

---

# **UNIT4: DATA VISUALIZATION AND INTERPRETATION**

---

<b>Structure</b>	<b>Page Nos.</b>
4.0 Introduction	
4.1 Objectives	
4.2 Different types of plots	
4.3 Histograms	
4.4 Box plots	
4.5 Scatter plots	
4.6 Heat map	
4.7 Bubble chart	
4.8 Bar chart	
4.9 Distribution plot	
4.10 Pair plot	
4.11 Line graph	
4.12 Pie chart	
4.13 Doughnut chart	
4.14 Area chart	
4.15 Summary	
4.16 Answers	
4.17 References	

---

## **4.0 INTRODUCTION**

---

The previous units of this course covers details on different aspects of data analysis, including the basics of data science, basic statistical concepts related to data science and data pre-processing. This unit explains the different types of plots for data visualization and interpretation. This unit covers the details of the plots for data visualization and further discusses their constructions and discusses the various use cases associated with various data visualization plots. This unit will help you to appreciate the real-world need for a workforce trained in visualization techniques and will help you to design, develop, and interpret visual representations of data. The unit also defines the best practices associated with the construction of different types of plots.

---

## **4.1 OBJECTIVES**

---

After going through this unit, you will be able to:

- Explain the key characteristics of various types of plots for data visualization;
- Explain how to design and create data visualizations;
- Summarize and present the data in meaningful ways;
- Define appropriate methods for collecting, analysing, and interpreting the numerical information.

---

## **4.2 DIFFERENT TYPES OF PLOTS**

---

As more and more data are available to us today, there are several varieties of charts and graphs than before. In reality, the amount of data that we produce, acquire, copy, and use now will be nearly doubled by 2025. Data visualisation is therefore crucial and serves as a powerful tool for organisations. One can benefit from graphs and charts in the following ways:

- Encouraging the group to act proactively.
- Showcasing progress toward the goal to the stakeholders

- Displaying core values of a company or an organization to the audience.

Moreover, data visualisation can bring heterogeneous teams together around new objectives and foster the trust among the team members. Let us discuss about various graphs and charts that can be utilized in expression of various aspects of businesses.

## 4.3 HISTOGRAMS

A histogram visualises the distribution of data across distinct groups with continuous classes. It is represented with set of rectangular bars with widths equal to the class intervals and areas proportional to frequencies in the respective classes. A histogram may hence be defined as a graphic of a frequency distribution that is grouped and has continuous classes. It provides an estimate of the distribution of values, their extremes, and the presence of any gaps or out-of-the-ordinary numbers. They are useful in providing a basic understanding of the probability distribution.

**Constructing a Histogram:** To construct a histogram, the data is grouped into specific class intervals, or “bins” and plotted along the x-axis. These represent the range of the data. Then, the rectangles are constructed with their bases along the intervals for each class. The height of these rectangles is measured along the y-axis representing the frequency for each class interval. It's important to remember that in these representations, every rectangle is next to another because the base spans the spaces between class boundaries.

**Use Cases:** When it is necessary to illustrate or compare the distribution of specific numerical data across several ranges of intervals, histograms can be employed. They can aid in visualising the key meanings and patterns associated with a lot of data. They may help a business or organization in decision-making process. Some of the use cases of histograms include-

- Distribution of salaries in an organisation
- Distribution of height in one batch of students of a class, student performance on an exam,
- Customers by company size, or the frequency of a product problem.

### Best Practices

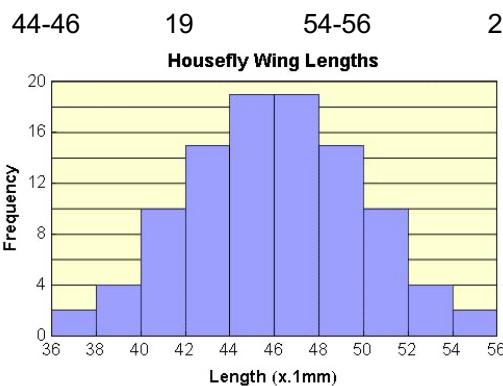
- **Analyse various data groups:** The best data groupings can be found by creating a variety of histograms.
- **Break down compartments using colour:** The same chart can display a second set of categories by colouring the bars that represent each category.

### Types of Histogram

**Normal distribution:** In a normal distribution, the probability that points will occur on each side of the mean is the same. This means that points on either side of the mean could occur.

Example: Consider the following bins shows the frequency of length of wings of housefly in 1/10 of millimetre.

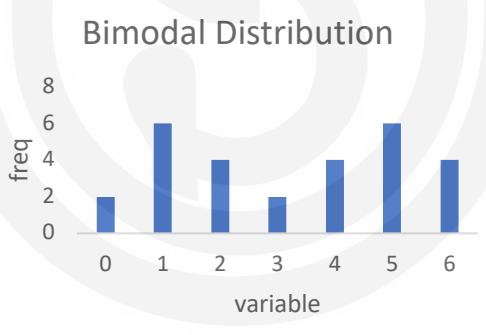
Bin	Frequency	Bin	Frequency
36-38	2	46-48	19
38-40	4	48-50	15
40-42	10	50-52	10
42-44	15	52-54	4



**Bimodal Distribution:** This distribution has two peaks. In the case of a bimodal distribution, the data must be segmented before being analysed as normal distributions in their own right.

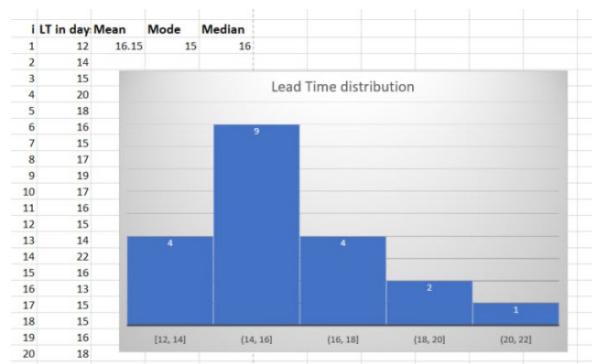
**Example:**

Variable	Frequency
0	2
1	6
2	4
3	2
4	4
5	6
6	4



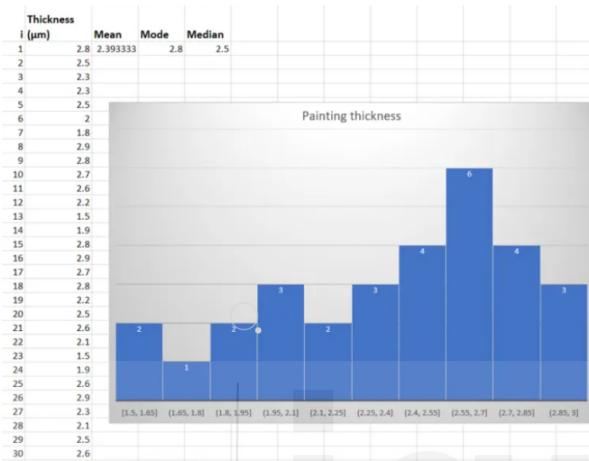
**Right-skewed distribution:** A distribution that is skewed to the right is sometimes referred to as a positively skewed distribution. A right-skewed distribution is one that has a greater percentage of data values on the left and a lesser percentage on the right. Whenever the data have a range boundary on the left side of the histogram, a right-skewed distribution is frequently the result.

**Example:**



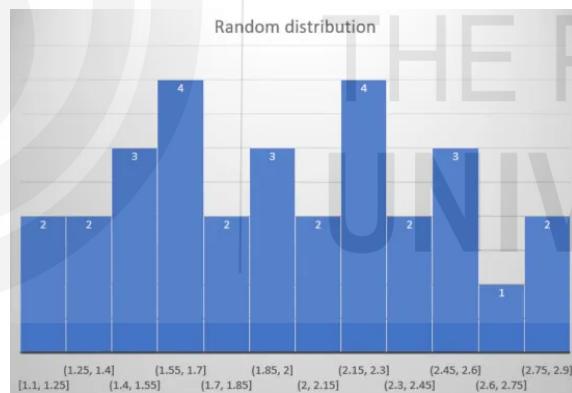
**Left-skewed distribution:** A distribution that is skewed to the left is sometimes referred to as a negatively skewed distribution. A distribution that is left-skewed will have a greater proportion of data values on the right side of the distribution and a lesser proportion of data values on the left. When the data have a range limit on the right side of the histogram, a right-skewed distribution commonly results. An alternative name for this is a right-tailed distribution.

Example:



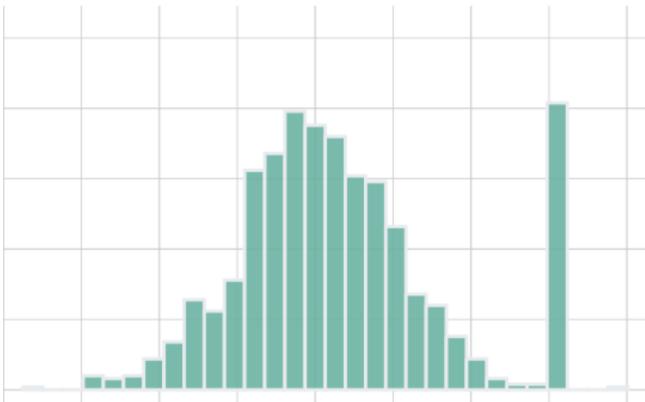
**A random distribution:** A random distribution is characterised by the absence of any clear pattern and the presence of several peaks. When constructing a histogram using a random distribution, it is possible that several distinct data attributes will be blended into one. As a result, the data ought to be partitioned and investigated independently.

Example:



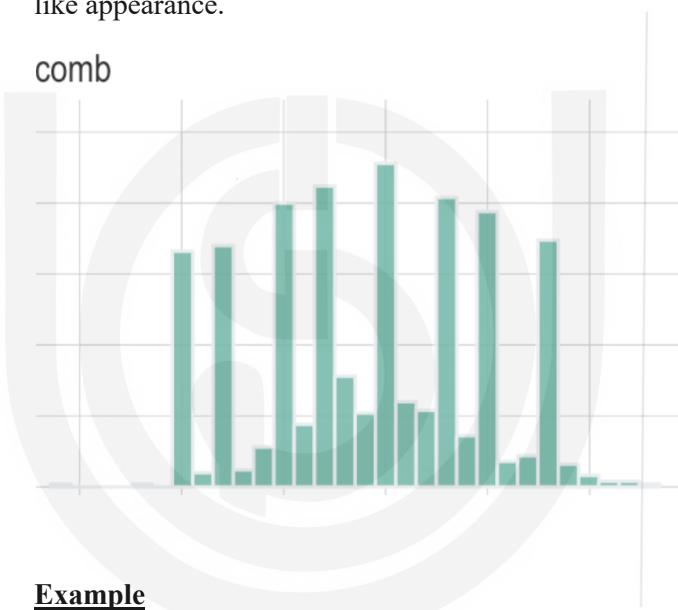
**Edge Peak Distribution:** When there is an additional peak at the edge of the distribution that does not belong there, this type of distribution is called an edge peak distribution. Unless you are very positive that your data set has the expected number of outliers, this almost always indicates that you have plotted (or collected) your data incorrectly (i.e. a few extreme views on a survey).

edge peak



**Comb Distribution:** Because the distribution seems to resemble a comb, with alternating high and low peaks, this type of distribution is given the name "comb distribution." Rounding off an object might result in it having a comb-like form. For instance, if you are measuring the height of the water to the nearest 10 centimetres but your class width for the histogram is only 5 centimetres, you may end up with a comb-like appearance.

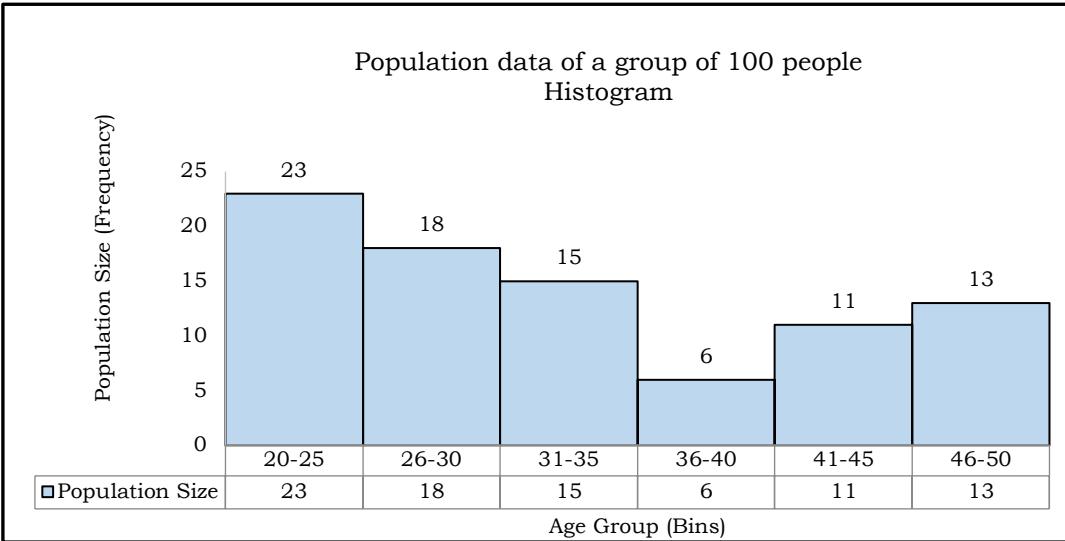
comb



### Example

Histogram for the population data of a group of 86 people:

Age Group (in years)	Population Size
20-25	23
26-30	18
31-35	15
36-40	6
41-45	11
46-50	13
<b>TOTAL</b>	<b>86</b>



### Check Your Progress 1

- What is the difference between a Bar Graph and a Histogram?
- .....
- .....

- Draw a Histogram for the following data:

Class Interval	Frequency
0 – 10	35
10 – 20	70
20 – 30	20
30 – 40	40
40 – 50	50

- Why is histogram used?
- .....
- .....

- What do histograms show?
- .....
- .....

---

## 4.4 BOX PLOTS

---

When displaying data distributions using the five essential summary statistics of minimum, first quartile, median, third quartile, and maximum, box-and-whisker plots, also known as boxplots, are widely employed. It is a visual depiction of data that aids in determining how widely distributed or how much the data values change. These boxplots make it simple to compare the distributions since it makes the centre, spread, and overall range understandable. They are utilised for data analysis wherein the graphical representations are used to determine the following:

1. Shape of Distribution
2. Central Value
3. Variability of Data

**Constructing a Boxplot:** The two components of the graphic are described by their names: the box, which shows the median value of data along with the first and third quartiles (25 percentile and 75 percentile), and the whiskers, which shows the remaining data. The 3rd quartile's difference from the first quartile of data is called the interquartile range. The highest and minimum points in the data can also be displayed using the whiskers. The points beyond  $1.5 \times \text{interquartile range}$  can be identified as suspected outliers.

**Use Cases:** A boxplot is frequently used to demonstrate whether a distribution is skewed and whether the data set contains any potential outliers, or odd observations. Boxplots are also very useful for comparing or involving big data sets. Examples of box plots include plotting the:

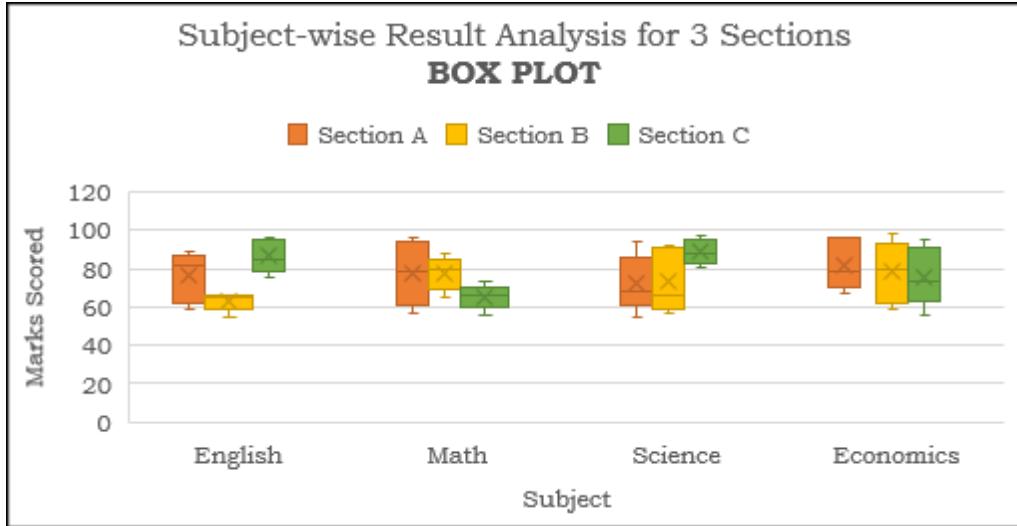
- Gas efficiency of vehicles
- Time spent reading across readers

### **Best Practices**

- **Cover the points within the box:** This aids the viewer in concentrating on the outliers.
- **Box plot comparisons between categorical dimensions:** Box plots are excellent for quickly comparing dataset distributions.

### **Example**

Subject	Section A	Section B	Section C
English	59	65	82
Math	96	73	66
Science	78	57	81
Economics	96	79	73
English	65	55	94
Math	78	65	56
Science	68	61	85
Economics	96	98	56
English	85	63	85
Math	93	88	68
Science	94	66	94
Economics	67	59	86
English	82	66	96
Math	64	79	63
Science	55	90	97
Economics	73	89	95
English	89	66	75
Math	57	81	73
Science	67	92	88
Economics	78	65	69



The boxplots clearly shows that Section B has performed poorly in English, whereas section C has performed poorly in Maths. Section A has mostly balanced performance, but the marks of the students are most dispersed.

### Check Your Progress 2

1. How to correctly interpret a boxplot?
- .....
- .....

2. What are the most important parts of a box plot?
- .....
- .....

3. What is the uses of box plot?
- .....
- .....

4. How do you describe the distribution of a box plot?
- .....
- .....

---

## 4.5 SCATTER PLOTS

---

Scatter plot is the most commonly used chart when observing the relationship between two quantitative variables. It works particularly well for quickly identifying possible correlations between different data points. The relationship between multiple variables can be efficiently studied using scatter plots, which show whether one variable is a good predictor of another or whether they normally fluctuate independently. Multiple distinct data points are shown on a single graph in a scatter plot. Following that, the chart can be enhanced with analytics like trend lines or cluster analysis. It is especially useful for quickly identifying potential correlations between data points.

**Constructing a Scatter Plot:** Scatter plots are mathematical diagrams or plots that rely on Cartesian coordinates. In this type of graph, the categories being compared are represented by the circles on the graph (shown by the colour of the circles) and the

numerical volume of the data (indicated by the circle size). One colour on the graph allows you to represent two values for two variables related to a data set, but two colours can also be used to include a third variable.

**Use Cases:** Scatter charts are great in scenarios where you want to display both distribution and the relationship between two variables.

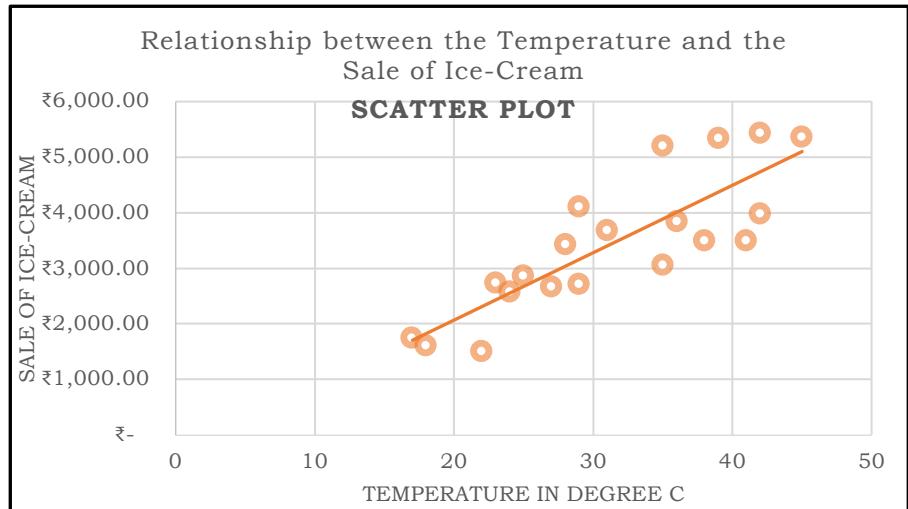
- Display the relationship between time-on-platform (How Much Time Do People Spend on Social Media) and churn (the number of people who stopped being customers during a set period of time).
- Display the relationship between salary and years spent at company

### **Best Practices**

- **Analyze clusters to find segments:** Based on your chosen variables, cluster analysis divides up the data points into discrete parts.
- **Employ highlight actions:** You can rapidly identify which points in your scatter plots share characteristics by adding a highlight action, all the while keeping an eye on the rest of the dataset.
- **mark customization:** individual markings Add a simple visual hint to your graph that makes it easy to distinguish between various point groups.

### **Example**

Temperature (in deg C)	Sale of Ice-Cream
17	₹ 1,750.00
18	₹ 1,603.00
22	₹ 1,500.00
29	₹ 2,718.00
27	₹ 2,667.00
28	₹ 3,422.00
31	₹ 3,681.00
23	₹ 2,734.00
24	₹ 2,575.00
25	₹ 2,869.00
35	₹ 3,057.00
36	₹ 3,846.00
38	₹ 3,500.00
41	₹ 3,496.00
42	₹ 3,984.00
29	₹ 4,109.00
39	₹ 5,336.00
35	₹ 5,197.00
42	₹ 5,426.00
45	₹ 5,365.00



Please note that a linear trendline has been fitted to scatter plot, indicating a positive change in sales of ice-cream with increase in temperature.

### Check Your Progress 3

- What are the characteristics of a scatter plot?
- .....
- .....

- What components make up a scatter plot?
- .....
- .....

- What is the purpose of a scatter plot?
- .....
- .....

- What are the 3 types of correlations that can be inferred from scatter plots?
- .....
- .....

## 4.6 HEAT MAP

Heatmaps are two-dimensional graphics that show data trends through colour shading. They are an example of part to whole chart in which values are represented using colours. A basic heat map offers a quick visual representation of the data. A user can comprehend complex data sets with the help of more intricate heat maps. Heat maps can be presented in a variety of ways, but they all have one thing in common: **they all make use of colour to convey correlations between data values**. Heat maps are more frequently utilised to present a more comprehensive view of massive amounts of data. It is especially helpful because colours are simpler to understand and identify than plain numbers.

Heat maps are highly flexible and effective at highlighting trends. Heatmaps are naturally self-explanatory, in contrast to other data visualisations that require

interpretation. The greater the quantity/volume, the deeper the colour (the higher the value, the tighter the dispersion, etc.). Heat Maps dramatically improve the ability of existing data visualisations to quickly convey important data insights.

**Use Cases:** Heat Maps are primarily used to better show the enormous amounts of data contained inside a dataset and help guide users to the parts of data visualisations that matter most.

- Average monthly temperatures across the years
- Departments with the highest amount of attrition over time.
- Traffic across a website or a product page.
- Population density/spread in a geographical location.

### **Best Practices**

- **Select the proper colour scheme:** This style of chart relies heavily on colour, therefore it's important to pick a colour scheme that complements the data.
- **Specify a legend:** As a related point, a heatmap must typically contain a legend describing how the colours correspond to numerical values.

### **Example**

Region-wise monthly sale of a SKU (stock-keeping unit)

ZONE	MONTH											
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
NORTH	75	84	61	95	77	82	74	92	58	90	54	83
SOUTH	50	67	89	61	91	77	80	72	82	78	58	63
EAST	62	50	83	95	83	89	72	96	96	81	86	82
WEST	69	73	59	73	57	61	58	60	97	55	81	92

The distribution of sales is shown in the sample heatmap above, broken down by zone and spanning a 12-month period. Like in a typical data table, each cell displays a numeric count, but the count is also accompanied by a colour, with higher counts denoting deeper hues.

### **Check Your Progress 4**

1. What type of input is needed for a heat map?

.....

.....

2. What kind of information does a heat map display?

.....

.....

3. What can be seen in heatmap?

.....

.....

---

## **4.7 BUBBLE CHART**

---

Bubble diagrams are used to show the relationships between different variables. They

are frequently used to represent data points in three dimensions, specifically when the bubble size, y-axis, and x-axis are all present. Using location and size, bubble charts demonstrate relationships between data points. However, bubble charts have a restricted data size capability since too many bubbles can make the chart difficult to read.

Although technically not a separate type of visualisation, bubbles can be used to show the relationship between three or more measurements in scatter plots or maps by adding complexity. By altering the size and colour of circles, large amounts of data are presented concurrently in visually pleasing charts.

**Constructing a Bubble Chart:** For each observation of a pair of numerical variables (A, B), a bubble or disc is drawn and placed in a Cartesian coordinate system horizontally according to the value of variable A and vertically according to the value of variable B. The area of the bubble serves as a representation for a third numerical variable (C). Using various colours in various bubbles, you may even add a fourth dataset (D: numerical or categorical).

By using location and proportions, bubble charts are frequently used to compare and illustrate the relationships between circles that have been classified. Bubble Charts' overall image can be utilised to look for patterns and relationships.

**Use Cases:** Usually, the positioning and ratios of the size of the bubbles/circles on this chart are used to compare and show correlations between variables. Additionally, it is utilised to spot trends and patterns in data.

- AdWords' analysis: CPC vs Conversions vs share of total conversions
- Relationship between life expectancy, GD per capita and population size

#### **Best Practices:**

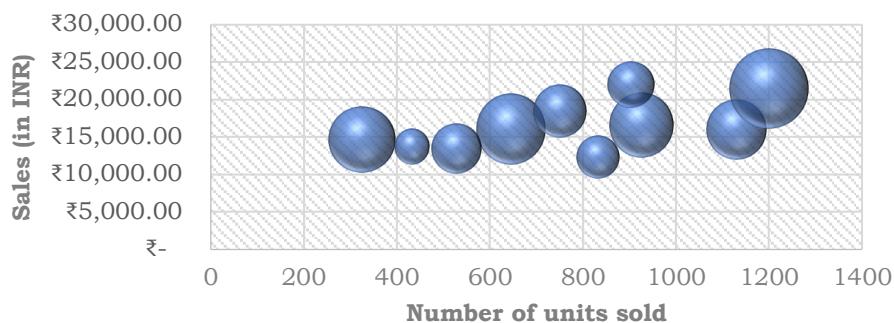
- **Add colour:** A bubble chart can gain extra depth by using colour.
- Set bubble size in appropriate proportion.
- **Overlay bubbles on maps:** From bubbles, a viewer can immediately determine the relative concentration of data. These are used as an overlay to provide the viewer with context for geographically-related data.

#### **Example**

Item Code	Units Sold	Sales (in Rs.)	Profit %
PC001	325	₹ 14,687.00	22%
PC002	1130	₹ 16,019.00	18%
PC003	645	₹ 16,100.00	25%
PC004	832	₹ 12,356.00	9%
PC005	1200	₹ 21,500.00	32%
PC006	925	₹ 16,669.00	21%
PC007	528	₹ 13,493.00	13%
PC008	750	₹ 18,534.00	14%
PC009	432	₹ 13,768.00	6%
PC0010	903	₹ 22,043.00	11%

The three variables in this example are sales, profits, and the number of units sold. Therefore, all three variables and their relationship can be displayed simultaneously using a bubble chart.

### Sales and Profit versus the Quantity sold **BUBBLE CHART**



#### Check Your Progress 5

1. What is bubble chart?

.....

.....

2. What is a bubble chart used for?

.....

.....

3. What is the difference between scatter plot and bubble chart?

.....

.....

4. What is bubble size in bubble chart?

.....

.....

## 4.8 BAR CHART

A bar chart is a graphical depiction of numerical data that uses rectangles (or bars) with equal widths and varied heights. In the field of statistics, bar charts are one of the methods for handling data.

**Constructing a Bar Chart:** The x-axis corresponds to the horizontal line, and the y-axis corresponds to the vertical line. The y-axis represents frequency in this graph. Write the names of the data items whose values are to be noted along the x-axis that is horizontal.

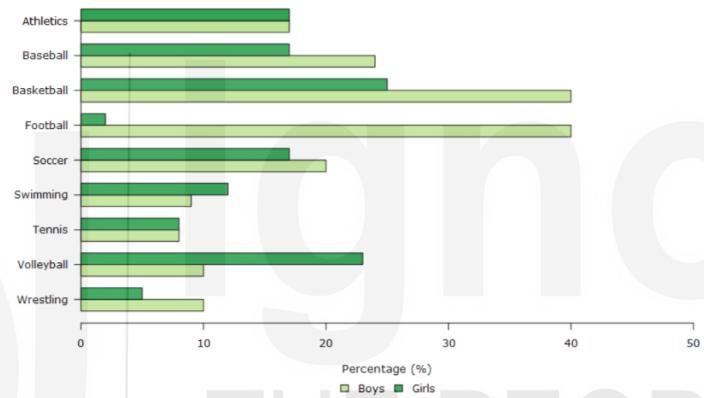
Along the horizontal axis, choose the uniform width of bars and the uniform gap between the bars. Pick an appropriate scale to go along the y-axis that runs vertically so that you can figure out how high the bars should be based on the values that are presented. Determine the heights of the bars using the scale you selected, then draw the bars using that information.

**Types of Bar chart:** Bar Charts are mainly classified into two types:

**Horizontal Bar Charts:** Horizontal bar charts are the type of graph that are used when the data being analysed is to be depicted on paper in the form of horizontal bars with their respective measures. When using a chart of this type, the categories of the data are indicated on the y-axis.

### Example:

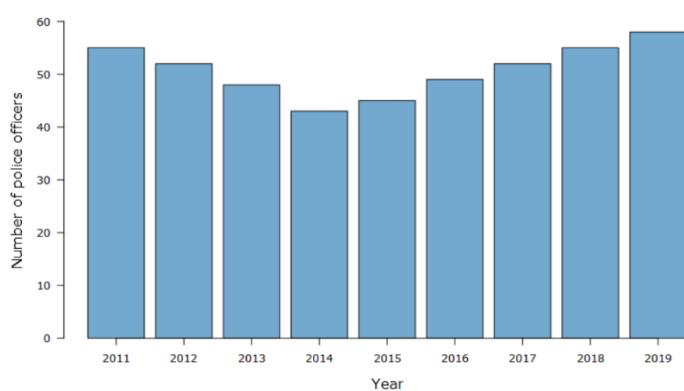
Sport	Percentage of boys (%)	Percentage of girls (%)
Athletics	17	17
Baseball	24	17
Basketball	40	25
Football	40	2
Soccer	20	17
Swimming	9	12
Tennis	8	8
Volleyball	10	23
Wrestling	10	5



**Vertical Bar Charts:** A vertical bar chart displays vertical bars on graph (chart) paper. These rectangular bars in a vertical orientation represent the measurement of the data. The quantities of the variables that are written along the x-axis are represented by these rectangular bars.

### Example:

Year	Number of police officers
2011	55
2012	52
2013	48
2014	43
2015	45
2016	49
2017	52
2018	55
2019	58

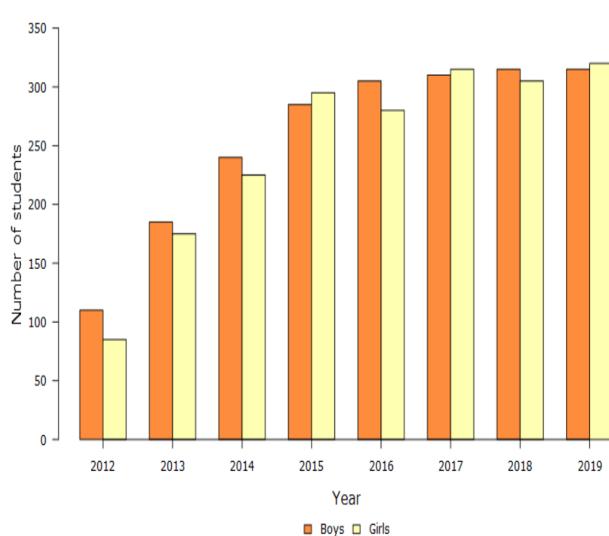


We can further divide bar charts into two basic categories:

**Grouped Bar Charts:** The grouped bar graph is also referred to as the clustered bar graph (graph). It is valuable for at least two separate types of data. The horizontal (or vertical) bars in this are categorised according to their position. If, for instance, the bar chart is used to show three groups, each of which has numerous variables (such as one group having four data values), then different colours will be used to indicate each value. When there is a close relationship between two sets of data, each group's colour coding will be the same.

#### Example:

Year	Number of boys	Number of girls
2012	110	85
2013	185	175
2014	240	225
2015	285	295
2016	305	280
2017	310	315
2018	315	305
2019	315	320

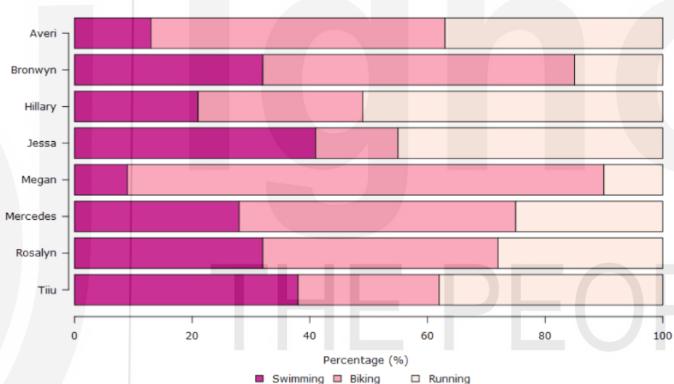


**Stacked Bar Charts:** The composite bar chart is also referred to as the stacked bar chart. It illustrates how the overall bar chart has been broken down into its

component pieces. We utilise bars of varying colours and clear labelling to determine which category each item belongs to. As a result, in a chart with stacked bars, each parameter is represented by a single rectangular bar. Multiple segments, each of a different colour, are displayed within the same bar. The various components of each separate label are represented by the various segments of the bar. It is possible to draw it in either the vertical or horizontal plane.

### Example:

First name	Percentage of time spent swimming (%)	Percentage of time spent cycling (%)	Percentage of time spent running (%)
Averi	13	50	37
Bronwyn	32	53	15
Hillary	21	28	51
Jessa	41	14	45
Megan	9	81	10
Mercedes	28	47	25
Rosalyn	32	40	28
Tiuu	38	24	38



**Use cases:** Bar charts are typically employed to display quantitative data. The following is a list of some of the applications of the bar chart-

- In order to clearly illustrate the relationships between various variables, bar charts are typically utilised. When presented in a pictorial format, the parameters can be more quickly and easily envisioned by the user.
- Bar charts are the quickest and easiest way to display extensive amounts of data while saving time. It is utilised for studying trends over extended amounts of time.

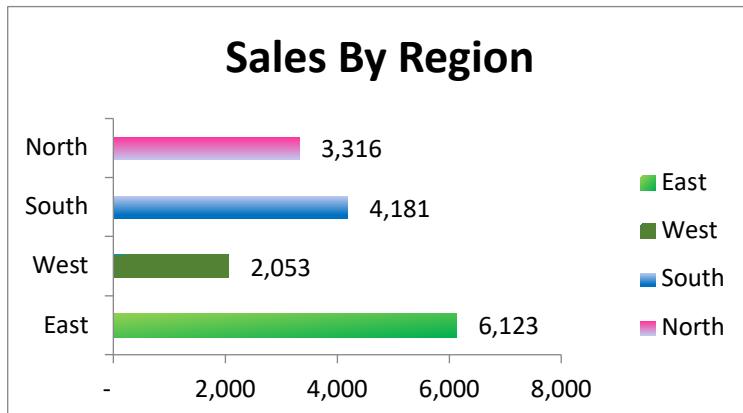
### Best Practices:

- Use a common zero valued baseline
- Maintain rectangular forms for your bars
- Consider the ordering of category level and use colour wisely.

### Example:

Region	Sales
East	6,123
West	2,053

South	4,181
North	3,316



#### Check your progress 6:

1. When should we use bar chart?

.....

.....

2. What are the different types of bar chart?

.....

.....

3. Draw a vertical bar chart.

.....

.....

4. Draw a horizontal bar chart.

Use the following data to answer the question 3 and 4:

Month	January	February	March
Number of visitors	150	300	250

## 4.9 DISTRIBUTION PLOT

Visually assessing the distribution of sample data, distribution charts do this by contrasting the actual distribution of the data with the theoretical values expected from a certain distribution. In addition to more traditional hypothesis tests, distribution plots can be used to establish whether the data from the sample follows a particular distribution. The distribution plot is useful for analysing the relationship between the range of a set of numerical data and its distribution. The values of the data are represented as points along an axis.

**Constructing a Distribution Plot:** You must utilise one or two dimensions, together with one measure, in a distribution plot. You will get a single line visualisation if you only use one dimension. If you use two dimensions, each value of the outer, second dimension will produce a separate line.

**Use Cases:** Distribution of a data set shows the frequency of occurrence of each

possible outcome of a repeatable event observed many times. For instance:

- Height of a population.
- Income distribution in an economy
- Test scores listed by percentile.

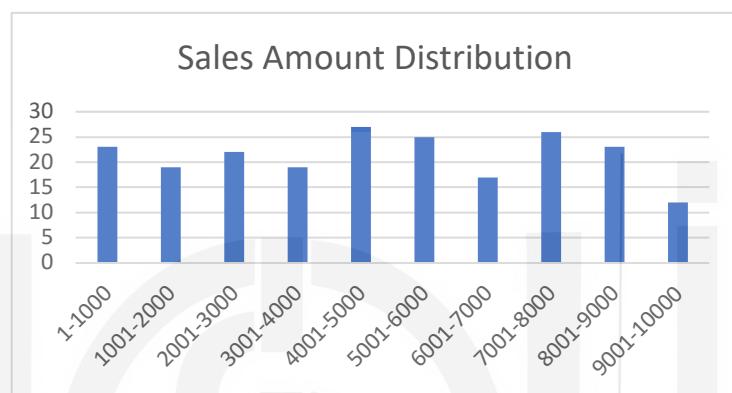
**Best Practices:**

- It is advisable to have equal class widths.
- The class intervals should be mutually exclusive and non-overlapping.
- Open-ended classes at the lower and upper limits (e.g., <10, >100) should be avoided.



## Example

Sales Amount	No. of Clients
1-1000	23
1001-2000	19
2001-3000	22
3001-4000	19
4001-5000	27
5001-6000	25
6001-7000	17
7001-8000	26
8001-9000	23
9001-10000	12
<b>Grand Total</b>	<b>213</b>



### Check your progress 7:

Q.1 What is the distribution plot?

.....  
.....

Q.2 When should we use distribution plot?

.....  
.....

Q.3 What do distribution graphs show?

.....  
.....

## 4.10 PAIR PLOT

The pairs plot is an extension of the histogram and the scatter plot, which are both fundamental figures. The scatter plots on the upper and lower triangles show the relationship (or lack thereof) between two variables. The histogram along the diagonal gives us the ability to see the distribution of a single variable, while the scatter plots on the upper and lower triangles show the relationship (or lack thereof) between two variables.

A pair plot can be utilised to gain an understanding of the optimum collection of characteristics to describe a relationship between two variables or to create clusters that are the most distinct from one another. Additionally, it is helpful to construct some straightforward classification models by drawing some straightforward lines or making linear separations in our data set.

**Constructing a Pair Plot:** If you have m attributes in your dataset, it creates a figure with m x m subplots. Each attribute's univariate histograms (distributions) make up the main-diagonal subplots. For a non-diagonal subplot, assume a position (i, j). The dataset's samples are all plotted using a coordinate system with the characteristics i and j as the axes. In other words, it projects the dataset on these two attributes only. This is particularly interesting to visually inspect how the samples are spread with respect to these two attributes ONLY. The "shape" of the spread can give you valuable insight on the relation between the two attributes.

**Use Cases:** A pairs plot allows us to see both distribution of single variables and relationships between two variables. It helps to identify the most distinct clusters or the optimum combination of attributes to describe the relationship between two variables.

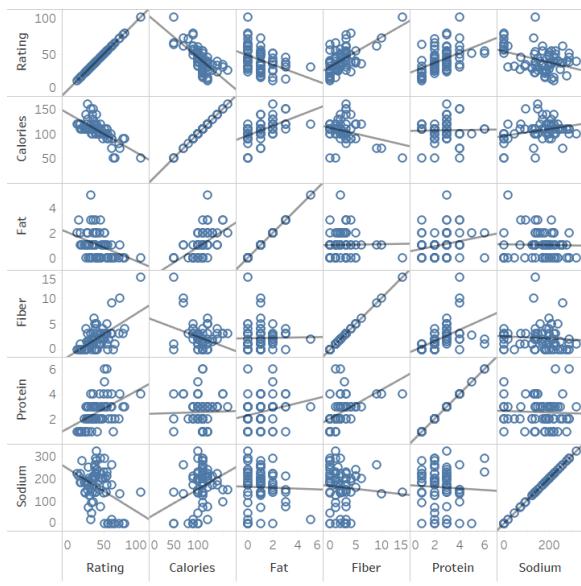
- By creating some straightforward linear separations or basic lines in our data set, it also helps to create some straightforward classification models.
- Analysing socio-economic data of a population.

### Best Practices:

- Use a different colour palette.
- For each colour level, use a different marker.

### Example:

calories	protein	fat	sodium	fiber	rating
70	4	1	130	10	68.40297
120	3	5	15	2	33.98368
70	4	1	260	9	59.42551
50	4	0	140	14	93.70491
110	2	2	180	1.5	29.50954
110	2	0	125	1	33.17409
130	3	2	210	2	37.03856
90	2	1	200	4	49.12025
90	3	0	210	5	53.31381
120	1	2	220	0	18.04285
110	6	2	290	2	50.765
120	1	3	210	0	19.82357
110	3	2	140	2	40.40021
110	1	1	180	0	22.73645
110	2	0	280	0	41.44502
100	2	0	290	1	45.86332
110	1	0	90	1	35.78279
110	1	1	180	0	22.39651
110	3	3	340	4	40.44877
110	2	0	220	1	46.89564
100	2	1	140	2	36.1762
100	2	0	190	1	44.33086
110	2	1	125	1	32.20758
110	1	0	200	1	31.43597
100	3	0	0	3	58.34514
120	3	2	160	5	40.91705
120	3	0	240	5	41.01549
110	1	1	135	0	28.02577
100	2	0	45	0	35.25244
110	1	1	280	0	23.80404
100	3	1	140	3	52.0769
110	3	0	170	3	53.37101
120	3	3	75	3	45.81172
120	1	2	220	1	21.87129
110	3	1	250	1.5	31.07222
110	1	0	180	0	28.74241
110	2	1	170	1	36.52368
140	3	1	170	2	36.47151
110	2	1	260	0	39.24111
100	4	2	150	2	45.32807
110	2	1	180	0	26.73452
100	4	1	0	0	54.85092
150	4	3	95	3	37.13686
150	4	3	150	3	34.13977
160	3	2	150	3	30.31335
100	2	1	220	2	40.10597
120	2	1	190	0	29.92429
140	3	2	220	3	40.69232
90	3	0	170	3	59.64284
130	3	2	170	1.5	30.45084
120	3	1	200	6	37.84059
100	3	0	320	1	41.50354
50	1	0	0	0	60.75611
50	2	0	0	1	63.00565
100	4	1	135	2	49.51187
100	5	2	0	2.7	50.82839
120	3	1	210	5	39.2592
100	3	2	140	2.5	39.7034
90	2	0	0	2	55.33314
110	1	0	240	0	41.99893
110	2	0	290	0	40.56016
80	2	0	0	3	68.23589
90	3	0	0	4	74.47295
90	3	0	0	3	72.80179
110	2	1	70	1	31.23005
110	6	0	230	1	53.13132
90	2	0	15	3	59.36399
110	2	1	200	0	38.83975
140	3	1	190	4	28.59279
100	3	1	200	3	46.65884
110	2	1	250	0	39.10617
110	1	1	140	0	27.7533
100	3	1	230	3	49.78745
100	3	1	200	3	51.59219
110	2	1	200	1	36.18756



The pair plot can be interpreted as follows:

Along the boxes of the diagonal, the variable names are displayed.

A scatterplot of the correlation between each pairwise combination of factors is shown in each of the remaining boxes. For instance, a scatterplot of the values for rating and sodium can be seen in the matrix's box in the top right corner. A scatterplot of values for rating, that is positively connected with rating, and so forth may be seen in the box in the upper left corner. We can see the association between each pair of variables in our dataset from this single visualisation. For instance, calories and rating appear to have a negative link but protein and fat appear to be unrelated.

#### Check your progress 8:

1. Why pair plot is used?

.....

2. How do you read a pairs plot?

.....

3. What does a pair plot show?

.....

.....

---

## 4.11 LINE GRAPH

---

A graph that depicts change over time by means of points and lines is known as a line graph, line chart, or line plot. It is a graph that shows a line connecting a lot of points or a line that shows how the points relate to one another. The graph is represented by the line or curve that connects successive data points to show quantitative data between two variables that are changing. The values of these two variables are compared along a vertical axis and a horizontal axis in linear graphs.

One of the most significant uses of line graphs is tracking changes over both short and extended time periods. It is also used to compare the changes that have taken place for diverse groups over the course of the same time period. It is strongly advised to use a line graph rather than a bar graph when working with data that only has slight

fluctuations.

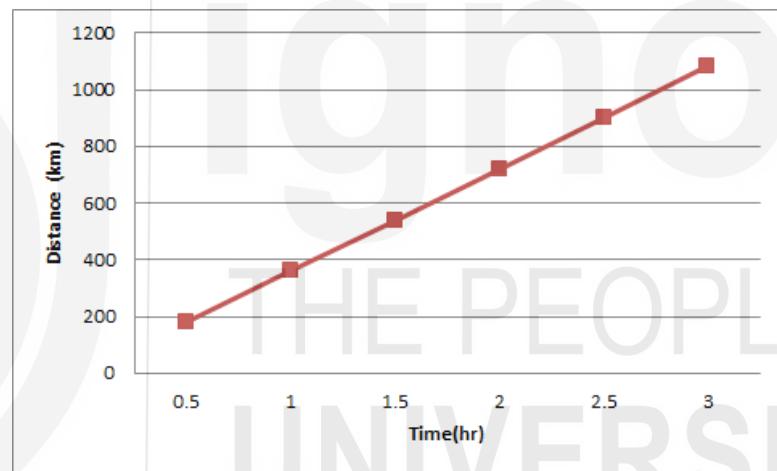
As an illustration, the finance department of a company would want to visualise how its current cash balance has changed over time. If so, they will plot the points over the horizontal and vertical axis using a line graph. It typically refers to the time period that the data span.

Following are the types of line graphs:

1. **Simple Line Graph:** Only a single line is plotted on the graph.

**Example:**

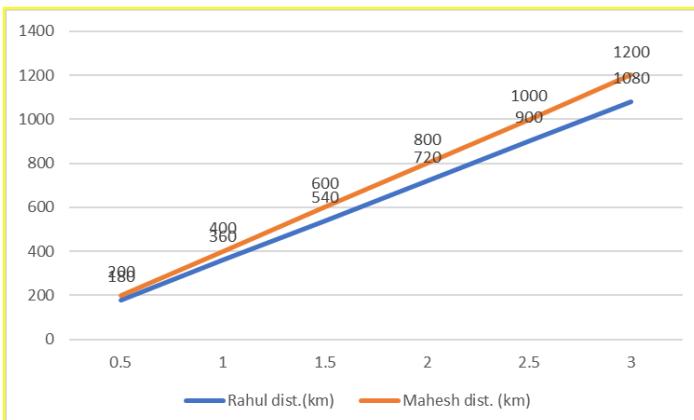
Time (hr)	Distance (km)
0.5	180
1	360
1.5	540
2	720
2.5	900
3	1080



2. **Multiple Line Graph:** The same set of axes is used to plot several lines. An excellent way to compare similar objects over the same time period is via a multiple line graph.

**Example:**

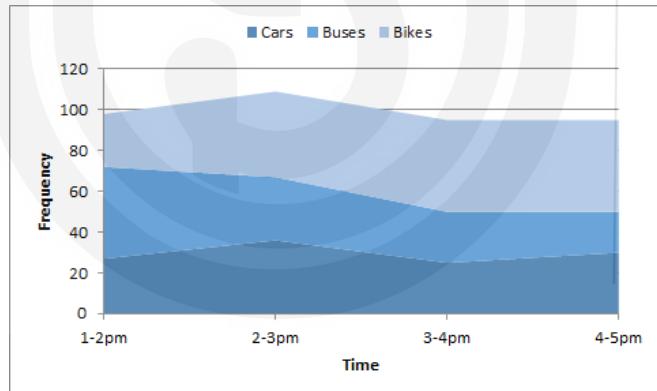
Time(hr)	Rahul dist.(km)	Mahesh dist. (km)
0.5	180	200
1	360	400
1.5	540	600
2	720	800
2.5	900	1000
3	1080	1200



3. **Compound Line Graph:** Whenever one piece of information may be broken down into two or more distinct pieces of data. A compound line graph is the name given to this particular kind of line graph. To illustrate each component that makes up the whole, lines are drawn. The line at the top displays the total, while the line below displays a portion of the total. The size of each component can be determined by the distance that separates every pair of lines.

**Example:**

Time	Cars	Buses	Bikes
1-2pm	37	45	42
2-3pm	44	34	26
3-4pm	23	39	27
4-5pm	29	41	48



**Constructing a line graph:** When we have finished creating the data tables, we will then use those tables to build the linear graphs. These graphs are constructed by plotting a succession of points, which are then connected together with straight lines to offer a straightforward method for analysing data gathered over a period of time. It provides a very good visual format of the outcome data that was gathered over the course of time.

**Use cases:** Tracking changes over both short and long time periods is an important application of line graphs. Additionally, it is utilised to compare changes over the same time period for various groups. Anytime there are little changes, using a line graph rather than a bar graph is always preferable.

- Straight line graphs can be used to explain potential future contract markets and business prospects.
- To determine the precise strength of medications, a straight-line graph is employed in both medicine and pharmacy.
- The government uses straight line graphs for both research and

budgetary planning.

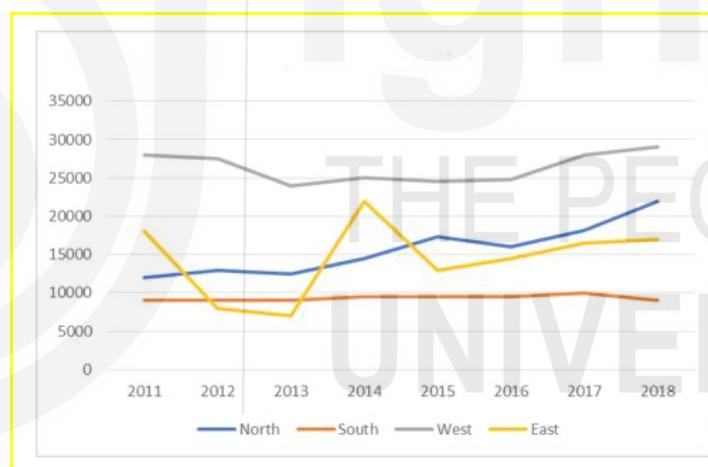
- Chemistry and biology both use linear graphs.
- To determine whether our body weight is acceptable for our height, straight line graphs are employed.

### Best Practices

- Only connecting adjacent values along an interval scale should be done with lines.
- In order to provide correct insights, intervals should be of comparable size.
- Select a baseline that makes sense for your set of data; a zero baseline might not adequately capture changes in the data.
- Line graphs are only helpful for comparing data sets if the axes have the same scales.

### Example:

Sales	2011	2012	2013	2014	2015	2016	2017	2018
North	12000	13000	12500	14500	17300	16000	18200	22000
South	9000	9000	9000	9500	9500	9500	10000	9000
West	28000	27500	24000	25000	24500	24750	28000	29000
East	18000	8000	7000	22000	13000	14500	16500	17000



### Check your progress 9:

Q.1 What is the line graph?

.....  
.....

Q.2 Where can we use line graph?

.....  
.....

Q.3 Draw a line chart from the following information:

A	B	C	D	E	F	G	H	I	J	K	L	M
Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Affordable Segment	173	153	195	147	120	144	148	109	174	130	172	131
Luxury Segment	189	189	105	112	173	109	151	197	174	145	177	161
Super Luxury Segment	185	185	126	134	196	153	112	133	200	145	167	110

## 4.12 PIE CHART

A pie chart, often referred to as a circle chart, is a style of graph that can be used to summarise a collection of nominal data or to show the many values of a single variable (e.g. percentage distribution). Such a chart resembles a circle that has been divided into a number of equal halves. Each segment corresponds to a specific category. The overall size of the circle is divided among the segments in the same proportion as the category's share of the whole data set.

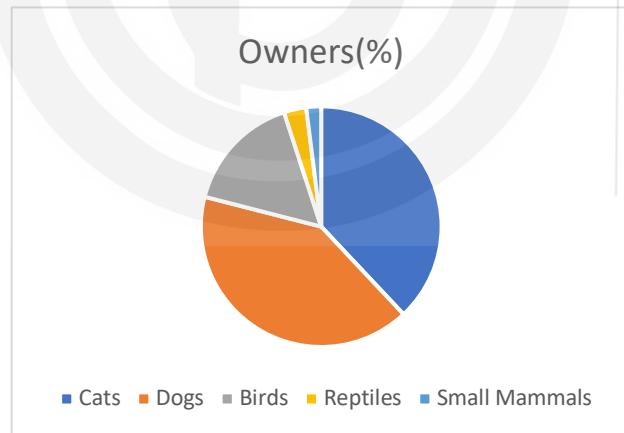
A pie chart often depicts the individual components that make up the whole. In order to bring attention to a particular piece of information that is significant, the illustration may, on occasion, show a portion of the pie chart that is cut away from the rest of the diagram. This type of chart is known as an exploded pie chart.

**Types of a Pie chart:** There are mainly two types of pie charts one is 2D pie chart and another is 3D pie chart. This can be further classified into flowing categories:

- 1. Simple Pie Chart:** The most fundamental kind of pie chart is referred to simply as a pie chart and is known as a simple pie chart. It is an illustration that depicts a pie chart in its most basic form.

**Example:**

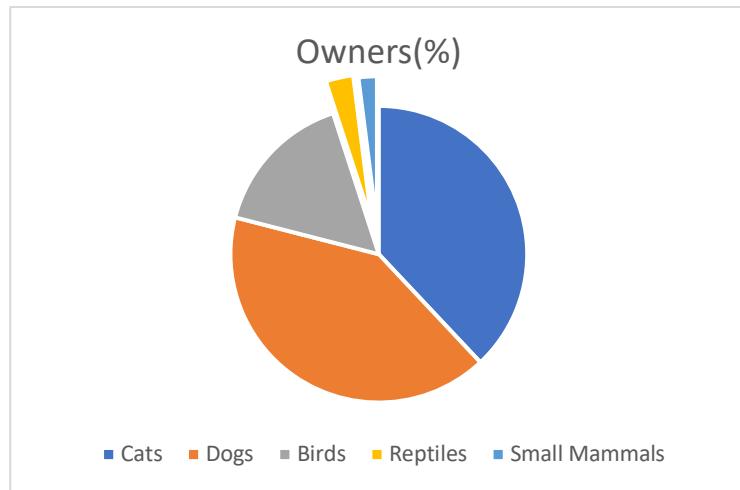
Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



- 2. Exploded Pie Chart:** To create an exploding pie chart, you must first separate the pie from the chart itself, as opposed to merging the two elements together. It is common practise to do this in order to draw attention to a certain section or slice of a pie chart.

**Example:**

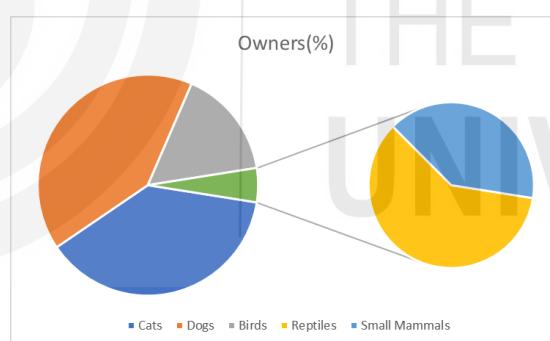
Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



**3. Pie of Pie:** The pie of pie method is a straightforward approach that enables more categories to be represented on a pie chart without producing an overcrowded and difficult-to-read graph. A pie chart that is generated from an already existing pie chart is referred to as a "pie of pie".

Example:

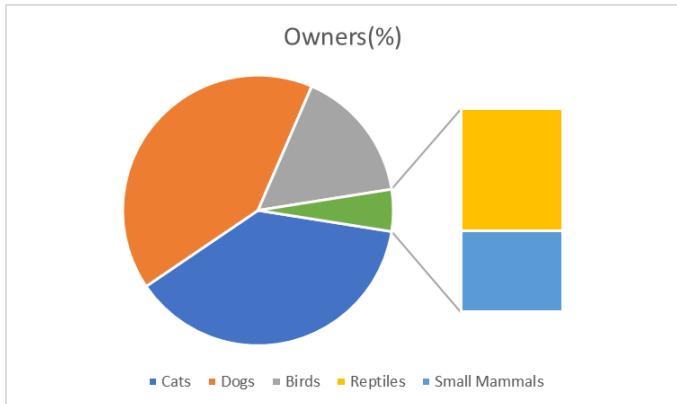
Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



- 3. Bar of Pie:** A bar of pie is an additional straightforward method for showing additional categories on a pie chart while minimising space consumption on the pie chart itself. The expansion that was developed from the already existing pie chart was a bar graph rather than a pie of pie, despite the fact that both serve comparable objectives.

Example:

Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



ignou  
THE PEOPLE'S  
UNIVERSITY

**Constructing a Pie chart:** “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

Therefore, the pie chart formula is given as  $(\text{Given Data}/\text{Total value of Data}) \times 360^\circ$

**Use cases:** If you want your audience to get a general idea of the part-to-whole relationship in your data, and comparing the exact sizes of the slices is not as critical to you, then you should use pie charts. And indicate that a certain portion of the whole is disproportionately small or large.

- Voting preference by age group
- Market share of cloud providers

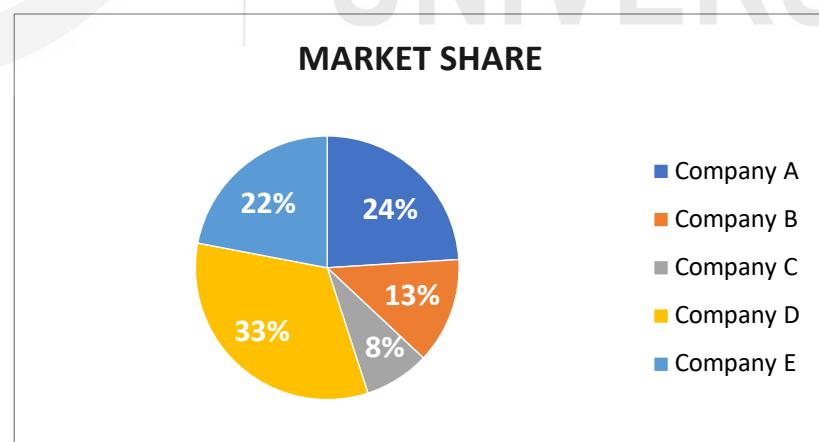
### Best Practices

- **Fewer pie wedges are preferred:** The observer may struggle to interpret the chart's significance if there are too many proportions to compare. Similar to this, keep the overall number of pie charts on dashboards to a minimum.

**Overlay pies on maps:** Pie charts can be used to further deconstruct geographic tendencies in your data and produce an engaging display.

### Example

COMPANY	MARKET SHARE
Company A	24%
Company B	13%
Company C	8%
Company D	33%
Company E	22%



### Check your progress 10:

Q1. What is the pie chart?

.....

.....

Q2. What are the different type of pie charts?

.....  
.....

Q.3 Draw a pie chart from the following information:

**Table: Favorite Type of Movie**

Comedy	Action	Romance	Drama	SciFi
4	5	6	1	4

---

## 4.13 DOUGHNUT CHART

---

Pie charts have been superseded by a more user-friendly alternative called a doughnut chart, which makes reading pie charts much simpler. It is recognised that these charts express the relationship of 'part-to-whole,' which is when all of the parts represent one hundred percent when collected together. It presents survey questions or data with a limited number of categories for making comparisons.

In comparison to pie charts, they provide for more condensed and straightforward representations. In addition, the center hole can be used to assist in the display of relevant information. You might use them in segments, where each arc would indicate a proportional value associated with a different piece of data.

**Constructing a Doughnut chart:** A doughnut chart, like a pie chart, illustrates the relationship of individual components to the whole, but unlike a pie chart, it can display more than one data series at the same time. A ring is added to a doughnut chart for each data series that is plotted within the chart itself. The beginning of the first data series can be seen near the middle of the chart. A specific kind of pie chart called a doughnut chart is used to show the percentages of categorical data. The amount of data that falls into each category is indicated by the size of that segment of the donut. The creation of a donut chart involves the use of a string field and a number, count of features, or rate/ratio field.

There are two types of doughnut chart one is normal doughnut chart and another is exploded doughnut chart. Exploding doughnut charts, much like exploded pie charts, highlight the contribution of each value to a total while emphasising individual values. However, unlike exploded pie charts, exploded doughnut charts can include more than one data series.

**Use cases:** Doughnut charts are good to use when comparing sets of data. By using the size of each component to reflect the percentage of each category, they are used to display the proportions of categorical data. A string field and a count of features, number, rate/ratio, or field are used to make a doughnut chart.

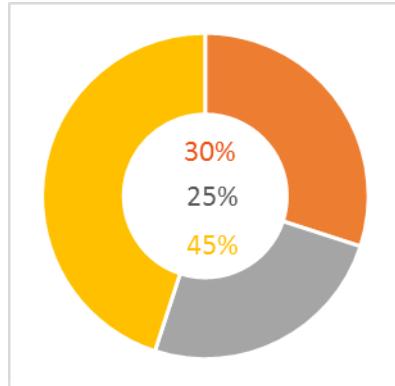
- Android OS market share
- Monthly sales by channel

### Best Practices

- Stick to five slices or less because thinner and long-tail slices become unreadable and uncomparable.
- Use this chart to display one point in time with the help of the filter legend.
- Well-formatted and informative labels are essential because the information conveyed by circular shapes alone is not enough and is imprecise.
- It is a good practice to sort the slices to make it more clear for comparison.

### Example:

Project Status	
Completed	30%
Work in progress	25%
Incomplete	45%



### Check your progress 11:

Q1. What is the doughnut chart?

.....  
.....

Q2 What distinguishes a doughnut chart from a pie chart?

.....  
.....

Q3 Draw a doughnut chart from the following information:

Product	2020	2021
x	40	50
y	30	60
z	60	70

---

## 4.14 AREA CHART

---

An area chart, a hybrid of a line and bar chart, shows the relationship between the numerical values of one or more groups and the development of a second variable, most often the passage of time. The inclusion of shade between the lines and a baseline, similar to a bar chart's baseline, distinguishes a line chart from an area chart. An area chart has this as its defining feature.

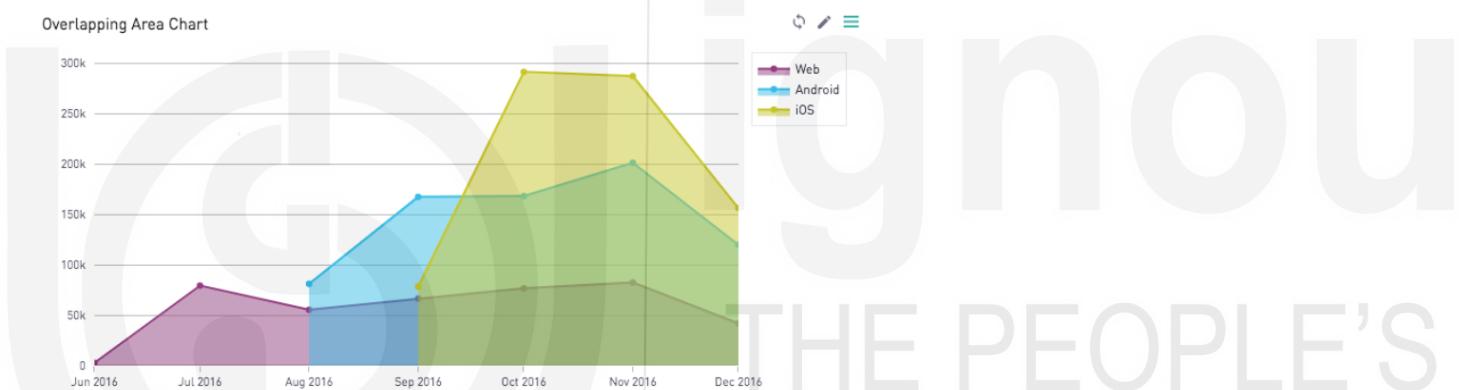
### Types of Area Chart:

**Overlapping area chart:** An overlapping area chart results if we wish to look at how the values of the various groups compare to one another. The conventional line chart serves as the foundation for an overlapping area chart. One point is plotted for each group at each of the horizontal values, and the height of the point indicates the group's value on the vertical axis variable.

All of the points for a group are connected from left to right by a line. A zero baseline is supplemented by shading that is added by the area chart between each line. Because the shading for different groups will typically overlap to some degree, the shading itself incorporates a degree of transparency to ensure that the lines delineating each group may be seen clearly at all times.

The shading brings attention to group that has the highest value by highlighting group's pure hue. Take care that one series is not always higher than the other, as this could cause the plot to become confused with the stacked area chart, which is the other form of area chart. In circumstances like these, the most prudent course of action will consist of sticking to the traditional line chart.

Months (2016)	Web	Android	iOS
June	0	-	
July	70k	-	
Aug	55k	80k	
Sep	60k	165k	80k
Oct	70k	165k	295k
Nov	80k	200k	290k
Dec	40k	125k	155k

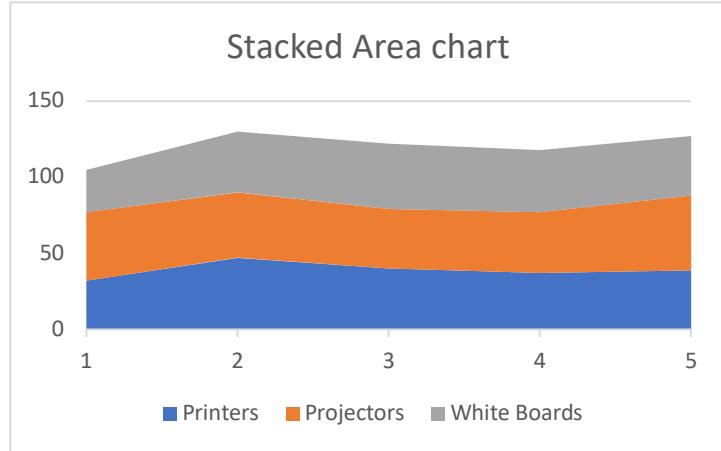


**Stacked area chart:** The stacked area chart is what is often meant to be conveyed when the phrase "area chart" is used in general conversation. When creating the chart of overlapping areas, each line was tinted based on its vertical value all the way down to a shared baseline. Plotting lines one at a time creates the stacked area chart, which uses the height of the most recent group of lines as a moving baseline. Therefore, the total that is obtained by adding up all of the groups' values will correspond to the height of the line that is entirely piled on top.

When you need to keep track of both the total value and the breakdown of that total by groups, you should make use of a stacked area chart. This type of chart will allow you to do both at the same time. By contrasting the heights of the individual curve segments, we are able to obtain a sense of how the contributions made by the various subgroups stack up against one another and the overall sum.

#### Example:

A	B	C	D
	Printers	Projectors	White Boards
2017	32	45	28
2018	47	43	40
2019	40	39	43
2020	37	40	41
2021	39	49	39



**Use Cases:** In most cases, many lines are drawn on an area chart in order to create a comparison between different groups (also known as series) or to illustrate how a whole is broken down into its component pieces. This results in two distinct forms of area charts, one for each possible application of the chart.

- **Magnitude of a single quantitative variable's trend** - An increase in a public company's revenue reserves, programme enrollment from a qualified subgroup by year, and trends in mortality rates over time by primary causes of death are just a few examples.
- **Comparison of the contributions made by different category members (or group)**- the variation in staff sizes among departments, or support tickets opened for various problems.
- Birth and death rates over time for a region, the magnitudes of cost vs. revenue for a business, the magnitudes of export vs. import over time for a country

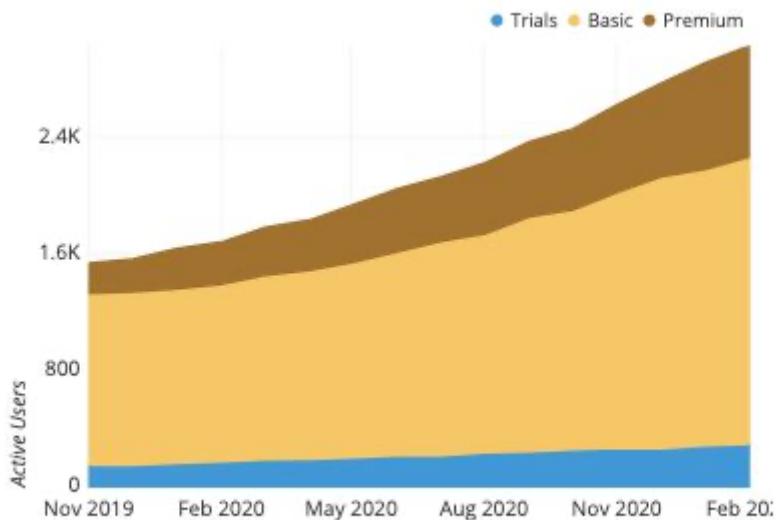
#### Best Practices:

- To appropriately portray the proportionate difference in the data, start the y-axis at 0.
- To boost readability, choose translucent, contrasting colours.
- Keep highly variable data at the top of the chart and low variable data at the bottom during stacking.
- If you need to show how each value over time contributes to a total, use a stacked area chart.
- However, it is recommended to utilise 100% stacked area charts if you need to demonstrate a part to whole relationship in a situation where the cumulative total is unimportant.

#### Example:

### Example of data structure

MONTH	TRIALS	BASIC	PREMIUM
2019-11	154	1180	201
2019-12	157	1186	219
2020-01	170	1195	270
2020-02	180	1213	285



The above Stacked area chart is belonging to tele-service offered by various television based applications. In this data, there are different type of subscribers who are using the services provided by tele-applications in different months.

#### **Check your progress 12:**

Q1. What is area chart?

.....  
.....

Q2. What are types of area charts?

.....  
.....

Q3. Draw an area chart from the following information:

	Product A	Product B	Product C
2017	2000	600	75
2018	2200	450	85
2019	2100	500	125
2020	3000	750	123

---

## **4.15 SUMMARY**

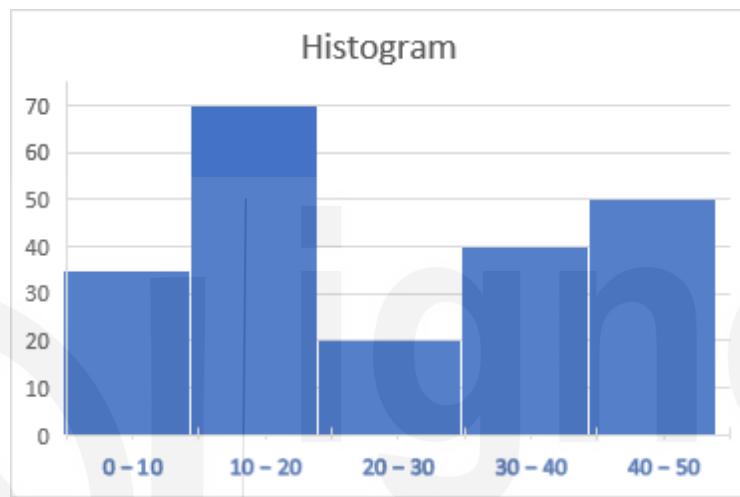
This Unit introduces you to some of the basic charts that are used in data science. The Unit defines the characteristics of Histograms, which are very popular in univariate frequency analysis of quantitative variables. It then discusses the importance and various terms used in the box plots, which are very useful while comparing quantitative variable over some qualitative characteristic. Scatter plots are used to visualise the relationships between two quantitative variables. The Unit also discusses about the heat map, which are excellent visual tools for comparing values. In case three variables are to be compared then you may use bubble charts. The unit also highlights the importance of bar charts, distribution plots, pair plots and line graphs. In addition, it highlights the importance of Pie chart, doughnut charts and area charts for visualising different kinds of data. In addition, there are many different kinds of charts that are used in different analytical tool. You may read about them from refferences.

## 4.16 ANSWERS

### Check Your Progress 1

i. A bar graph is a pictorial representation using vertical and horizontal bars in a graph. The length of bars are proportional to the measure of data. It is also called bar chart. A histogram is also a pictorial representation of data using rectangular bars, that are adjacent to each other. It is used to represent grouped frequency distribution with continuous classes.

ii.



- iii. It is used to summarise continuous or discrete data that is measured on an interval scale. It is frequently used to conveniently depict the key characteristics of the data distribution.
- iv. A histogram is a graphic depiction of data points arranged into user-specified ranges. The histogram, which resembles a bar graph in appearance, reduces a data series into an intuitive visual by collecting numerous data points and organising them into logical ranges or bins.

### Check Your Progress 2

1. Follow these instructions to interpret a boxplot. :

Step 1: Evaluate the major characteristics. Look at the distribution's centre and spread. Examine the potential impact of the sample size on the boxplot's visual appeal.

Step 2: Search for signs of anomalous or out-of-the-ordinary data. Skewed data suggest that data may not be normal. Other situations in your data may be indicated by outliers.

Step 3: Evaluate and compare groups. Evaluate and compare the centre and spread of groups if your boxplot contains them.

2. A boxplot is a common method of showing data distribution based on a five-number summary ("minimum," first quartile ("Q1"), median ("Q3"), and "maximum"). You can learn more about your outliers' values from it.
3. Box plots are generally used for 3 purposes -
  - Finding outliers in the data
  - Finding the dispersion of data from a median
  - Finding the range of data
4. The box plot distribution will reveal the degree to which the data are clustered, how skewed they are, and also how symmetrical they are.
  - Positively Skewed: The box plot is positively skewed if the distance from the median to the maximum is greater than the distance from the median to the minimum.
  - Negatively skewed: Box plots are said to be negatively skewed if the distance from the median to the minimum is higher than the distance from the median to the maximum.
  - Symmetric: When the median of a box plot is equally spaced from both the maximum and minimum values, the box plot is said to be symmetric.

### Check Your Progress 3

1.
  - The most practical method for displaying bivariate (2-variable) data is a scatter plot.
  - A scatter plot can show the direction of a relationship between two variables when there is an association or interaction between them (positive or negative).
  - The linearity or nonlinearity of an association or relationship can be ascertained using a scatter plot.
  - A scatter plot reveals anomalies, questionably measured data, or incorrectly plotted data visually.
2.
  - The Title- A brief description of what is in your graph is provided in the title.
  - The Legend- The meaning of each point is explained in the legend.
  - The Source- The source explains how you obtained the data for your graph.
  - Y-Axis.
  - The Data.
  - X-Axis.
3. A scatter plot is composed of a horizontal axis containing the measured values of one variable (independent variable) and a vertical axis representing the measurements of the other variable (dependent variable). The purpose of the scatter plot is to display what happens to one variable when another variable is changed.
4.
  - Positive Correlation.
  - Negative Correlation.
  - No Correlation (None)

### Check Your Progress 4

1. Three main types of input exist to plot a heatmap: wide format, correlation matrix, and long format.

**Wide format:** The wide format (or the untidy format) is a matrix where each row is an individual, and each column is an observation. In this case, the heatmap makes a visual representation of the matrix: each square of the heatmap represents a cell. The color of the cell changes according to its value.

**Correlation matrix:** Suppose you measured several variables for n individuals. A common task is to check if some variables are correlated. You can easily calculate the correlation between each pair of variables, and plot this as a heatmap. This lets you discover which variable is related to the other.

**Long format:** In the tidy or long format, each line represents an observation. You have 3 columns: individual, variable name, and value (x, y and z). You can plot a heatmap from this kind of data.

2. A heat map is a two-dimensional visualisation of data in which colours stand in for values. A straightforward heat map offers a quick visual representation of the data. The user can comprehend complex data sets with the help of more intricate heat maps.
3. Using one variable on each axis, heatmaps are used to display relationships between two variables. You can determine if there are any trends in the values for one or both variables by monitoring how cell colours vary across each axis.

### Check Your Progress 5

1. A bubble chart is a variant of a scatter chart in which the data points are swapped out for bubbles, with the size of the bubbles serving as a representation of an additional dimension of the data. A bubble chart horizontal and vertical axes are both value axes.
2. To identify whether at least three numerical variables are connected or exhibit a pattern, bubble charts are utilised. They could be applied in specific situations to compare categorical data or demonstrate trends across time.
3. In scatter charts, one numeric field is displayed on the x-axis and another on the y-axis, making it simple to see the correlation between the two values for each item in the chart. A third numerical field in a bubble chart regulates the size of the data points.
4. Any bubbles between 0 and 5 pts on this scale will appear at 5 pt, and all the bubbles on your chart will be between 5 and 20 pts. To construct a chart that displays many dimensions, combine bubble size with colour by value.

### Check Your Progress 6

**Answer 1:**

In the process of statistics development, bar charts are typically employed to display the data. The following is a list of some of the applications of the bar chart:

To clearly illustrate the relationships between various variables, bar charts are typically utilised. When presented in a pictorial format, the parameters can be more quickly and easily envisioned by the user.

Bar charts are the quickest and easiest way to display extensive amounts of data while also saving time.

The method of data representation that is most commonly utilised. As a result, it is utilised in a variety of different sectors.

When studying trends over extended amounts of time, it is helpful to have this information.

**Answer 2:**

Charts are primarily divided into two categories:

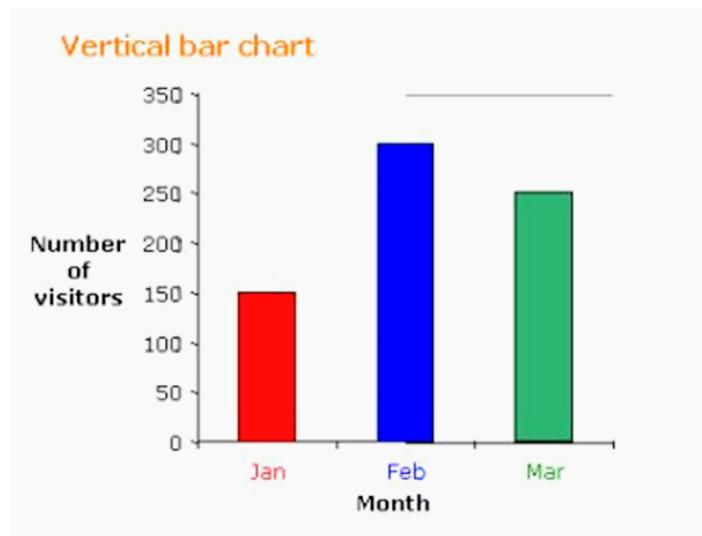
Horizontal Bar Charts:

Vertical Bar Charts

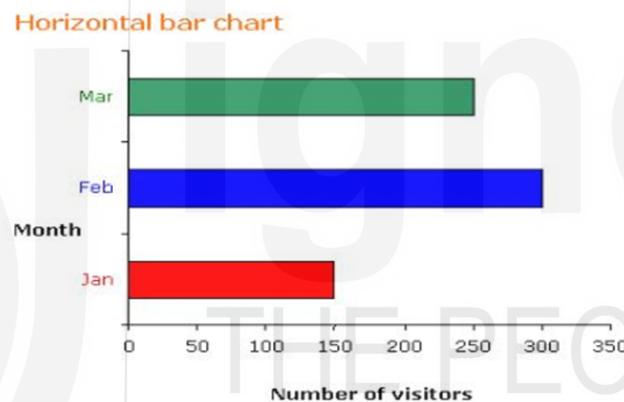
We can further divide into two types:

Grouped Bar Charts  
Stacked Bar Charts

### **Answer 3:**



### **Answer4:**



### **Check Your Progress 7:**

1. For visually assessing the distribution of sample data, you can draw distribution charts. Using these charts, you can contrast the actual distribution of the data with the theoretical values expected from a certain distribution.
2. The distribution plot is useful for analysing the relationship between the range of a set of numerical data and its distribution. You are only allowed to use one or two dimensions and one measure when creating a distribution graphic.
3. These graphs show - how the data is distributed; how the data is composed; how values relate to one another.

### **Check Your Progress 8:**

1. We can visualise pairwise relationships between variables in a dataset using pair plots. By condensing a lot of data into a single figure, this gives the data a pleasant visual representation and aids in our understanding of the data.
2. A scatter plot of a and b, one of a and c, and finally one of a and d are shown in the first line. b and a (symmetric to the first row) are in the second row, followed by b and c, b and d, and so on. In pairs, no sums, mean squares,

or other calculations are performed. That is in your data frame if you discover it in your pairings plot.

3. Pair plots are used to determine the most distinct clusters or the best combination of features to describe a connection between two variables. By creating some straightforward linear separations or basic lines in our data set, it also helps to create some straightforward classification models.

### Check Your Progress 9:

1. A graph that depicts change over time by means of points and lines is known as a line graph, line chart, or line plot. It is a chart that depicts a line uniting numerous points or a line that illustrates the relation between the points. The line or curve used to depict quantitative data between two changing variables in the graph combines a sequence of succeeding data points to create a representation of the graph.

2. Tracking changes over a short as well as a long period of time is one of the most important applications of line graphs. Additionally, it is utilised to compare the modifications that have occurred for various groups throughout the course of the same period of time. When dealing with data that has only minor variations, using a line graph rather than a bar graph is strongly recommended. For instance, the finance team at a corporation may wish to chart the evolution of the cash balance that the company now possesses throughout the course of time.

3.



### Check Your Progress 10:

1. A pie chart, often referred to as a circle chart, is a style of graph that can be used to summarise a collection of nominal data or to show the many values of a single variable. (e.g. percentage distribution).

2. There are mainly two types of pie charts one is 2D pie chart and another is 3D pie chart. This can be further classified into flowing categories:

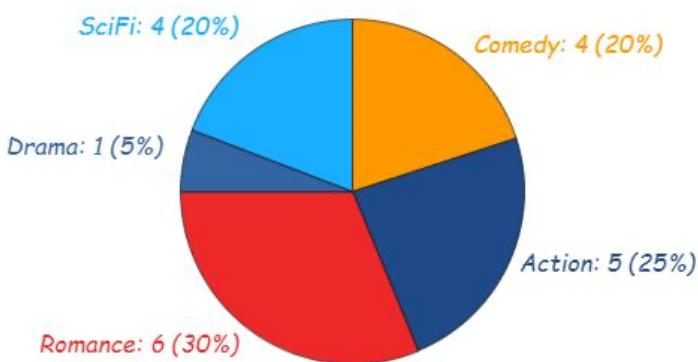
1. Simple Pie Chart

2. Exploded Pie Chart

3. Pie of Pie

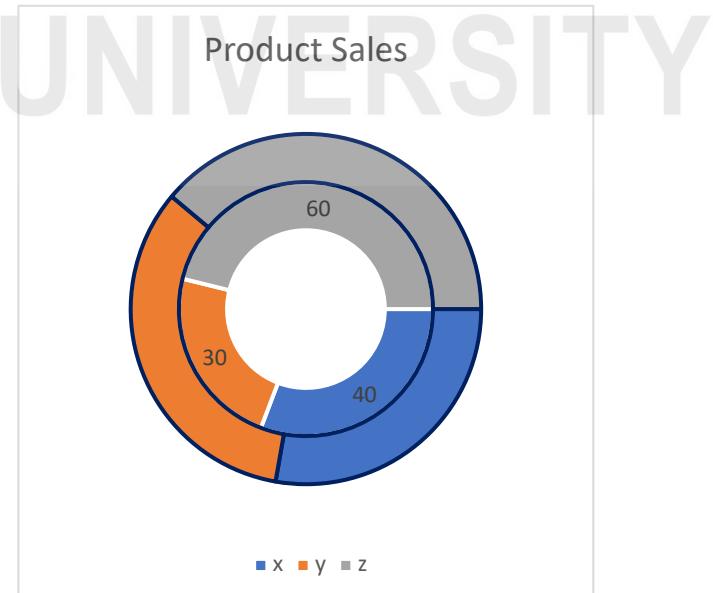
#### 4. Bar of Pie

3.



#### Check Your Progress 11:

1. Pie charts have been superseded by a more user-friendly alternative called a doughnut chart, which makes reading pie charts much simpler. It is recognised that these charts express the relationship of 'part-to-whole,' which is when all of the parts represent one hundred percent when collected together. In comparison to pie charts, they provide for more condensed and straightforward representations.
2. A donut chart is similar to a pie chart, with the exception that the centre is cut off. When you want to display particular dimensions, you use arc segments rather than slices. Just like a pie chart, this form of chart can assist you in comparing certain categories or dimensions to the greater overall; nevertheless, it has a few advantages over its pie chart counterpart.
- 3.



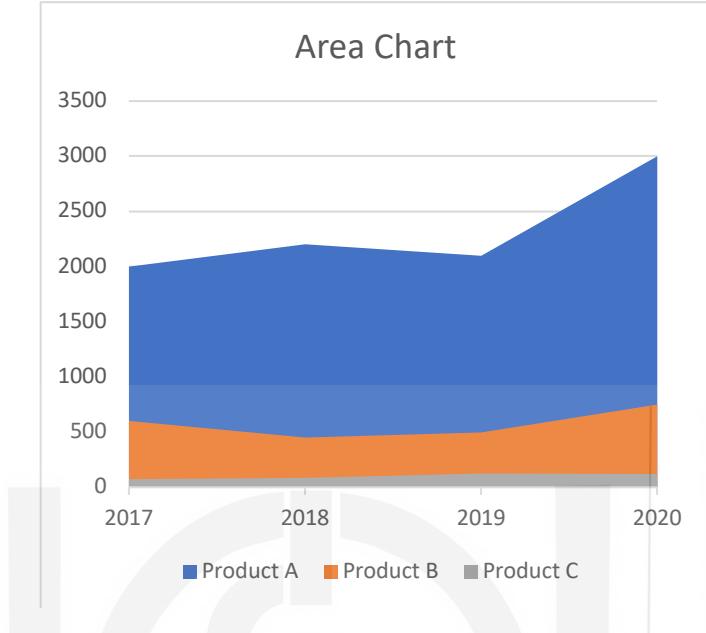
#### Check Your Progress 12:

1. An area chart shows how the numerical values of one or more groups change in proportion to the development of a second variable, most frequently the passage of time.

It combines the features of a line chart and a bar chart. A line chart can be differentiated from an area chart by the addition of shading between the lines and a baseline, just like in a bar chart. This is the defining characteristic of an area chart.

## 2. Overlapping area chart and Stacked area chart

3.



## 4.17 REFERENCES

- Useful Ways to Visualize Your Data (With Examples). Pdf
- Data Visualization Cheat Sheet. Pdf
- Which chart or graph is right for you? Pdf
- <https://www.excel-easy.com/examples/frequency-distribution.html>
- <https://analyticswithlohr.com/2020/09/15/556/>
- <https://www.fusioncharts.com/line-charts>
- <https://evolytics.com/blog/tableau-201-make-stacked-area-chart/>
- <https://chartio.com/learn/charts/area-chart-complete-guide/>
- <https://www.lifewire.com/exploding-pie-charts-in-excel-3123549>
- <https://www.formpl.us/resources/graph-chart/line/>
- <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch9/bargraph-diagrammebarres/5214818-eng.htm>
- <https://sixsigmamania.com/?p=475>
- <https://study.com/academy/lesson/measures-of-dispersion-variability-and-skewness.html>