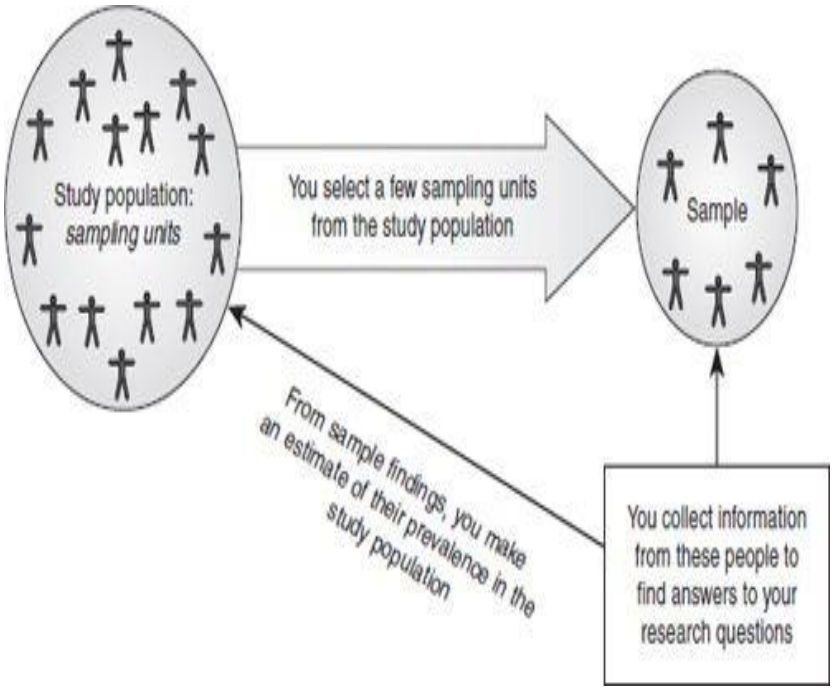
4.1 Sampling, therefore, is the process of selecting a few (a sample) from a bigger group (the sampling population) to become the basis for estimating or predicting the prevalence of an unknown piece of information, situation or outcome regarding the bigger group. A sample is a subgroup of the population you are interested in.

Let us take a very simple example to explain the concept of sampling. Suppose you want to estimate the average age of the students in your class. There are two ways of doing this. The first method is to contact all students in the class, find out their ages, add them up and then divide this by the number of students (the procedure for calculating an average). The second method is to select a few students from the class, ask them their ages, add them up and then divide by the number of students you have asked. From this you can make an estimate of the average age of the class.



Sampling terminology

Let us, again, consider the examples used above where our main aims are to find out the average age of the class, the average income of the families living in the city and the likely election outcome for a particular state or country. Let us assume that you adopt the sampling method – that is, you select a few students, families or electorates to achieve these aims. In this process there are a number of aspects:

The class, families living in the city or electorates from which you select your sample are called the population or study population, and are usually denoted by the letter N.

The small group of students, families or electors from whom you collect the required information to estimate the average age of the class, average income or the election outcome is called the sample.

The number of students, families or electors from whom you obtain the required information is called the sample size and is usually denoted by the letter n.

The way you select students, families or electors is called the sampling design or sampling strategy.

Each student, family or elector that becomes the basis for selecting your sample is called the sampling unit or sampling element.

A list identifying each student, family or elector in the study population is called the sampling frame. If all elements in a sampling population cannot be individually identified, you cannot have a sampling frame for that study population.

Your findings based on the information obtained from your respondents (sample) are called sample statistics. Your sample statistics become the basis of estimating the prevalence of the above characteristics in the study population.

What is a sample?

A sample is a finite part of a statistical population whose properties are studied to gain information about the whole (Webster, 1985). When dealing with people, it can be defined as a set of respondents(people) selected from a larger population for the purpose of a survey.

A population is a group of individuals persons, objects, or items from which samples are taken for measurement for example a population of presidents or professors, books or students.

What is sampling?

Sampling is the act, process, or technique of selecting a suitable sample, or a representative part of a population for the purpose of determining parameters or characteristics of the whole population. Sampling is concerned with the selection of a subset of individuals from within a population to estimate characteristics of the whole population.

What is the purpose of sampling?

To draw conclusions about populations from samples, we must use inferential statistics which enables us to determine a population`s characteristics by directly observing only a portion (or sample) of the population. We obtain a sample rather than a complete enumeration (a census) of the population for many reasons. Obviously, it is cheaper to observe a part rather than the whole, but we should prepare ourselves to cope with the dangers of using samples. In this tutorial, we will investigate various kinds of sampling procedures. Some are better than others but all may yield samples that are inaccurate and unreliable. We will learn how to minimize these dangers. Ok

Researchers rarely survey the entire population because the cost of a census is too high. The three main advantages of sampling are that the cost is lower, data collection is faster, and since the data set is smaller it is possible to ensure homogeneity and to improve the accuracy and quality of the data.

In business and medical research, sampling is widely used for gathering information about a population.

Why sampling?

Sampling is done because you usually cannot gather data from the entire population. Even in relatively small populations, the data may be needed urgently, and including everyone in the population in your data collection may take too long.

- **The time factor.**

A sample may provide you with needed information quickly. For example, covid -19 has broken out in our country; the disease is killing within hours nobody knows what it is. So medical field are required to conduct quick tests to help people and save the situation. If you try a census of those affected, they may be long dead when you arrive with your results. In such a case just a few of those already infected could be used to provide the required information.

- **The very large populations**

Many populations about which inferences must be made are quite large. For example, Consider the population of high school seniors in United States of America, a group numbering 4,000,000. The responsible agency in the government has to plan for how they will be absorbed into the different departments and even the private sector. The employers would like to have specific knowledge about the student`s plans in order to make compatible plans to absorb them during the coming year. But the big size of the population makes it physically impossible to conduct a census. In such a case, selecting a representative sample may be the only way to get the information required from high school seniors.

- **The partly accessible populations**

There are some populations that are so difficult to get access to that only a sample can be used. Like people in prison, like crashed aeroplanes in the deep seas. The inaccessibility may be economic or time related. Like a particular study population may be so costly to reach like the population of planets that only a sample can be

used. In other cases, a population of some events may be taking too long to occur that only sample information can be relied on.

Sample Design

It is a definite plan for obtaining a sample from a given population. It refers to the procedure, adopted by a researcher for selecting items for a sample.

Characteristics of a good sample design

(a) Sample design must result in a truly representative sample.

(b) Sample design must be such which results in a small sampling error.

(c) Sample design must be viable in the context of funds available for the research study.

(d) Sample design must be such so that systematic bias can be controlled in a better way.

(e) Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

4.2 Steps in sampling design

1. Define the population.
2. Determining the sampling frame
3. Define the sample unit.

4. Selecting the sampling technique.

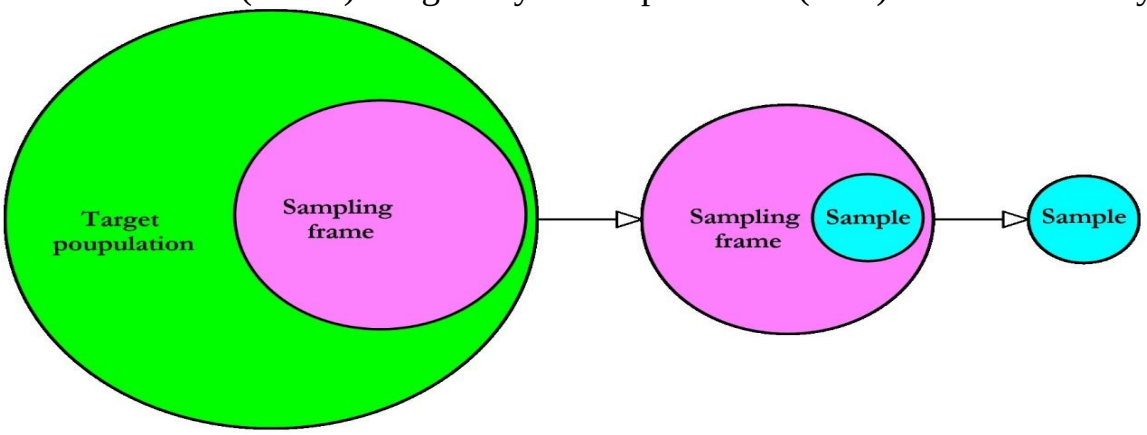
5. Determining the sample size.

6. Execution of sampling process.

7. Selection of the sample.

Step 1: Define the Population

A population must be defined in terms of elements, sampling units, extent and time. In relation to these constituent parts, Eg; the population of purchasing agent is. (element) purchasing agents in (sampling unit) companies and governmental agencies that have (extent) bought any of our products (time) in the last three years



Target population - Population of interest in which the researcher wants to generalize the results of study

Sampling frame - Part of accessible target popultion for study (the sample can be taken from accessible target popultion and not from the entire target population unless researcher has access to entire target population)

Sample - Actual units selected for the study

Step 2: Specify the Sampling Frame

If a probability sample is to be taken, a sampling frame is required. A sampling frame is a means of representing the elements of the population. A sampling frame maybe a telephone book, a city directory, an employee roster, a listing of all students attending a university, or a list of all possible phone numbers.

Maps also serve frequently as sampling frames. A sample of areas within a city may be taken and another sample of households may then be taken within each area. City blocks are sometimes sampled and all households on each sample block given instructions as to how to take "random walks" from the intersection and select the households to be interviewed.

A perfect sampling frame is one in which every element of the population is represented once but only once.

Step 3: Specify Sampling Unit

The sampling unit is the basic unit containing the elements of the population to be sampled. It may be the element itself or a unit in which the element is contained. For example, if one wanted a sample of males over 13 years of age, it might be possible to sample them directly. In this case, the sampling unit would be identical with the element. However, it might be easier to select households as the sampling unit and

interview all males over 13 years of age in each household. Here the sampling unit and the population element are not the same.

The sampling unit selected is often dependent upon the sampling frame. If a relatively complete and accurate listing of elements is available – register of purchasing agents, for example – one may well want to sample them directly. If no such register is available, one may need to sample companies as the basic sampling unit.

Step 4: Selection of Sampling Method

The sampling method is the way the sample units are to be selected. Five basic choices must be made in deciding on a sampling method:

Probability versus non probability,

Single unit versus cluster of units,

Unstratified versus stratified,

Equal unit probability versus unequal unit probability, and

Single stage versus multistage.

Step 5: Determination of the Sample Size

The determination of the proper sample size has traditionally been taught by one method in statistics classes and often practiced by an entirely different approach in the field. The reason for this is that traditional sampling theory generally ignores the

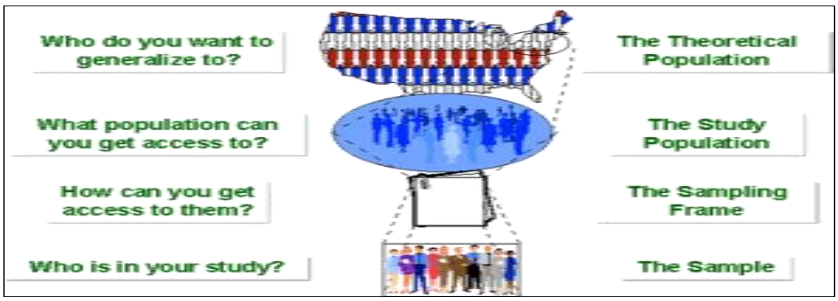
concept of the cost versus the value of the information to be provided by various sized samples. Practitioners have been forced to deal with the realities of sampling economics regardless of whether theory recognizes them.

Step 6: Specify the Sampling Plan

The sampling plan involves the specification of how each of the decisions made thus far is to be implemented. It may have been decided that the household will be the element and the block the sampling unit. How is a household defined operationally? How is the interviewer to be instructed to distinguish between families and households in instances where two families and some distant relatives of one of them are sharing the same apartment? How is the interviewer to be instructed to take a systematic sample of households on the block? What should the interviewer do when a housing unit selected is vacant? What is the callback procedure for households at which no one is at home? What age respondent speaking for the household is acceptable?

Step 7: Select the Sample

The final step in the sampling process is the actual selection of the sample elements. This requires a substantial amount of office and fieldwork, particularly if personal interviews are involved.



4.3 Types of sampling techniques

Representation Basis / Element Selection Technique	**Probability**	**Non- probability**
Unrestricted	Simple random sampling	Convenience sampling

Restricted	(Complex probability sampling) ❖ Stratified sampling ❖ Systematic ❖ Cluster	Purposive sampling ❖ Judgement ❖ Quota ❖ Snow ball

4.4 Probability sampling scheme is one in which every unit in the population has a chance (greater than zero) of being selected in the sample, and this probability can be accurately determined. The combination of these traits makes it possible to produce unbiased estimates of population totals, by weighting sampled units according to their probability of selection.

Probability sampling includes: Simple Random Sampling, Systematic Sampling, Stratified Sampling, Probability Proportional to Size Sampling, and Cluster or Multistage Sampling. These various ways of probability sampling have two things in common:

1. Every element has a known nonzero probability of being sampled and
2. involves random selection at some point.

Non probability sampling is any sampling method where some elements of the population have no chance of selection (these are sometimes referred to as 'out of

coverage'/'under covered'), or where the probability of selection can't be accurately determined. It involves the selection of elements based on assumptions regarding the population of interest, which forms the criteria for selection. Hence, because the selection of elements is nonrandom, non probability sampling does not allow the estimation of sampling errors. These conditions give rise to exclusion bias, placing limits on how much information a sample can provide about the population. Information about the relationship between sample and population is limited, making it difficult to extrapolate from the sample to the population.

Non probability sampling methods include accidental sampling, quota sampling and purposive sampling. In addition, non response effects may turn any probability design into a non probability design if the characteristics of non response are not well understood, since non response effectively modifies each element's probability of being sampled.

Probability sampling techniques

Simple random sampling

It is defined as one in which each element of the population has an equal and independent chance of being selected. Simple random sampling is a probability sampling procedure that gives every element in the target population, and each possible sample of a given size, an equal chance of being selected. As such, it is an equal probability selection method (EPSEM).

Three techniques are typically used in carrying out

The lottery method, **a table of random numbers**, **and randomly generated numbers using a computer program** (i.e., random number generator). In using the lottery method (also referred to as the "blind draw method" and the "hat model"), the numbers representing each element in the target population are placed on chips (i.e., cards, paper, or some other objects). The chips are then placed in a container and thoroughly mixed. Next, blindly select chips from the container until the desired sample size has been obtained.

Disadvantages of this method of selecting the sample are that it is time-consuming, and is limited to small populations.

A table of random numbers may also be used. The numbers in a table of random numbers are not arranged in any particular pattern. They may be read in any manner, i.e., horizontally, vertically, diagonally, forward, or backward. In using a table of

random numbers, the researcher should blindly select a starting point and then systematically proceed down (or up) the columns of numbers in the table. The number of digits that are used should correspond to the total size of the target population. Every element whose assigned number matches a number the researcher comes across is selected for the sample. Numbers the researcher comes across that do not match the numbers assigned the elements in the target population are ignored. As in using the lottery method, using a table of random numbers is a **tedious, time-consuming process, and is not recommended for large populations.**

Instead, statistical software should be used for large populations. Most statistical software and spreadsheet software have routines for generating random numbers. Elements of the populations

whose assigned numbers match the numbers generated by the software are included in the sample. One may select a number from a table of random numbers for use as the starting number for the process.

Procedures -

1. Lottery method
2. Random number tables
   - A table of random digits arranged in rows and columns.
   - After assigning an identification number to each member of the population, numbers in the random numbers table are used to select those who will be in the sample.

	1	2	3	4	5	6	7	8	9	10
1	49486	93775	88744	80091	92732	38532	41506	54131	44804	43637
2	94860	36746	04571	13150	65383	44616	97170	25057	02212	41930
3	10169	95685	47585	53247	60900	20097	97962	04267	29283	07550
4	12018	45351	15671	23026	55344	54654	73717	97666	00730	89083
5	45611	71585	61487	87434	07498	60596	36255	82880	84381	30433
6	89137	30984	18842	69619	53872	95200	76474	67528	14870	59628
7	94541	12057	30771	19598	96069	10399	50649	41909	09994	75322
8	89920	28843	87599	30181	26839	02162	56676	39342	95045	60146
9	32472	32796	15255	39636	90819	54150	24064	50514	15194	41450
10	63958	47944	82888	66709	66525	67616	75709	56879	29649	07325

Stratified sampling

Stratified random sampling is also referred to as proportional random sampling.

In stratified random sampling, the subjects are initially grouped into different classifications such as gender, level of education, or socioeconomic status. It's important to note that these classifications should not have any overlapping subjects.

From here, researchers randomly select the final list of subjects from the different defined categories to ensure a well rounded sample.

This method of probability sampling is best used when the goal of the research is to study a particular subgroup within a greater population. It also results in more precise statistical outcomes than simple random sampling.

Stratified random sampling creates layers within a sample that are extremely accurate when it comes to representing the layers with the population, but it too can be time consuming and tedious while creating larger samples.

It is used in situation where the population can be easily divided into groups or strata which are distinctly different from each other, but the elements within a group are homogeneous with respect to some characteristics.

Students of a college can be divided into strata on the basis of gender, courses offered, age etc Gender based -male and female. Splitting subjects into mutually exclusive groups and then using simple random sampling to choose members from groups.

Members of these groups should be distinct so that every member of all groups get equal opportunity to be selected using simple probability. This sampling method is also called “random quota sampling.”

Two types-

1. Proportionate Stratified sampling
2. Disproportionate Stratified sampling

Proportionate Stratified sampling

The number of units selected from each stratum is proportionate to the share of stratum in the population.

In a college there are total 2500 students out of which 1500 students are enrolled in under graduate courses and 1000 are enrolled in post graduate courses. If a sample of 100 is to be chosen using proportionate stratified sampling. Find out how many students under graduate courses and in post graduate courses? 60 from under graduate courses and 40 from post graduate courses.

Example consider income distribution of households from a heterogeneous population consisting of rural(70%)/ urban(30%) and male (60%) and female (40%)respondents , how to select 1000 households?

Urban		Rural	
Male	Female	Male	Female

.60*.30=.18	.30*.40=.12	.70*.60=.42	.70*.40=.28
.18*1000=180	.12*1000=120	.42*1000=420	.28*1000=280

180+120+420+280=1000

Representation of the subgroups can be proportionate or disproportionate. For example, if you wanted to sample 100 farmers from a population of farmers in which 90% are male and 10% are female, a *proportionate* stratified sample would select 90 males and 10 females. But you may want to know more about the women farmers than is possible in a sample of only ten subjects. So you can select a *disproportionate* stratified sample

Disproportionate Stratified sampling

(variances among strata affect sample size for each stratum)

To compare the differences among strata then it becomes necessary to draw equal units from all strata irrespective of their share in population. It is used if ----

1) some strata are too small

2) some strata are more important than others

3) some strata are more diversified than others

Example consider income distribution of households from a heterogeneous population consisting of rural(70%)/ urban(30%) and male (60%) and female (40%)respondents ,how to select 1000 households with an estimated variance of 25,9,16 and 4

Stratum	**Stratum population proportion**	**Variance**	**Std. deviation**		**Sample size**
Urban male	**.18**	**25**	**5**	**.18*5=.90**	**.9*1000/3.2=281**
Urban female	**.12**	**16**	**4**	**.12*4=.48**	**.48*1000/3.2=150**
Rural male	**.42**	**9**	**3**	**.42*3=1.26**	**1.26*1000/3.2=394**

Rural female	**.28**	**4**	**2**	**.28*2=.56**	**.56*1000/3.2=175**
Total				**3.20**	**1000**

Cluster sampling

It is used when there is no satisfactory sampling frame available therefore a list has to be constructed. Groups of elements are heterogeneous in nature within the group**.** It is a way to select participants randomly that are spread out geographically. For example, if you wanted to choose 100 participants from the entire population of the U.S., it is likely impossible to get a complete list of everyone. Instead, the researcher randomly selects areas (i.e., cities or counties) and randomly selects from within those boundaries.

Cluster sampling usually analyzes a particular population in which the sample consists of more than a few elements, for example, city, family, university, etc. Researchers then select the clusters by dividing the population into various smaller sections.

Difference between Cluster sampling and Stratified sampling

Stratified sampling has intra group homogeneity and inter group heterogeneity.

Eg, a college comprising number of students from different departments.

Cluster sampling has intra group heterogeneity.

Cluster Sampling – Area Method

- **Drawing the area sample:**
    - Divide the geo area into sectors (subareas) and give them names/numbers, determine how many sectors are to be sampled (typically a judgment call), randomly select these subareas. Do either a census or a systematic draw within each area.
    - To determine the total geo area estimate add the counts in the subareas together and multiply this number by the ratio of the total number of subareas divided by number of subareas.

Sometimes it is more cost-effective to select respondents in groups ('clusters'). Sampling is often clustered by geography, or by time periods. (Nearly all samples are in some sense 'clustered' in time - although this is rarely taken into account in the analysis.) For instance, if surveying households within a city, we might choose to select 100 city blocks and then interview every household within the selected blocks.

Clustering can reduce travel and administrative costs. In the example above, an interviewer can make a single trip to visit several households in one block, rather than having to drive to a different block for each household.

It also means that one does not need a sampling frame listing all elements in the target population. Instead, clusters can be chosen from a cluster-level frame, with an element-level frame created only for the selected clusters. In the example above, the

sample only requires a block-level city map for initial selections, and then a household-level map of the 100 selected blocks, rather than a household-level map of the whole city.

Cluster sampling generally increases the variability of sample estimates above that of simple random sampling, depending on how the clusters differ between themselves, as compared with the within-cluster variation. For this reason, cluster sampling requires a larger sample than SRS to achieve the same level of accuracy - but cost savings from clustering might still make this a cheaper option.

Cluster sampling is commonly implemented as multistage sampling. This is a complex form of cluster sampling in which two or more levels of units are embedded one in the other. The first stage consists of constructing the clusters that will be used to sample from. In the second stage, a sample of primary units is randomly selected from each cluster (rather than using all units contained in all selected clusters). In following stages, in each of those selected clusters, additional samples of units are selected, and so on. All ultimate units (individuals, for instance) selected at the last step of this procedure are then surveyed. This technique, thus, is essentially the process of taking random subsamples of preceding random samples.

Multistage sampling

Multi-stage sampling involves a combination of two or more of the probability sampling methods outlined above.With more advanced research, using just one form

of probability sampling does not ensure the randomization necessary to ensure confidence in results.

By combining various probability sampling techniques at various stages of research initiatives, researchers are able to maintain confidence that they are mitigating biases as much as possible.

It can substantially reduce sampling costs, where the complete population list would need to be constructed (before other sampling methods could be applied). By eliminating the work involved in describing clusters that are not selected, multistage sampling can reduce the large costs associated with traditional cluster sampling.

Systematic sampling

- It is used in those cases where complete list of the population from which sample is to be drawn is available. It involves the ordering of the universe. The ordering may be alphabetical, numerical and geographical etc.
- The items in the population are included in the intervals of magnitude ie sampling interval- 'n'
- Sampling interval (SI) = population list size (N) divided by a pre-determined sample size (n)

How to draw:

1) calculate SI,

2) select a number between 1 and SI randomly,

3) go to this number as the starting point and the item on the list here is the first in the sample,

4) add SI to the position number of this item and the new position will be the second sampled item,

5) continue this process until desired sample size is reached.

Example There are 1000 items in the population and we want to draw a sample of 200 items.

- What is the sampling interval?
- 1000/200= 5
- It means the items selected will be a gap of 5.
- If first selected is 3rd item then what will be the next? 3rd , 8th etc.

Merit	**Demerit**
Less time Suitable when the formal list of population unit is available	Chance for bias Lack complete randomness, only first unit is selected at random.

When to use probability sampling?

Use probability sampling in these instances:

1. When you want to reduce the sampling bias: This sampling method is used when the bias has to be minimum. The selection of the sample largely determines the quality of the research's inference. How researchers select their sample largely

determines the quality of a researcher's findings. Probability sampling leads to higher quality findings because it provides an unbiased representation of the population.

2. When the population is usually diverse: Researchers use this method extensively as it helps them create samples that fully represent the population. Say we want to find out how many people prefer medical tourism over getting treated in their own country. This sampling method will help pick samples from various socio-economic strata, background, etc. to represent the broader population.

3. To create an accurate sample: Probability sampling help researchers create accurate samples of their population. Researchers use proven statistical methods to draw a precise sample size to obtained well-defined data.

4.5 Non-probability sampling designs

What is non-probability sampling?

Definition: Non-probability sampling is defined as a sampling technique in which the researcher selects samples based on the subjective judgment of the researcher rather than random selection. It is a less stringent method. This sampling method depends heavily on the expertise of the researchers. It is carried out by observation, and researchers use it widely for qualitative research.

Non-probability sampling is a sampling method in which not all members of the population have an equal chance of participating in the study, unlike probability

sampling. Each member of the population has a known chance of being selected. Non-probability sampling is most useful for exploratory studies like a pilot survey (deploying a survey to a smaller sample compared to pre-determined sample size). Researchers use this method in studies where it is impossible to draw random probability sampling due to time or cost considerations. Non-probability sampling designs do not follow the theory of probability in the choice of elements from the sampling population. Non-probability sampling designs are used when the number of elements in a population is either unknown or cannot be individually identified. In such situations the selection of elements is dependent upon other considerations. There are five commonly used non-random designs, each based on a different consideration, which are commonly used in both qualitative and quantitative research. These are:

1. **Quota sampling**
2. **Convenience Sampling**
3. **Accidental sampling;**
4. **Judgemental sampling or purposive sampling;**
5. **Expert sampling;**
6. **Snowball (Referral) sampling.**

What differentiates these designs being treated as quantitative or qualitative is the predetermined sample size. In quantitative research you use these designs to select a predetermined number of cases (sample size), whereas in qualitative research you do not decide the number of respondents in advance but continue to select additional cases till you reach the data saturation point. In addition, in qualitative

research, you will predominantly use judgemental and accidental sampling strategies to select your respondents. Expert sampling is very similar to judgemental sampling except that in expert sampling the sampling population comprises experts in the field of enquiry. You can also use quota and snowball sampling in qualitative research but without having a predetermined number of cases in mind (sample size).

- **Quota sampling:** Quota sampling is quite frequently used in marketing research. It involves the fixation of certain quotas, which are to be fulfilled by the interviewers. Suppose, 2,00,000 students are appearing for a competitive examination. We need to select 1% of them based on quota sampling. The classification of quota may be as follows:

Example: Classification of Samples

Category	Quota
General merit	1,000
Sport	600
NRI	100
SC/ST	300
Total	2,000

Quota sampling involves the following steps:

The population is divided into segments on the basis of certain characteristics. Here, the segments are termed as cells.

2. A quota of unit is selected from each cell.

Samples that set a specific number of certain types of individuals to be interviewed. This type of sampling is most appropriate where one or more of the subgroups is very small in comparison to other groups, or where the target of the study is a specific and oversampling of a group may provide more accurate results

- Often used to ensure that convenience samples will have desired proportion of different respondent classes. The population is divided into cells on the basis of relevant control characteristics.
- A quota of sample units is established for each cell.
- 50 women, 50 men

In quota sampling, the population is first segmented into mutually exclusive sub-groups, just as in stratified sampling. Then judgement is used to select the subjects or units from each segment based on a specified proportion. For example, an interviewer may be told to sample 200 females and 300 males between the age of 45 and 60.

It is this second step which makes the technique one of non-probability sampling. In quota sampling the selection of the sample is non-random. For example interviewers might be tempted to interview those who look most helpful. The problem is that these samples may be biased because not everyone gets a chance of selection. This random element is its greatest weakness and quota versus probability has been a matter of controversy for many years.

Merit

Less cost and time to collect data.

Demerit

There is more chance for bias in sample selection.

Convenience Sampling Method

Convenience samples*:* samples drawn at the convenience of the interviewer. People tend to make the selection at familiar locations and to choose respondents who are like themselves.

Error occurs

1) in the form of members of the population who are infrequent or nonusers of that location.

2) who are not typical in the population

Accidental sampling

Accidental sampling is also based upon convenience in accessing the sampling population. Whereas quota sampling attempts to include people possessing an obvious/visible characteristic, accidental sampling makes no such attempt. You stop collecting data when you reach the required number of respondents you decided to have in your sample.

This method of sampling is common among market research and newspaper reporters. It has more or less the same advantages and disadvantages as quota sampling but, in addition, as you are not guided by any obvious characteristics, some people contacted may not have the required information.

Accidental sampling (sometimes known as grab, convenience or opportunity sampling) is a type of non probability sampling which involves the sample being drawn from that part of the population which is close to hand. That is, a population is

selected because it is readily available and convenient. It may be through meeting the person or including a person in the sample when one meets them or chosen by finding them through technological means such as the internet or through phone. The researcher using such a sample cannot scientifically make generalizations about the total population from this sample because it would not be representative enough. For example, if the interviewer were to conduct such a survey at a shopping center early in the morning on a given day, the people that he/she could interview would be limited to those given there at that given time, which would not represent the views of other members of society in such an area, if the survey were to be conducted at different times of day and several times per week. This type of sampling is most useful for pilot testing. Several important considerations for researchers using convenience samples include:

1. Are there controls within the research design or experiment which can serve to lessen the impact of a non-random convenience sample, thereby ensuring the results will be more representative of the population?

2. Is there good reason to believe that a particular convenience sample would or should respond or behave differently than a random sample from the same population?

3. Is the question being asked by the research one that can adequately be answered using a convenience sample?

In social science research, snowball sampling is a similar technique, where existing study subjects are used to recruit more subjects into the sample. Some variants of snowball sampling, such as respondent driven sampling, allow calculation of selection probabilities and are probability sampling methods under certain conditions.

Judgemental or purposive sampling

The primary consideration in purposive sampling is your judgement as to who can provide the best information to achieve the objectives of your study. You as a researcher only go to those people who in your opinion are likely to have the required information and be willing to share it with you.

This type of sampling is extremely useful when you want to construct a historical reality, describe a phenomenon or develop something about which only a little is known. This sampling strategy is more common in qualitative research, but when you use it in quantitative research you select a predetermined number of people who, in your judgement, are best positioned to provide you the needed information for your study.

Judgment samples*:* samples that require a judgment or an "educated guess" on the part of the interviewer as to who should represent the population. Also, "judges" (informed individuals) may be asked to suggest who should be in the sample.

- Subjectivity enters in here, and certain members of the population will have a smaller or no chance of selection compared to others .

Purposive sampling: where the researcher chooses a sample based on their knowledge about the population and the study itself. The study participants are chosen based on the study's purpose. There are several types of purposive sampling. For a full list, advantages and disadvantages of the method, see the article: Purposive Sampling.

- **Expert sampling**

The only difference between judgemental sampling and expert sampling is that in the case of the former it is entirely your judgement as to the ability of the respondents to contribute to the study. But in the case of expert sampling, your respondents must be known experts in the field of interest to you. This is again used in both types of research but more so in qualitative research studies. When you use it in qualitative research, the number of people you talk to is dependent upon the data saturation point whereas in quantitative research you decide on the number of experts to be contacted without considering the saturation point.

You first identify persons with demonstrated or known expertise in an area of interest to you, seek their consent for participation, and then collect the information either individually or collectively in the form of a group.

- **Referral samples (snowball samples):** samples which require respondents to provide the names of additional respondents
    - Members of the population who are rare ,less known, disliked, or whose opinions conflict with the respondent have a low probability of being selected.

- *Example a researcher wants to study about homeless people in a locality. There is no accessible sampling frame for the population from which sample is to be taken. The best way is to go to the area and identify one or two. Through them the researcher can establish contact with other homeless people in the vicinity.*

- Snowball sampling is the process of selecting a sample using networks. To start with, a few individuals in a group or organisation are selected and the required information is collected from them. They are then asked to identify other people in the group or organisation, and the people selected by them become a part of the sample. Information is collected from them, and then these people are asked to identify other members of the group and, in turn, those identified become the basis of further data collection (Figure 12.7). This process is continued until the required number or a saturation point has been reached, in terms of the information being sought.

- This sampling technique is useful if you know little about the group or organization you wish to study, as you need only to make contact with a few individuals, who can then direct you to the other members of the group. This method of selecting a sample is useful for studying communication patterns, decision making or diffusion of knowledge within a group. There are disadvantages to this technique, however. The choice of the entire sample rests upon the choice of individuals at the

first stage. If they belong to a particular faction or have strong biases, the study may be biased. Also, it is difficult to use this technique when the sample becomes fairly large.

Stratified Sampling	**Quota Sampling**
Stratified sampling and quota sampling are similar in that:	
Population is divided into categories; elements are then selected from each category.	Population is divided into categories; elements are then selected from each category.
Purpose is to select a representative sample and/or facilitate subgroup analyses.	Purpose is to select a representative sample and/or facilitate subgroup analyses.
Stratified sampling and quota sampling are dissimilar in that:	
Elements within each category are selected using simple random sampling, and as a result:	Elements within each category are selected using availability sampling, and as a result:

A sampling frame is required.	A sampling frame is not required.
Random sampling error can be estimated.	Random sampling error cannot be estimated
Selection bias is minimized.	Selection bias is not minimized.
Purpose is to reduce sampling error	

Comparison of Stratified Sampling and Cluster Sampling	
The population is separated into strata, and then sampling is conducted within each stratum.	The population is separated into clusters, and then clusters are sampled.
Analysis of individual strata is permitted in addition to analysis of the total sample.	Analysis of individual categories (clusters) are permitted in addition to analysis of the total sample.
In order to minimize sampling error,	In order to minimize sampling error,

within group differences among strata should be minimized, and between/group differences among strata should be maximized.	within group differences should be consistent with those in the population, and between-group differences among the clusters should be minimized.
A sampling frame is needed for the entire target population.	In single-state cluster sampling, a sampling frame is needed only for the clusters. In two-stage and multistage cluster sampling, a sampling frame of individual elements is needed only for the elements in the clusters selected at the final stage.
Main purpose: increase precision and representation.	Main purpose: decrease costs and increase operational efficiency.
More precision compared to simple random sampling.	Lower precision compared to simple random sampling.
Common stratification variables: age, gender, income, race.	Common classification variables: geographical area, school, grade level.

Difference between Probability Sampling &Non-Probability Sampling

Probability Sampling	Non-Probability Sampling

➢ Every element in the target population or universe [sampling frame] has equal probability of being chosen in the sample for the survey being conducted. ➢ Scientific, operationally convenient and simple in theory. ➢ Parametric tests are mostly used ➢ Results may be generalized.	➢ Every element in the universe [sampling frame] does not have equal probability of being chosen in the sample. ➢ Operationally convenient and simple in theory. ➢ Non -Parametric tests are mostly used ➢ Results may not be generalized.

Summary

In quantitative research you are guided by your desire to select a random sample, whereas in qualitative research you are guided by your judgement as to who is likely to provide you with the 'best' information.

In quantitative studies you collect information from a predetermined number of people but, in qualitative research, you do not have a sample size in mind. Data collection based upon a predetermined sample size and the **saturation point** distinguishes their use in quantitative and qualitative research.

Guidelines for choosing type of probability sample design

- **Simple random sampling**. Consider choosing simple random sampling if one has access to a complete and accurate sampling frame of the target population that is complete and accurate but does not contain auxiliary information that may be used for stratification purposes.
- **Stratified sampling**. Consider choosing stratified sampling if:

• It is possible to divide a population into two or more homogeneous strata and construct a sampling frame for each stratum.

• One has access to a sampling frame of the target population that is complete and accurate and contains auxiliary information that may be used for stratification purposes.

• Some subgroups of the population are vastly different from other subgroups.

• It is very important to minimize sampling error.

• There is a concern about under representing smaller subgroups.

• The population is heterogeneous.

• There is a desire to use different selection methods for different strata.

• It is likely that answers to the research questions of a study are likely to be different for different subgroups.

• It is useful when each stratum needs to be reported separately.

• Comparative analysis of strata is desired.

- **Proportionate stratified sampling**. Consider choosing proportionate stratified sampling if sub groups of approximately the same size are to be investigated or compared.
- **Disproportionate stratified sampling**. Consider choosing disproportionate stratified sampling if:• Subgroups of vastly different sizes are to be investigated or compared.

• It is important to include a large number of elements from a small segment of the population.

• One is primarily interested in key similarities and differences among strata.

• Some observations are limited or hard to obtain.

• It is important to make statistically valid statements about subgroups.

• Subgroups of the population have different variances for the variables of interest.

• Costs of data collection are different across population subgroups.

- **Systematic sampling**. Consider choosing systematic sampling if:

• It is difficult to identify items using a simple random sampling method.

• It is important to use a probability sampling procedure that can be easily implemented.

• A sampling frame is not available or impractical to prepare, but a stream of representative elements of the population is available.

• The listing of the population is essentially random or can be randomized.

- **Cluster sampling**. Consider choosing cluster sampling if:

• It is important to minimize data collection costs and there are substantial fixed costs associated with each data collection location.

• A sampling frame of individual population elements is not available but a sampling frame of clusters of elements is available.

• Travel costs can be substantially reduced.

%%%%%%%%%%%%%%%%%%