

RESEARCH METHODOLOGY

UNIT- 2

Syllabus - UNIT-II : SAMPLING AND DATA COLLECTION

Sampling: Meaning, definition, need and types. Sampling errors - Merits and demerits of Sampling. Data collection: Sources of data; Primary and Secondary data. Procedure for data collection, Tool of data collection - Questionnaire – Interview-Schedule.

Expected Learning Outcome:

- 1. Define principal concepts of sampling.**
- 2. Explains the advantages of sampling.**
- 3. Lists the stages of sampling process.**
- 4. Categorizes and defines the sampling methods.**

SAMPLING

SAMPLING - Meaning:

When researching an aspect of the human mind or behavior, researchers simply cannot collect data from every single individual in most cases. Instead, they choose a smaller sample of individuals that represent the larger group. If the sample is truly representative of the population in question, researchers can then take their results and generalize them to the larger group.

In statistics, a sample is a subset of a population that is used to represent the entire group as a whole. When doing research, it is often impractical to survey every member of a particular population because the sheer number of people is simply too large. Sampling is a process, which allows us to study a small group of people from the large group to

derive inferences that are likely to be applicable to all the people of the large group. Sometimes it is not feasible to study a whole group

Sample Vs Sampling

Sample is the subset of the population. The process of selecting a sample is known as sampling. Number of elements in the sample is the sample size. The difference lies between the above two is whether the sample selection is based on randomization or not.

Definition:

Sampling is a technique of selecting individual members or a subset of the population to make statistical inferences from them and estimate characteristics of the whole population.

Different sampling methods are widely used by researchers in market research so that they do not need to research the entire population to collect actionable insights. It is also a time-convenient and a cost-effective method and hence forms the basis of any research design. Sampling techniques can be used in a research survey software for optimum derivation.

For example, if a drug manufacturer would like to research the adverse side effects of a drug on the country's population, it is almost impossible to conduct a research study that involves everyone. In this case, the researcher decides a sample of people from each demographic and then researches them, giving him/her indicative feedback on the drug's behavior.

NEED OF SAMPLING:

Everyone who has ever worked on a research project knows that resources are limited; time, money and people never come in an unlimited supply. For that reason, most research projects aim to gather data from a sample of people, rather than from the entire population (the census being one of the few exceptions). This is because sampling allows researchers to:

Save Time

Contacting everyone in a population takes time. And, invariably, some people will not respond to the first effort at contacting them, meaning researchers have to invest more time for follow-up. Random sampling is much faster than surveying everyone in a population, and obtaining a non-random sample is almost always faster than random sampling. Thus, sampling saves researchers lots of time.

Save Money

The number of people a researcher contacts is directly related to the cost of a study. Sampling saves money by allowing researchers to gather the same answers from a sample that they would receive from the population.

Non-random sampling is significantly cheaper than random sampling, because it lowers the cost associated with finding people and collecting data from them. Because all research is conducted on a budget, saving money is important.

Collect Richer Data

Sometimes, the goal of research is to collect a little bit of data from a lot of people (e.g., an opinion poll). At other times, the goal is to collect a lot of information from just a few people (e.g., a user study or ethnographic interview). Either way, sampling allows researchers to ask participants more questions and to gather richer data than does contacting everyone in a population.

Features of Good Sampling:

1. **Representative Character:** very important feature of sampling is, it must be representative of population. Since the samples are taken from population they are expected to be representative but the way samples selected makes difference as far as representative character is concerned.
2. **Small sampling error:** Sample design must be such which results in a small sampling error.
3. **Adequate number of samples:** entire study is governed by the representative units. Since the opinions are very subjective in social sciences it is necessary that their number should be adequate.

4. **Economic Viability:** Sample design must be viable in the context of funds available for the research study.

Types of sampling: sampling methods

Sampling in market research is of two types –

1. Probability sampling and non-probability sampling. Let's take a closer look at these two methods of sampling.

Probability sampling: Probability sampling is a sampling technique where a researcher sets a selection of a few criteria and chooses members of a population randomly. All the members have an equal opportunity to be a part of the sample with this selection parameter.

Non-probability sampling: In non-probability sampling, the researcher chooses members for research at random. This sampling method is not a fixed or predefined selection process. This makes it difficult for all elements of a population to have equal opportunities to be included in a sample.

In this blog, we discuss the various probability and non-probability sampling methods that you can implement in any market research study.

Types of probability sampling

Probability sampling is a sampling technique in which researchers choose samples from a larger population using a method based on the theory of probability. This sampling method considers every member of the population and forms samples based on a fixed process.

For example, in a population of 1000 members, every member will have a 1/1000 chance of being selected to be a part of a sample. Probability sampling eliminates bias in the population and gives all members a fair chance to be included in the sample.

There are four types of probability sampling techniques:

Simple random sampling: One of the best probability sampling techniques that helps in saving time and resources, is the Simple Random Sampling method. It is a reliable method of obtaining information where every single member of a population is chosen randomly, merely by chance. Each individual has the same probability of being chosen to be a part of a sample.

For example, in an organization of 500 employees, if the HR team decides on conducting team building activities, it is highly likely that they would prefer picking chits out of a bowl. In this case, each of the 500 employees has an equal opportunity of being selected.

Cluster sampling: Cluster sampling is a method where the researchers divide the entire population into sections or clusters that represent a population. Clusters are identified and included in a sample based on demographic parameters like age, sex, location, etc. This makes it very simple for a survey creator to derive effective inference from the feedback.

For example, if the United States government wishes to evaluate the number of immigrants living in the Mainland US, they can divide it into clusters based on states such as California, Texas, Florida, Massachusetts, Colorado, Hawaii, etc. This way of conducting a survey will be more effective as the results will be organized into states and provide insightful immigration data.

Systematic sampling: Researchers use the systematic sampling method to choose the sample members of a population at regular intervals. It requires the selection of a starting point for the sample and sample size that can be repeated at regular intervals. This type of sampling method has a predefined range, and hence this sampling technique is the least time-consuming.

Stratified random sampling: Stratified random sampling is a method in which the researcher divides the population into smaller groups that don't overlap but represent the

entire population. While sampling, these groups can be organized and then draw a sample from each group separately.

For example, a researcher looking to analyze the characteristics of people belonging to different annual income divisions will create strata (groups) according to the annual family income.. By doing this, the researcher concludes the characteristics of people belonging to different income groups. Marketers can analyze which income groups to target and which ones to eliminate to create a roadmap that would bear fruitful results.

Uses of probability sampling

There are multiple uses of probability sampling. They are:

Reduce Sample Bias: Using the probability sampling method, the bias in the sample derived from a population is negligible to non-existent. The selection of the sample mainly depicts the understanding and the inference of the researcher. Probability sampling leads to higher quality data collection as the sample appropriately represents the population.

Diverse Population: When the population is vast and diverse, it is essential to have adequate representation so that the data is not skewed towards one demographic. For example, if Square would like to understand the people that could make their point-of-sale devices, a survey conducted from a sample of people across the US from different industries and socio-economic backgrounds helps.

Create an Accurate Sample: Probability sampling helps the researchers plan and create an accurate sample. This helps to obtain well-defined data.

Types of non-probability sampling

The non-probability method is a sampling method that involves a collection of feedback based on a researcher or statistician's sample selection capabilities and not on a fixed selection process. In most situations, the output of a survey conducted with a non-probable sample leads to skewed results, which may not represent the desired target population. But, there are situations such as the preliminary stages of research or cost constraints for conducting research, where non-probability sampling will be much more useful than the other type.

Four types of non-probability sampling explain the purpose of this sampling method in a better manner:

Convenience sampling: This method is dependent on the ease of access to subjects such as surveying customers at a mall or passers-by on a busy street. It is usually termed as convenience sampling, because of the researcher's ease of carrying it out and getting in touch with the subjects. Researchers have nearly no authority to select the sample elements, and it's purely done based on proximity and not representativeness. This non-probability sampling method is used when there are time and cost limitations in collecting feedback. In situations where there are resource limitations such as the initial stages of research, convenience sampling is used.

For example, startups and NGOs usually conduct convenience sampling at a mall to distribute leaflets of upcoming events or promotion of a cause – they do that by standing at the mall entrance and giving out pamphlets randomly.

Judgmental or purposive sampling: Judgmental or purposive samples are formed by the discretion of the researcher. Researchers purely consider the purpose of the study, along with the understanding of the target audience. For instance, when researchers want to understand the thought process of people interested in studying for their master's

degree. The selection criteria will be: “Are you interested in doing your masters in ...?” and those who respond with a “No” are excluded from the sample.

Snowball sampling: Snowball sampling is a sampling method that researchers apply when the subjects are difficult to trace. For example, it will be extremely challenging to survey shelter Less people or illegal immigrants. In such cases, using the snowball theory, researchers can track a few categories to interview and derive results. Researchers also implement this sampling method in situations where the topic is highly sensitive and not openly discussed—for example, surveys to gather information about HIV Aids. Not many victims will readily respond to the questions. Still, researchers can contact people they might know or volunteers associated with the cause to get in touch with the victims and collect information.

Quota sampling: In Quota sampling, the selection of members in this sampling technique happens based on a pre-set standard. In this case, as a sample is formed based on specific attributes, the created sample will have the same qualities found in the total population. It is a rapid method of collecting samples.

Uses of non-probability sampling

Non-probability sampling is used for the following:

Create a hypothesis: Researchers use the non-probability sampling method to create an assumption when limited to no prior information is available. This method helps with the immediate return of data and builds a base for further research.

Exploratory research: Researchers use this sampling technique widely when conducting qualitative research, pilot studies, or exploratory research.

Budget and time constraints: The non-probability method when there are budget and time constraints, and some preliminary data must be collected. Since the survey design is not rigid, it is easier to pick respondents at random and have them take the survey or questionnaire.

How do you decide on the type of sampling to use?

For any research, it is essential to choose a sampling method accurately to meet the goals of your study. The effectiveness of your sampling relies on various factors. Here are some steps expert researchers follow to decide the best sampling method.

- a) Jot down the research goals. Generally, it must be a combination of cost, precision, or accuracy.
- b) Identify the effective sampling techniques that might potentially achieve the research goals.
- c) Test each of these methods and examine whether they help in achieving your goal.
- d) Select the method that works best for the research.

Select your respondents

SAMPLING ERROR

“Sampling error is the error that arises in a data collection process as a result of taking a sample from a population rather than using the whole population.”

Sampling Error - Definition

Sampling error is defined as the amount of inaccuracy in estimating some value, which occurs due to considering a small section of the population, called the sample, instead of the whole population. It is also called an error. Sample surveys take into account the study of a tiny segment of a population, so, there is always a particular amount of inaccuracy in the information obtained. This inaccuracy can be defined as error variance or sampling error.

The measure of the sampling error can be calculated for particular sample size and design. This measure is termed as the correctness of the sampling plan. Sampling error is also due to the concept called sampling bias. This error is considered a systematic error.

Sampling error is one of two reasons for the difference between an estimate of a population parameter and the true, but unknown, value of the population parameter. The other reason is non-sampling error. Even if a sampling process has no non-sampling errors then estimates from different random samples (of the same size) will vary from sample to sample, and each estimate is likely to be different from the true value of the population parameter.

The sampling error for a given sample is unknown but when the sampling is random, for some estimates (for example, sample mean, sample proportion) theoretical methods may be used to measure the extent of the variation caused by sampling error.”

Non-sampling error:

“Non-sampling error is the error that arises in a data collection process as a result of factors other than taking a sample.

Non-sampling errors have the potential to cause bias in polls, surveys or samples.

There are many different types of non-sampling errors and the names used to describe them are not consistent. Examples of non-sampling errors are generally more useful than using names to describe them.

Sampling error is the statistical error that results when an analyst selects a sample that is not representative of the population as a whole. In other words, it is the difference between the observed value of a sample statistic (mean, variance, or standard deviation) and the actual but unknown population parameter. For example, we calculate the sampling error for variance as:

A Breakdown of Sampling Error

A sampling exercise that involves the selection of a few elements to represent the entire population is always susceptible to both sampling errors and non-sampling errors. While sampling errors result from the data collection process, non-sampling errors are unrelated to the sample and are systematic. A good example of a non-sampling error, for instance during the administration of a questionnaire, would manifest in the form of asking “leading” questions or phrasing the question in a manner that dictates the respondent to give a particular response.

Administrative Data

Administrative data are an important source of information for social science research. For example, school records have been used to track trends in student academic performance. Administrative data generally refers to data collected as part of the management and operations of a publicly funded program or service. Today, use of administrative data is becoming increasingly common in research about child care and early education. These data often are a relatively cost-effective way to learn more about the individuals and families using a particular service or participating in a particular program, but they do have some important limitations.

The advantages and disadvantages of using administrative data are described here. Issues pertaining to the access to such data are discussed. Terms relating to administrative data and its use in research studies are defined in the Research Glossary.

Advantages of Administrative Data

1. Administrative data make possible analyses at the state and local levels that are rarely possible using national survey data.
2. Such data often contain detailed, accurate measures of participation in various social programs. They typically include large numbers of cases, making possible many different types of analyses.

3. Data on the same individuals and/or same programs over a long period of time can be used for longitudinal and trend studies.
4. Potential for linking data from several programs in order to get a more complete picture of individuals and the services received.
5. At the state level, such data provide effective ways for assessing state-specific programs and can be useful for several forms of program evaluation.
6. The large sample sizes allow small program effects to be more easily detected, and permit effects to be estimated for different groups.
7. It is less expensive to obtain administrative data than to collect data directly on the same group.

Limitations of Administrative Data

Administrative data are collected to manage services and comply with government reporting regulations. Because the original purpose of the data is not research, this presents several challenges.

The administrative data only describe the individuals or families using a service and provide no information about similar people who do not use the service.

The potential observation period for any subject being studied (e.g., a person, a family, a child care program) is limited to the period of time that the subject is using the service for which the data are being collected.

Generally, only those services that are publicly funded are included in the administrative data. For example, a researcher cannot rely on subsidy data to learn about all child care providers in the state or on non-subsidized forms of child care being used to augment child care that is subsidized.

Many variables used in administrative data are not updated regularly, so it is important to learn how and when each variable is collected. For instance, an "earnings" variable in administrative data for subsidized child care generally is entered at the time that

eligibility is determined and then updated when eligibility is redetermined. When this is the case, there is no way to know, using administrative data alone, what a family earns in the months between eligibility determination and redetermination.

Important variables needed for a particular research study may not be collected in administrative data.

Because the data are limited to data on program participants, information on those eligible for the program but who are not enrolled is often not available. Thus, administrative data may not be especially useful for estimating certain characteristics such as participation rates.

Measurement error can pose a substantial challenge to analysts using administrative data. Factors affecting measurement error include:

Data that were improperly entered at the agency

Incomplete or inaccurate data items, particularly those items not required by the agency for management or reporting purposes

Missing values on variables that have been overwritten by updated versions when cases are reviewed

Procedures for accessing the data for research purposes can be time consuming and difficult. Protecting the privacy of program participants and the confidentiality of the data when they are used for research is a major concern to program officials.

Researchers interested in using administrative data for the purpose of research should expect to invest considerable time learning about the details of the administrative data system, the specific data elements being used, the data entry process and standards, and changes in the data system and data definitions over time. It also takes time to transform administrative data into research datasets that can be used in statistical analyses.

Elements of Sampling

Elements - Individual members of a population whose characteristics are to be measured

Population - The set of individuals to which study findings will be generalized

Sampling - Not necessary when all units in the population are identical or resources are unlimited

Sampling Frame - List of the study population

Enumeration Units - Units that contain one or more elements and that are listed in a sampling frame

Sampling Units - Units listed at each stage of a multistage sampling design.

Common types of Sampling Error

Population Specification Error - this error occurs when the researcher does not understand who they should survey. For example, imagine a survey about breakfast cereal consumption. Who to survey? It might be the entire family, the mother, or the children. The mother might make the purchase decision, but the children influence her choice.

Sample Frame Error - A frame error occurs when the wrong sub-population is used to select a sample. A classic frame error occurred in the 1936 presidential election between Roosevelt and Landon. The sample frame was from car registrations and telephone directories. In 1936, many Americans did not own cars or telephones, and those who did were largely Republicans. The results wrongly predicted a Republican victory.

Selection Error - this occurs when respondents self-select their participation in the study – only those that are interested respond. Selection error can be controlled by going extra lengths to get participation. A typical survey process includes initiating pre-survey

contact requesting cooperation, actual surveying, and post-survey follow-up. If a response is not received, a second survey request follows, and perhaps interviews using alternate modes such as telephone or person-to-person.

Non-Response - Non-response errors occur when respondents are different than those who do not respond. This may occur because either the potential respondent was not contacted or they refused to respond. The extent of this non-response error can be checked through follow-up surveys using alternate modes.

Sampling Errors - These errors occur because of variation in the number or representativeness of the sample that responds. Sampling errors can be controlled by

- (1) Careful sample designs,
- (2) Large samples (check out our online sample size calculator), and
- (3) Multiple contacts to assure representative response.

Type I error

The first kind of error is the rejection of a true null hypothesis as the result of a test procedure. This kind of error is called a type I error (false positive) and is sometimes called an error of the first kind.

Type II error

The second kind of error is the failure to reject a false null hypothesis as the result of a test procedure. This sort of error is called a type II error (false negative) and is also referred to as an error of the second kind.

How to Reduce Sampling Error?

There are two methods by which this sampling error can be reduced. The methods are

1. Increasing sample size
2. Stratification

Increasing Sample Size

From a population, we can select any sample of any size. The size depends on the experiment and the situation. If the size of the sample increases, the chance of occurrence of the sampling error will be less. There will be no error if the sample size and the population size coincide. Hence, sampling error is in inverse proportion to the sample size.

Stratification

If all the population units are homogeneous or the population has the same characteristic feature, it's very easy to get a sample. The sample can be taken as a representative of the entire population. But if the population is not homogeneous (i.e population with the different characteristic features); it is impossible to get a perfect sample. In such conditions, to get a better representative, the sample design is altered. The population is classified into different groups called strata, that contain similar units. From each of these strata, a sub-sample is selected in a random manner. Thus, all the groups are defined in the sample, the sampling error is reduced. Hence, the sub-sample size from each stratum is in proportion with the stratum size.

Merits and Demerits of Sampling Method of Data Collection

The methods, using which, we can get the samples; below are given its merits and demerits on the whole.

Merits:

1. Economical:

It is economical, because we have not to collect all data. Instead of getting data from 5000 farmers, we get it from 50-100 only.

2. Less Time Consuming:

As no of units is only a fraction of the total universe, time consumed is also a fraction of total time. Number of units is considerably small, hence the time.

3. Reliable:

If sample is taken judiciously, the results are very reliable and accurate.

4. Organisational Convenience:

As samples are taken and the number of units is smaller, the better (Trained) enumerators can be employed by the organisation.

5. More Scientific:

According to Prof R.A. Fisher, “The sample technique has four important advantages over census technique of data collection. They are Speed, Economy, Adaptability and Scientific approach.”

It is based on certain laws such as:

- (a) Law of Statistical Regularity
- (b) Law of Inertia of Large numbers
- (c) Law of Persistence
- (d) Law of Validity.

6. Detailed Enquiry:

A detailed study can be undertaken in case of the units included in the sample. Size of sample can be taken according to time and money available with the investigator.

7. Indispensable Method:

If universe is bigger, there remains no option but to proceed for this method. It is specially used for infinite, hypothetical and perishable universes.

Demerits:

1. Absence of Being Representative:

Methods, such as purposive sampling may not provide a sample, that is representative.

2. Wrong Conclusion:

If the sample is not representative, the results will not be correct. These will lead to the wrong conclusions.

3. Small Universe:

Sometimes universe is so small that proper samples cannot be taken not of it. Number of units are so less.

4. Specialised Knowledge:

It is a scientific method. Therefore, to get a good and representative sample, one should have special knowledge to get good sample and to perform proper analysis so that reliable result may be achieved.

5. Inherent defects:

The results which are achieved though the analysis of sampling data may not be accurate as this method have inherent defects. There is not even a single method of sampling which has no **demerit**.

6. Sampling Error:

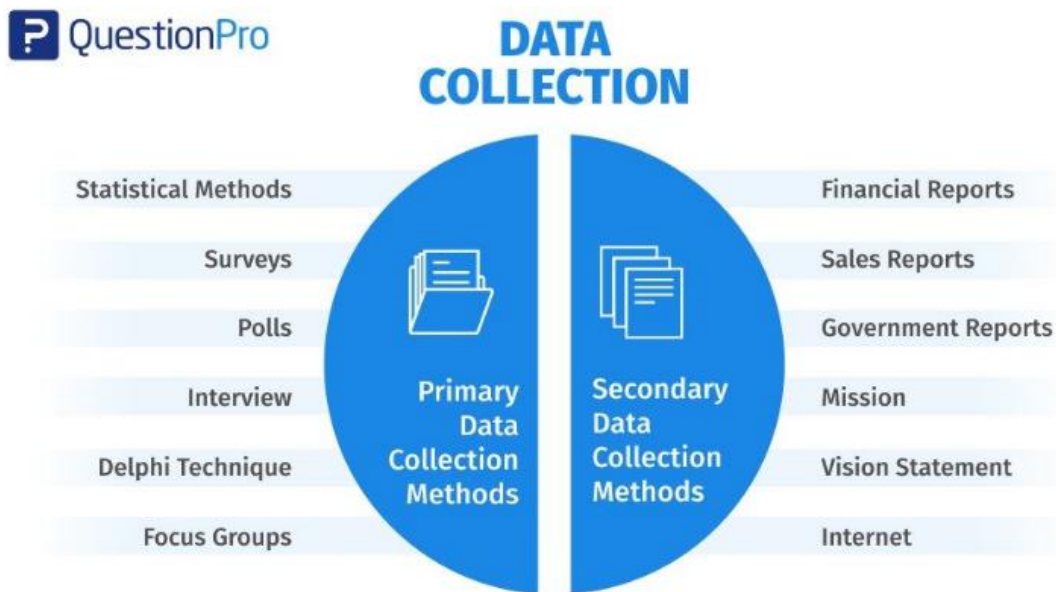
This method of sampling has many errors.

7. Personal Bias:

As in many cases the investigator, chooses samples, such as convenience method, chances of personal bias creep in.

DATA COLLECTION

Data Collection Methods: Definition, Examples and Sources



Data is a collection of facts, figures, objects, symbols, and events gathered from different sources. Organizations collect data to make better decisions. Without data, it would be difficult for organizations to make appropriate decisions, and so data is collected at

various points in time from different audiences. For instance, before launching a new product, an organization needs to collect data on product demand, customer preferences, competitors, etc. In case data is not collected beforehand, the organization's newly launched product may lead to failure for many reasons, such as less demand and inability to meet customer needs.

Although data is a valuable asset for every organization, it does not serve any purpose until analyzed or processed to get the desired results.

Data collection is a process of collecting information from all the relevant sources to find answers to the research problem, test the hypothesis and evaluate the outcomes.

SOURCES OF DATA:

Sources of Data can be classified into 2 types. Statistical sources refer to data that are gathered for some official purposes and incorporate censuses and officially administered surveys. Non-statistical sources refer to the collection of data for other administrative purposes or for the private sector.

What are the different sources of data?

Following are the two sources of data:

1. Internal Source

When data are collected from reports and records of the organisation itself, it is known as the internal source.

For example, a company publishes its 'Annual Report' on Profit and Loss, Total Sales, Loans, Wages etc.

2. External Source

When data are collected from outside the organisation, it is known as the external source.

For example, if a Tour and Travels Company obtains information on 'Karnataka Tourism' from Karnataka Transport Corporation, it would be known as external sources of data.

Types of Data

A) Primary Data

Primary data means 'First-hand information' collected by an investigator.

It is collected for the first time.

It is original and more reliable.

For example Population census conducted by the government of India after every 10 years.

B) Secondary Data

Secondary data refers to 'Second-hand information'.

These are not originally collected rather obtained from already published or unpublished sources.

For example the Address of a person taken from the Telephone Directory or Phone number of a company taken from 'Just Dial'.

Students can also refer to Meaning and Sources of Secondary Data

Methods of Collecting Primary Data

Direct Personal Investigation

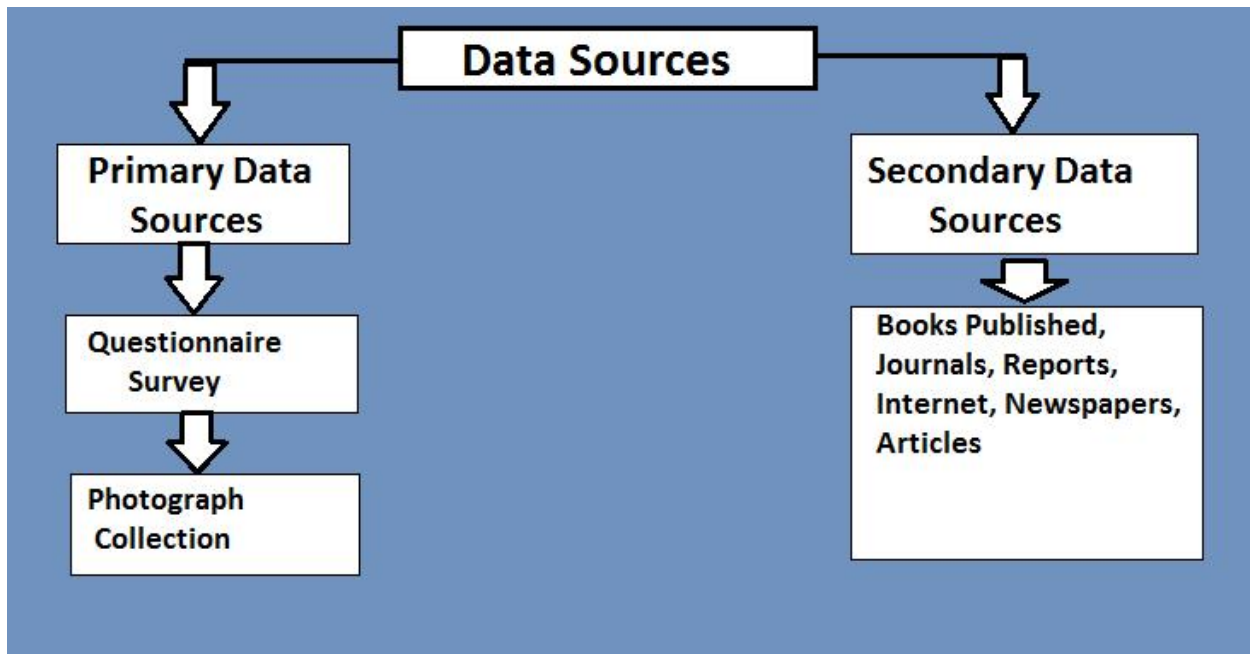
Indirect Oral Investigation

Information Through Correspondents

Telephonic Interview

Mailed Questionnaire

The questionnaire filled by enumerators



Procedure for data collection

. There are various steps commonly involved in the Data Collection process.

1. Identifying opportunities/issues for data collection

It involves identifying the particular opportunities/issues for collecting data. An in-depth analysis of external and internal processes is vital for getting a clear picture of all that occurs within an organization. By identifying and exploring specific issues in the data collection process, you can get the flexibility to make decisions about when and how you should collect information.

2. Set objectives

Once you consider the important questions, you have to set clear goals individualized for every issue based on the collection analysis and techniques. Every opportunity or issue is

reviewed in this step from the external and internal evaluation that is carried out earlier. A specific issue is chosen to start with.

3. Collecting Data

Once every question is clearly defined and the goals are properly set, information is collected. A committee is set up and specific people are chosen. These people are likely to be held accountable for every important decision that is taken regarding the collection process including finances, coordination, logistics, and designs. The logistics, resources, and technology needed for the implementation of data collection, as well as the people in charge of the data collection, are identified and determined.

4. Data Analysis and interpretation

Based on the qualitative or quantitative data collection techniques, data analysis could be intimidating at times. Businesses need to determine whether there are sufficient expertise and internal capacity for in-house data interpretation or whether a professional consultant should be hired for the task.

Often times, traditional data collection does not progress as fast as one would generally prefer, which is why it is important that CrowdForce provides access to fast and reliable data. Our network of field agents makes it possible to collect real-time data in different locations simultaneously. The hassle associated with data collection is drastically reduced as we combine steps 3 & 4 to reduce the amount of time used to gather the required data.

TOOLS OF DATA COLLECTION / INSTRUMENTS FOR DATA COLLECTION

Data collection is an important step in the research process. The instrument you choose to collect the data will depend on the type of data you plan on collecting (qualitative or quantitative) and how you plan to collect it.

A number of common data-collecting instruments are used in construction research:

1. Questionnaires
2. Interviews
3. Observations
4. Archival documents and government sources
5. Laboratory experiments
6. Quasi experiment

Primary Data Collection Methods

Primary data is collected from the first-hand experience and is not used in the past. The data gathered by primary data collection methods are specific to the research's motive and highly accurate.

Primary data collection methods can be divided into two categories: quantitative methods and qualitative methods.

Quantitative Methods:

Quantitative techniques for market research and demand forecasting usually make use of statistical tools. In these techniques, demand is forecast based on historical data. These methods of primary data collection are generally used to make long-term forecasts. Statistical methods are highly reliable as the element of subjectivity is minimum in these methods.

Time Series Analysis

The term time series refers to a sequential order of values of a variable, known as a trend, at equal time intervals. Using patterns, an organization can predict the demand for its products and services for the projected time.

Smoothing Techniques

In cases where the time series lacks significant trends, smoothing techniques can be used. They eliminate a random variation from the historical demand. It helps in identifying patterns and demand levels to estimate future demand. The most common methods used in smoothing demand forecasting techniques are the simple moving average method and the weighted moving average method.

Barometric Method

Also known as the leading indicators approach, researchers use this method to speculate future trends based on current developments. When the past events are considered to predict future events, they act as leading indicators.

Qualitative Methods:

Qualitative methods are especially useful in situations when historical data is not available. Or there is no need of numbers or mathematical calculations. Qualitative research is closely associated with words, sounds, feeling, emotions, colors, and other elements that are non-quantifiable. These techniques are based on experience, judgment, intuition, conjecture, emotion, etc.

Quantitative methods do not provide the motive behind participants' responses, often don't reach underrepresented populations, and span long periods to collect the data. Hence, it is best to combine quantitative methods with qualitative methods.

Surveys

Surveys are used to collect data from the target audience and gather insights into their preferences, opinions, choices, and feedback related to their products and services. Most survey maker software often a wide range of question types to select.

You can also use a ready-made survey template to save on time and effort. Online surveys can be customized as per the business's brand by changing the theme, logo, etc. They can be distributed through several distribution channels such as email, website, offline app, QR code, social media, etc. Depending on the type and source of your audience, you can select the channel.

Once the data is collected, survey software can generate various reports and run analytics algorithms to discover hidden insights. A survey dashboard can give you the statistics related to response rate, completion rate, filters based on demographics, export and sharing options, etc. You can maximize the effort spent on online data collection by integrating survey builder with third-party apps.

Polls

Polls comprise of one single or multiple choice question. When it is required to have a quick pulse of the audience's sentiments, you can go for polls. Because they are short in length, it is easier to get responses from the people.

Similar to surveys, online polls, too, can be embedded into various platforms. Once the respondents answer the question, they can also be shown how they stand compared to others' responses.

Interviews

In this method, the interviewer asks questions either face-to-face or through telephone to the respondents. In face-to-face interviews, the interviewer asks a series of questions to the interviewee in person and notes down responses. In case it is not feasible to meet the person, the interviewer can go for a telephonic interview. This form of data collection is suitable when there are only a few respondents. It is too time-consuming and tedious to repeat the same process if there are many participants.

Delphi Technique

In this method, market experts are provided with the estimates and assumptions of forecasts made by other experts in the industry. Experts may reconsider and revise their estimates and assumptions based on the information provided by other experts. The consensus of all experts on demand forecasts constitutes the final demand forecast.

Focus Groups

A small group of people, around 8-10 members, discuss the common areas of the problem. Each individual provides his insights on the issue concerned. A moderator regulates the discussion among the group members. At the end of the discussion, the group reaches a consensus.

Questionnaire

A questionnaire is a printed set of questions, either open-ended or closed-ended. The respondents are required to answer based on their knowledge and experience with the issue concerned. The questionnaire is a part of the survey, whereas the questionnaire's end-goal may or may not be a survey.

Secondary Data Collection Methods

Secondary data is the data that has been used in the past. The researcher can obtain data from the sources, both internal and external, to the organization.

Internal sources of secondary data:

1. Organization's health and safety records
2. Mission and vision statements
3. Financial Statements
4. Magazines
5. Sales Report
6. CRM Software
7. Executive summaries

External sources of secondary data:

1. Government reports
2. Press releases
3. Business journals
4. Libraries
5. Internet

The secondary data collection methods, too, can involve both quantitative and qualitative techniques. Secondary data is easily available and hence, less time-consuming and expensive as compared to the primary data. However, with the secondary data collection methods, the authenticity of the data gathered cannot be verified.

Secondary Data Collection Methods

Secondary data is data collected by someone other than the actual user. It means that the information is already available, and someone analyses it. The secondary data includes magazines, newspapers, books, journals etc. It may be either published data or unpublished data.

1. Published data are available in various resources including
2. Government publications
3. Public records
4. Historical and statistical documents
5. Business documents
6. Technical and trade journals
7. Unpublished data includes
8. Diaries
9. Letters
10. Unpublished biographies etc.

Questionnaire method of data collection

A questionnaire is a research instrument consisting of a series of questions for the purpose of gathering information from respondents. Questionnaires can be thought of as a kind of written interview. They can be carried out face to face, by telephone, computer or post. Questionnaires provide a relatively cheap, quick and efficient way of obtaining large amounts of information from a large sample of people. Questionnaire is as an instrument for research, which consists of a list of questions, along with the choice of answers, printed or typed in a sequence on a form used for acquiring specific information

from the respondents. The questionnaire is prepared in such a way that it translates the required information into a series of questions, that informants can and will answer.

Characteristics of a Good Questionnaire

The following are characteristics of good questionnaires:

1. It should consist of a well-written list of questions.
2. The questionnaire should deal with an important or significant topic to create interest among respondents.
3. It should seek only that data which cannot be obtained from other sources.
4. It should be as short as possible but should be comprehensive.
5. It should be attractive.
6. Directions should be clear and complete.
7. It should be represented in good psychological order proceeding from general to more specific responses.
8. Double negatives in questions should be avoided.
9. Putting two questions in one question also should be avoided. Every question should seek to obtain only one specific information.
10. It should be designed to collect information which can be used subsequently as data for analysis.

Format of Questions in Questionnaires

The questions asked can take two forms:

Restricted questions, also called closed-ended, ask the respondent to make choices — yes or no, check items on a list, or select from multiple choice answers. Restricted questions are easy to tabulate and compile.

Unrestricted questions are open-ended and allow respondents to share feelings and opinions that are important to them about the matter at hand.

Unrestricted questions are not easy to tabulate and compile, but they allow respondents to reveal the depth of their emotions.

If the objective is to compile data from all respondents, then sticking with restricted questions that are easily quantified is better.

If degrees of emotions or depth of sentiment is to be studied, then develop a scale to quantify those feelings.

Advantages of Questionnaire

1. One of the greatest benefits of questionnaires lies in their uniformity — all respondents see exactly the same questions.
2. It is an inexpensive method, regardless of the size of the universe.
3. Free from the bias of the interviewer, as the respondents answer the questions in his own words.
4. Respondents have enough time to think and answer.
5. Due to its large coverage, respondents living in distant areas can also be reached conveniently.
6. Comparability

Limitations of Questionnaire

The risk of collection of inaccurate and incomplete information is high in the questionnaire, as it might happen that people may not be able to understand the question correctly.

The main demerits of this system can also be listed here:

1. Low rate of return of the duly filled in questionnaires; bias due to no-response is often Indeterminate.
2. It can be used only when respondents are educated and cooperating.
3. The control over questionnaire may be lost once it is sent.

4. There is inbuilt inflexibility because of the difficulty of amending the approach once
5. questionnaires have been dispatched.
6. There is also the possibility of ambiguous replies or omission of replies altogether to certain
7. Questions; interpretation of omissions is difficult.
8. It is difficult to know whether willing respondents are truly representative.
9. This method is likely to be the slowest of all.

Interview Schedules

Data Collection through Schedules – Very similar to the Questionnaire method. The main difference is that a schedule is filled by the trained enumerator who is specially appointed for the purpose. Enumerator goes to the respondents, asks them the questions from the Performa in the order listed, and records the responses in the space provided.

Open-ended – Questions in which the respondent answers in his own words.

Closed-ended (or Fixed Alternative) – Question in which respondent selects one or more options from pre-determined set of responses.

Simple dichotomy → Closed ended question with only two response alternatives

Multiple Choice → Closed ended question with more than two response alternatives.

Determinant choice – Multiple choice question in which respondent must select only one of the response alternatives.

Checklist question - Multiple choice question in which respondent can select more than one of the response alternatives.

Investigator - One who conducts the investigation i.e. statistical enquiry and seeks information is known as Investigator. It can be an individual person or an organization.

Enumerators- Enumerators are the persons who help the Investigators in the collection of data.

Informant - Informants are the respondents who supply the information to the investigator or enumerators.

Interview Method of Data Collection

Types of interviews

An interview is generally a qualitative research technique which involves asking open-ended questions to converse with respondents and collect elicited data about a subject. The interviewer in most cases is the subject matter expert who intends to understand respondent opinions in a well-planned and executed series of questions and answers. Interviews are similar to focus groups and surveys when it comes to garnering information from the target market but are entirely different in their operation – focus groups are restricted to a small group of 6-10 individuals whereas surveys are quantitative in nature. Interviews are conducted with a sample from a population and the key characteristic they exhibit is their conversational tone.

Fundamental Types of Interviews in Research

A researcher has to conduct interviews with a group of participants at a juncture in the research where information can only be obtained by meeting and personally connecting with a section of their target audience. Interviews offer the researchers with a platform to prompt their participants and obtain inputs in the desired detail. There are three fundamental types of interviews in research:

Structured Interviews:

Structured interviews are defined as research tools that are extremely rigid in their operations and allow very little or no scope of prompting the participants to obtain and analyze results. It is thus also known as a standardized interview and is significantly

quantitative in its approach. Questions in this interview are pre-decided according to the required detail of information.

Structured interviews are excessively used in survey research with the intention of maintaining uniformity throughout all the interview sessions.

They can be closed-ended as well as open-ended – according to the type of target population. Closed-ended questions can be included to understand user preferences from a collection of answer options whereas open-ended can be included to gain details about a particular section in the interview.

Advantages of structured interviews:

Structured interviews focus on the accuracy of different responses due to which extremely organized data can be collected. Different respondents have different type of answers to the same structure of questions – answers obtained can be collectively analyzed.

1. They can be used to get in touch with a large sample of the target population.
2. The interview procedure is made easy due to the standardization offered by structured interviews.
3. Replication across multiple samples becomes easy due to the same structure of interview.
4. As the scope of detail is already considered while designing the interview, better information can be obtained and the researcher can analyze the research problem in a comprehensive manner by asking accurate research questions.
5. Since the structure of the interview is fixed, it often generates reliable results and is quick to execute.
6. The relationship between the researcher and the respondent is not formal due to which the researcher can clearly understand the margin of error in case the

respondent either degrades to be a part of the survey or is just not interested in providing the right information.

Disadvantages of structured interviews:

1. Limited scope of assessment of obtained results.
2. The accuracy of information overpowers the detail of information.
3. Respondents are forced to select from the provided answer options.
4. The researcher is expected to always adhere to the list of decided questions irrespective of how interesting the conversation is turning out to be with the participants.
5. A significant amount of time is required for a structured interview.
6. Learn more: Market Research

Semi-Structured Interviews:

Semi-structured interviews offer a considerable amount of leeway to the researcher to probe the respondents along with maintaining basic interview structure. Even if it is a guided conversation between researchers and interviewees – an appreciable flexibility is offered to the researchers. A researcher can be assured that multiple interview rounds will not be required in the presence of structure in this type of research interview.

Keeping the structure in mind, the researcher can follow any idea or take creative advantage of the entire interview. Additional respondent probing is always necessary to garner information for a research study. The best application of semi-structured interview is when the researcher doesn't have time to conduct research and requires detailed information about the topic.

Advantages of semi-structured interviews:

Questions of semi-structured interviews are prepared before the scheduled interview which provides the researcher with time to prepare and analyze the questions.

1. It is flexible to an extent while maintaining the research guidelines.
2. Researchers can express the interview questions in the format they prefer, unlike the structured interview.
3. Reliable qualitative data can be collected via these interviews.
4. Flexible structure of the interview.

Disadvantages of semi-structured interviews:

Participants may question the reliability factor of these interviews due to the flexibility offered.

Comparing two different answers becomes difficult as the guideline for conducting interviews is not entirely followed. No two questions will have the exact same structure and the result will be an inability to compare and infer results.

Unstructured Interviews:

Also called as in-depth interviews, unstructured interviews are usually described as conversations held with a purpose in mind – to gather data about the research study. These interviews have the least number of questions as they lean more towards a normal conversation but with an underlying subject.

The main objective of most researchers using unstructured interviews is to build a bond with the respondents due to which there are high chances that the respondents will be 100% truthful with their answers. There are no guidelines for the researchers to follow and so, they can approach the participants in any ethical manner to gain as much information as they possibly can for their research topic.

Since there are no guidelines for these interviews, a researcher is expected to keep their approach in check so that the respondents do not sway away from the main research motive. For a researcher to obtain the desired outcome, he/she must keep the following factors in mind:

Intent of the interview.

The interview should primarily take into consideration the participant's interest and skills.

All the conversations should be conducted within permissible limits of research and the researcher should try and stick by these limits.

The skills and knowledge of the researcher should match the purpose of the interview.

Researchers should understand the do's and don'ts of unstructured interviews.

Advantages of Unstructured Interviews:

Due to the informal nature of unstructured interviews – it becomes extremely easy for researchers to try and develop a friendly rapport with the participants. This leads to gaining insights in extreme detail without much conscious effort.

The participants can clarify all their doubts about the questions and the researcher can take each opportunity to explain his/her intention for better answers.

There are no questions which the researcher has to abide by and this usually increases the flexibility of the entire research process.

Disadvantages of Unstructured Interviews:

As there is no structure to the interview process, researchers take time to execute these interviews.

The absence of a standardized set of questions and guidelines indicates that the reliability of unstructured interviews is questionable.

In many cases, the ethics involved in these interviews are considered borderline upsetting.

Learn more: [Qualitative Market Research](#)

Methods of Research Interviews:

There are three methods to conduct research interviews, each of which is peculiar in its application and can be used according to the research study requirement.

Personal Interviews:

Personal interviews are one of the most used types of interviews, where the questions are asked personally directly to the respondent. For this, a researcher can have a guide online surveys to take note of the answers. A researcher can design his/her survey in such a way that they take notes of the comments or points of view that stands out from the interviewee.

Advantage:

1. More information and that too in greater depth can be obtained.
2. Interviewer by his own skill can overcome the resistance, if any, of the respondents; the interview method can be made to yield an almost perfect sample of the general population.
3. There is greater flexibility under this method as the opportunity to restructure questions is always there, specially in case of unstructured interviews.
4. Observation method can as well be applied to recording verbal answers to various questions.
5. Personal information can as well be obtained easily under this method.

6. Samples can be controlled more effectively as there arises no difficulty of the missing returns; non-response generally remains very low.
7. The interviewer can usually control which person(s) will answer the questions. This is not possible in mailed questionnaire approach. If so desired, group discussions may also be held.
8. The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions than would be the case if mailed questionnaire is used.
9. The language of the interview can be adopted to the ability or educational level of the person interviewed and as such misinterpretations concerning questions can be avoided.
10. The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results.

Disadvantages:

1. It is a very expensive method, especially when large and widely spread geographical sample is taken.
2. There remains the possibility of the bias of interviewer as well as that of the respondent; there also remains the headache of supervision and control of interviewers.
3. Certain types of respondents such as important officials or executives or people in high income groups may not be easily approachable under this method and to that extent the data may prove inadequate.
4. This method is relatively more-time-consuming, specially when the sample is large and recalls upon the respondents are necessary.

5. The presence of the interviewer on the spot may over-stimulate the respondent, sometimes even to the extent that he may give imaginary information just to make the interview interesting.
6. Under the interview method the organisation required for selecting, training and supervising the field-staff is more complex with formidable problems.
7. Interviewing at times may also introduce systematic errors.
8. Effective interview presupposes proper rapport with respondents that would facilitate free and frank responses. This is often a very difficult requirement.

Telephonic Interviews:

Telephone interviews: This method of collecting information consists in contacting respondents on telephone itself. It is not a very widely used method, but plays important part in industrial surveys, particularly in developed regions. The chief merits of such a system are:

Advantages:

1. To find the interviewees it is enough to have their telephone numbers on hand.
2. They are usually lower cost.
3. The information is collected quickly.
4. Having a personal contact can also clarify doubts, or give more details of the questions.

Disadvantages:

1. It is more flexible in comparison to mailing method.
2. It is faster than other methods i.e., a quick way of obtaining information.

3. It is cheaper than personal interviewing method; here the cost per response is relatively low.
4. Recall is easy; callbacks are simple and economical.
5. There is a higher rate of response than what we have in mailing method; the non-response is generally very low.
6. Replies can be recorded without causing embarrassment to respondents.
7. Interviewer can explain requirements more easily.
8. At times, access can be gained to respondents who otherwise cannot be contacted for one reason or the other.
9. No field staff is required.
10. Representative and wider distribution of sample is possible.

Email or Web Page Interviews:

Online research is growing more and more because consumers are migrating to a more virtual world and it is best for each researcher to adapt to this change.

The increase in people with Internet access has made it popular that interviews via email or web page stand out among the types of interviews most used today. For this nothing better than an online survey.

More and more consumers are turning to online shopping, which is why they are a great niche to be able to carry out an interview that will generate information for the correct decision making.

Advantages of email surveys:

Speed in obtaining data

The respondents respond according to their time, at the time they want and in the place they decide.

Online surveys can be mixed with other research methods or using some of the previous interview models. They are tools that can perfectly complement and pay for the project.

A researcher can use a variety of questions, logics, create graphs and reports immediately.

Undoubtedly, the objective of the research will set the pattern of what types of interviews are best for data collection. Based on the research design, a research can plan and test the questions, for instance, if the questions are the correct and if the survey flows in the best way.

In addition there are other types of research that can be used under specific circumstances, for example in the case of no connection or adverse situations to carry out surveyors, in these types of occasions it is necessary to conduct a field research, which can not be considered an interview if not rather a completely different methodology.

To summarize the discussion, an effective interview will be one that provides researchers with the necessary data to know the object of study and that this information is applicable to the decisions researchers make.

Key Terms to Remember

- **Sampling:** A statistical method is a miniature picture or cross section of the entire group or aggregate from which the sample” - P.V.Young
- **Population:** Population consists of whole set of data or information from the entire universe, which is considered to be the whole source of information.
- **Sampling frame:** Sampling frame is a defined part of a population.
- **Random Sampling:** every item in the population has a chance of being selected as a sample.
- **Non random:** samples are judged on the basis of personal knowledge, opinion, interest etc. of the researcher

Questions

1. Define Sampling, Sample and Population.
2. What are the types of sampling?
3. Explain the features of good sampling.
4. How the sample size is determined?
5. Undertake any research problem. Consider the type of population and calculate the adequacy of sample size using appropriate formula.
6. Study the research articles in the area of your interest. Study how the respective researcher has selected the sample (type of population, sampling method for respective study).

MCQ – Questions

1. Sample is a sub-set of:
 - A. **Population**
 - B. Data
 - C. Set
 - D. Distribution
2. List of all the units of the population is called:
 - A. Random sampling
 - B. Bias
 - C. **Sampling frame**
 - D. Probability sampling
3. Any measure of the population is called:
 - A. Finite
 - B. **Parameter**
 - C. Without replacement
 - D. Random
4. If all the units of a population are surveyed, it is called:
 - A. Random sample
 - B. Random sampling
 - C. Sampled population
 - D. **Complete enumeration**
5. Probability distribution of a statistics is called:
 - A. Sampling
 - B. Parameter
 - C. Data
 - D. **Sampling distribution**

6. The difference between a statistic and the parameter is called:
- A. Probability
 - B. **Sampling error**
 - C. Random
 - D. Non-random
7. Standard deviation of sampling distribution of a statistic is called:
- A. Serious error
 - B. Dispersion
 - C. **Standard error**
 - D. Difference
8. A distribution formed by all possible values of a statistics is called
- A. Binomial distribution
 - B. Hypergeometric distribution
 - C. Normal distribution
 - D. **Sampling distribution**
9. In probability sampling, probability of selecting an item from the population is known and is:
- A. Equal to zero
 - B. **Non zero**
 - C. Equal to one
 - D. All of the above
10. A population about which we want to get some information is called:
- A. Finite population
 - B. Infinite population
 - C. Sampling population
 - D. **Target population**

11. Study of population is called:

- A. Parameter
- B. Statistic
- C. Error
- D. **Census**

12. Sampling based upon equal probability is called:

- A. Probability sampling
- B. Systematic sampling
- C. **Simple random sampling**
- D. Stratified random sampling

13. In sampling with replacement, an element can be chosen:

- A. Less than once
- B. **More than once**
- C. Only once
- D. Difficult to tell

14. Standard deviation of sample mean without replacement will be _____ standard deviation of sample mean with replacement:

- A. **Less than**
- B. More than
- C. 2 times
- D. Equal to

15. In sampling without replacement, an element can be chosen:

- A. Less than once
- B. More than once
- C. **Only once**
- D. Difficult to tell

16. Which of the following statement is true?
- A. Standard error is always one
 - B. Standard error is always zero
 - C. Standard error is always negative
 - D. **Standard error is always positive**
17. Random sampling is also called:
- A. **Probability sampling**
 - B. Non-probability sampling
 - C. Sampling error
 - D. Random error
18. Non-random sampling is also called:
- A. Biased sampling
 - B. **Non-probability sampling**
 - C. Random sampling
 - D. Representative sample
19. Sampling error can be reduced by:
- A. Non-probability sampling
 - B. Increasing the population
 - C. Decreasing the sample size
 - D. **Increasing the sample size**
20. A Plan for obtaining a sample from a population is called:
- A. Population design
 - B. **Sampling design**
 - C. Sampling frame
 - D. Sampling distribution

21. The difference between the expected value of a statistic and the value of the parameter being estimated is called a:
- A. Sampling error
 - B. Non-sampling error
 - C. Standard error
 - D. **Bias**
22. The standard deviation of any sampling distribution is called:
- A. **Standard error**
 - B. Non-sampling error
 - C. Type- I error
 - D. Type II-error
23. The standard error increases when sample size:
- A. Increases
 - B. **Decreases**
 - C. is Fixed
 - D. is More than 30
24. The mean of the sample means is exactly equal to the:
- A. Sample mean
 - B. **Population mean**
 - C. Weighted mean
 - D. Combined mean
25. When a random sample is drawn from each stratum, it is known as:
- A. Simple random sampling
 - B. **Stratified random sampling**
 - C. Probability sampling
 - D. Purposive sampling

26. Suppose a finite population has 6 items and 2 items are selected at random without replacement, then all possible samples will be:

- A. 6
- B. 12
- C. **15**
- D. 36

27. In random sampling, the probability of selecting an item from the population is:

- A. Unknown
- B. **Known**
- C. Un-decided
- D. One