

fluctuations.

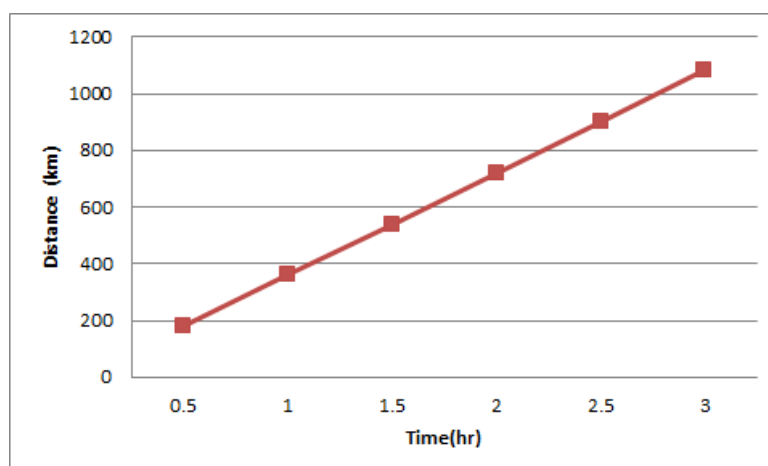
As an illustration, the finance department of a company would want to visualise how its current cash balance has changed over time. If so, they will plot the points over the horizontal and vertical axis using a line graph. It typically refers to the time period that the data span.

Following are the types of line graphs:

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

Example:

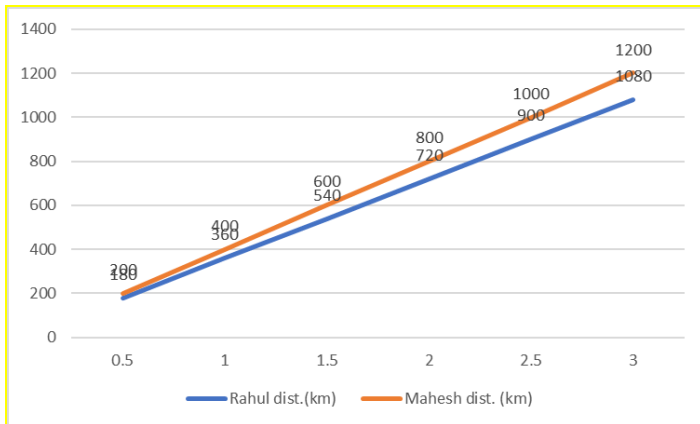
Time (hr)	Distance (km)
0.5	180
1	360
1.5	540
2	720
2.5	900
3	1080



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:

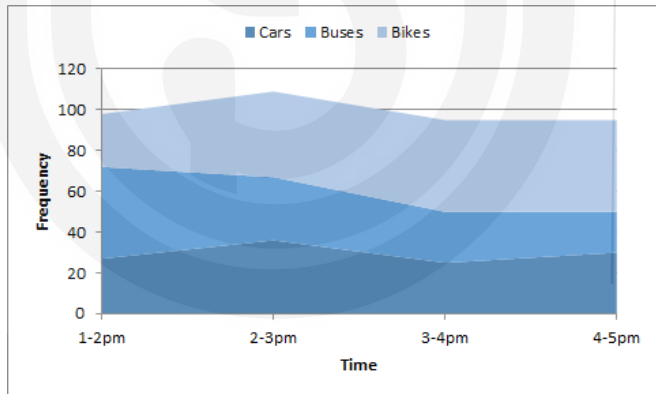
Time(hr)	Rahul dist.(km)	Mahesh dist. (km)
0.5	180	200
1	360	400
1.5	540	600
2	720	800
2.5	900	1000
3	1080	1200



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.

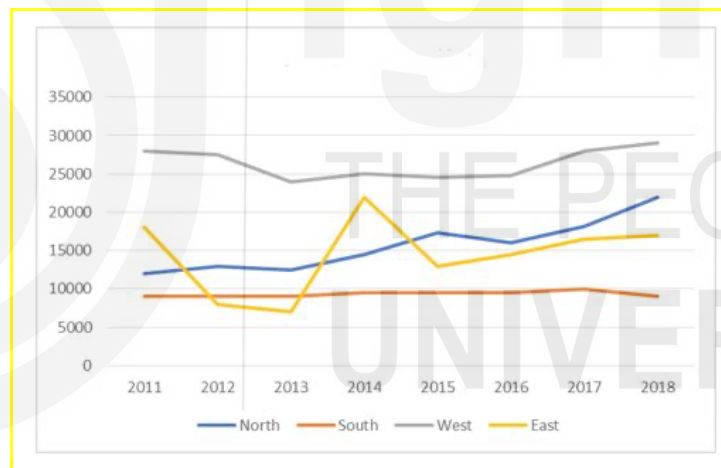
- Chemistry and biology both use linear graphs.
- To determine whether our body weight is acceptable for our height, straight line graphs are employed.

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.

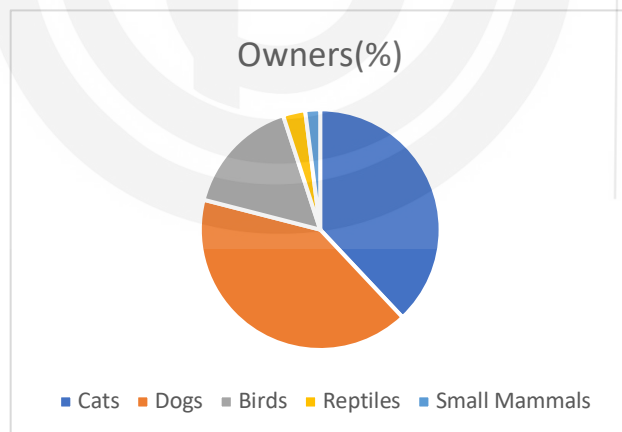
A pie chart often depicts the individual components that make up the whole. In order to bring attention to a particular piece of information that is significant, the illustration may, on occasion, show a portion of the pie chart that is cut away from the rest of the diagram. This type of chart is known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

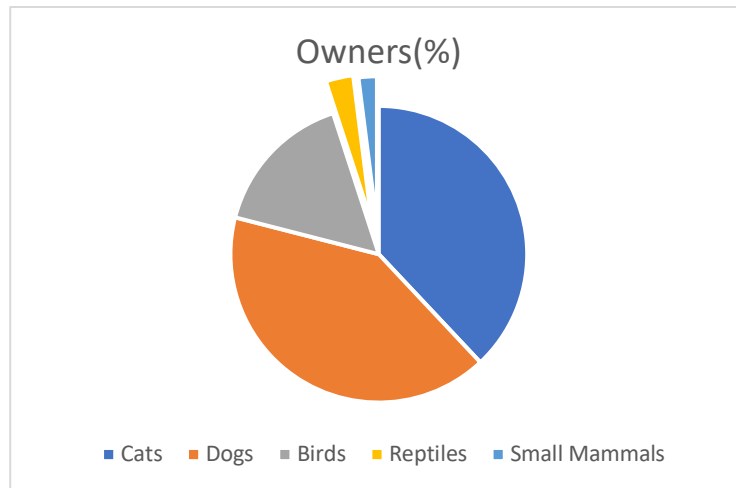
Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



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:

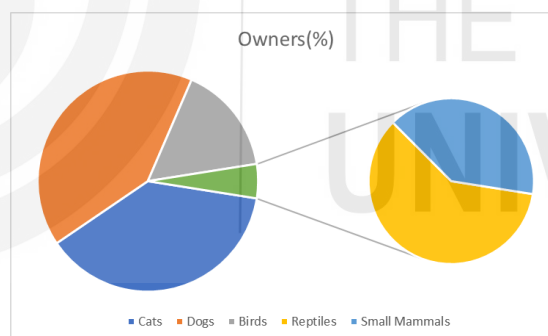
Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



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:

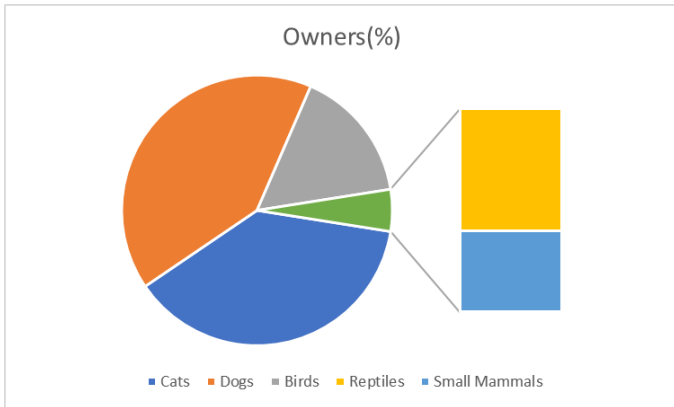
Pets	Owners (%)
Cats	38
Dogs	41
Birds	16
Reptiles	3
Small Mammals	2



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.

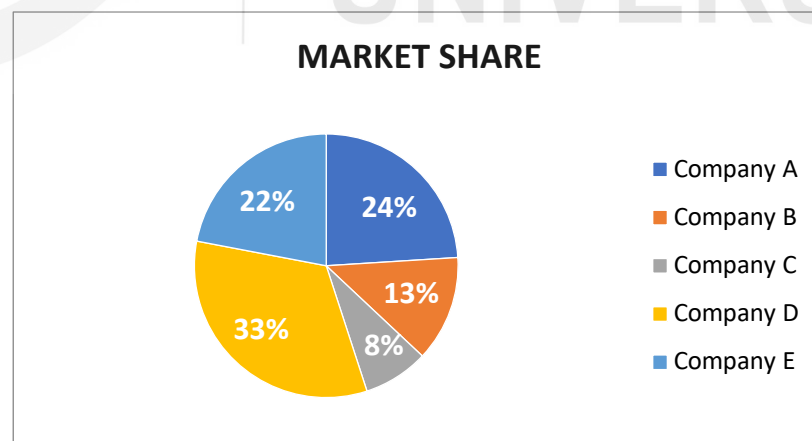
- Voting preference by age group
- Market share of cloud providers

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:

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.

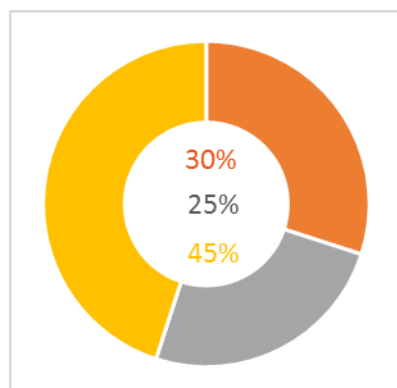
- Android OS market share
- Monthly sales by channel

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?

.....
.....

Q.2 What distinguishes a doughnut chart from a pie chart?

.....
.....

Q.3 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:

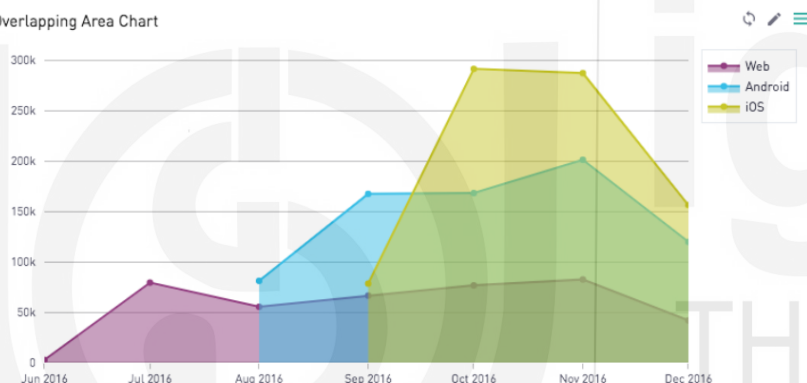
Overlapping area chart: An overlapping area chart results if we wish to look at how the values of the various groups compare to one another. The conventional line chart serves as the foundation for an overlapping area chart. One point is plotted for each group at each of the horizontal values, and the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

Overlapping Area Chart

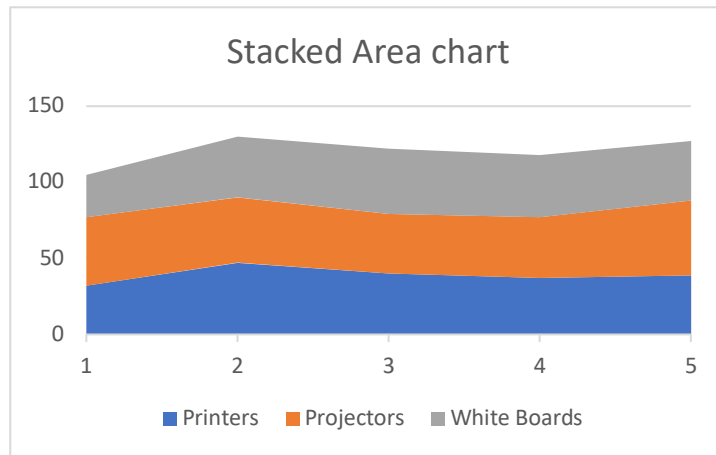


Stacked area chart: The stacked area chart is what is often meant to be conveyed when the phrase "area chart" is used in general conversation. When creating the chart of overlapping areas, each line was tinted based on its vertical value all the way down to a shared baseline. Plotting lines one at a 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

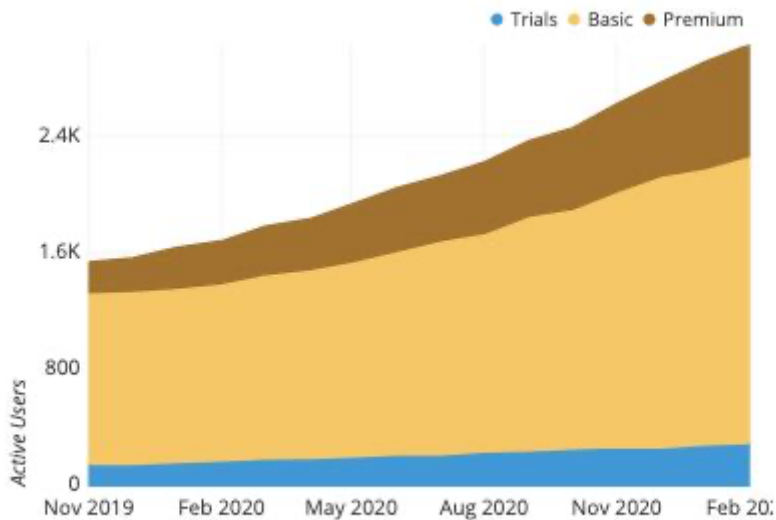
Best Practices:

- 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

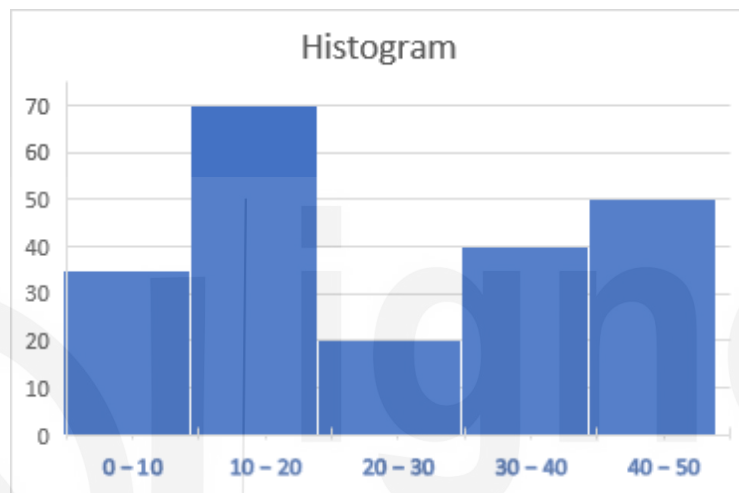
This Unit introduces you to some of the basic charts that are used in data science. The Unit defines the characteristics of Histograms, which are very popular in univariate frequency analysis of quantitative variables. It then discusses the importance and various terms used in the box plots, which are 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

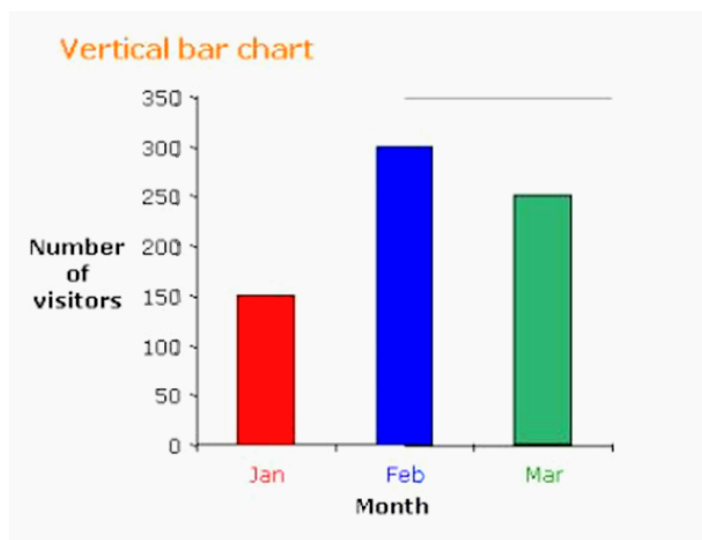
We can further divide into two types:

Grouped Bar Charts
Stacked Bar Charts

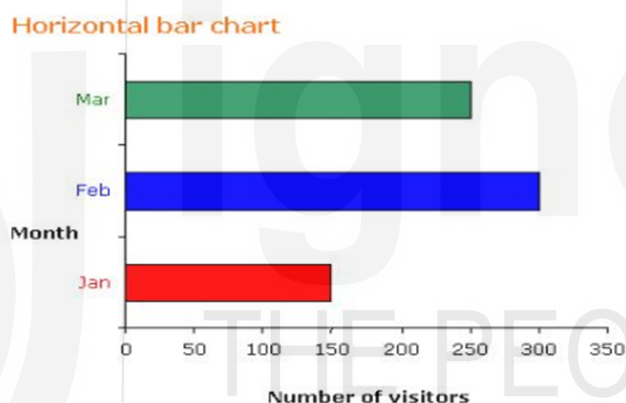


ignou
THE PEOPLE'S
UNIVERSITY

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:

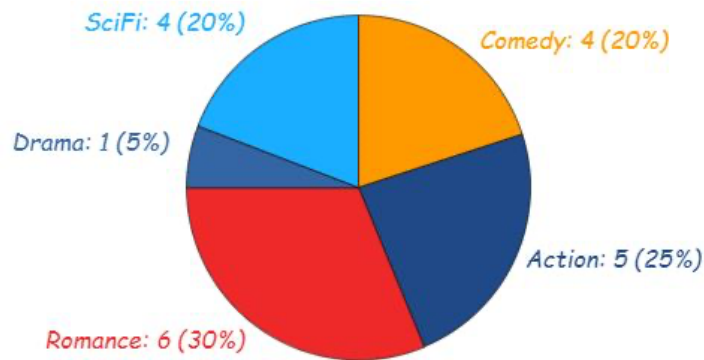
1. A pie chart, often referred to as a circle chart, is a style of graph that can be used to summarise a collection of nominal data or to show the many values of a single variable. (e.g. percentage distribution).

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

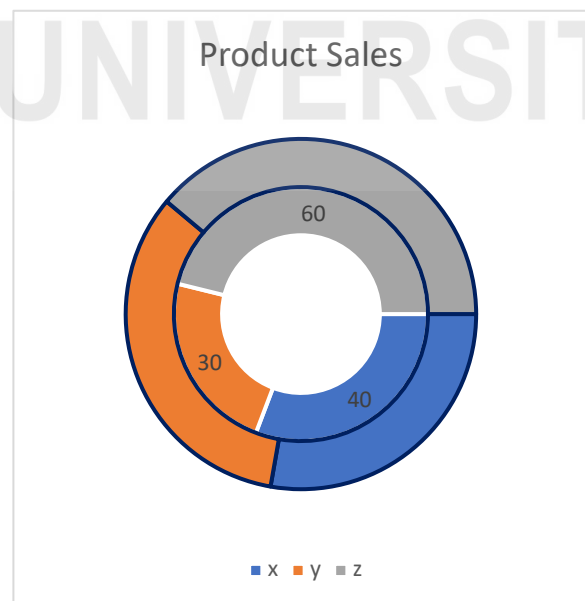
3.



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.



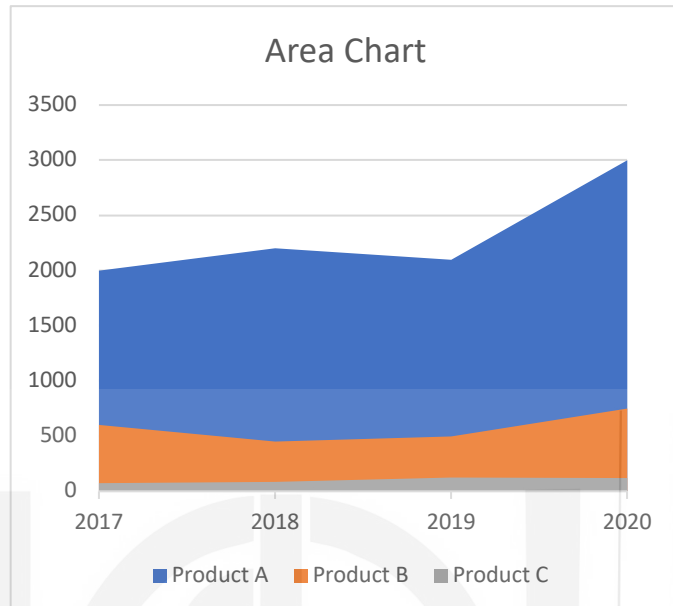
Check Your Progress 12:

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-diagrammeabarres/5214818-eng.htm>
- <https://sixsigmamania.com/?p=475>
- <https://study.com/academy/lesson/measures-of-dispersion-variability-and-skewness.html>