listening can help the companies to improve the quality of the product by observing the opinions expressed by the customers.

Online platforms help the company find out how they are performing against competitors.

Thus, it is clear that media listening is the act of using a tool to monitor what is being said on internet.

Means of conducting media listening

1.Email

2.Survey

3.Comment Boxes

4.Customer Feedback on social media

5. On-site Activity

6.Exploratory interviews

Survey Instrument(Questionnaire)

A questionnaire is a formalized set of questions designed to obtain information from the respondents or the subjects selected for a study.

It is an instrument that is generally handed over to the respondents and filled in by them with no help from the interviewer or any other person.

A questionnaire has three specific characteristics

It must translate the information needed into a set of specific questions that the respondents can and will answer

It must uplift, motivate, and encourage the respondents to become involved in the interview and complete the interview.

It should minimize the response as well as other types of errors.

Advantages of Questionnaire

1. Questionnaires are inexpensive

2. Questionnaires are practical

3. Questionnaires offer a quick way to get results

4. Scalability

5. Comparability

6. Easy Analysis and visualization

7. Questionnaires offer actionable data

8. Respondent anonymity

9. Questionnaires don't have time constraints

10. Questionnaires can cover every aspect of a topic

Disadvantages of questionnaire

1. Dishonest answers

2. Unanswered questions

3. Differences in understanding and interpretation

4. Hard to convey feelings and emotions

5. Some questions are difficult to analyse

6. Respondents may have a hidden agenda

7. Lack of personalization

8. Unconscientious responses

9. Accessibility issues

10. Questionnaire or survey fatigue

Questionnaire Designing

Preparation/ construction of questionnaire is a lengthy and time-consuming activity

Preparing a good questionnaire is an art

There are two types of questionnaires

A. Consumer questionnaire: useful for collecting information from consumers

B. Trade questionnaire: useful for collecting information from traders and dealers

Steps involved in Questionnaire Designing

Preliminary decisions

Informal interviewing

Drafting questionnaire

Piloting questionnaire

Deciding final form of questionnaire

Final questionnaire

1.Preliminary decisions:

Data need, sources of data, sample layout, questionnaire need to be decided

2.Informal interviewing:

Interview is arranged with few people to test the points and to find out how few people thinks about the issues and their reactions.

3.Drafting questionnaire:

All efforts to be made to draft an ideal questionnaire. Questionnaire must be logical and should facilitate tabulation and analysis of the information should be simple and quick

4.Piloting questionnaire

Also called as pre testing questionnaire. Pre testing is done with a small group of samples taken from entire sample. It is done in order to remove possible defects and deficiencies in the questionnaire.

5.Deciding final form of questionnaire

In this step of , final touch is given in the light of the experience gained in pretesting. All necessary corrections and revisions are incorporated.

6.Final questionnaire

Preparing a complete, compact and perfect questionnaire.

Features of questionnaire

Short &Simple

Easily understandable

Specific and relevant questions

Logical arrangement of questionnaire

Interesting to respondents

Adequate space for answers

Purpose of enquiry should be clear

Avoiding lengthy, personal confusing questions

Types of questions in the questionnaire

1.Close-ended Questions/Structured questions

2. Open-ended questions/Unstructured questions

3.Dichotomous or Two –way questions or "Yes" or "No" questions

4. Multiple choice questions

Close –ended Questions/Structured questions

These ones are used to collect quantitative data which is information recorded as a count or numerical value.

The data is quantifiable which means it can be used for mathematical calculations or statistical analysis. In essence, it answers the question of how much, how many, or how often.

Standardized questionnaires are best used when we are already formed an initial hypothesis or built out a prototype for a product. It is used to stress test the assumptions, designs, use cases, etc. before going further with product development. Because of its clear focus, the questions asked are narrow in scope and solicit specific information.

Closed-ended questions are questions that can only be answered by selecting from a limited number of options, usually multiple-choice, 'yes' or 'no', or a rating scale (e.g., from strongly agree to strongly disagree).

Closed-ended questions give limited insight, but can easily be analysed for quantitative data.

For example, one of the most popular closed questions in marketing is the Net Promoter Score® (NPS) question, which asks people "How likely are you to recommend this

product/service on a scale from 0 to 10?" and uses numerical answers to calculate overall score trends

closed-ended question can be asked when we want answers that can be plotted on a graph and used to show trends and percentages

Open-ended questions/Unstructured questions

Open-ended questions: Open up a conversation with this question. These are good survey questions to get more meaningful answers from as people have the opportunity to give more feedback through a text box.

These are also known as unstructured questionnaires. They're used to collect qualitative data which is information that can be observed and recorded but isn't numerical in nature. It's used to approximate and characterize.

Exploratory questionnaires are ideal in the early stages and want to learn more about a topic before designing a solution or hypothesis. For example, in the early stages of product development and don't know enough about the market then exploratory questionnaires are ideal.

Open-ended questions

Open-ended survey questions require respondents to type their answer into a comment box and don't provide specific pre-set answer options. Responses are then viewed individually or by text analysis tools.

When it comes to analysing data, open-ended questions aren't the best option. It's not easy to quantify written answers which is why text boxes are better for providing qualitative data. Allowing the respondents to offer feedback in their own words could help uncover opportunities that may have otherwise overlooked.

However, for analyse the data , engage in some quantitative marketing research and utilize closed questions.

Examples of Closed-end & Open-end questions

Closed-ended question example	Open-ended question example
Would you recommend our product/service?	What were the main reasons you chose our product/service?
Did you experience good customer service?	How did you feel about our customer service?
Would you consider using our product/service again?	What would make you use our product/service again?
Did you like our product/service?	What is the most important feature of our product/service for you?
Are you interested in buying product/service today?	Why are you looking for product/service today?
Are you happy with your experience with us?	How would you describe your experience with us?
Did you find what you were looking for today?	How can we help you find what you are looking for today?

Dichotomous or Two-way questions

The dichotomous question is a question that can have two possible answers.

Dichotomous questions are usually used in a survey that asks for a Yes/No, True/False, Fair/Unfair or Agree/Disagree answers.

They are used for a clear distinction of qualities, experiences, or respondent's opinions They are useful for the collection of simple factual information.

Dichotomous questions (Yes/No) may seem simple, but they have few problems both on the part of the survey respondent and in terms of analysis.

Yes/No questions often force customers to choose between options that may not be that simple and

may lead to a customer deciding on an option that doesn't truly capture their feelings

The benefits of dichotomous questions are that they are easy and short. Dichotomous questions have the advantage to ease responses and ease the analysis of the data.

Multiple-choice questions

Multiple choice questions are the most popular survey question type.

They allow the respondents to select one or more options from a list of answers that are given. They're intuitive, easy to use in different ways, help produce easy-to-analyse data, and provide mutually exclusive choices. Because the answer options are fixed, and respondents have an easier survey-taking experience.

Researcher will get structured survey responses that produce clean data for analysis. Multiple answer multiple choice questions are commonly shown with square checkboxes. They allow respondents to check off all the choices that apply to them

A common drawback of multiple-choice questions is that they force to limit responses to a predetermined list of options. Choice is given but it is restricted to few options only. This can cause bias in the results.

One can solve this problem by adding another answer option or comment field

Rating Scales

Three-point, five-point, and seven-point scales are all included in the umbrella term "rating scale". A rating scale provides more than two options, in which the respondent can answer in neutrality over a question being asked.

Examples:

1. Three-point Scales

Good - Fair – Poor

Agree – Undecided - Disagree

Extremely- Moderately - Not at all

Too much - About right - Too little

2. Five-point Scales (e.g., Likert Scale)

Strongly Agree – Agree – Undecided / Neutral - Disagree - Strongly Disagree

Always – Often – Sometimes – Seldom – Never

Extremely – Very - Moderately – Slightly - Not at all

Excellent - Above Average – Average - Below Average - Very Poor

3. Seven-point Scales

Exceptional – Excellent – Very Good – Good – Fair – Poor – Very Poor

Very satisfied - Moderately satisfied - Slightly satisfied – Neutral - Slightly dissatisfied - Moderately Dissatisfied- Very dissatisfied

Likert scale

A Likert scale is a question which is a five-point or seven-point scale. The choices range from Strongly Agree to Strongly Disagree so the survey maker can get a holistic view of people's opinions. All Likert scales also include a mid-point e.g., neither agree nor disagree, for those who are neutral on the subject matter.

It was developed in 1932 by the social psychologist Renzi's Likert. Over the years, they've evolved to become a favourite amongst survey makers as they obtain definite opinions, impressions, and approaches from the respondents.

The Likert scale question is used to understand the level of agreement that the respondents have with a particular statement. The range provided in this scale is used to gain insights about respondent feelings and opinions. Agreement, frequency, likelihood, quality, or importance can be measured using a Likert scale with corresponding anchors. Scales can be either a unipolar Likert scale or a bipolar Likert scale.

Unipolar Likert scale & Bipolar Likert Scale

Unipolar Likert scale

Unipolar scales are more contoured, allowing users to instead focus on a single item's absence or presence. The scale measures the ordinal data, but most of the time, unipolar scales generate more accurate answers. An example of a unipolar satisfaction scale is: not at all satisfied, slightly satisfied, moderately satisfied, very satisfied, and completely satisfied. A unipolar Likert scale question type indicates a respondent to think of the presence or absence of a quality or trait. It is arranged on a 5-point scale.

Bipolar Likert Scale

A bipolar scale indicates a respondent to balance two different qualities, defining the relative proportion of those qualities. Where a unipolar scale has one "pole," a bipolar scale has two opposites. For example, a common bipolar scale includes the following choices: completely dissatisfied, mostly dissatisfied, somewhat dissatisfied, neither satisfied nor dissatisfied, somewhat satisfied, mostly satisfied, and completely satisfied. That is a scale with 0 in the middle (-3, -2, -1, 0, 1, 2, 3).

Likert scale

This scale has gained tremendous popularity in online surveys and is used in every study, such as customer satisfaction, employee engagement, or employee satisfaction. One can bifurcate the Likert scale into two types: Odd Likert Scale and the Even Likert Scale.

Decide wisely, which type of Likert scale would provide the best results. Analysis of the target audience, evaluating the purpose of survey research is critical in determining the type of Likert scale question. Depending on whether you'd want to prompt the respondents to give responses that you wish to or wish to provide a neutral option, the respondents can select if they do not have any bias towards the other answers. The odd Likert scale question offers a central point for the respondents to choose if they're neutral.

The even Likert Scale questions have options without a midpoint due to which the respondents will be forced to choose from the provided answer options. The midpoint in the Odd Likert Scale will be interpreted differently by different respondents, but it will never be completely biased.

Semantic Differential Scales

A semantic differential scale is only used in specialist surveys in order to gather data and interpret based on the connotative meaning of the respondent's answer. It uses a pair of clearly opposite words, and can either be marked or unmarked.

Data Analysis & Interpretation

Data processing

Data processing is, generally, "the collection and manipulation of items of data to produce meaningful information"

Data processing occurs when data is collected and translated into usable information. Usually performed by a data scientist or team of data scientists, it is important for data processing to be done correctly as not to negatively affect the end product, or data output.

Data processing starts with data in its raw form and converts it into a more readable format (graphs, documents, etc.), giving it the form and context necessary to be interpreted by computers and utilized by employees throughout an organization

Data processing cycle-Features/Stages

Six stages of data processing

1. Data collection

Collecting data is the first step in data processing. Data is pulled from available sources, including data lakes and data warehouses. It is important that the data sources available are trustworthy and well-built so the data collected (and later used as information) is of the highest possible quality.

2. Data preparation

Once the data is collected, it then enters the data preparation stage. Data preparation, often referred to as "pre-processing" is the stage at which raw data is cleaned up and organized for the following stage of data processing. During preparation, raw data is diligently checked for any errors. The purpose of this step is to eliminate bad data (redundant, incomplete, or incorrect data) and begin to create high-quality data for the best business intelligence.

3. Data input

The clean data is then entered into its destination (perhaps a CRM like Salesforce or a data warehouse like Redshift), and translated into a language that it can understand. Data input is the first stage in which raw data begins to take the form of usable information.

4. Processing

During this stage, the data inputted to the computer in the previous stage is actually processed for interpretation. Processing is done using machine learning algorithms, though the process itself may vary slightly depending on the source of data being processed (data lakes, social networks, connected devices etc.) and its intended use (examining advertising patterns, medical diagnosis from connected devices, determining customer needs, etc.).

5. Data output/interpretation

The output/interpretation stage is the stage at which data is finally usable to non-data scientists. It is translated, readable, and often in the form of graphs, videos, images, plain text, etc.). Members of the company or institution can now begin to self-serve the data for their own data analytics projects

6. Data storage

The final stage of data processing is storage. After all of the data is processed, it is then stored for future use. While some information may be put to use immediately, much of it will serve a purpose later on. Plus, properly stored data is a necessity for compliance with data protection legislation. When data is properly stored, it can be quickly and easily accessed by members of the organization when needed.

Essentials of processing data

Reliability & Validity

Stability

Completeness

Accuracy

Uniformity

Flexibility

Realistic

Acceptance

Use of resources

Consistency

Processing of data--editing, coding, classification and tabulation

After collecting data, the method of converting raw data into meaningful statement; includes data processing, data analysis, and data interpretation and presentation.

Data reduction or processing mainly involves various manipulations necessary for preparing the data for analysis.

The process (of manipulation) could be manual or electronic.

It involves editing, categorizing the open-ended questions, coding, computerization and preparation of tables and diagrams.

Editing Of Data

Editing is the process of eliminating errors in the raw data collected so that the data may subsequently undergo tabulation and classification.

Data editing is possible at two levels

Micro-editing i.e., editing of data at record level

Macro-editing i.e., editing of data for removing errors in data collected.

Need or significance of data editing

To Check accuracy of data

To bring consistency in the information collected through interviews of respondents

To remove errors

To improve quality of data

To evaluate the performance of field staff

To take uniform decisions on incorrect answers.

Procedure/Process of editing data

Five types of answers are usually dealt with.

1.The incorrect answer

2.The incomplete answer

3.The inconsistent answer

4. The 'Don't know' answer or 'No reply' answer

5.The answer not available in the standard format.

Editing Methods

Interactive editing

The term interactive editing is commonly used for modern computer-assisted manual editing. Most interactive data editing tools applied at National Statistical Institutes (NSIs) allow one to check the specified edits during or after data entry, and if necessary to correct erroneous data immediately.

Several approaches can be followed to correct erroneous data:

Re contact the respondent

Compare the respondent's data to his data from previous year

Compare the respondent's data to data from similar respondents • Use the subject matter knowledge of the human editor

Interactive editing is a standard way to edit data.

Selective editing

Selective editing is an umbrella term for several methods to identify the influential errors, and outliers. Selective editing techniques aim to apply interactive editing to a well-chosen subset of the records, such that the limited time and resources available for interactive editing are allocated to those records where it has the most effect on the quality of the final estimates of publication figures.

In selective editing, data is split into two streams:

The critical stream

The non-critical stream

The critical stream consists of records that are more likely to contain influential errors. These critical records are edited in a traditional interactive manner. The records in the non-critical stream which are unlikely to contain influential errors are not edited in a computer assisted manner.

Macro editing

There are two methods of macro editing:

Aggregation method

This method is followed in almost every statistical agency before publication: verifying whether figures to be published seem plausible. This is accomplished by comparing quantities in publication tables with same quantities in previous publications.

Distribution method

Data available is used to characterize the distribution of the variables. Then all individual values are compared with the distribution. Records containing values that could be considered uncommon (given the distribution) are candidates for further inspection and possibly for editing.

Automatic editing

In automatic editing records are edited by a computer without human intervention[7]. Prior knowledge on the values of a single variable or a combination of variables can be formulated as a set of edit rules which specify or constrain the admissible values

Essentials of good data editing

Quick & Accurate

Economical

Flexible

Avoid bias

Uniformity

Consistency

Completeness

Reliability

Coding of Data

Coding is the procedure of classifying the answers to a question into meaningful categories.

It helps in arranging or grouping of varies responses into manageable categories for the purpose of classification, tabulation and analysis of data.

Numeric coding simplifies the researcher’s task in converting a nominal variable like gender to a 1 or 2.

Coding

Data coding involves assigning a number to the participants responses so, they can be entered into data base.

In coding, categories are the partitions of a data set of as given variable. For instance, if the variable is gender, the categories are male and female.

Categorization is the process of using rules to partition a body of data.

Both closed and open questions must be coded.

Types of Code

Numerical Codes

Alphabetical Codes

Alpha- Numerical Code

A Code is a number/symbol which stands for each type of reply. Coding means assigning a number, symbol or other mark to every answer.

Coding can be done along with editing or it may be done after editing.

Essentials of Good Coding

Should cover all type of answers given by the respondents

Coding categories should not overlap

A code sheet containing the code numbers should be prepared for a quick reference

Work should be done by a responsible person to avoid mistakes

Appropriate use of numbers, alphabets, symbols must be used.

Tabulation Of Data

Tabulation is a technique or process of summarising and arranging the data in a compact form for further analysis

It is sorting and counting of data and allocating the same to relevant classification of categories.

Arrangement of data is done in rows & columns

Need & Importance of Tabulation of Data

Brings orderly arrangement of data

Brings accuracy and speed in arranging the data

Facilitates correct analysis and interpretation of data

Saves time during analysis of data

Needed for drawing diagrams and charts

Methods of Tabulation of data

There are two methods

1.Manual/Hand Tabulation:

Counting is done by hand. Very old method. Done with the help of tally bars.

2. Machine/Mechanical/Electronic Tabulation

Now a days computer processing is done as computers handle large number of complex data speedily & accurately

Graphical Presentation of Data

A picture is said to be more effective than words for describing a particular thing

A graphic representation is the geometrical image of a set of data

It is a mathematical picture

It enables us to think about a statistical problem in visual terms

It is an effective and economic device for the presentation , understanding and interpretation of the collected data

IMPOTANCE OF GRAPHICAL REPRESENTATION

It is used to make the data understandable to common man

It helps in easy and quick understanding of data

Data displayed by graphical representation can be memorised for a long time

Can be compared at a glance.

TYPES OF GRAPHICAL REPRESENTATION

Line graph

Line graphs are simple Mathematical graphs that are drawn on the graph paper by plotting the data connecting one variable on the horizontal X- axis and other variable of data on the vertical Y-axis.

Bar graph

In bar graphs data is represented by bars. The bars can be made in any direction i.e., vertical or horizontal

Histogram

A histogram is essentially a bar graph of a frequency distribution. It can be constructed for equal as well as unequal class intervals

Area of any rectangle of a histogram is proportional to the frequency of that class.

When do we use histogram?

When data are given in the form of frequencies.

When class interval has to be displayed by a diagram.

When we need to calculate the Mode of a distribution graphically.

PIE DIAGRAM

It is a circle in which different components are represented through the sections or portions of a circle.

To draw a pie diagram, first the value of each category is expressed as a percentage of the total and then the angle 360º is divided in the same percentages

When do we use pie diagram?

When the data are given in percentage.

When different aspect of a variable is to be displayed.

When the data are to be compared normally.

Analysis of Data

David J. Luck and Ronald S Rubin

" The refinement and manipulation of data that prepares them for application of logical inference"

Data analysis is a process of inspecting, cleansing, transforming and modelling data with the goal of discovering useful information, informing conclusions and supporting decision-making

Purpose of Data Analysis

To ensure that necessary information is available to solve a problem

To convert huge data into meaningful form for interpretation

To confirm that date is legitimate

To verify the reliability of measuring instruments

To understand the result from surveys

Purpose of Data Analysis

To provide information on data gaps for designing and redesigning surveys to formulate quality objectives

To perform data quality assessment

To publish or summarise results of data analysis in official releases

To see that the data falls within parameters.

Types of Data Analysis

Text Analysis

Text Analysis is also referred to as Data Mining. It is a method to discover a pattern in large data sets using databases or data mining tools. It used to transform raw data into business information. Business Intelligence tools are present in the market which is used to take strategic business decisions. Overall, it offers a way to extract and examine data and deriving patterns and finally interpretation of the data

Statistical Analysis

Statistical Analysis shows "What happen?" by using past data in the form of dashboards. Statistical Analysis includes collection, Analysis, interpretation, presentation, and modelling of data. It analyses a set of data or a sample of data. There are two categories of this type of Analysis - Descriptive Analysis and Inferential Analysis.

Descriptive Analysis

analyses complete data or a sample of summarized numerical data. It shows mean and deviation for continuous data whereas percentage and frequency for categorical data.

Inferential Analysis

analyses sample from complete data. In this type of Analysis, you can find different conclusions from the same data by selecting different samples.

Diagnostic Analysis/Causal Analysis

Diagnostic Analysis shows "Why did it happen?" by finding the cause from the insight found in Statistical Analysis. This Analysis is useful to identify behaviour patterns of data. If a new problem arrives in your business process, then you can look into this Analysis to find similar patterns of that problem.

And it may have chances to use similar prescriptions for the new problems.

Predictive Analysis

Predictive Analysis shows "what is likely to happen" by using previous data. The simplest example is like if last year I bought two dresses based on my savings and if this year my salary is increasing double then I can buy four dresses. But of course, it's not easy like this because you have to think about other circumstances like chances of prices of clothes is increased this year or maybe instead of dresses you want to buy a new bike, or you need to buy a house!

So here, this Analysis makes predictions about future outcomes based on current or past data. Forecasting is just an estimate. Its accuracy is based on how much detailed information you have and how much you dig in it.

Prescriptive Analysis

Prescriptive Analysis combines the insight from all previous Analysis to determine which action to take in a current problem or decision. Most data driven companies are utilizing Prescriptive Analysis because predictive and descriptive Analysis are not enough to improve

data performance. Based on current situations and problems, they analyse the data and make decisions

Interpretation of Data

Data interpretation is the process of reviewing data through some predefined processes which will help assign some meaning to the data and arrive at a relevant conclusion.

It involves taking the result of data analysis, making inferences on the relations studied, and using them to conclude.

Essentials of Interpretation of data

Simplicity

Continuous

Complete in all respects

Balance

Flexibility

Selectivity

Acceptance

Use of resources

Utility of Action

Effective control

Importance/Significance of Interpretation of Data

Condensed data

Part of research process

Logical reasoning

Creative process

Accuracy & Simplicity

Decision making

Validate the findings

Predict the future

Use of hypotheses

Findings & recommendations

Descriptive analysis

A descriptive analysis is an important first step for conducting statistical analyses.

It gives an idea of the distribution of the data, helps to detect outliers and typos, and enable you identify associations among variables, thus making you ready to conduct further statistical analyses

Descriptive research has three main purposes:

To Describe, To Explain and To Validate findings.

Descriptive statistics is useful because it allows to take a large amount of data and summarize it.

For example, let's say you had data on the incomes of one million people. No one is going to want to read a million pieces of data; if they did, they wouldn't be able to get any useful information from it. On the other hand, if you summarize it, it becomes useful: an average wage, or a median income, is much easier to understand than reams of data

Descriptive statistics are one of the fundamental "must know" with any set of data. It gives a general idea of trends in the data including:

The mean, mode, median and range.

Variance and standard deviation.

Skewness.

Count, maximum and minimum

Descriptive statistics can be further broken down into several sub-areas, like:

Measures of central tendency.

Measures of Variability or Dispersion

Measures of Diversion from Normality/ Shapes of Distributions

Measures of Probability.

Measures of Central Tendency

Mean: The mean is the sum of the values in the set divided by the number of items in the set. There is basically no difference between the mean, the arithmetic mean, and the average. In statistics, the term 'mean' is used instead of 'average'

For example, what is the mean for 2, 4, and 6?

$(2 + 4 + 6)/3 = 12/3$

$(2 + 4 + 6)/3 = 4$

Mode: The mode of a set of data is the value or values that occur more often.

For example, the mode of 6, 5, 2, 3, 5, 1, 8 is 5 because 5 appears more often than the other values.

Median: The median of an ordered set of data is the value in the middle.

For example, the median of 4, 6, 8, 10, 12 is 8 since 8 is the value in the middle

Measures of Variability or Dispersion

Variability is the measure of Spread of a data set.

The commonly used measures are range variance and standard variance.

Range is the difference between the largest and smallest measures.

Variability relates to how spread out a group data is i.e., variability measures how much the score differs from each other.

Variability is also called as Dispersion or Spread

Measures of Measures of Diversion from Normality/ Shapes of Distributions

In statistics, normality tests are used to determine whether a data set is modelled for normal distribution. Many statistical functions require that a distribution be normal or nearly normal.

There are both graphical and statistical methods for evaluating normality:

Graphical methods include the histogram and normality plot

Statistically, two numerical measures of shape – skewness and excess kurtosis – can be used to test for normality. If skewness is not close to zero, then data set is not normally distributed

Skewness

In statistics, skewness is a measure of the asymmetry of the probability distribution of a random variable about its mean. In other words, skewness tells you the amount and direction of skew (departure from horizontal symmetry). The skewness value can be positive or negative, or even undefined. If skewness is 0, the data are perfectly symmetrical, although it is quite unlikely for real-world data.

Skewness lets you test by how much the overall shape of a distribution deviates from the shape of the normal distribution

Kurtosis

The term ‘Kurtosis’ refers to ‘peakiness’ or the flatness of a frequency distribution as compared with the normal. The measure of Kurtosis describes the degree of concentration of frequencies in a given distribution.

Kurtosis is a measure of how differently shaped are the tails of a distribution as compared to the tails of the normal distribution. While skewness focuses on the overall shape, Kurtosis focuses on the tail shape

Testing of Hypothesis

A Hypothesis test is a statistical test that is used to determine whether there is enough evidence in a sample of data to infer that a certain condition is true for the entire population.

An assumption about a population parameter or about probability distribution of a random variable regarding population is known as Statistical hypothesis. Simply, the hypothesis is an assumption which is tested to determine the relationship between two data sets.

For example, waiting time for getting into a bus in 15 minutes

A hypothesis test examines two opposing hypotheses about a population: the null hypothesis and the alternative hypothesis

Null hypothesis (H0):

The null hypothesis states that a population parameter is equal to a value. The null hypothesis is often an initial claim that researchers specify using previous research or knowledge

Alternative Hypothesis (H1):

The alternative hypothesis states that the population parameter is different than the value of the population parameter in the null hypothesis. The alternative hypothesis is what you might believe to be true or hope to prove true.

In hypothesis testing, two opposing hypotheses about a population are formed Viz. Null Hypothesis (H0) and Alternative Hypothesis (H1). The Null hypothesis is the statement which asserts that there is no difference between the sample statistic and population parameter and is the one which is tested, while the alternative hypothesis is the statement which stands true if the null hypothesis is rejected.

The null hypothesis is opposite of alternative hypothesis. If we can prove that the null hypothesis is false , then we can conclude that the research or alternative hypothesis is true.

One-Tailed Test & Two-Tailed Test

There are two forms of research hypothesis.

The critical or rejection region can be one side of the distribution or both sides of the distribution depending on the nature of the research hypothesis.

Directional research hypothesis or one tail test has one critical region. On the contrary, the non-directional research hypothesis or two tail tests has two rejection regions on both sides of the distribution.

One-Tailed Hypothesis Tests

One-tailed hypothesis tests are also known as directional and one-sided tests because you can test for effects in only one direction. When you perform a one tailed test, the entire significance level percentage goes into the extreme end of one tail of the distribution.

A one-tailed test is a statistical test in which the critical area of a distribution is one-sided so that it is either greater than or less than a certain value, but not both

An example can be whether a machine produces more than one-percent defective products. In this situation, if the estimated value exists in one of the one-sided critical areas, depending on the direction of interest (greater than or less than), the alternative hypothesis is accepted over the null hypothesis

Two-tailed test

A test that is conducted to show whether the mean of the sample is significantly greater than and significantly less than the mean of a population is considered a two-tailed test

A two-tailed test is appropriate if the estimated value is greater or less than a certain range of values, for example, whether a test taker may score above or below a specific range of scores

Example:

A medical researcher is interested in finding out whether a new medication will have any undesirable side effects. He knows that the mean pulse rate for the population under study is 82 beats per minute. Hence, he tries to find whether the pulse rate will increase, decrease, or remain unchanged after a patient takes the medication?

Type 1 Error, Type 2 Error

In every study, there is a chance for error. In other words, there is always the possibility of reaching the wrong conclusion. There are two different errors that we can make, Type 1 error and Type 2 error.

Type 1 Error: We may reject the null hypothesis when it is, in fact, true. This occurs when we think there is a difference between our groups, but there really isn't. We can minimise this possibility by selecting an appropriate alpha level .

Type 2 Error: Type 2 error occurs when we fail to reject a null hypothesis when it is, in fact, false (i.e., believing that the groups do not differ, when in fact they do).

Unfortunately, these two errors are inversely related. As we try to control for a Type 1 error, we actually increase the likelihood that we will commit a Type 2 error.

Hypothesis Testing Procedure

Set up a Hypothesis

Set up a suitable Significance Level

Determining a suitable Test Statistic

Determining the Critical Region

Performing computations

Decision-making

Understanding t-Tests: 1-sample, 2-sample, and Paired t-Test

In statistics, t-tests are a type of hypothesis test that allows you to compare means. They are called t-tests because each t-test boils your sample data down to one number, the t-value. If you understand how t-tests calculate t-values, you're well on your way to understanding how these tests work

Note: The One Sample T Test can only compare a single sample mean to a specified constant. It cannot compare sample means between two or more groups. If you wish to

compare the means of multiple groups to each other, you will likely want to run an Independent Samples to Test (to compare the means of two groups) or a One-Way ANOVA (to compare the means of two or more groups).

The One Sample T Test

The One Sample T Test determines whether the sample mean is statistically different from a known or hypothesized population mean. The One Sample T Test is a parametric test.

This test is also known as:

Single Sample t Test

The variable used in this test is known as:

Test variable

In a One Sample t Test, the test variable is compared against a "test value", which is a known or hypothesized value of the mean in the population.

Analysis of variance ANOVA

Analysis of variance (ANOVA) is a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) used to analyse the differences among group means in a sample. ANOVA was developed by the statistician Ronald Fisher. The ANOVA is based on the law of total variance, where the observed variance in a particular variable is partitioned into components attributable to different sources of variation. In its simplest form, ANOVA provides a statistical test of whether two or more population means are equal, and therefore generalizes the t-test beyond two means.

ANOVA is a form of statistical hypothesis testing heavily used in the analysis of experimental data. A test result (calculated from the null hypothesis and the sample) is called statistically significant if it is deemed unlikely to have occurred by chance, assuming the truth of the null hypothesis. A statistically significant result, when a probability (p-value) is less than a pre-specified threshold (significance level), justifies the rejection of the null hypothesis, but only if a probability of the null hypothesis is not high.

F-Test

An-test is any statistical test in which the test statistic has an F-distribution under the null hypothesis.

It is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled.

Exact "F-tests" mainly arise when the models have been fitted to the data using least squares. The name was coined by George W. Snedecor, in honour of Sir Ronald A. Fisher. Fisher initially developed the statistic as the variance ratio in the 1920s.

Common examples of the use of F-tests include the study of the following cases:

The hypothesis that the means of a given set of normally distributed populations, all having the same standard deviation are equal. This is perhaps the best-known Ftest, and plays an important role in the analysis of variance (ANOVA).

The hypothesis that a proposed regression model fits the data well.

The hypothesis that a data set in a regression analysis follows the simpler of two proposed linear models that are nested within each other.

In addition, some statistical procedures, such as Scheffé's method for multiple comparisons adjustment in linear models, also use F-tests.

Chi-Square Test

A chi-squared test, also written as$\chi 2$ test, is a statistical hypothesis test that is valid to perform when the test statistic is chi-squared distributed under the null hypothesis, specifically Pearson's chi-squared test and variants thereof. Pearson's chi-squared test is used to determine whether there is a statistically significant difference between the expected frequencies and the observed frequencies in one or more categories of a contingency table.

In the standard applications of this test, the observations are classified into mutually exclusive classes. If the null hypothesis that there are no differences between the classes in the population is true, the test statistic computed from the observations follows a $\chi 2$ frequency distribution. The purpose of the test is to evaluate how likely the observed frequencies would be assuming the null hypothesis is true.

Test statistics that follow a $\chi 2$ distribution occur when the observations are independent and normally distributed, which assumptions are often justified under the central limit theorem. There are also $\chi 2$ tests for testing the null hypothesis of independence of a pair of random variables based on observations of the pairs.

Chi-squared tests often refers to tests for which the distribution of the test statistic approaches the $\chi 2$ distribution asymptotically, meaning that the sampling distribution (if the null hypothesis is true) of the test statistic approximates a chi squared distribution more and more closely as sample sizes increase.

T Test, F Test and Z test

T-test is a small sample test. It was developed by William Gosset in 1908 It is also called students to test(pen name)

Size of sample is small (n<30)

Z-test was given by Fisher. T-test is used when correlation coefficient of population Is zero. But if population Coeff. Correlation is not zero then z-test is used.

In z-test sample size is large (n>30)

F-test also given by Fisher. F-test is used to the two independent estimation of population variance.

Advanced Statistical Techniques

Correlation & Regression Analysis

Correlation is a statistical technique that can show whether and how strongly pairs of variables are related. For example, height and weight are related; taller people tend to be heavier than shorter people. The relationship isn't perfect

Correlation is a statistic that measures the degree to which two variables move in relation to each other.

In finance, the correlation can measure the movement of a stock with that of a benchmark index, such as the S&P 500.

Correlation measures association, but doesn't show if x causes y or vice versa, or if the association is caused by a third–perhaps unseen–factor.

Correlation

Correlation shows the strength of a relationship between two variables and is expressed numerically by the correlation coefficient. The correlation coefficient's values range between -1.0 and 1.0.

A perfect positive correlation means that the correlation coefficient is exactly 1. This implies that as one security moves, either up or down, the other security moves in lockstep, in the same direction.

A perfect negative correlation means that two assets move in opposite directions, while a zero correlation implies no linear relationship at all.

When two sets of data are strongly linked together, we say they have a High Correlation.

The word Correlation is made of Co- (meaning "together"), and Relation

Correlation is Positive when the values increase together, and

Correlation is Negative when one value decreases as the other increases

Correlation Is Not Causation

A common saying is "Correlation Is Not Causation".

What it really means is that a correlation does not prove one thing causes the other:

One thing might cause the other

The other might cause the first to happen • They may be linked by a different thing

Or it could be random chance!

There can be many reasons the data has a good correlation.

Example

A few years ago, a survey of employees found a strong positive correlation between "Studying an external course" and Sick Days.

Does this mean:

Studying makes them sick?

Sick people study a lot?

Or did they lie about being sick so they can study more?

Without further research we can't be sure why.

Scatter Plot

A scatter plot (also called a scatterplot, scatter graph, scatter chart, scattergram, or scatter diagram) is a type of plot or mathematical diagram to display values for typically two variables for a set of data. If the points are coded (colour/shape/size), one additional variable can be displayed. The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis.

Types of Correlation

1.Perfect Correlation: Positive & Negative

2.Strong Correlation: Positive & Negative

3.Weak Correlation: Positive & Negative

4. No Correlation

5.Non-Linear:

Some uses of Correlations

Prediction

If there is a relationship between two variables, we can make predictions about one from another.

Validity

Concurrent validity (correlation between a new measure and an established measure). •

Reliability

Test-retest reliability (are measures consistent).

Inter-rater reliability (are observers consistent).

Theory verification

Predictive validity.

Correlation Coefficient

Correlation Coefficient measures the degree of linear relationship between the variables. It is introduced by Karl Pearson.

Karl Pearson's Coefficient of Correlation is an extensively used mathematical method in which the numerical representation is applied to measure the level of relation between linear related variables. The coefficient of correlation is expressed by "r".

Karl Pearson Correlation Coefficient Formula

Value of Coefficient Correlation

When a correlation coefficient is (1) that means every increase in one variable, there is a positive increase in other fixed proportion. For instance, shoe sizes change according to the length of the foot and are (almost) perfect correlation.

When a correlation coefficient is (-1) that means every positive increase in one variable, there is a negative decrease in other fixed proportion.

When a correlation coefficient is (0) for every increase, it means there is no positive or negative increase, and the two variables are not related.

Regression

Regression is a statistical method used in finance, investing, and other disciplines that attempts to determine the strength and character of the relationship between one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

Correlation & Regression are closely related. Correlation is only measure of association and is of no use in prediction. Regression is more powerful than correlation

Regression methods are useful in developing quantitative relationship between variables which can be used in prediction.

It plays an important role in human activities, as it is a powerful tool which used to forecast the past, present and future events on the basis of present or past events.

Types of regression

The two basic types of regression are

simple linear regression and

multiple linear regression, although there are non-linear regression methods for more complicated data

and analysis. Simple linear regression uses one independent variable to explain or predict the outcome of the dependent variable Y, while multiple linear regression uses two or more independent variables to predict the outcome.

Uses of Regression

Regression can help finance and investment professionals as well as professionals in other businesses. Regression can also help predict sales for a company based on weather, previous sales, GDP growth, or other types of conditions.

The capital asset pricing model (CAPM) is an often-used regression model in finance for pricing assets and discovering costs of capital.

Formula for Regression

Simple linear regression: $Y = a + bX + u$

Multiple linear regression: $Y = a + b1X1 + b2X2 + b3X3 + ... + btXt + u$

Where:

Y = the variable that you are trying to predict (dependent variable).

X = the variable that you are using to predict Y (independent variable).

a = the intercept.

b = the slope.

u = the regression residual.

Application of Regression

Regression is often used to determine how many specific factors such as the price of a commodity, interest rates, particular industries, or sectors influence the price movement of an asset. The aforementioned CAPM is based on regression, and it is utilized to project the expected returns for stocks and to generate costs of capital. A stock's returns are regressed against the returns of a broader index, such as the S&P 500, to generate a beta for the particular stock.

Beta is the stock's risk in relation to the market or index and is reflected as the slope in the CAPM model. The return for the stock in question would be the dependent variable Y, while the independent variable X would be the market risk premium.

Additional variables such as the market capitalization of a stock, valuation ratios, and recent returns can be added to the CAPM model to get better estimates for returns. These additional factors are known as the Fama-French factors, named after the professors who developed the multiple linear regression model to better explain asset returns.1

Discriminant Analysis

Discriminant analysis is statistical technique used to classify observations into non-overlapping groups, based on scores on one or more quantitative predictor variables.

For example, a doctor could perform a discriminant analysis to identify patients at high or low risk for stroke. The analysis might classify patients into high- or low-risk groups, based on personal attributes (e.g., chololesterol level, body mass) and/or lifestyle behaviours (e.g., minutes of exercise per week, packs of cigarettes per day).

During a study, there are often questions that strike the researcher that must be answered.

These questions include questions like ‘are the groups different?’, ‘on what variables, are the groups most different?’, ‘can one predict which group a person belongs to using such variables?’ etc.

In answering such questions, discriminant analysis is quite helpful.

Discriminant analysis is a technique that is used by the researcher to analyse the research data when the criterion or the dependent variable is categorical and the predictor or the independent variable is interval in nature.

The term categorical variable means that the dependent variable is divided into a number of categories. For example, three brands of computers, Computer A, Computer B and Computer C can be the categorical dependent variable.

Discriminant analysis (DA) is a multivariate technique used to classify a given set of objects, individuals or entities into two or more groups or categories based on a given set of predetermined variables relating to their characteristics, types, or any other attributes.

This technique can also be used to identify which variables contribute to making the classification.

The DA indicates which of the predictors are the most differentiating (highest discriminant weights); in other words, which dimensions distinguish best among these consumer segments and why respondents fall into one group versus another group.

In sum, the DA is a technique for classification, differentiation, and profiling

Objective of discriminant analysis

The objective of discriminant analysis is to develop discriminant functions that are nothing but the linear combination of independent variables that will discriminate between the categories of the dependent variable in a perfect manner. It enables the researcher to examine whether significant differences exist among the groups, in terms of the predictor variables. It also evaluates the accuracy of the classification.

Discriminant analysis is described by the number of categories that is possessed by the dependent variable.

Establishing a discriminate function.

This function (a mathematical equation) is used to discriminate between individuals in the population and allocate each one of them to a group within the population. The function is established on the basis of a series of measurements (predictor variables) or observations made on the individuals.

To know whether significant differences exist among the groups, in terms of the predictor variables.

Finding out which predictor variables contribute to most of the inter-group differences.

Evaluating the accuracy of classification.

Some Fields of Application of DA

Agriculture research: e.g., to analyse technology adoption behaviour of farmers • Financial research: e.g., to analyse default behaviour of borrowers

Human resources: e.g., to identify the characteristics of high performers • Market research: to analyse market segmentation Examples: A classic example of DA is for a bank's loan department to develop a credit risk profile for current customers to predict the outcome (credit risk or not) for new loan applications.

Other examples would be to predict whether new customers will buy or not, be brand loyal or not, if a sales strategy will have a low, medium, or high success rate, or into which segment new customers will fall.

Applications of discriminant analysis

There are many examples that can explain when discriminant analysis fits.

It can be used to know whether heavy, medium and light users of soft drinks are different in terms of their consumption of frozen foods.

In the field of psychology, it can be used to differentiate between the price sensitive and non-price sensitive buyers of groceries in terms of their psychological attributes or characteristics.

In the field of business, it can be used to understand the characteristics or the attributes of a customer possessing store loyalty and a customer who does not have store loyalty

Powerful tool for analysing and describing group differences and for classifying cases into groups formed on the basis of their similarities and differences on multiple variables.

In the field of psychology, it can be used for differentiate between price sensitive and non-price sensitive buyers of groceries in terms of their psychological attributes.

In the business, it can be used for understanding customers store loyalty.

It is useful in arts , humanities, social sciences and can be used to tackle the problems of selection , placement and taxonomy.

Difference between multiple regression and DA:

It is similar to regression equation in which there is one dependent variable and one or more independent variables.

However, it differs from the multiple regression with respect to the measurement of the dependent variable.

In regression analysis, the dependent variable is measured in metric scale whereas in the discriminant analysis it is measured in nominal (non-metric) scale.

Difference between Metric and Non-metric Variables

Metric variable: Variable with a constant unit of measurement. If a metric variable is scaled from 1 to 9, the difference between 1 and 2 is the same as that between 8 and 9.

Nonmetric variable: Variable with values that serve merely as a label or means of identification, also referred to as a categorical, nominal, binary, qualitative, or taxonomic variable. The number on a football jersey is an example.

Some Important Concepts Used in DA

Discriminant function: It is the equation representing output of a discriminant analysis using a particular process.

Discriminating variables: These are the variables that have been identified to discriminate between two or more groups.

Discriminant score: It is similar to the estimated value of dependent variable in a multiple regression. Putting the values of the predictor variables in the discriminant function and multiplying with the coefficients, we get discriminant scores.

Centroid: Centroid or group centroid is the mean value of discriminant scores for a particular group. The number of centroids is equal to the number of groups included in the analysis.

Standardized canonical discriminant function coefficients: These are similar to partial regression coefficients of multiple regression and reveal relative importance of discriminating variables.

Structure matrix: It reports the discriminant loadings of the variables on the discriminant function. The discriminant loadings are the simple Pearson correlations between the predictor variables and the discriminant function. Usually, any predictor with a loading of 0.30 or more is considered to be central in defining the discriminant dimension.

Canonical discriminant function coefficients: These are unstandardized discriminant function coefficients, and are used to write down the discriminant function.

Cluster Analysis

Cluster analysis is a class of techniques that are used to classify objects or cases into relative groups called clusters. Cluster analysis is also called classification analysis or numerical taxonomy. In cluster analysis, there is no prior information about the group or cluster membership for any of the objects.

Cluster Analysis has been used in marketing for various purposes. Segmentation of consumers in cluster analysis is used on the basis of benefits sought from the purchase of the product. It can be used to identify homogeneous groups of buyers.

Cluster analysis involves formulating a problem, selecting a distance measure, selecting a clustering procedure, deciding the number of clusters, interpreting the profile clusters and finally, assessing the validity of clustering.

The variables on which the cluster analysis is to be done should be selected by keeping past research in mind. It should also be selected by theory, the hypotheses being tested, and the judgment of the researcher. An appropriate measure of distance or similarity should be selected; the most commonly used measure is the Euclidean distance or its square.

Typical research questions the cluster analysis answers

Medicine – What are the diagnostic clusters? To answer this question the researcher would devise a diagnostic questionnaire that includes possible symptoms (for example, in psychology, anxiety, depression etc.). The cluster analysis can then identify groups of patients that have similar symptoms.

Marketing – What are the customer segments? To answer this question a market researcher may conduct a survey covering needs, attitudes, demographics, and behaviour of customers. The researcher then may use cluster analysis to identify homogenous groups of customers that have similar needs and attitudes.

Education – What are student groups that need special attention? Researchers may measure psychological, aptitude, and achievement characteristics. A cluster analysis then may identify what homogeneous groups exist among students (for example, high achievers in all subjects, or students that excel in certain subjects but fail in others).

Biology – What is the taxonomy of species? Researchers can collect a data set of different plants and note different attributes of their phenotypes. A cluster analysis can group those observations into a series of clusters and help build a taxonomy of groups and subgroups of similar plants

Cluster Analysis

Clustering procedures in cluster analysis may be hierarchical, non-hierarchical, or a two-step procedure. A hierarchical procedure in cluster analysis is characterized by the development of a tree like structure. A hierarchical procedure can be agglomerative or divisive. Agglomerative methods in cluster analysis consist of linkage methods, variance methods, and centroid methods. Linkage methods in cluster analysis are comprised of single linkage, complete linkage, and average linkage.

The non-hierarchical methods in cluster analysis are frequently referred to as K means clustering. The two-step procedure can automatically determine the optimal number of clusters by comparing the values of model choice criteria across different clustering solutions. The choice of clustering procedure and the choice of distance measure are interrelated. The relative sizes of clusters in cluster analysis should be meaningful. The clusters should be interpreted in terms of cluster centroids.

Applications of cluster analysis

Medicine: What are diagnostic clusters?

Marketing: What are customer segments?

Education: What are student groups that need special attention?

Biology: What is the taxonomy of species?

Factor Analysis

Factor Analysis (FA) is an exploratory data analysis method used to search influential underlying factors or latent variables from a set of observed variables. It helps in data

interpretations by reducing the number of variables. It extracts maximum common variance from all variables and puts them into a common score.

Factor analysis is widely utilized in market research, advertising, psychology, finance, and operation research. Market researchers use factor analysis to identify price sensitive customers, identify brand features that influence consumer choice, and helps in understanding channel selection criteria for the distribution channel.

Factor analysis is a linear statistical model. It is used to explain the variance among the observed variable and condense a set of the observed variable into the unobserved variable called factors. Observed variables are modelled as a linear combination of factors and error terms (Source). Factor or latent variable is associated with multiple observed variables, who have common patterns of responses. Each factor explains a particular amount of variance in the observed variables. It helps in data interpretations by reducing the number of variables.

Factor analysis is a method for investigating whether a number of variables of interest X1, X2,……., Xl, are linearly related to a smaller number of unobservable factors F1, F2,..……, Fk.

Assumptions:

There are no outliers in data.

Sample size should be greater than the factor.

There should not be perfect multicollinearity.

There should not be homoscedasticity between the variables.

Types of Factor Analysis

Exploratory Factor Analysis: It is the most popular factor analysis approach among social and management researchers. Its basic assumption is that any observed variable is directly associated with any factor.

Confirmatory Factor Analysis (CFA): Its basic assumption is that each factor is associated with a particular set of observed variables. CFA confirms what is expected on the basic.

How does factor analysis work?

The primary objective of factor analysis is to reduce the number of observed variables and find unobservable variables. These unobserved variables help the market researcher to conclude the survey. This conversion of the observed variables to unobserved variables can be achieved in two steps:

Factor Extraction: In this step, the number of factors and approach for extraction selected using variance partitioning methods such as principal components analysis and common factor analysis.

Factor Rotation: In this step, rotation tries to convert factors into uncorrelated factors — the main goal of this step to improve the overall interpretability. There are lots of rotation

methods that are available such as: Varimax rotation method, Quart Imax rotation method, and Promax rotation method.

Terminology used in factor analysis

What is a factor?

A factor is a latent variable which describes the association among the number of observed variables. The maximum number of factors are equal to a number of observed variables. Every factor explains a certain variance in observed variables. The factors with the lowest amount of variance were dropped. Factors are also known as latent variables or hidden variables or unobserved variables or Hypothetical variables.

What are the factor loadings?

The factor loading is a matrix which shows the relationship of each variable to the underlying factor. It explains how closely variables are related. It shows the correlation coefficient for observed variable and factor. It shows the variance explained by the observed variables.

What are Eigen values?(latent roots)-sum of squared values of factor loadings. • Eigenvalues represent variance explained each factor from the total variance. It is also known as characteristic roots. Eigen value indicates the relative importance of each factor.

What are Communalities?

Communalities are the sum of the squared loadings for each variable. It represents the common variance. It ranges from 0-1 and value close to 1 represents more variance.

What is Factor Rotation?

Rotation is a tool for better interpretation of factor analysis. It re-distributes the communalities with a clear pattern of loadings. Different rotations reveal different structures in the data.

Choosing the Number of Factors

Kaiser criterion is an analytical approach, which is based on the more significant proportion of variance explained by factor will be selected. The eigenvalue is a good criterion for determining the number of factors. Generally, an eigenvalue greater than 1 will be considered as selection criteria for the feature.

The graphical approach is based on the visual representation of factors' eigenvalues also called scree plot. This scree plot helps us to determine the number of factors where the curve makes an elbow.

Factor Analysis Test

Adequacy Test

Before you perform factor analysis, you need to evaluate the “factorability” of our dataset. Factorability means "can we found the factors in the dataset?". There are two methods to check the factorability or sampling adequacy:

Bartlett's Test

Kaiser-Meyer-Olkin Test

Bartlett's test of sphericity checks whether or not the observed variables intercorrelate at all using the observed correlation matrix against the identity matrix. If the test found statistically insignificant, you should not employ a factor analysis.

In this Bartlett 's test, the p-value is 0. The test was statistically significant, indicating that the observed correlation matrix is not an identity matrix.

Kaiser-Meyer-Olkin (KMO) Test measures the suitability of data for factor analysis. It determines the adequacy for each observed variable and for the complete model. KMO estimates the proportion of variance among all the observed variable. Lower proportion id more suitable for factor analysis. KMO values range between 0 and 1. Value of KMO less than 0.6 is considered inadequate. The overall KMO for our data is 0.84, which is excellent. This value indicates that you can proceed with your planned factor analysis

Pros and Cons of Factor Analysis

Factor analysis explores large dataset and finds interlinked associations. It reduces the observed variables into a few unobserved variables or identifies the groups of inter-related variables, which help the market researchers to compress the market situations and find the hidden relationship among consumer taste, preference, and cultural influence. Also, it helps in improve questionnaire in for future surveys. Factors make for more natural data interpretation.

Results of factor analysis are controversial. Its interpretations can be debatable because more than one interpretation can be made of the same data factors. After factor identification and naming of factors requires domain knowledge.

Merits of factor analysis

Condense and simplify data

Helpful in identifying important and interesting, relationship among the data that were not easy to see from the data alone

The analysis can reveal the latent factors(underlying factors that are not directly observed) that determine relationships among several variables concerning a research study.

Limitations of factor analysis

Involves laborious computations

Results of single factor analysis is considered generally less reliable.

Atleast need to be done twice. If similar results are obtained, then confidence is increased

Complicated analysis tool, which can be used only when one has thorough knowledge and enough experience in handling this tool

Applications of Factor Analysis

Investment: Creating a portfolio of diversified stocks

Human resources: Staffing of a company

Insurance: Actuarial tables and statistics to create policies

Restaurants: Demographics and target customers to create menu.

Multidimensional scaling(MDS)

Multidimensional scaling is a visual representation of distances or dissimilarities between sets of objects. “Objects” can be colours, faces, map coordinates, political persuasion, or any kind of real or conceptual stimuli . Objects that are more similar (or have shorter distances) are closer together on the graph than objects that are less similar (or have longer distances). As well as interpreting dissimilarities as distances on a graph, MDS can also serve as a dimension reduction technique for high dimensional data.

From a non-technical point of view, the purpose of multidimensional scaling (MDS) is to provide a visual representation of the pattern of proximities (i.e., similarities or distances) among a set of objects. For example, given a matrix of perceived similarities between various brands of air fresheners, MDS plots the brands on a map such that those brands that are perceived to be very similar to each other are placed near each other on the map, and those brands that are perceived to be very different from each other are placed far away from each other on the map

Applications of MDS

Applied in variety of disciplines, including psychology, economics, sociology, political science, anthropology and organisational behaviour.

MDS can be used to understand wide range of phenomenon- ranging from perceptions of nations to visual patterns

In market research MDS is often used to plot data such as perceptions about products or brands. This will display both the number and nature of the data set in an easy way to interpret, visual way. MDS is ideally suited to study conflict & negotiation.

Research Report

Research Report

Research reports are recorded data prepared by researchers or statisticians after analysing information gathered by conducting organized research, typically in the form of surveys or qualitative methods

Reports usually are spread across a vast horizon of topics but are focused on communicating information about a particular topic and a very niche target market. The primary motive of research reports is to convey integral details about a study for marketers to consider while designing new strategies.

Definition

Oxford dictionary, "a statement or record of ascertained facts"

In the ordinary language, a report is a summary statement/document of facts, and figures prepared for the purpose of information, analysis and following up of action.

According to American Marketing Society, the purpose of report is to "convey to the interested persons whole result of the study in sufficient detail and so arranged as to enable each reader to comprehend the data and determine himself the validity of conclusions"

Features of Research Report

Precision

In a good report, the report writer is very clear about the exact and definite purpose of writing the report. His investigation, analysis, recommendations and others are directed by this central purpose. Precision of a report provides the unity to the report and makes it a valuable document for best usage.

Accuracy of Facts

Information contained in a report must be based on accurate fact. Since decisions are taken on the basis of report information, any inaccurate information or statistics will lead to wrong decision. It will hamper to achieve the organizational goal.

Relevancy

The facts presented in a report should not be only accurate but also be relevant. Irrelevant facts make a report confusing and likely to be misleading to make proper decision.

Reader-Orientation

While drafting any report, it is necessary to keep in mind about the person who is going to read it. That's why a good report is always reader oriented. Reader's knowledge and level of understanding should be considered by the writer of report.

Well reader-oriented information qualifies a report to be a good one.

Simple Language

This is just another essential features of a good report. A good report is written in a simple language avoiding vague and unclear words. The language of the report should not be

influenced by the writer's emotion or goal. The message of a good report should be self-explanatory.

Conciseness

A good report should be concise but it does not mean that a report can never be long. Rather it means that a good report or a business report is one that transmits maximum information with minimum words. It avoids unnecessary detail and includes everything which are significant and necessary to present proper information.

Unbiased Recommendation

Recommendation on report usually make effect on the reader mind. So, if recommendations are made at the end of a report, they must be impartial and objective. They should come as logical conclusion for investigation and analysis.

Clarity

Clarity depends on proper arrangement of facts. A good report is absolutely clear. Reporter should make his purpose clear, define his sources, state his findings and finally make necessary recommendation. To be an effective communication through report, A report must be clear to understand for making communication success.

Attractive Presentation

Presentation of a report is also a factor which should be consider for a good report. A good report provides a catchy and smart look and creates attention of the reader. Structure, content, language, typing and presentation style of a good report should be attractive to make a clear impression in the mind of its reader.

Importance of Research Report

Decision-Making Tool

Today’s complex business organizations require thousands of information. Reports provide the required information a large number of important decisions in business or any other area are taken on the basis of the information presented in the reports. This is one of the great importance of the report.

Investigation

Whenever there is any problem, a committee or commission or study group investigates the problem to find out the reason behind the problem and present the findings with or without the recommendation in the form of a report. It is another importance of the report.

Evaluation

Large scale organizations are engaged in multidimensional activities. It is not possible for a single top executive to keep a personal watch on what others are doing. So, the executive depends on reports to evaluate the performance of various departments or units.

Quick Location

There is no denying the fact that business executives need information for quick decision-making. As top executives are found to be busy for various purposes), they need vital sources of information. Such sources can be business reports.

Development of skill

Report writing skill develops the power of designing, organization coordination, judgment, and communication.

Neutral presentation of facts

Facts are required to be presented in a neutral way; such presentation is ensured through a report as it investigates, explains, and evaluates any facts independently.

Professional Advancement

The report also plays a major role in professional achievement. For promotion to the rank-and-file position, satisfactory job performance is enough to help a person. But for promotion to a high-level position, intellectual ability is highly required. Such ability can be expressed through the report submitted to a higher authority

Importance of Research Report

Facilitates policy decisions and follow up measures

Acts as a document of permanent record

Presents data, conclusions and recommendations in a preside form

Gives publicity to research works

Acts as a communication channel

Guides management in planning, decision making and coordination.

Can be used as a reference material for future.

Structure of Research Report

(Layout/Components/Contents)

A research report should be a complete, compact, self-sufficient as well as self-explanatory document.

Research Report should be drafted as per standard layout. Layout means systematic way of presenting data. Normally layout contains four components. They are:

Introduction

Main body of the report

Conclusions & Recommendations

Appendices

Normal contents of research report

A. Introduction:

1.Title of the Report

2.Letter of Transmittal: Whom it is directed, reasons, official authorisation

3.Title Page: First page-

4.Index of contents

5.Acknowledgements

6.Description of problem

7.Purpose & Scope of the study

8.Research Methodology used

9.Summary of findings

B. Main body of research report:

1.Analysis of the problem

2.Methodology used for analysis

3.Presentations of findings & Limitations

C. Conclusions:

D. Recommendations:

E. Appendices/Annexures:

Appendices include forms and questionnaire issued, maps, tables, charts and diagrams, technical details of sample design, procedures and methodology of study. Etc.,

F. Footnotes and Bibliography

G. Index and Glossary:

Meanings and definition of difficult words or terms used in the report.

Types of Report

A. On the Basis of Medium of Communication:

1.Oral Report

2..Written Report

B. On the Basis of Purpose:

1.Report for publication

2.Report for Executives(Popular Report)

3.Report for detailed Study

C. On another basis

1.Interim Report

2.Summary Report

3.Research Abstract

4.Research Article

1.Oral Report:

It is the presentation of information (findings, conclusions and recommendations) through the spoken words(Verbal presentation)

The purpose of oral presentation is to deliver the findings and implications in concise, organised and interesting fashion.

It is desirable to distribute a handout containing statistical data to the audience while presenting the oral data.

Oral reporting acts as a good supplement to written reporting.

2.Written Report

Written report is presented in the form of self-contained and comprehensive document with necessary technical, commercial and other details.

Research Reports are normally prepared in written form in convenience and ready reference of the sponsoring agency.

Benefits

Used as a reference document

Facilitates execution of recommendations • All details of report are given in compact form

Company executives prefers written reports.

3.Report for Publication

This report is for the publication in the press, magazine, journal, etc in the form of article.

It is a brief/summary type report which gives important details, findings and recommendations.

It is for the general information

The writer must consider the level of audience while preparing such report.

4.Technical Report

A technical report is a comprehensive and detailed report

It contains various details such as statement of objectives and problems, methodology and procedures followed, findings and conclusions, recommendations, appendices, summary and bibliography.

It is a complete report and can be used as self-explanatory document.

Mainly useful for those interested in the detailed study of the report.

5.Report of the Executives(Popular Report)

A popular report is submitted to company executives for consideration and decision making.

It is a brief report which gives a short visual summary of the results of the research work undertaken.

Company executives prefer such summary report.

It is intended for decision makers.

6.Interim Report

Interim report is submitted after the completion of some work, but not the full

This happens when the work involved in the research project is large in volume and may requires many years for completion.

The interim report contains narration of what has done so far and outcome of the work completed

Government Committees or Commissions may prepare and submit interim report

7.Summary Report

A summary report is normally for the benefit of general public.

Such reports are desirable for any study whose findings are of general interest.

Prepared in simple language

Charts and tables are used for quick understandings

Published in newspapers and magazines and related to educational or social or cultural matters or even in economic and social matters.

8.Research Abstract

Summary of Technical report already prepared.

Normally engineering students or research scholars prepare research abstract before submitting their thesis.

It contains details of research work in brief and compact manner.

Such abstract is sent by research scholars to the university and the university sends it to examiners or referees for scrutiny.

9.Research Article

It is a brief report/summary document prepared for publication in research or professional journal.

It relates to a specific subject and contains various details in a compact manner.

Complete but compact documents.

Useful to the students as well as research institutions.

Process in report writing

Research Report is the outcome of the research activity. It contains all details of the research work in a compact manner.

Its and final and major job.

Skilled job requires experience, communication skills, maturity related detailed information relating to research project undertaken.

Report is the outcome of lengthy research process.

Steps in research writing process

1.Planning for writing research report

2.Deciding the target audience

3.Selecting suitable format

4.Logicl arrangement

5.Drafting & redrafting of report

6. Approval of the report

7.Printing & binding of report

8.Submission of report

9.Feedback of such report

Essentials of a good research report

Firm foundation

Effective organisation

Suitable Framework

Tailored to suit the needs of the reader

Informative, concise, compact, accurate & reliable document

Clarity

Objectivity

Logical & attractive arrangement of facts

Use of visual devices

Secrecy

References

Impersonal style

Attractive graphic presentation

Impressive language

Guidelines for good /impressive report writing

1.Based on sound foundation and detailed study of the problem

2.Recommendations as per conclusions drawn

3.Simple & lucid style of writing

4.As per well planned framework

5.Due considerations to readers.

6. Brief & compact

7.Adequate attention to editorial, presentation & physical appearance

8. Attractive & persuasive presentation

9.Accuracy in report writing

10.Effective organisation of the report

11. Selectivity in presentation

12.Attention to the needs of the sponsoring organisation

13. Inclusion of essential element in the report

14.Read,think,plan,write and rewrite.