



# CHOOSING THE RIGHT SAMPLING TECHNIQUE FOR YOUR RESEARCH

Awanis Ku Ishak, PhD

SBM

## Sampling...

The process of selecting a number of individuals for a study in such a way that the individuals represent the larger group from which they were selected.

## Purpose of Sampling

To gather data about the population in order to make an inference that can be generalized to the population

## **Sample...**

**...the representatives selected for a study whose characteristics exemplify the larger group from which they were selected**

## **Population**

**...the larger group from which individuals are selected to participate in a study**

# Define the Target Population

The **target population** is the collection of elements or objects that possess the information sought by the researcher and about which inferences are to be made.

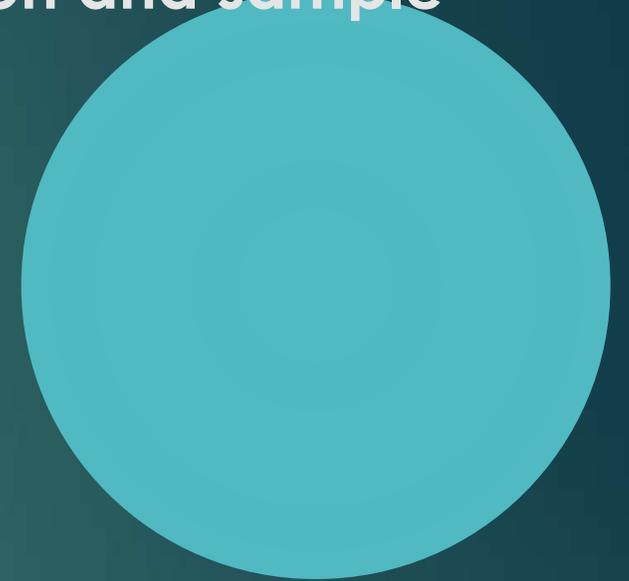
The target population should be defined in terms of elements, sampling units, extent, and time.

- ▶ **An element** is the object about which or from which the information is desired, e.g., the respondent.
- ▶ **A sampling unit is an element**, or a unit containing the element, that is available for selection at some stage of the sampling process.
- ▶ **Extent** refers to the geographical boundaries.
- ▶ **Time** is the time period under consideration.

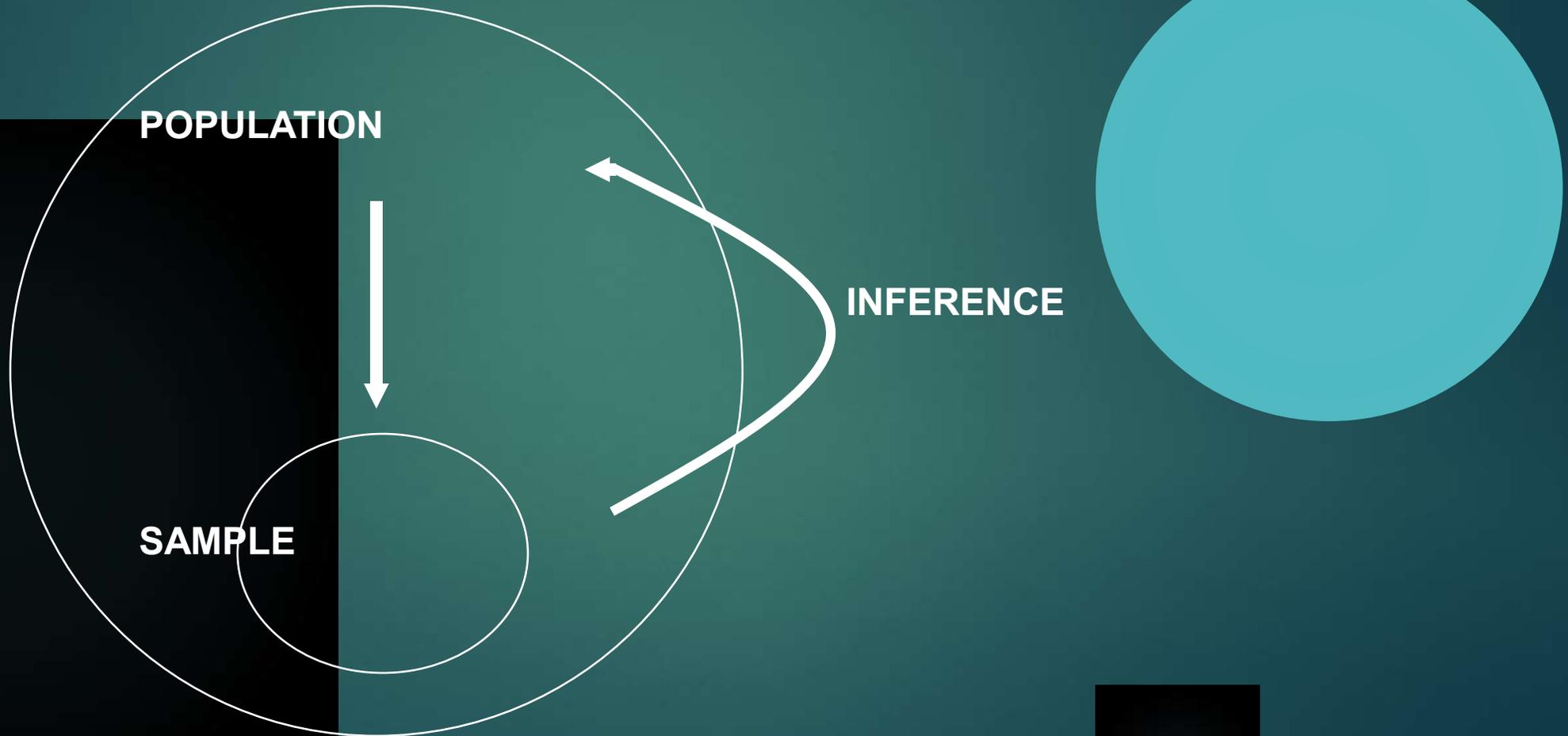
# Define the Target Population

Important factors in determining target population and sample size :

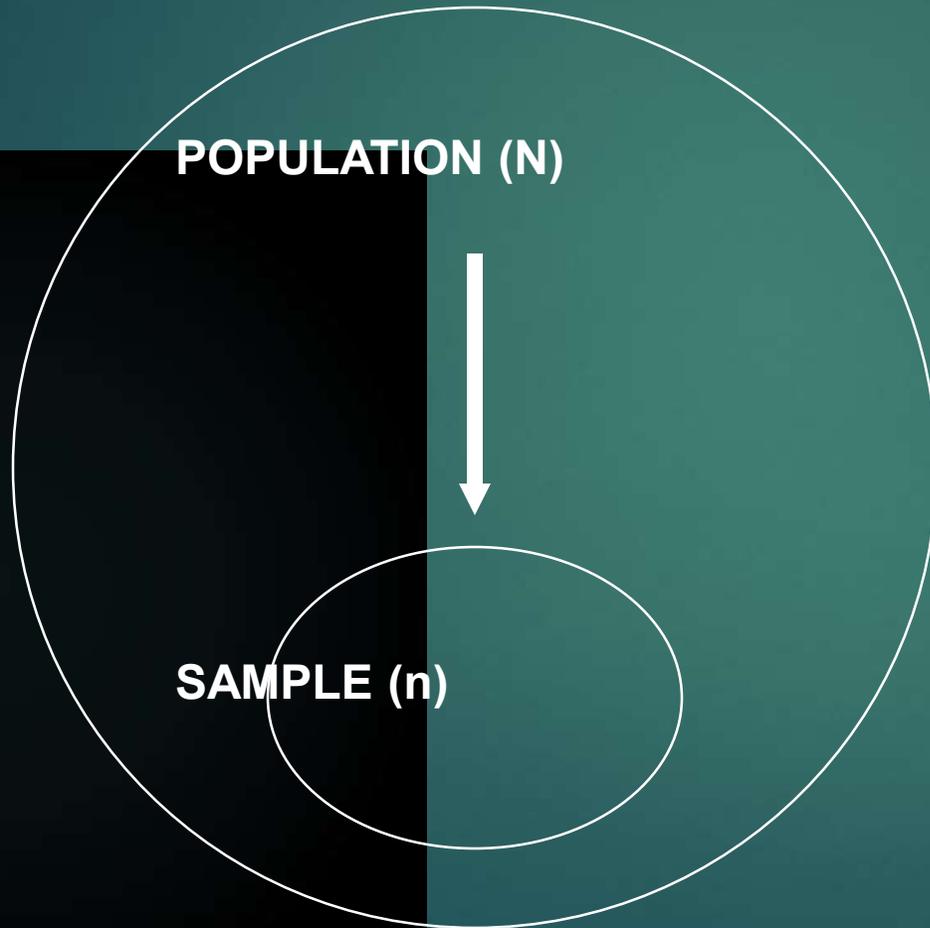
- ▶ the importance of the decision
- ▶ the nature of the research
- ▶ the number of variables
- ▶ the nature of the analysis
- ▶ sample sizes used in similar studies
- ▶ incidence rates
- ▶ completion rates
- ▶ resource constraints



# The sampling process...

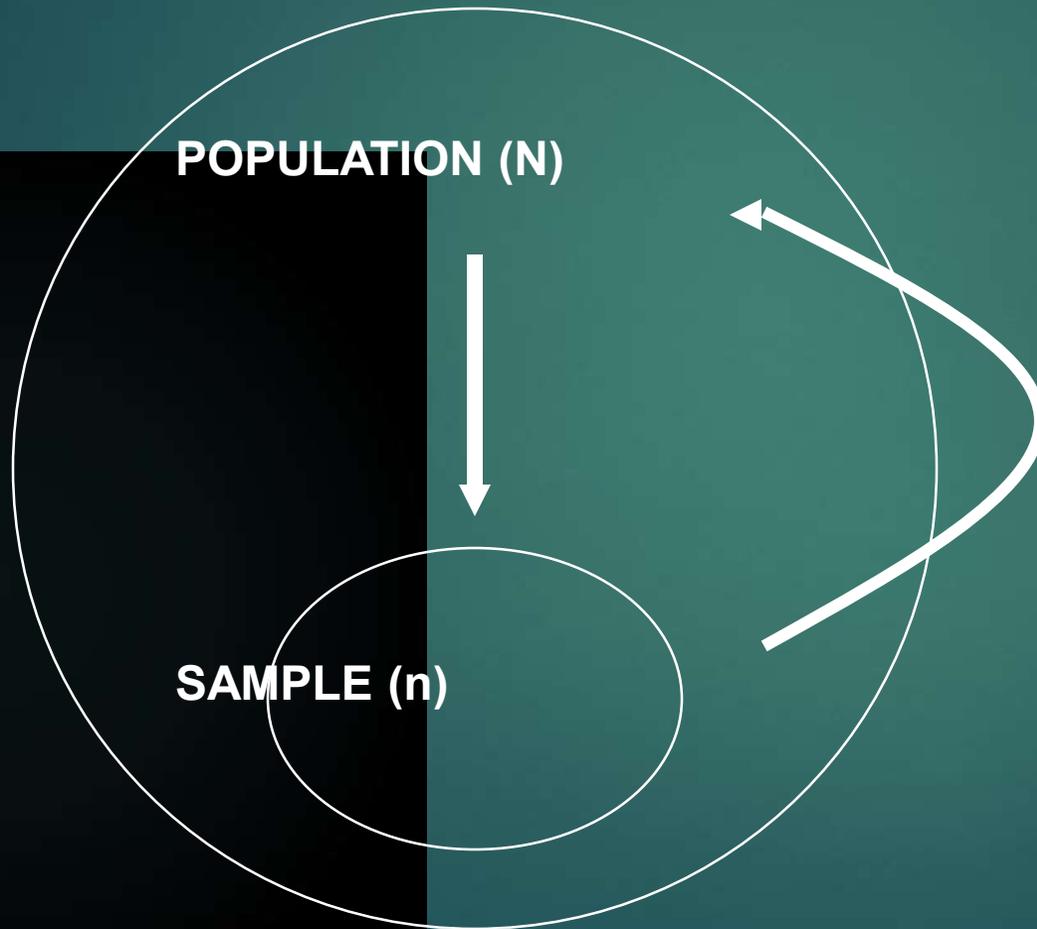


# Regarding the sample...

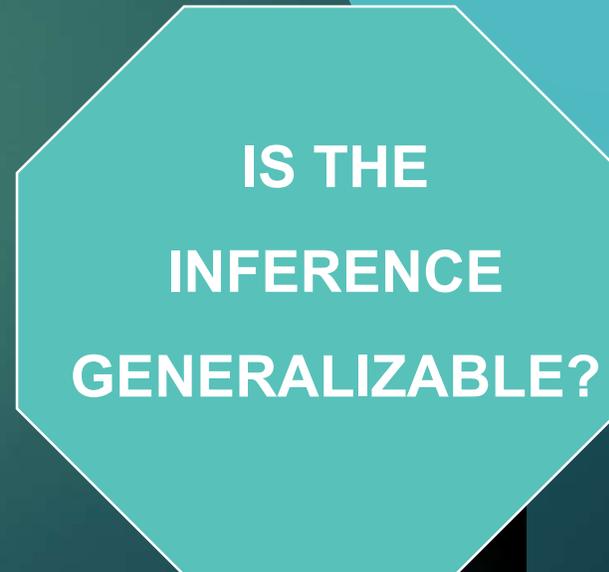


IS THE SAMPLE  
REPRESENTATIVE?

# Regarding the inference...



INFERENCE



# Mistakes to be conscious of...

1. Sampling error
2. Sampling bias

...which threaten to render a study's findings invalid

## **Sampling error...**

**...the chance and random variation in variables that occurs when any sample is selected from the population**

**...sampling error is to be expected**

## **Sampling bias...**

**...nonrandom differences, generally the fault of the researcher, which cause the sample is over-represent individuals or groups within the population and which lead to invalid findings**

**...sources of sampling bias include the use of volunteers and available groups**

# Control for sampling bias and error...

- Be aware of the sources of sampling bias and identify how to avoid it.
- Decide whether the bias is so severe that the results of the study will be seriously affected
- In the final report, document awareness of bias, rationale for proceeding, and potential effects

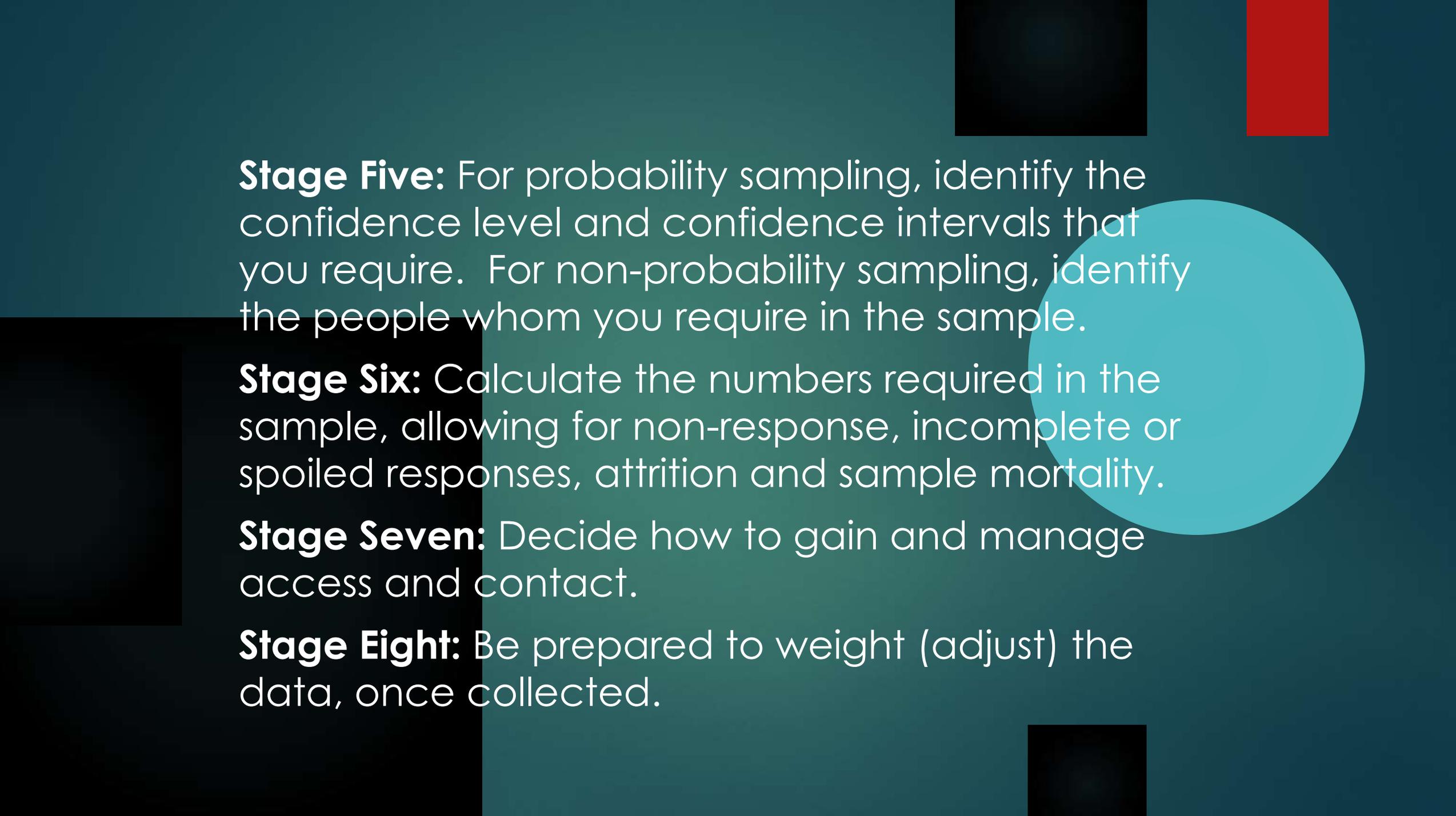
# SAMPLING DESIGN

**Stage One:** Decide whether you need a sample, or whether it is possible to have the whole population.

**Stage Two:** Identify the population, its important features (the sampling frame) and its size.

**Stage Three:** Identify the kind of sampling strategy you require (e.g. which variant of probability, non-probability, or mixed methods sample you require).

**Stage Four:** Ensure that access to the sample is guaranteed. If not, be prepared to modify the sampling strategy.



**Stage Five:** For probability sampling, identify the confidence level and confidence intervals that you require. For non-probability sampling, identify the people whom you require in the sample.

**Stage Six:** Calculate the numbers required in the sample, allowing for non-response, incomplete or spoiled responses, attrition and sample mortality.

**Stage Seven:** Decide how to gain and manage access and contact.

**Stage Eight:** Be prepared to weight (adjust) the data, once collected.

# Determine the sample size...

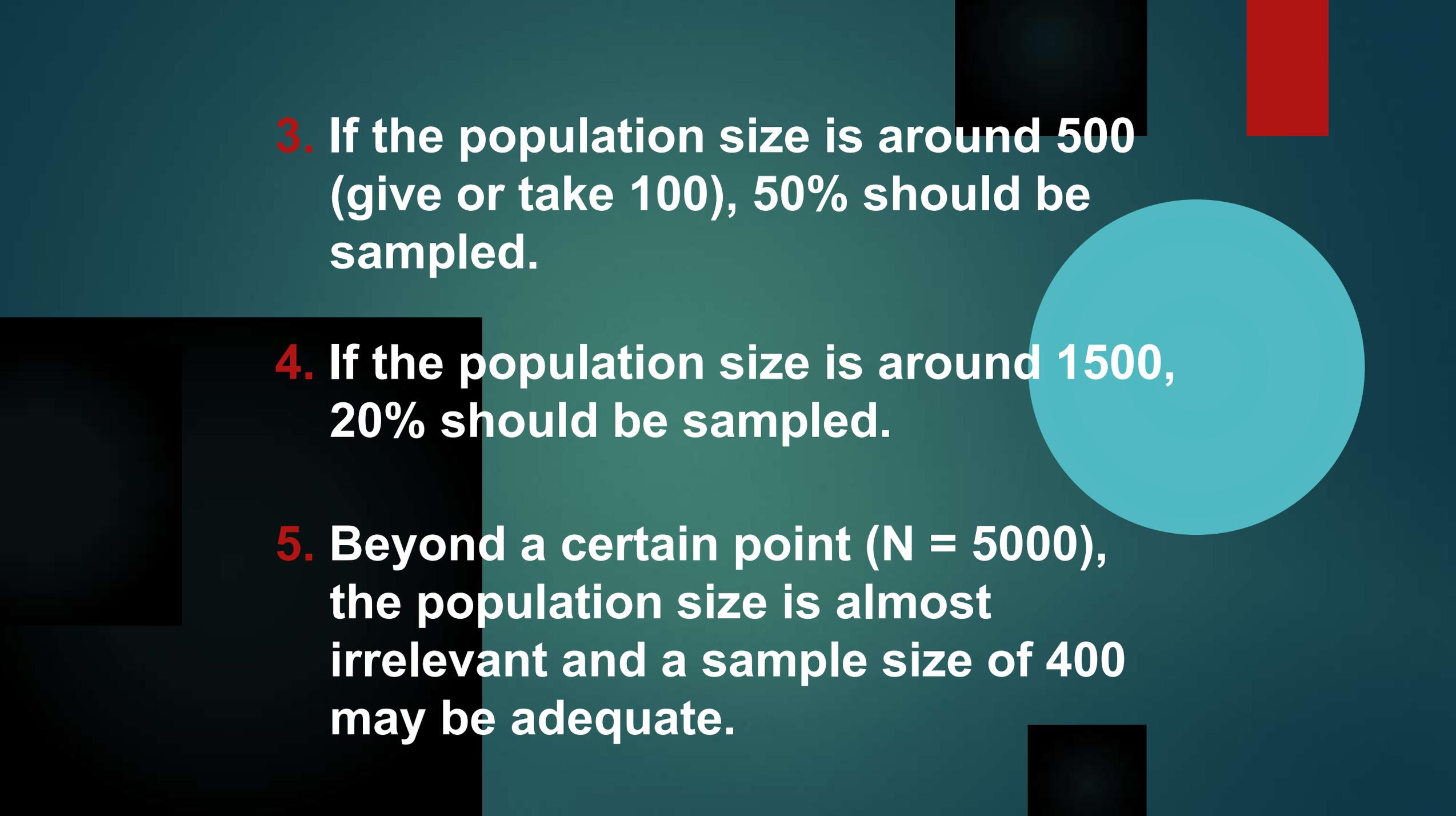
- The size of the sample influences both the representativeness of the sample and the statistical analysis of the data

**...larger samples are more likely to detect a difference between different groups**

**...smaller samples are more likely not to be representative**

# Rules of thumb for determining the sample size...

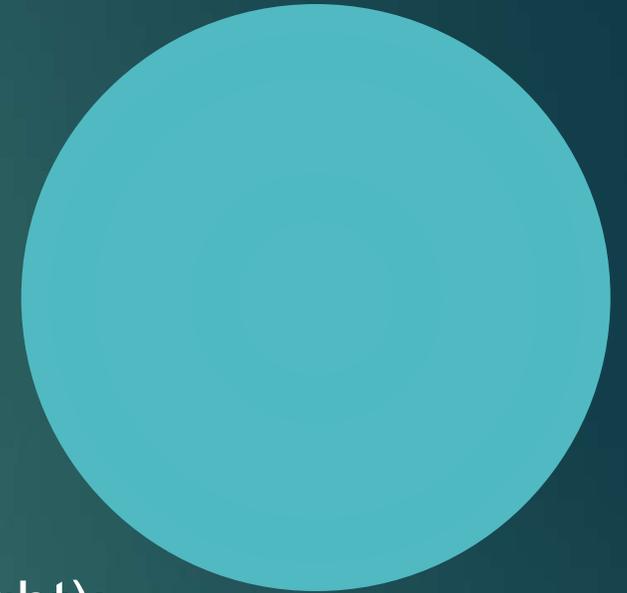
1. The larger the population size, the smaller the percentage of the population required to get a representative sample
2. For smaller samples ( $N < 100$ ), there is little point in sampling. Survey the entire population.

- 
- 3.** If the population size is around 500 (give or take 100), 50% should be sampled.
  - 4.** If the population size is around 1500, 20% should be sampled.
  - 5.** Beyond a certain point ( $N = 5000$ ), the population size is almost irrelevant and a sample size of 400 may be adequate.

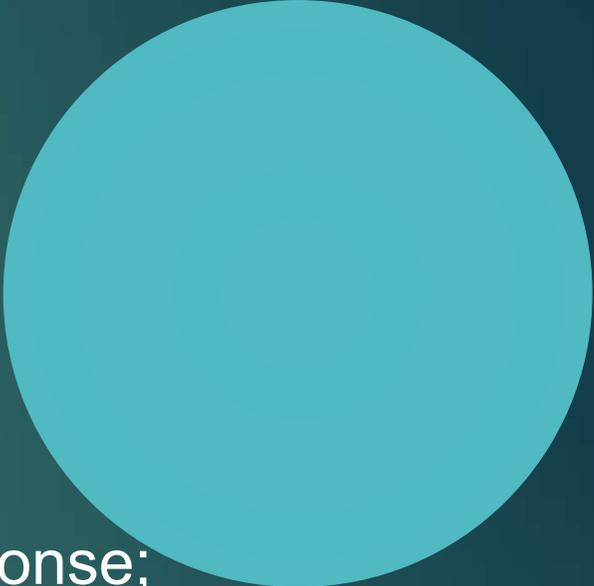
# HOW LARGE MUST MY SAMPLE BE?

It all depends on:

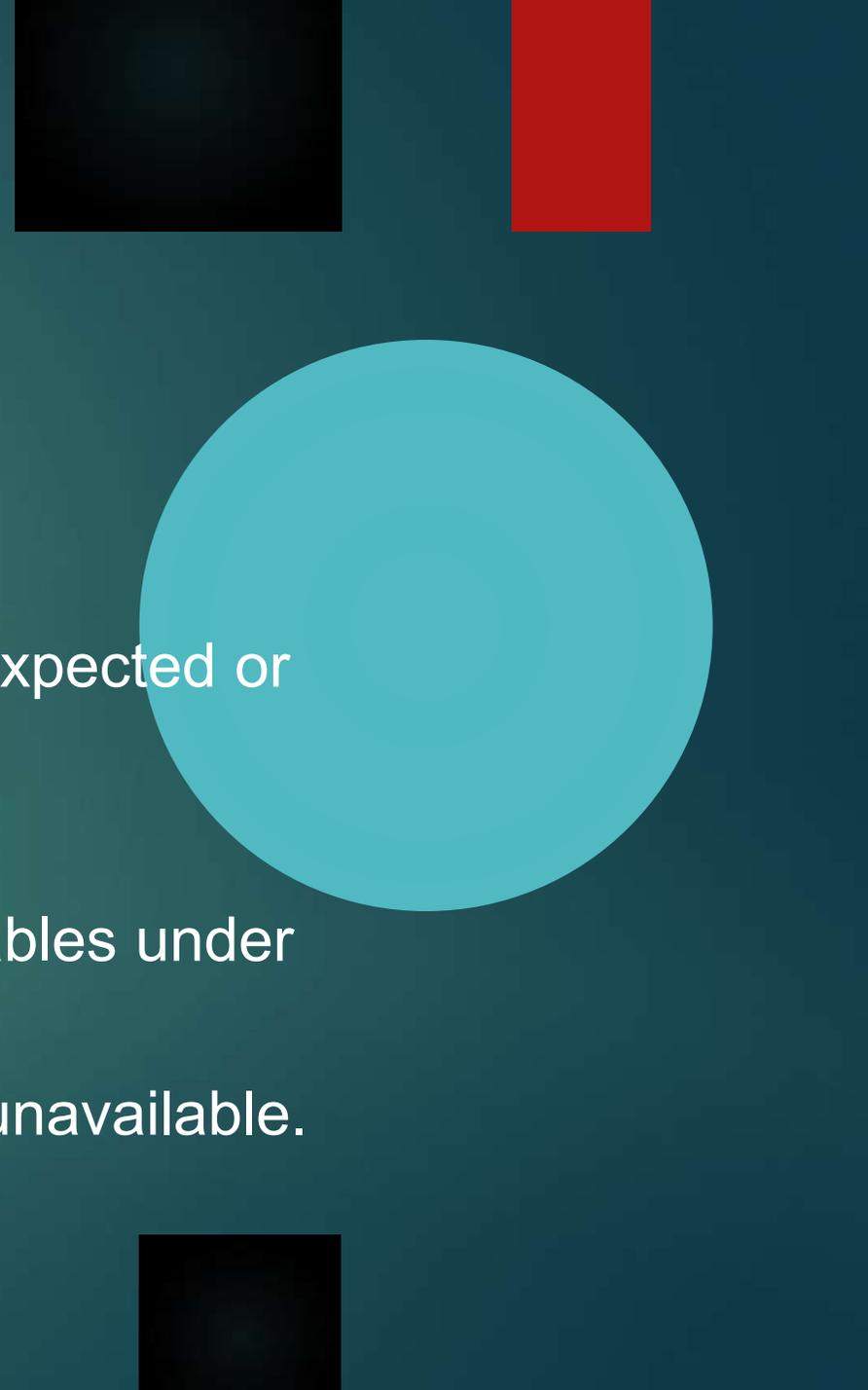
- The research purposes, questions and design;
- The population size;
- The confidence level and confidence interval required;
- The likely response rate;
- The accuracy required (the smallest sampling error sought);
- The kinds of variables to be used (categorical, continuous);
- The statistics to be used;



# HOW LARGE MUST MY SAMPLE BE?

- The number of strata required;
  - The number of variables included in the study;
  - The variability of the factor under study;
  - The kind(s) of sample;
  - The representativeness of the sample;
  - The allowances to be made for attrition and non-response;
  - The need to keep proportionality in a proportionate sample;
  - The kind of research that is being undertaken (qualitative/quantitative/mixed methods).
- 

# SAMPLE SIZE



Large samples are preferable when:

- ▶ there are many variables;
- ▶ only small differences or small relationships are expected or predicted;
- ▶ the sample will be broken down into subgroups;
- ▶ the sample is heterogeneous in terms of the variables under study;
- ▶ reliable measures of the dependent variable are unavailable.

# SAMPLE SIZE

- ▶ Ensure a sufficiently large sample for each variable.
- ▶ Samples in qualitative research must be large enough to generate 'thick descriptions'.
- ▶ A large sample does not guarantee representativeness; representativeness depends on the **sampling strategy**.
- Sample size also depends on the heterogeneity or homogeneity of the population: if it is highly homogeneous then a smaller sample may be possible.

# SAMPLE SIZE

- ▶ Sample size depends on the style of research (e.g. surveys may require large samples, ethnographies may require smaller samples).
- ▶ Sample size depends on the numbers of variables to be used, the kinds of variables, and the statistics to be calculated.
- ▶ Sample size depends on the scales being used in measurement (the larger the scale, the larger the sample).

# Select the sample...

- ▶ A process by which the researcher attempts to ensure that the sample is representative of the population from which it is to be selected

**...requires identifying the sampling method that will be used**

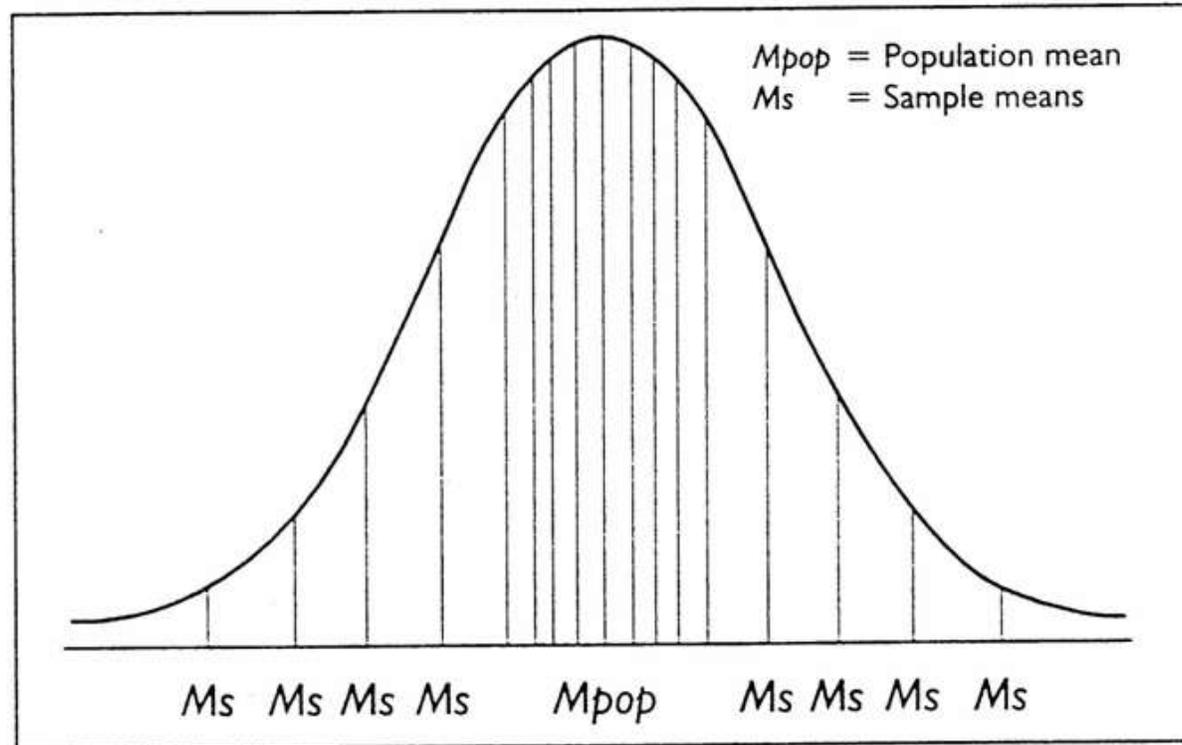
# SAMPLING ERROR OF THE SAMPLE

- ▶ Sampling error is the difference between the sample mean and the population mean, due to the chance selection of individuals.
- ▶ Sampling error reduces as the sample size increases.
- ▶ Samples of  $>25$  usually yield a normal sampling distribution of the mean.

# SAMPLING ERROR

Sample size depends on the margin of error and the confidence levels that the researcher is prepared to tolerate.

Distribution of sample means showing the spread of a selection of sample means around the population mean



# SAMPLE SIZE, CONFIDENCE LEVELS AND SAMPLING ERROR

N = Population; S = Sample

Note: As the population increases, the proportion of the population in the sample decreases.

N	S (95%)	S (99%)
50	44	50
100	79	99
200	132	196
500	217	476
1,000	278	907
2,000	322	1,661
5,000	357	3,311



# THE REPRESENTATIVENESS OF THE SAMPLE

- ▶ What is being represented (e.g. groups, variables, spread of population).
- ▶ If the sample has unequal sub-groups, then it may be necessary equalize the sample by weighting, to represent more fairly the population.

# ACCESS TO THE SAMPLE

- ▶ Is access to the sample permitted, practicable, realistic?
- ▶ Who will give/withhold/deny permission to access the sample?
- ▶ Who are the 'gatekeepers'?

# SAMPLING STRATEGIES

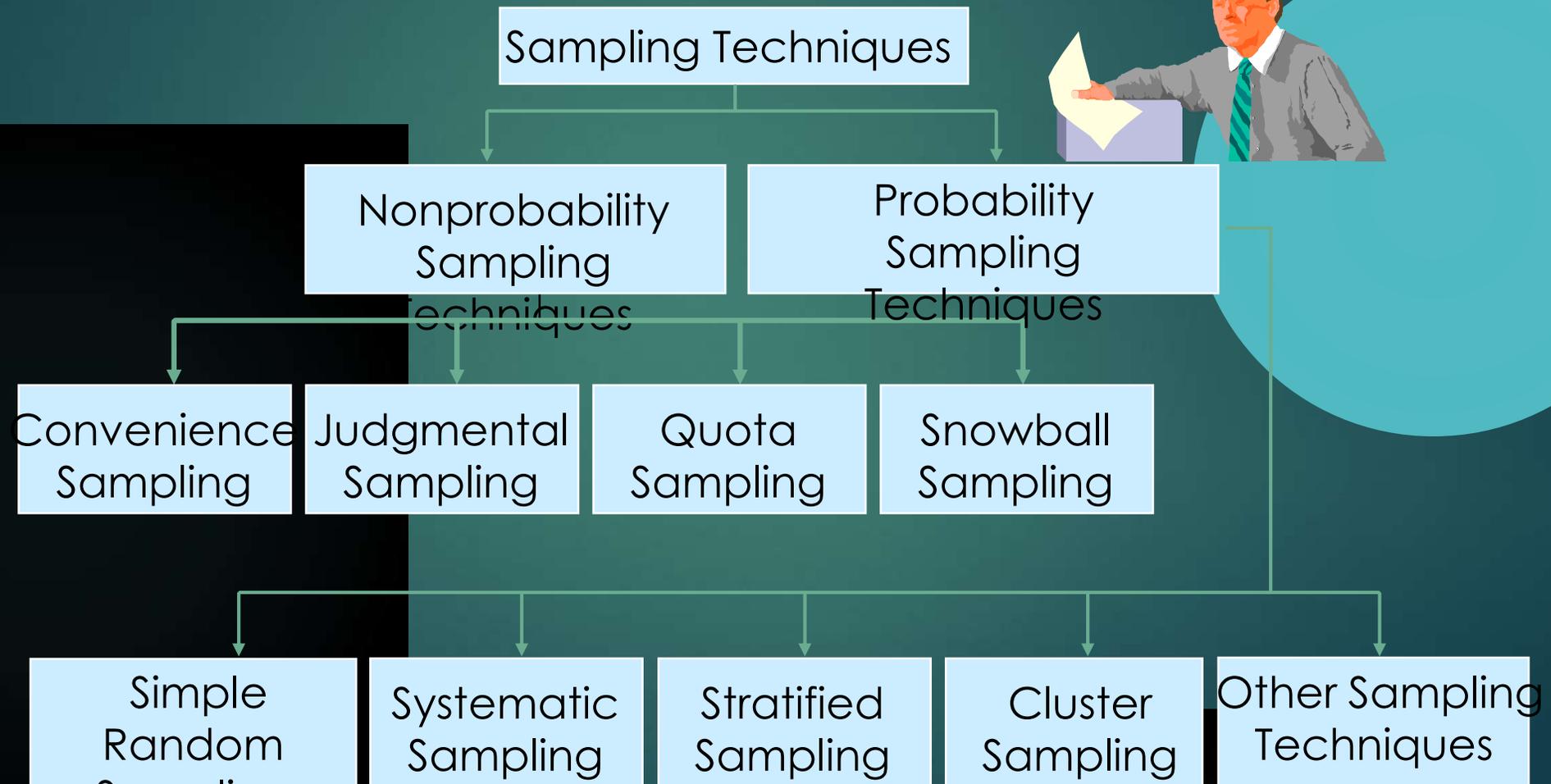
- ▶ Probability sample
- ▶ Non-probability sample



# Probability versus Nonprobability

- ▶ **Probability Samples:** each member of the population has a known non-zero probability of being selected
  - ▶ Methods include random sampling, systematic sampling, and stratified sampling.
- ▶ **Nonprobability Samples:** members are selected from the population in some nonrandom manner
  - ▶ Methods include convenience sampling, judgment sampling, quota sampling, and snowball sampling

# Classification of Sampling Techniques





# COMPARISON BETWEEN PROBABILITY AND NON PROBABILITY SAMPLING

# Comparison between Probability and Non Probability Sampling

Domain of comparison	Probability Sampling	Non Probability Sampling
Requirement of Resources	<b>More resources</b> in terms of time, cost and efforts	<b>Less resources</b> are required
Selection of Sample	<b>Random</b> ; No subjective judgment of researcher is involved	<b>Non random</b> , subjective judgements of researchers is involved
Quality of Inferences drawn	<b>Generalize</b> to the population	<b>Not generalizable</b> to the population
Best suited for researchers	To <b>understand a population</b>	To <b>understand or develop a concept or idea</b>
Applicable to the kind of population	<b>Elements which are finite in number</b> which precisely defined and a specific category	<b>Elements which are infinite</b> , too general category, nor precisely defined
Chances of error and biases	<b>Sampling errors and systematic biases are less likely to occur</b>	<b>Prone to encounter with systematic errors and sampling biases</b>
Types	Simple random sampling, systematic random sampling, Stratified random sampling, Cluster sampling, Multistage sampling	Volunteer sampling, Convenient sampling, Purposive sampling, Quota Sampling, Snowball sampling, Matched sampling, Genealogy based sampling



# TYPES OF PROBABILITY SAMPLING TECHNIQUES

# TYPES OF PROBABILITY SAMPLING TECHNIQUES

Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Simple Random	Every element has an equal chance of been selected. Exhaustive lists of elements are essential. Elements are selected randomly.	Exhaustive list of elements is produced. Each element is allotted a number. Numbers are randomly selected through lottery method or using computer generated random table.	Omits the chance of systematic errors and sampling biases. Representative sample is produced.	Difficult for very large population.
Systematic Random	Elements are selected at a regular interval. Exhaustive list may or may not be required. Is used when homogenous population is grouped within itself	Available lists are compiled to form a single list. Each element is given a number to select an appropriate interval. An element is selected at random and then every kth element from the first selection is included in the sample.	Ensures the extension of sample to the whole population make it possible to get a probable sample where list of elements cannot be	If the existing grouping is biased in some way the sample may not be representative. Difficult for very large population

# TYPES OF PROBABILITY SAMPLING TECHNIQUES

Name of Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Stratified	is used when a population is heterogeneous	first the population is divided into homogenous sub groups called as strata then elements are randomly selected from each stratum	ensures a representative sample for a heterogeneous population	it requires more resources in terms of time and efforts if the variable used for making strata is not appropriate depending on the research, the whole working may go in vain
Cluster	is used when the target population is homogenous but is spread over a wide geographical region instead of elements clusters are randomly selected a cluster is defined as a group of	first the population is divided into clusters each cluster is allotted a number then the decided number of clusters are selected randomly	make the probability sampling possible for a large population	there is possibility of systematic error if the selected clusters fail to capture the characteristic diversity of the target population, the sample cannot

# TYPES OF PROBABILITY SAMPLING TECHNIQUES

Name of Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Multistage sampling	it can be defined as sampling within the sample two or more probability sampling techniques are combined first a sample is extracted randomly and then from the selected sample another sample is extracted thus the to reach a final sample there are at least two stages	first the target population is divided into clusters clusters are randomly selected out of the selected clusters, there may be a formation of clusters or strata (in case of heterogeneity of population) now a random selection of clusters is done/ or there occurs a selection of elements from each strata the final units selected are investigated	a representative sample is produced for a population that is spread over a wide geographical region and is also heterogeneous	if the characteristic criterion used for the formation of strata at any stage is not appropriate, the sample cannot be representative of the population if also there occurs a systematic error in the selection of clusters, the results would not be able to generaliz



# TYPES OF NON PROBABILITY SAMPLING TECHNIQUES

# TYPES OF NON PROBABILITY SAMPLING TECHNIQUE

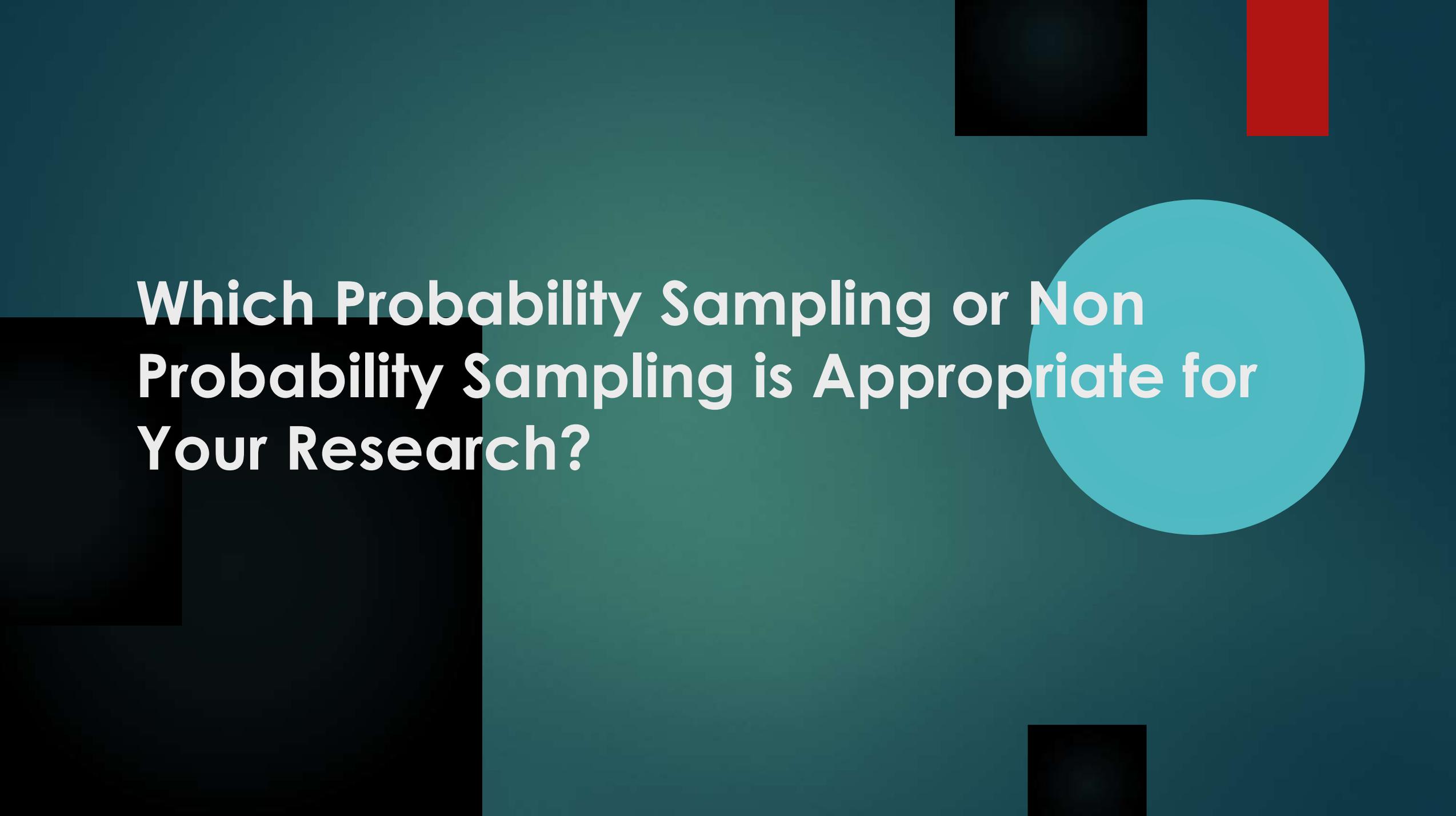
Name of Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Volunteer sampling	Participants self select themselves for the research	The researcher makes an announcement about his/ her research. Interested people approach the researcher	It helps in gathering a big amount of data in a very short time with small efforts	Sample is not representative of the population as the people who come to the researcher are those who already have an interest in the topic
Convenient Sampling	The elements of population who are easy or convenient to approach are selected It is used when the target population is defined in terms of a very broad category; for example, men, women, college students, business men etc.	Any member of the target population who is available at the moment is approached and is asked for participation in the research.	It is easy, inexpensive, consumes less efforts.	It is subjected to sampling biases and systematic errors.

# TYPES OF NON PROBABILITY SAMPLING TECHNIQUE

Name of Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Purposive sampling	Participants are selected with a purpose (a criterion of selection) in mind. Not any available person can be included but only those available are included who meet the criteria			
Quota sampling	It is used when population is heterogeneous.	Heterogeneous population is divided into homogenous sub groups. A quota for each sub group is set.	It captures the characteristic diversity of population. Ensures the representation from minority groups.	Sampling biases are possible to occur.
Snow ball Sampling	It is used when the target population is not readily available and difficult to approach.	First approached participant is asked to refer to the others. Next is asked to refer to another and so on.	It makes possible to approach a rarely available population.	Systematic errors are likely to occur because of chain networking.

# TYPES OF NON PROBABILITY SAMPLING TECHNIQUE

Name of Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Matched Sampling	Is used in experimental researches. A pair of two matched elements is selected.	One element is judged to be a part of the research. Then, another element is explored that resembles the first one on a variety of variables important for the research .	It provides a control group for the experiment.	The two elements must be matched on every possible influencing factor so it may be claimed that the intervention and not something else has produced the change.
Genealogy based Sampling	Instead of selecting household in an area, the members of the entire families are selected, whether or not living in the same house).	One element is persuaded for the research. Then, first participant is asked to refer to his/her close relatives.	Gives a reasonable cross section of the community by age and gender. Saves time and efforts.	

The background features a teal gradient with several geometric shapes: a black square in the top right, a red square in the top right, a large cyan circle on the right side, and a black square in the bottom right. The text is centered on the left side of the slide.

**Which Probability Sampling or Non  
Probability Sampling is Appropriate for  
Your Research?**

# Which Sampling Technique to Use in Research?

Choosing appropriate technique is important to make a research project beneficial, meaningful and successful.

Some consideration:

**Representative samples** produced by probability sampling. But some instances where it is impossible to use a probability sampling technique. Elements of target population may be so widely dispersed and/or apparently infinite in number that capturing a probability is not possible.

Probability techniques can be used only for **finite population**. Otherwise use non probability techniques.

Probability sampling is not possible for **not very precisely defined** population (too general a category).

## Prime goal of research also affects sampling technique.

- If it is meant at exploring an idea rather than understanding a population, use non-probability techniques;
- If it is intended to develop the understanding about the population, use probability sampling.
- Also, it depends on the availability of resources (i.e. time, cost and efforts) to use which sampling technique.
- Where enough resources are available, probability techniques are preferred; where resources are not enough non probability techniques are suggested.

# Which Probability Sampling is Appropriate for Your Research?

Use probability sampling technique on the basis of goal of your research and availability of your resources.

Appropriate type of the sampling technique to produce a representative sample.

Some guide lines to select an appropriate sampling technique:

- If you target population is homogenous\* and an exhaustive list of elements is available/ possible to prepare, **simple random sampling** is appropriate.
- If your target population is homogenous\* and list of elements cannot be produced for any reason, **systematic sampling** is best suited.
- **Systematic random sampling** has also an edge over simple random in the situation where the target population is homogenous but is already grouped within itself. The reason of this preference is the ability of systematic sampling to ensure the extension of sample to the whole population.

- If your target population is heterogeneous\* and is not very much widely spread geographically, **stratified random sampling** is used.
- If your target population is homogenous\* but is spread in a wide geographical area, **cluster sampling** is used.
- Sometimes it happens that the clusters are very big and inclusion of all the elements is not possible; in this case participants are randomly selected within the cluster (using simple random or systematic random sampling). This sampling technique is called as **double staged sampling**.
- If you target population is spread over a wide geographical area and is also heterogeneous\*, **multistage sampling** is most effective.

# Which Non-Probability Sampling is Appropriate for Your Research?

Use non-probability sampling technique on the basis of goal of your research and availability of your resources (i.e. time, money, and efforts):

Below are some guide lines:

- If you have shortage of time and/or efforts cannot be utilized in search of participants, **volunteer sampling** is most feasible.
- If your target population is a very general category: found everywhere and not precisely defined, **convenient sampling** is used.
- If your target population is heterogeneous and it is important to take representation from every sub group, **quota sampling** is the method to use.
- If your target population is the type of people who are not easily available, lesser in number and rare to find, snowball sampling technique is used.