BP801T.BIOSTATISTICS AND RESEARCH METHODOLOGY(Theory)

Unit-III

Introduction to research

Introduction to Research

Need for research

Need for design of Experiments

Experiential Design Technique

Plagiarism Graphs:

- Histogram
- Pie Chart
- Cubic Graph
- response surface plot
- Counter Plot graph

A detailed and careful study of something to find out more information about it

Research is the careful consideration of study regarding a particular concern or problem using scientific methods. According to the American sociologist Earl Robert Babbie, “research is a systematic inquiry to describe, explain, predict, and control the observed phenomenon. It involves inductive and deductive methods.

Research is "creative and systematic work undertaken to increase the stock of knowledge". It involves the collection, organization and analysis of evidence to increase understanding of a topic, characterized by a particular attentiveness to controlling sources of bias and error.

What is the purpose of research?

Exploratory: As the name suggests, researchers conduct exploratory studies to explore a group of questions. The answers and analytics may not offer a conclusion to the perceived problem. It is undertaken to handle new problem areas that haven't been explored before. This exploratory process lays the foundation for more conclusive data collection and analysis.

Descriptive: It focuses on expanding knowledge on current issues through a process of data collection. Descriptive research describe the behavior of a sample population. Only one variable is required to conduct the study. The three primary purposes of descriptive studies are describing, explaining, and validating the findings. For example, a study conducted to know if top-level management leaders in the 21st century possess the moral right to receive a considerable sum of money from the company profit.

Explanatory: Causal or explanatory research is conducted to understand the impact of specific changes in existing standard procedures. Running experiments is the most popular form. For example, a study that is conducted to understand the effect of rebranding on customer loyalty.

Research methods are broadly classified as **Qualitative** and **Quantitative**.
Both methods have distinctive properties and data collection methods.

Qualitative methods

Qualitative research is a method that collects data using conversational methods, usually open-ended questions. The responses collected are essentially non-numerical. This method helps a researcher understand what participants think and why they think in a particular way.

Types of qualitative methods include:

- One-to-one Interview
- Focus Groups
- Ethnographic studies
- Text Analysis
- Case Study

Quantitative methods

Quantitative methods deal with numbers and measurable forms. It uses a systematic way of investigating events or data. It answers questions to justify relationships with measurable variables to either explain, predict, or control a phenomenon.

Types of quantitative methods include:

- Survey research
- Descriptive research
- Correlational research

Design of experiments (DOE) is defined as a branch of applied statistics that deals with planning, conducting, analyzing, and interpreting controlled tests to evaluate the factors that control the value of a parameter or group of parameters.

Experimental design techniques

- The design of fundamental relationship under controlled situation is called experimental design.
- Controlling is nothing but holding one variable constant while other variables are free to vary within the experiment.
- Independent variables are manipulated in order to measure their effect on dependent variable.

- An experiment is an observation under controlled conditions.
- It is a design in which some of the variables being studied are manipulated to control the conditions.

Types

- Pre –experimental designs
- Quasi-experimental design
- Observational or true experimental design

Pre- experimental designs

- It is the design in which the basic experimental steps are followed without assigning any control group.
- The researcher studies a single group with no any comparison within this group.
- It is a preparatory to true experimental design.
- It is the simplest form of research design.

Pre experimental design only manipulation
no randomisation and no control

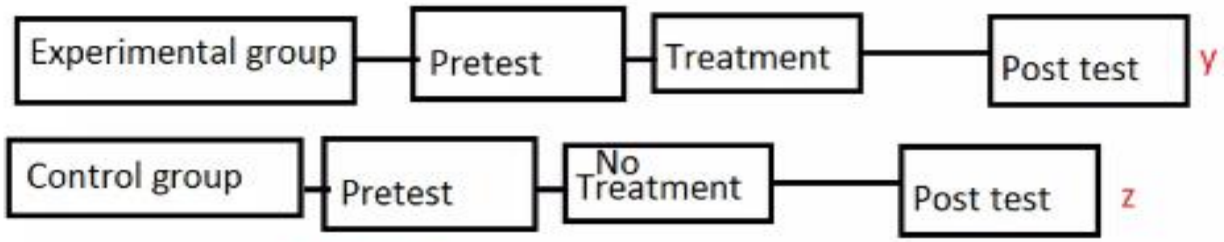
1. one shot case design

Pre experimental design only manipulation
no randomisation and no control

2.one group pretest post test design

Pre experimental design only manipulation
no randomisation and no control

Before after with control design



= [y-x] - [z-A]

Quasi experimental design

- The study subject are not randomly assigned to the groups.
- Manipulation of an independent variable.
- It differs from experimental research because there is no control, random selection and no active manipulation.
- It is performed when a control group cannot be created or random selection cannot be performed.
- It is used in certain medical and paramedical studies.

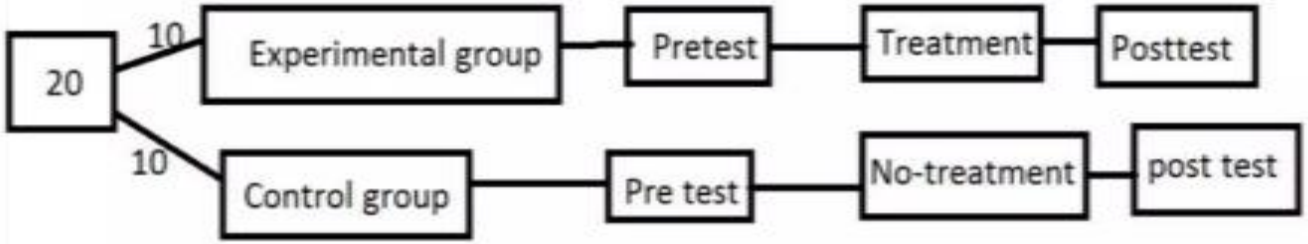
Quasi experimental design

Non randomised control group design

Manipulation- yes

Randomisation- No

Control group- yes



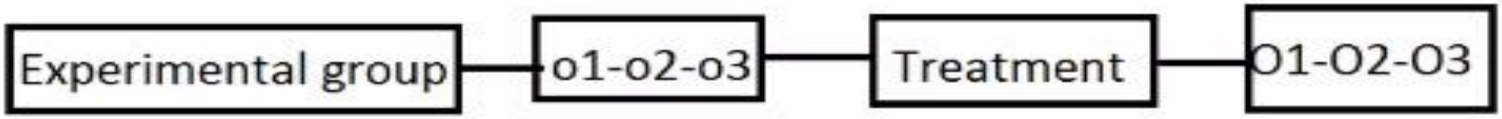
Quasi experimental design

Time series design- long term access of effect

Manipulation- yes

randomisation -no

Control-yes



before treatment i hr, 2 hr, 3hr access the pain level

Manipulation -pain killer injection is given

After treatment 1 hr, 2 hr, 3 hr access pain level.

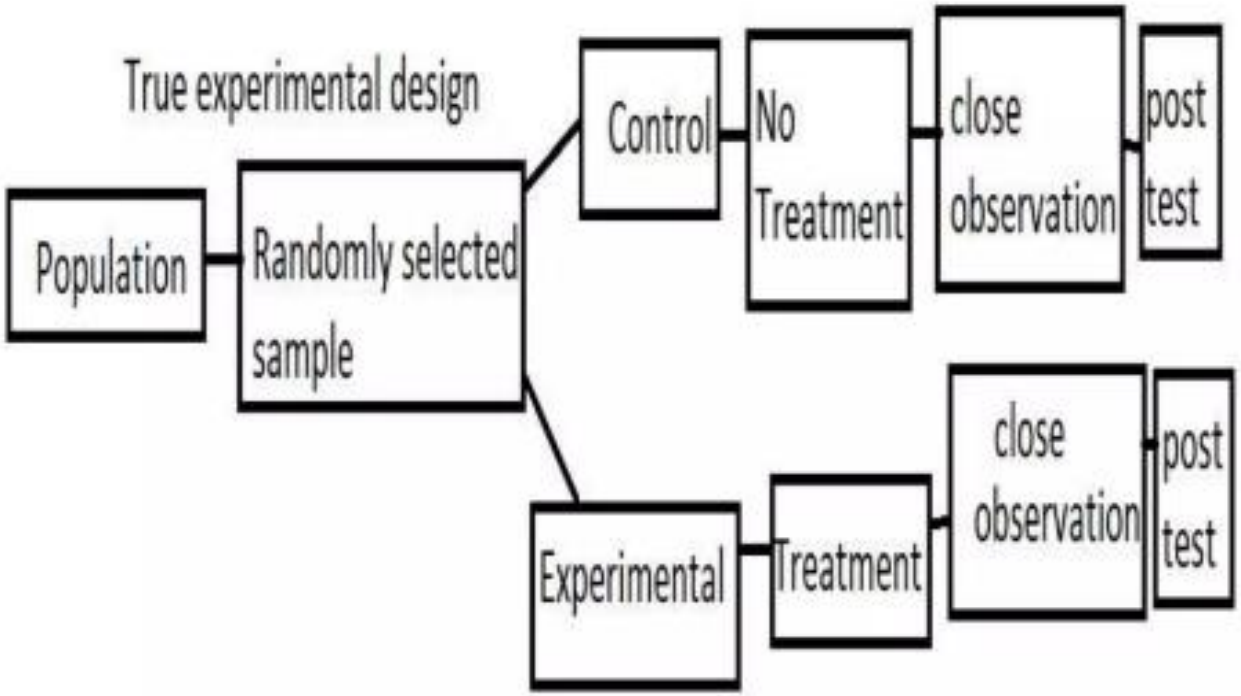
True experimental design

- It employs statistical analysis to support or reject a hypothesis.
- It has control group.
- It is the most accurate type of experimental research it supports or disproves the hypothesis using statistical analysis
- It is the only design where it cause and effect the relationship between the variables.

True experimental design satisfy the following

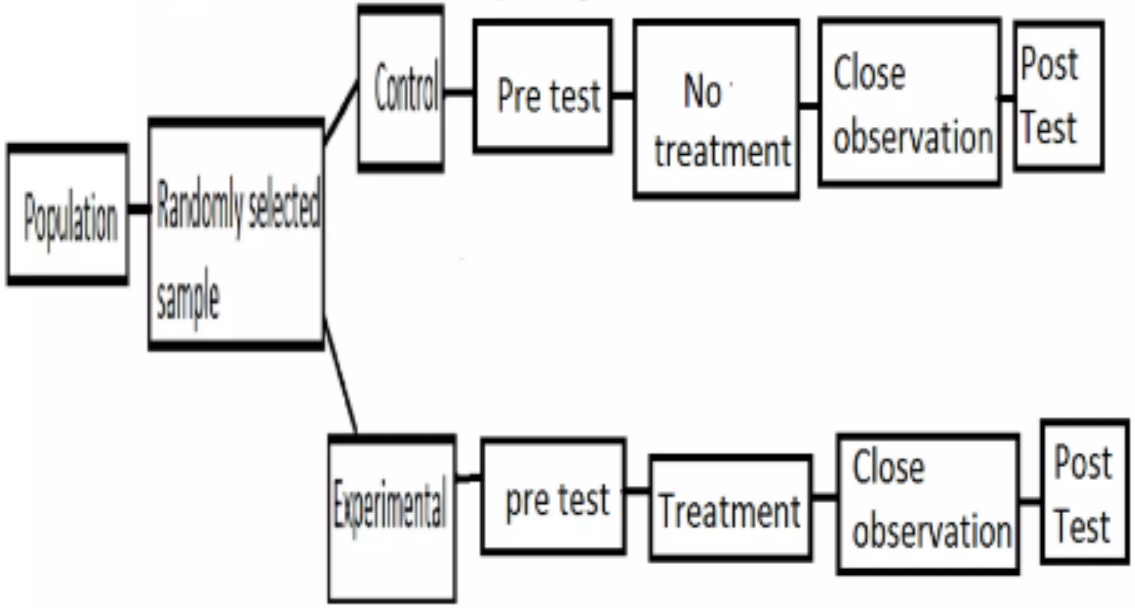
- There must be a control group which wont change
- An experimental group which will experience the changed variables.
- There must be a variables that can be manipulated by researcher.
- There must be a random distribution
- It can be further categorized as

Post test only control group design



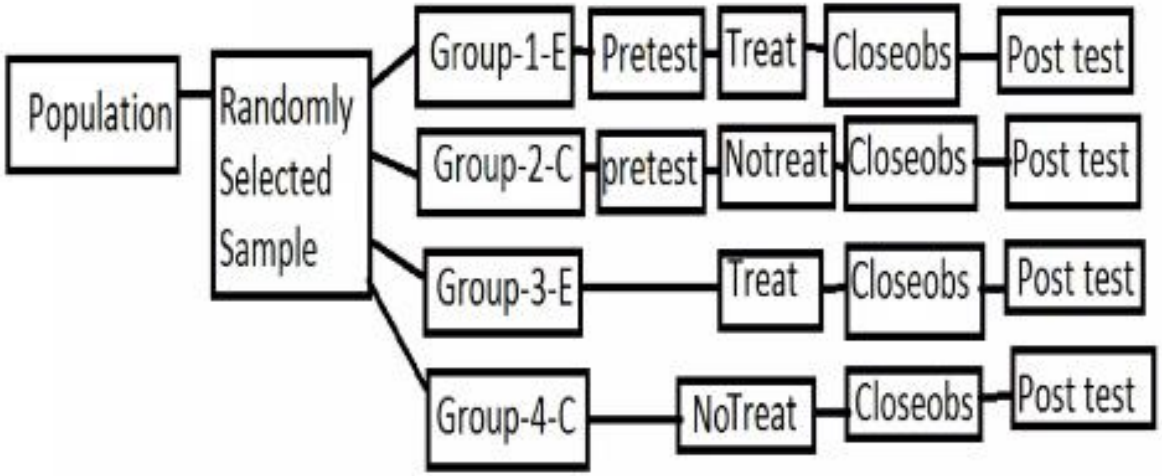
True experimental design

Pretest posttest control group design



True experiment design

Solomon Four Group Design



Plagiarism is presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into your work without full acknowledgement.

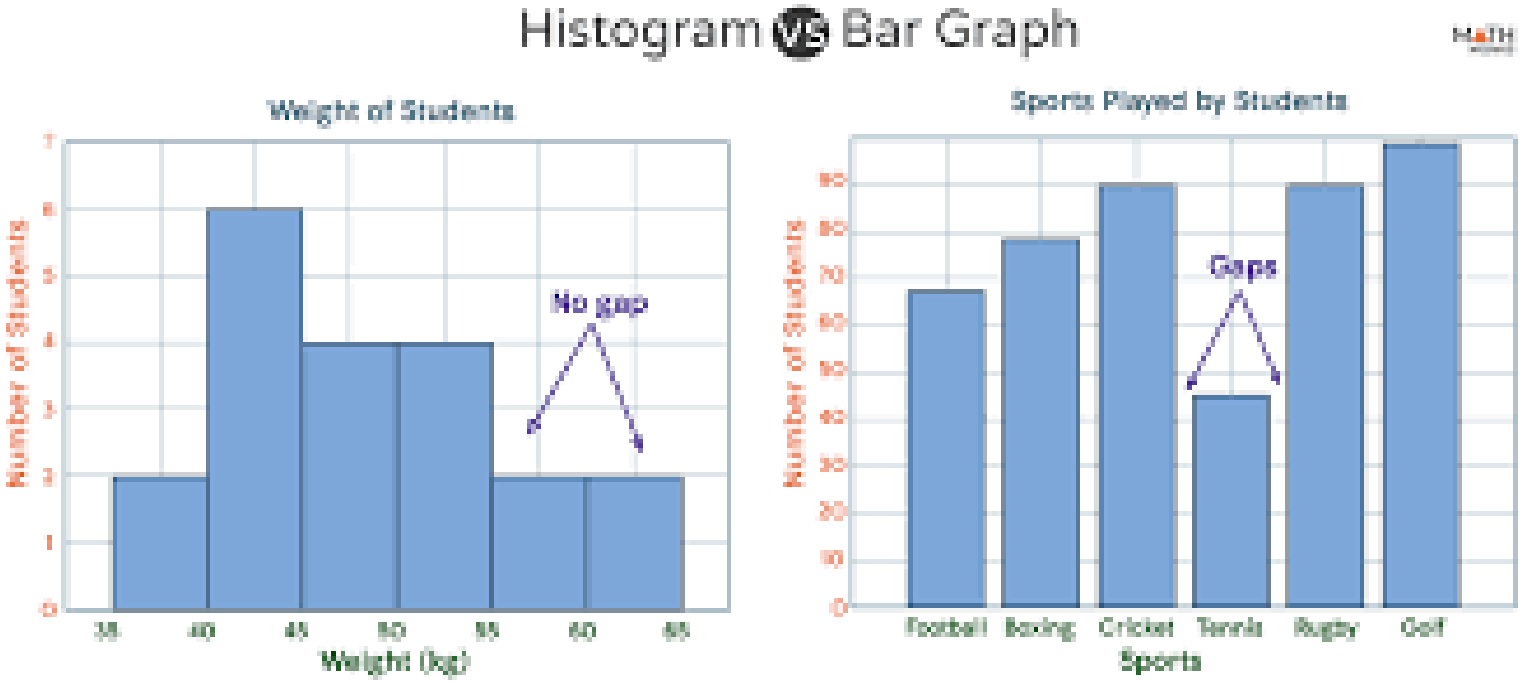
Plagiarism is the unethical practice of using words or ideas (either planned or accidental) of another author/researcher or your own previous works without proper acknowledgment.

How Can You Avoid Plagiarism in a Research Paper?

1. Paraphrase your content
2. Use quotes to indicate that the text has been taken from another paper.
3. Cite your Sources – Identify what does and does not need to be cited
4. Maintain records of the sources you refer to
5. Use plagiarism checkers (iThenticate or HelioBLAST (formerly eTBLAST)

What Is a Histogram? A histogram is **a graphical representation of data points organized into user-specified ranges**.

A **histogram** is an approximate representation of the distribution of numerical data. The term was first introduced by Karl Pearson.
A **histogram** is a graphical display of data with bars of different heights, where each bar groups numbers into ranges.

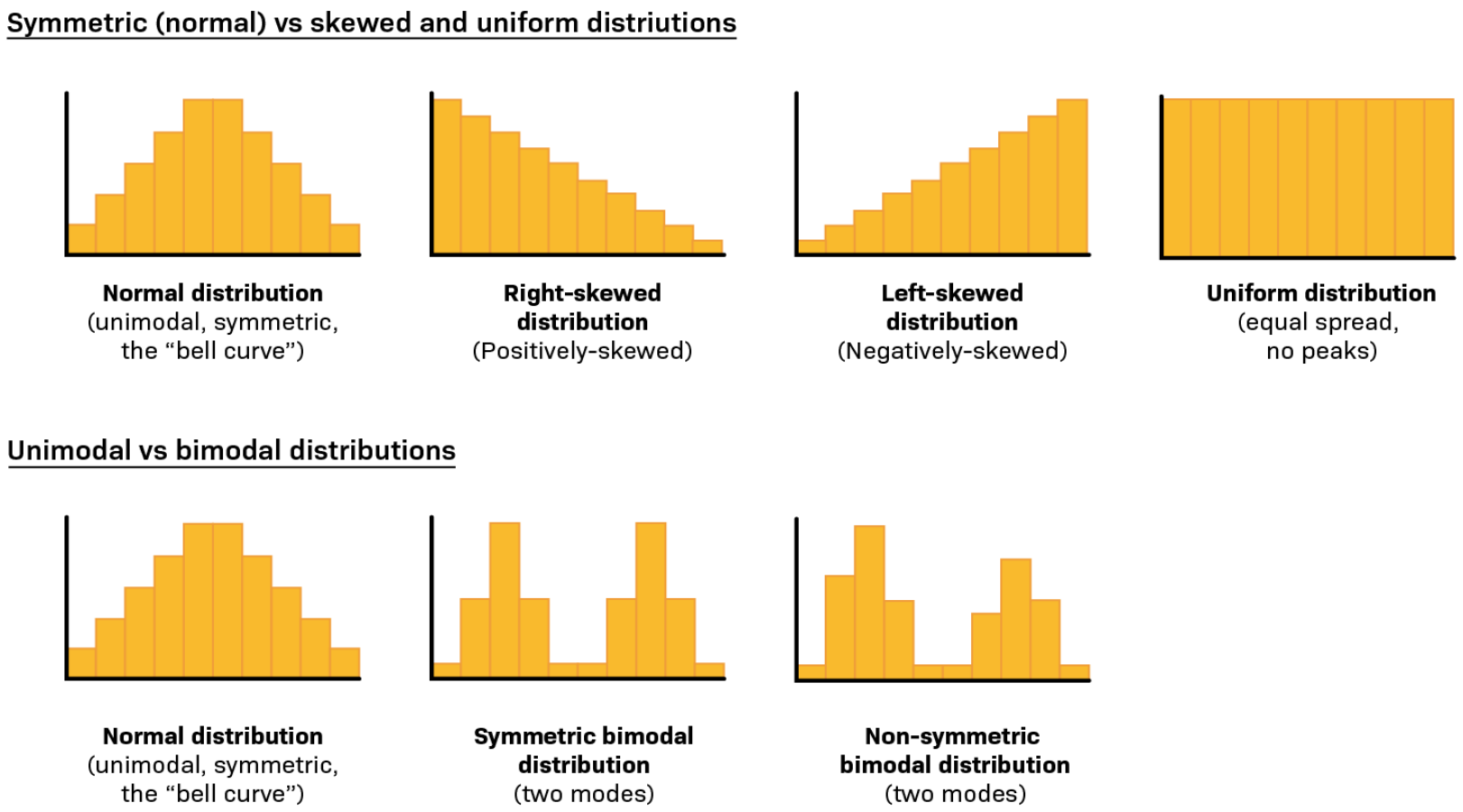


Histogram

Bar Graph

Types of Histogram

Uniform histogram.
Symmetric histogram.
Bimodal histogram.



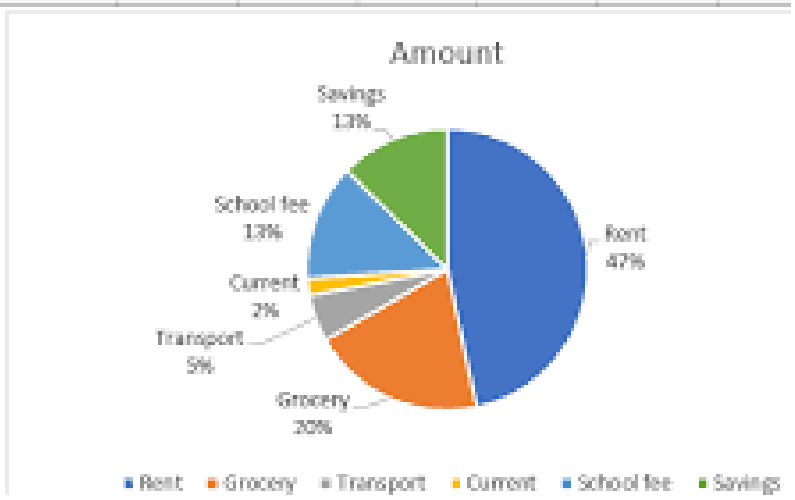
A **pie chart** is a type of graph that represents the data in the circular graph. The slices of pie show the relative size of the data, and it is a type of pictorial representation of data. A pie chart requires a list of categorical variables and numerical variables.

The "**pie chart"** is also known as a "circle chart", dividing the circular statistical graphic into sectors or sections to illustrate the numerical problems. Each sector denotes a proportionate part of the whole. To find out the composition of something, Pie-chart works the best at that time. In most cases, pie charts replace other graphs like the bar graph, line plots, histograms, etc.

The pie chart is an important type of data representation. It contains different segments and sectors in which each segment and sector of a pie chart forms a specific portion of the total(percentage). The sum of all the data is equal to 360°.
The total value of the pie is always 100%.
To work out with the percentage for a pie chart, follow the steps given below:
Categorize the data
Calculate the total
Divide the categories
Convert into percentages
Finally, calculate the degrees

Pie Chart Examples

	Expenses	Amount
2	Rent	7000
3	Grocery	3000
4	Transport	800
5	Current	300
6	School fee	2000
7	Savings	1900
8		
9		
10		
11		
12		
13		



There are mainly 7 types of pie charts using which one can represent the data. They are

3D pie chart and perspective pie cake,

Doughnut chart,

Exploded pie chart,

Polar area diagram,

Ring chart,

Spie chart,

Square chart.

NON-PARAMETRIC TEST

- 2 Definition
- Nonparametric satistics, also known as distribution-free statistics, are methods of testing hypotheses when the nature of the distributions are unknown.Some of nonparametric statistics: Sign test, Wilcoxon signed rank test, Wilcoxon rank sum test, Kruskal-Wallis test, Friedman test, Rank correlation

Mann Whitney U Test (Wilcoxon Rank Sum Test)

The modules on hypothesis testing presented techniques for testing the equality of means in two independent samples. An underlying assumption for appropriate use of the tests described was that the continuous outcome was approximately normally distributed or that the samples were sufficiently large (usually $n_1 \geq 30$ and $n_2 \geq 30$) to justify their use based on the Central Limit Theorem. When comparing two independent samples when the outcome is not normally distributed and the samples are small, a nonparametric test is appropriate.

A popular nonparametric test to compare outcomes between two independent groups is the Mann Whitney U test. The Mann Whitney U test, sometimes called the Mann Whitney Wilcoxon Test or the Wilcoxon Rank Sum Test, is used to test whether two samples are likely to derive from the same population (i.e., that the two populations have the same shape). Some investigators interpret this test as comparing the medians between the two populations. Recall that the parametric test compares the means (H_0: $\mu_1=\mu_2$) between independent groups.

In contrast, the null and two-sided research hypotheses for the *nonparametric test* are stated as follows:

H_0: The two populations are equal versus

H_1: The two populations are not equal.

This test is often performed as a two-sided test and, thus, the research hypothesis indicates that the populations are not equal as opposed to specifying directionality. A one-sided research hypothesis is used if interest lies in detecting a positive or negative shift in one population as compared to the other. The procedure for the test involves pooling the observations from the two samples into one combined sample, keeping track of which sample each observation comes from, and then ranking lowest to highest from 1 to n_1+n_2, respectively.

Kruskal Wallis Test

•The Kruskal Wallis test is the non parametric alternative to the One Way ANOVA. *Non parametric* means that the test doesn't assume your data comes from a particular distribution. The H test is used when the assumptions for ANOVA aren't met (like the assumption of normality). It is sometimes called the *one-way ANOVA on ranks*, as the ranks of the data values are used in the test rather than the actual data points.

•The test determines whether the medians of two or more groups are different. Like most statistical tests, you calculate a test statistic and compare it to a distribution cut-off point. The test statistic used in this test is called the **H statistic.** The hypotheses for the test are:

H_0: population medians are equal.
H_1: population medians are not equal.

The Kruskal Wallis test will tell you if there is a significant difference between groups. However, it won't tell you *which* groups are different. For that, you'll need to run a Post Hoc test.

Friedman Test

The **Friedman Test** is a non-parametric alternative to the Repeated Measures ANOVA.
It is used to determine whether or not there is a statistically significant difference between the means of three or more groups in which the same subjects show up in each group.

When to Use the Friedman Test

The Friedman Test is commonly used in two situations:

1. Measuring the mean scores of subjects during three or more time points.

For example, you might want to measure the resting heart rate of subjects one month before they start a training program, one month after starting the program, and two months after using the program. You can perform the Friedman Test to see if there is a significant difference in the mean resting heart rate of patients across these three time points.

2. Measuring the mean scores of subjects under three different conditions.

For example, you might have subjects watch three different movies and rate each one based on how much they enjoyed it. Since each subject shows up in each sample, you can perform a Friedman Test to see if there is a significant difference in the mean rating of the three movies.