ISBN: 978-81-971650-3-0

4. Sampling Design

Dr. Syed Mohammad Bakhteyar Fatmi

Associate Professor and Head,
Dept. of Psychology, Oriental College,
Patna (Bihar).

Abstract:

The methods for performing sampling are outlined in this publication. Additionally, in order to choose the best sample method for their research, researchers must be aware of the variations among the various types of sampling procedures and approaches. In this regard, the various sample methods and approaches are also presented in this study. An essential instrument for gathering data for legal study is the primary purpose of sampling design. The researcher must choose how to choose a sample for legal study before beginning any work. Sample design is another term for this sample selection process. A clear strategy for deciding what information is truly gathered so that the researcher can take a sample from a particular group is known as sample design. There are two main categories of sampling techniques: nonprobability, also known as non-random sampling, and probability, also known as random sampling. Each unit or component of the general population has an equal chance of being included in the sample when it is a probability sample. The nonprobability sampling method makes it impossible to ascertain if a population component might be represented in the sample. This paper will talk about. Sampling Scheme.

Keywords:

Sampling Design, Methods, Data Collecting, Research, Size, Population, Frame, Sampling Bias, Error, Non-Probability Error, Non-Response.

4.1 Introduction:

The methodical selection of components from a population of interest is known as sampling, and it allows a researcher to fairly generalize findings about the population by examining the sample. From a small number of people, like the nation's nuclear scientists, to a sizable population, like its kids who are in school, the population size varies greatly.

Since there are fewer scientists in the nation with nuclear science specializations, it is relatively easier for a researcher to identify the population for the study in the first scenario. In certain cases, researchers may gather data from the entire population, depending on time and finances. Data gathering from a group of elements taken from the population may be required instead due to operational, technical, and material constraints in research. Census data are those that are gathered from every segment of the population. Sample data are those that are gathered from a small number of chosen respondents. Here, the key question is how the researcher uses data from a sample to make population-level generalizations or interpretations.

In order to describe the characteristics of the population, sampling entails choosing some or all of the population's components. Therefore, sampling is the process of choosing certain components from a broad or small population.

Samples are the items chosen for methodical observation or data collecting using different techniques. The purpose of a sample is to provide information about the entire population by analyzing a limited portion of it. The sample size should not be the same as the population size. [1]

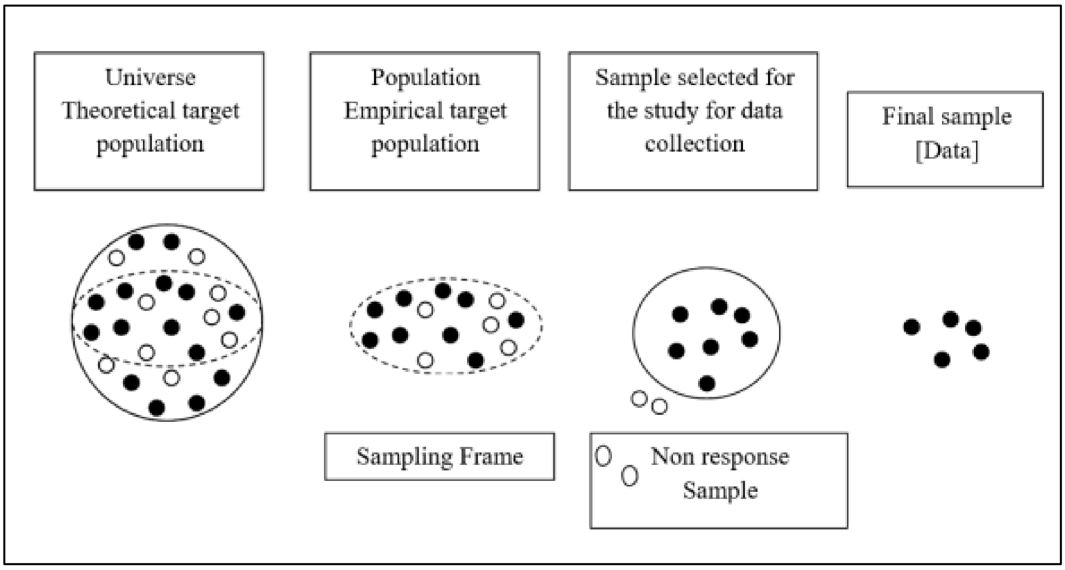

Figure 4.1: Illustration of population – Sample Design

It is typically not feasible for any researcher to examine every component from a very big population. Furthermore, gathering data from the whole population is not always necessary for sociological research.

It is actually thought to be pointless to gather statistics from the whole population. Sampling is done when it is neither practical or feasible to deal with every element in the population, or when gathering data from every element is not useful.

We will now talk about a different crucial and practical step in a quantitative research investigation. We refer to it as "sampling." Sampling is not exclusive to research. We also practice it, in a sense.

Every day, we employ rudimentary forms of sampling. For example, housewives test a few cooked rice grains to determine when it's ready to be served; it makes sense that it's not possible to test every grain in the cooking pot. The fact that we can generally draw some sort of general conclusion about the universe from a sample, or a small number of elements, is demonstrated by our daily experiences. Sampling is the process of choosing a subset of the world in order to make conclusions about the universe as a whole.

4.2 Basic Concepts and Terms:

Before reading about the sample procedure, it is necessary to comprehend the ideas and terminology that we will cover in the next part.

- **The Population:** The whole set of units from which a sample is to be drawn.
- **Sample:** The subset of the population chosen for study. It represents a portion of the populace. There are two possible approaches to selection: non-probability and probability.
- **Sampling Frame:** A list of every unit in the population that will be used to choose the sample
- **A Representative Sample:** Is one that fairly and microcosmically represents the population by reflecting the population.
- **Sampling Bias:** A distortion in the sample representatives that occurs when certain members of the population (or more specifically, the sampling frame) have little to no chance of being chosen for the sample.
- **Sampling Error:** The inaccuracy in study findings caused by the discrepancy between a sample and the population it was drawn from. This could happen even if a probability sample has been used.
- **Non-Probability Error:** Inaccuracy in study results resulting from variations between the population and sample, resulting from either non-response or deficiencies in the sampling frame, or from issues like poorly phrased interview questions, inadequate data processing, or bad interviewing.
- **Non-Response:** A non-sampling error source that is more likely to occur when people are being sampled. It happens when certain sample participants are unwilling to participate, are unreachable, or are unable to provide the necessary data for whatever reason (such as mental illness).
- **Census:** The count of the entire population. Therefore, data are considered census data if they are gathered with respect to all units in a population as opposed to a sample of those units. Generally speaking, "the census" refers to a national census, which is an exhaustive count of every person living in a nation state. In India, this type of enumeration takes place once every ten years. [2]

It is unlikely that the researcher should be able to gather data from every example in order to address the study questions.

Thus, choosing a sample is necessary. The population is the total set of instances that the researcher selects a sample from.

Because they lack the time and resources to analyze the entire population, researchers use sampling techniques to lower the number of cases.

The steps that are likely to be taken when doing sampling are shown in Figure 9.2.

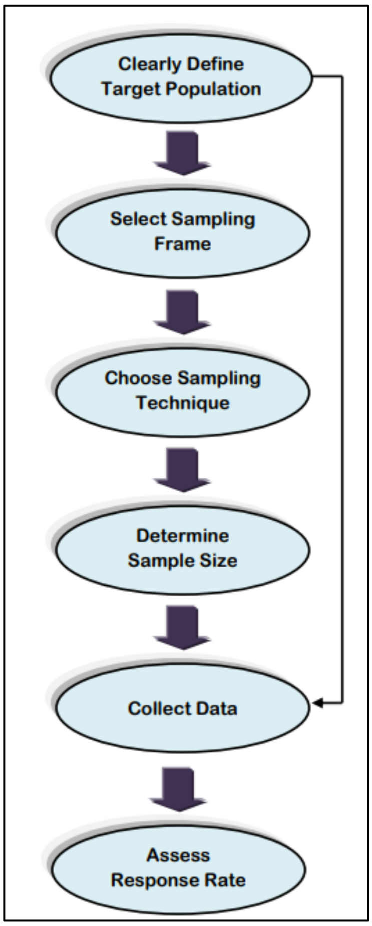

Figure 4.2: Sampling Process Steps

- **Stage 1:** Clearly Define the Aim and Target Clearly defining the target population is the first stage in the sampling procedure. Population is frequently associated with the total number of individuals residing in a given nation.
- **Stage 2:** Choose a Sampling Frame A list of the actual cases from which a sample will be taken is called a sampling frame. A representative sampling frame for the population is required.
- **Stage 3:** Select a Sampling Method It is important to understand what is meant by sampling and the reasons behind researchers' sample selection processes before delving into the different kinds of sampling techniques. Sampling is the process of selecting a subset from a given sampling frame or population as a whole. Drawing conclusions about a population or drawing broad conclusions in light of accepted theory are two uses for sampling. Essentially, the choice of sampling technique determines this. [3]

Sampling strategies can be broadly classified into two categories:

- Probability or random sampling
- Non- probability or non- random sampling

Selecting a broad sampling technique is necessary before deciding on a specific sort of sample technique.

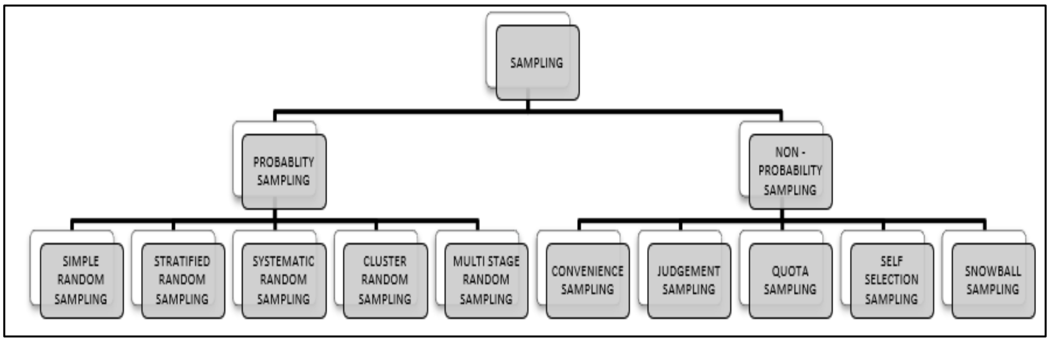

Figure 4.3: Shows The Various Types of Sampling Techniques.

- **Probability Sampling:**

By using probability sampling, you can make sure that each person in your sample has an equal chance of being chosen for your study. Probability sampling comes in four primary flavors: basic random, cluster, systematic, and stratified.

a. Simple Random Sampling:

Simple random sampling is exactly that—simple and random, as the name implies. This method gives you a completely random subset of your group, which you can choose using a random number generator or by drawing from a hat, for example. This enables you to use the information from the subset (sample) to make generalizations about the entire population.

b. Cluster Sampling:

Your population is split into subgroups for cluster sampling based on shared features with the entire population. For your sample, you randomly choose an entire subgroup rather than a few individuals.

Before you finally get at your study sample, the unit selection process may occasionally go through a number of stages. For this purpose, a state might be partitioned, for instance, into districts, districts into blocks, blocks into villages, and villages into distinct groups of individuals, with a random or quota sample being taken from each group. For instance, if 3 of a state's 15 districts were chosen at random, along with a total of 6 from each district, 10 towns from each block, and 20 households from each village, the total number of respondents would be 3,600. Large-scale surveys across vast regions are conducted using this design. One benefit is that just specific clusters need to have a comprehensive sample frame, as opposed to the full target area. Time and travel expenditure savings are also realized. If significant subgroups are overlooked, there is a chance that the target population will not be fully represented. [4]

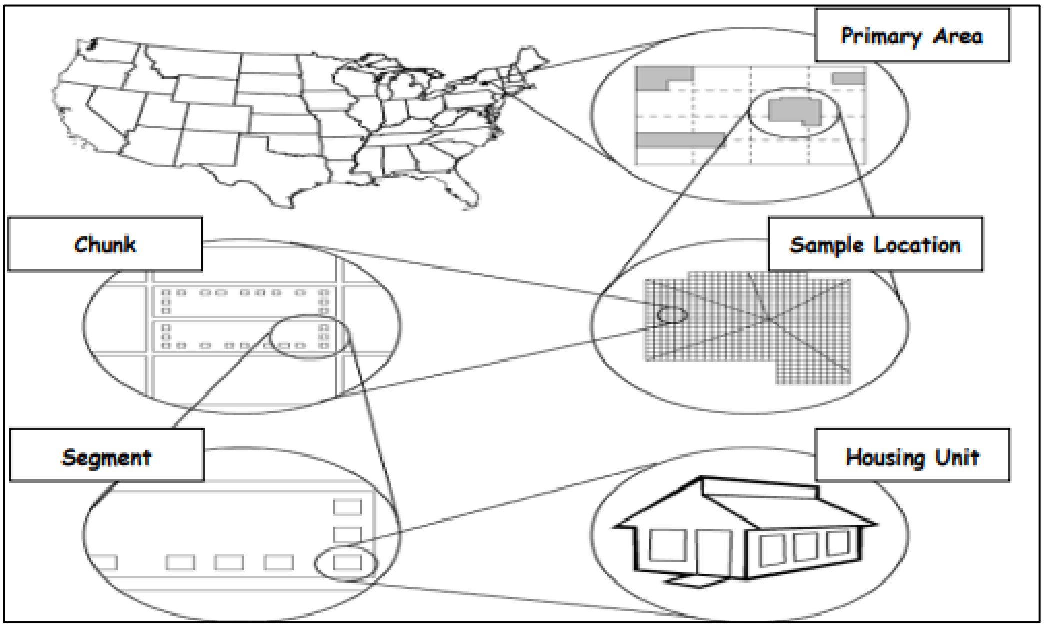

Figure 4.4: Cluster Sampling

c. Systematic Sampling:

Conducting systematic sampling is even simpler than basic random sampling. This method assigns a number to every member of the target population. Rather than drawing numbers at random, participants are selected on a regular basis. It's critical that the list has no hidden patterns that could distort the sample.

d. Stratified Sampling:

By using stratified random sampling, a population is divided into smaller groupings known as strata. The shared characteristics of the individuals, such as age, education level, or income, form the basis of the strata. When you think these parallels point to further parallels that will appeal to a larger segment of your population, you employ this approach.

e. Multi Stage Random Sampling:

This probability sampling method combines two or more of the previously listed sample approaches. There isn't a single kind of probability sampling that works well for the majority of intricate field or lab research projects. Most studies are carried out in phases, using a different random sample technique at each stage.

- **Non-Probability Sampling Methods:**

People are chosen using non-random criteria in a non-probability sample, so not every person has an equal chance of being included.

Though more prone to sampling bias, this kind of sample is more affordable and quicker to get. As a result, your ability to draw conclusions about the population may be more constrained than it would be with random samples. Making a non-probability sample as representative of the population as feasible is still important.

In exploratory and qualitative research, non-probability sampling methods are frequently employed. These studies focus on developing a preliminary understanding of a small or understudied population rather than testing a hypothesis about a large population. [5]

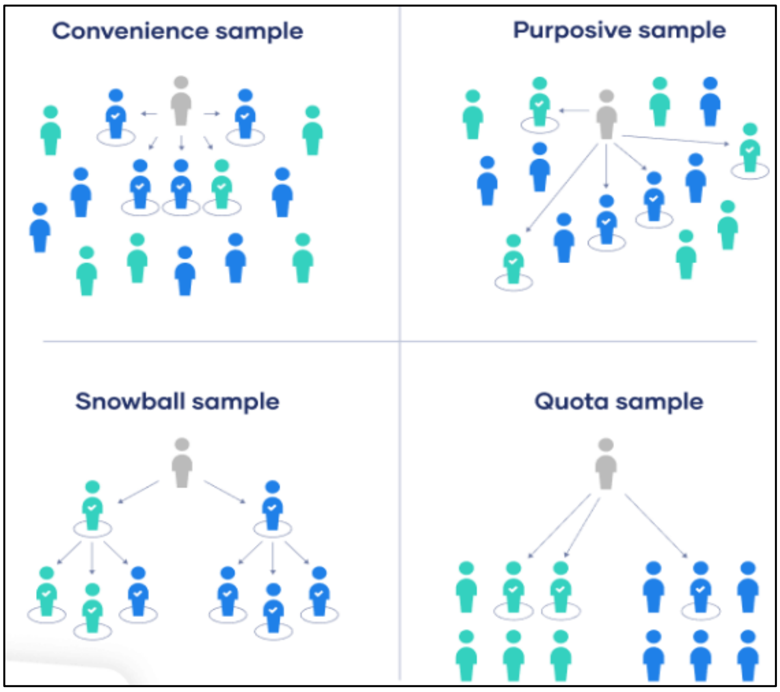

Figure 4.5: Non-Probability Sampling Methods

a. Convenience Sampling:

All that is included in a convenience sample are those who are most easily reached by the researcher.

Although this is a quick and low-cost method of collecting preliminary data, it cannot yield inferences that are applicable to the entire community because it is impossible to determine whether the sample is representative of it. Convenience samples are susceptible to biases in selection and sampling.

b. Voluntary Response Sampling:

A voluntary response sample is mostly focused on accessibility, much like a convenience sample. Rather than the researcher selecting and getting in touch with participants, participants volunteer themselves (e.g. by responding to a public online survey).

Because some people are naturally more likely than others to volunteer, voluntary answer samples are always biased in some way due to self-selection bias.

c. Purposive Sampling:

With this kind of sampling, sometimes referred to as judgement sampling, the researcher uses their knowledge to choose a sample that will be most helpful to their research goals.

It is frequently employed in qualitative research when the goal is to learn in-depth information on a particular phenomenon rather than draw general conclusions from statistics or when the population is small and focused. Strict inclusion criteria and justifications are necessary for a purposive sample to be effective.

Be cautious of observer bias influencing your arguments and always make sure to clearly state your inclusion and exclusion criteria.

d. Snowball Sampling:

Snowball sampling can be used to attract participants through other participants if the population is difficult to reach. The amount of people you can reach out to grows exponentially as you establish more connections. The drawback of this is that representativeness is also compromised because participant recruitment leaves you with no means of determining how representative your sample is. Bias in sampling may result from this.

e. Quota Sampling:

The non-random selection of a preset quantity or percentage of units is the foundation of quota sampling. We refer to this as a quota.

To recruit sample units, you first split the population into mutually exclusive subgroups, or strata, and keep doing so until your quota is reached. You chose the particular qualities that these units have in common before creating your strata. Controlling who or what goes into your sample is the goal of quota sampling.

- **Stage 4:** Establish the Sample Size A random sample must be sufficiently large to allow for generalization while preventing biases or errors in sampling. A number of factors determine what is sufficient, which frequently perplexes those doing surveys for the first time. This is due to the fact that the significance in this case lies not in the percentage of the research population that is sampled, but rather in the absolute size of the sample that is chosen in relation to the population's complexity, the researcher's objectives, and the types of statistical manipulation that will be applied during data analysis. The idea that the less likely results will be biased, the larger the sample, is true; nevertheless, when samples exceed a certain size, declining returns may occur, which must be weighed against the researcher's resources. To be blunt, sampling error is reduced by bigger sample sizes, but at a diminishing pace. The sample size can be found using a number of statistical formulas.

- **Stage 5:** Collect Information The following stage is data collection after the target population, sampling frame, sampling procedure, and sample size have been determined.
- **Stage 6:** Determine the Reaction Rate A study's response rate is the percentage of cases that consent to participate. These examples are drawn from the initial dataset. The majority of researchers never really get a 100% response rate. The respondent may have been located but the researchers are unable to get in touch with them, or they may have declined to respond, been deemed ineligible to respond, or been unable to respond. To sum up, each non-response has the potential to skew the final sample, hence response rate matters. In certain ways, sample bias can be lessened by carefully defining the sample, using the appropriate sampling technique, and producing a sizable sample. [6]

4.3 Sampling Error and Survey Bias:

There is usually some mistake in survey results. Sample and non-sampling errors are two categories into which total errors can be divided. Error in this context refers to both random and systematic biases.

a. Sampling Biases and Errors: The sample design causes sampling biases and errors. They consist of

- **Selection Bias:** When the findings are calculated using assumptions about selection probabilities rather than the actual probability.
- **Error in Random Sampling:** Variations in outcomes at random because sample members were chosen at random.

b. Non-Sampling Error: Non-sampling errors are additional errors that arise from issues with data processing, sample design, or collecting that may have an influence on the final survey estimates. Among them are

- **Over-coverage:** Data from non-population sources are included.
- **Under-coverage:** This is the result of some population members being underrepresented in the sample. Convenience samples frequently have an issue with under-coverage.
- **Measurement Error:** This happens when participants misinterpret a question or struggle to provide a response.
- **Processing Error:** Data coding errors.
- **Non-response:** Not getting all the data needed from every person chosen.

After sampling, the precise procedure used in sampling should be reviewed instead of the planned procedure in order to examine any potential impacts that divergences may have on analysis that comes after. One specific issue is the lack of reaction.

Unit non-response, which refers to not completing any portion of the survey, and item non-response, which is submission or participation in the survey but not completing one or more survey components or questions, are the two main categories of non-response. Many of the people who are listed as being in the sample for a survey may not want to participate, may not have the time to participate (opportunity cost), or the survey administrators may not have been able to get in touch with them. In this instance, there's a chance that variations between responders and non-respondents could skew estimations of population parameters.

Improving survey design, providing incentives, and carrying out follow-up studies that repeatedly try to get in touch with the unresponsive and describe their similarities and differences from the rest of the frame are common ways to solve this. When demographic standards are available, the impacts can also be reduced by imputing data based on responses to other questions or by weighing the data. [7]

Measurement Error-Related Bias: Bias can also result from subpar measurement procedures. The setting in which the survey is done, the questions posed, and the respondent's condition are all factors in the measurement process in survey research. The bias resulting from issues with the measurement procedure is referred to as response bias.

4.4 Characteristics of a Good Sample Design:

The characteristics of a good sample as follows;

- Sample design must result in a truly representative sample,
- Sample design must be such which results in a small sampling error,
- Sampling design must be viable in the context of funds available for the research study,
- Sample design must be such that systematic bias can be controlled in a better way, and
- 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.

4.4.1 Sample Size:

Determining the ideal sample size is a challenging question. Several limitations can be used to determine the sample size. The appropriate sample size is explained in this section of the dissertation.

The different aspects that affect sample size, such as financial resources and time available for data collection, statistical studies employed in the research, etc., may also be clear to the readers.

The process of deciding how many replicates or observations to include in a statistical sample is known as *sample size determination*. Any empirical study where drawing conclusions about the population from a sample is the aim must consider the sample size. In actuality, the cost of gathering data and the requirement for adequate statistical power are taken into consideration when determining the sample size utilized in a study. Multiple sample sizes may be used in complex research. For instance, in a stratified survey, various sample sizes would be used for each strata. Since all population statistics are gathered during a census, the sample size is equal to the size of the population. There could be varying sample sizes for each group in an experimental design where a study is split up into several treatment groups. Comprehending the differences between sampling and non-sampling error is crucial, as they can both lead to overall survey error.

In any sample survey, the crucial question of "How many participants should be chosen for a survey" should be addressed. The investigation's goals and conditions, however, must be taken into account before providing a response.

Both statistical and non-statistical factors play a role in the sample size selection process. The availability of resources, labor, finances, ethics, and sample frame are examples of non-statistical factors. In terms of statistics, the anticipated prevalence of eye issues in school-age children as well as the desired precision of the prevalence estimate will be taken into account. [8]

A. Factors That Influence Sample Size:

The following are some of the numerous variables that determine the "right" sample size for a certain application:

- Cost considerations (e.g., maximum budget, desire to minimize cost).
- Administrative concerns (e.g., complexity of the design, research deadlines).
- Minimum acceptable level of precision.
- Confidence level.
- Variability within the population or subpopulation (e.g., stratum, cluster) of interest.

B. Sampling Method:

These elements have complex relationships. The last portion of this lecture discusses a scenario that frequently arises with basic random samples, however an exhaustive examination of all the variations is outside the purview of this tutorial: How to determine the smallest sample size that offers the necessary level of accuracy.

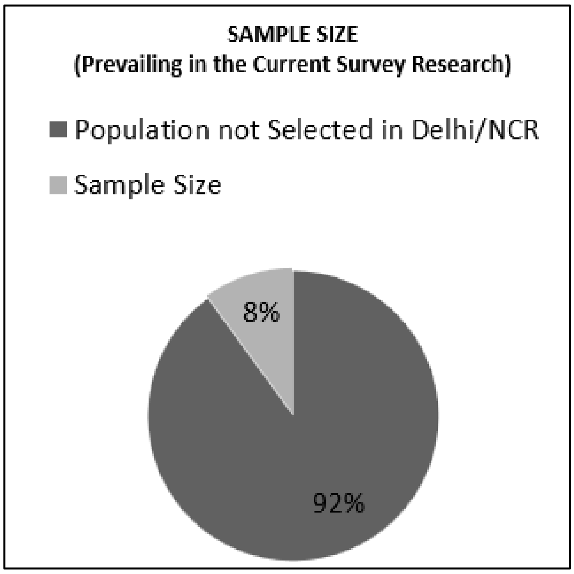

Figure 9.6: Pie Chart Depicting the Selected Sample Size [9]

Pie chart showing the chosen sample size shows the total population of Delhi/NCR, or 5000 people; the chosen sample size is 400 people who force the remaining 4600 people in the survey area to live as they see fit.

The following variables have been taken into consideration when selecting the sample:

- People who belong to a family of Chartered Accountants, Advocates or any other legal profession.
- People employed in various companies/Salaried persons.
- People running their own business/company.
- Shopkeepers and Retailers involved in Sale/Purchase of goods and commodities on a regular basis.
- People engaged in Service Sector.

Information about individuals who have not yet registered with the Income Tax Department and the U.P. Commercial Tax Department is archived.

Based on the previously mentioned Factors, the researcher has determined the chosen Sample Size. The quantity and choice of individuals or observations used in a study can serve as indicators of poor quality sampling. Acquiring a suitable sample size in both aspects is crucial for numerous reasons. The influence of outliers or extreme observations is largely mitigated by large sample sizes, which are also more typical of the population. To obtain findings among variables that are significantly different, a sufficiently high sample size is also required.

A big sample size expands the range of potential data and provides a clearer picture for analysis in qualitative research, where the objective is to "reduce the chances of discovery failure."

Finding the right sample size can be done in a variety of methods. Abbie Griffin and John Hauser discovered that "20–30 in-depth interviews are necessary to reveal 90–95% of the customer needs for the product categories studied” when conducting in-depth qualitative research.

Performing complex statistical calculations is typically required to determine the precise sample size required for a study. But the computed margin of error is used to choose a suitable sample size that is appropriate for the majority of investigations. [10]

4.5 Conclusion:

Finally, it can be concluded that the main costs and time savings associated with using a sample in research are avoided if an appropriate sampling strategy, a suitable sample size, and the necessary safety measures are taken to minimize sampling and measurement errors. A sample should also produce accurate and trustworthy results. The references listed below provide further information on sampling, as do many other library-available publications on statistics and qualitative research. The selection of sample size is a key component of the sampling strategy. Even though there might not be another option, it might be challenging and time-consuming to determine a suitable sample size from lists that already exist. The choice is heavily influenced by the target population and the nature of the study.

4.6 References:

1. Golafshani N. Understanding reliability and validity in qualitative research. *Qual Rep.* 2003; 8:597–607.
2. Peace KE, Parrillo AV, Hardy JC. “Assessing the validity of statistical inferences in public health research: An Evidence – based ‘Best Practice’ approach” *J Georgia Public Health Assoc.* 2008; 1:1.
3. Deng Y, Reeves M, Jacobs B, Birbeck L, Kothari R. IV tissue plasminogen activator use in acute stroke. *Neurology.* 2006; 66:306–12.
4. Salant, Priscilla, I. Dillman, and A. Don. How to conduct your own survey? No. 300.723 S3. 1994.
5. Hansen, Morris H., William N. Hurwitz, and William G. Madow. "Sample Survey Methods and Theory." (1953).
6. Major steps in selection of sample, available at http://www.sagepub.com/upmdata/40802_1.pdf. p. 5, visited on 22-10-2012
7. Jean Faugier and Mary Sargeant, ‘Sampling hard to reach populations’, Journal of Advanced Nursing’, 796 (1997).
8. Laura Blankertz, ‘American Journal of Evaluation, 1998 the value and practicality of deliberate sampling for heterogeneity: a critical multiplist perspective’ vol. 19(3), American Journal of Evaluation 309 (1998).
9. S.K. Varma & M. Afzal Wani (eds.), Legal Research and methodology, 319 (ILI Delhi. First Reprint, 2006)
10. Robert Magnani, Sampling Guide, December 1997, p.8.