3 Plotting Data using Matplotlib

Fastrack Revision

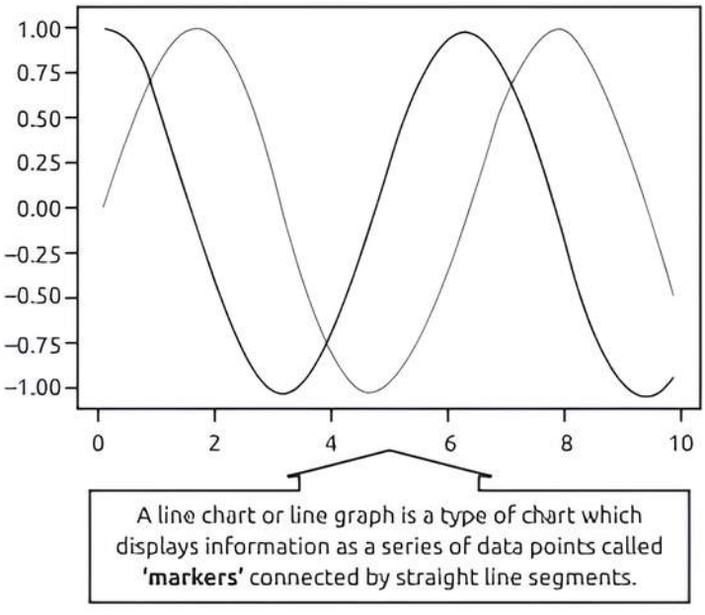
- **Data Visualisation:** It refers to the graphical or visual representation of information and data using visual elements like charts, graphs and maps, etc. The purpose of plotting data is to visualise variation or show relationships between variables.
- **Pyplot of Matplotlib Library:** The matplotlib is a Python library that provides many interfaces and functionality for 2D – graphics similar to MATLAB's in various forms. It provides both a very quick way to visualise data from Python and publication – quality figures in many formats.
- **Pyplot:** It is a collection of methods within matplotlib which allows user to construct 2D plots easily and interactively.
- **Installing and Importing Matplotlib:** It can be installed using the following pip command from the command prompt:

```
pip install matplotlib
```

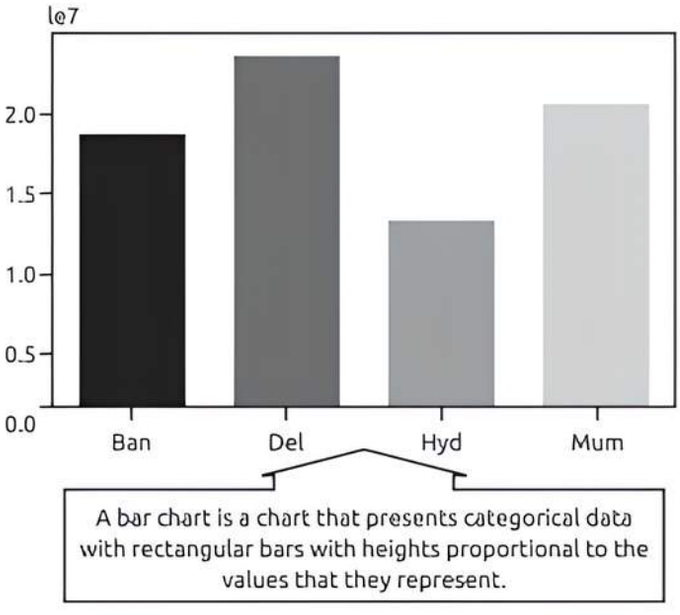
- For plotting using Matplotlib, we need to import its pyplot module using the following command:

```
import matplotlib.pyplot as plt
```

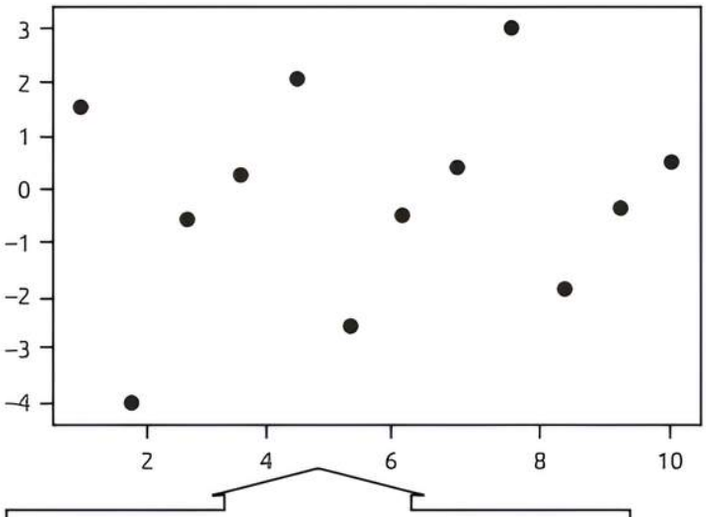
- **Plot ():** The pyplot module of matplotlib contains a collection of functions that can be used to work on a plot. The plot() function of the pyplot module is used to create a figure.
- **Figure:** A figure is the overall window where the outputs of pyplot functions are plotted. A figure contains a Notes plotting area, legend, axis labels, ticks, title, etc.
- **Boxplot Chart:** A boxplot is the visual representation of the statistical five number summary of a given data set. With pyplot, a boxplot is created using **boxplot ()** function.
- **Chart Types:** We can create different types of graphs and charts using Pyplot. Some commonly used charts types are:
  - **Line Chart:** A line chart or line graph is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. With pyplot, a line chart is created using **plot()** function.
  - **Bar Chart:** A bar chart or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. With pyplot, a bar chart is created using **bar()** and **barh()** functions.
  - **Scatter Plot:** The scatter plot is similar to a line chart, the major difference is that while line graph connects the data points with a line, scatter chart simply plots the data points to show the trend in the data. With pyplot, a scatter chart is created using **scatter()** function.
  - **Pie Chart:** A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. With pyplot, a pie chart is created using **pie()** function. But pie chart can plot only one data sequence unlike other chart types.
  - **Histogram Plot:** A histogram is a type of graph that provides a visual interpretation of numerical data by indicating the number of data points that lie within a range of values. With pyplot, a histogram is created using **hist()** function.



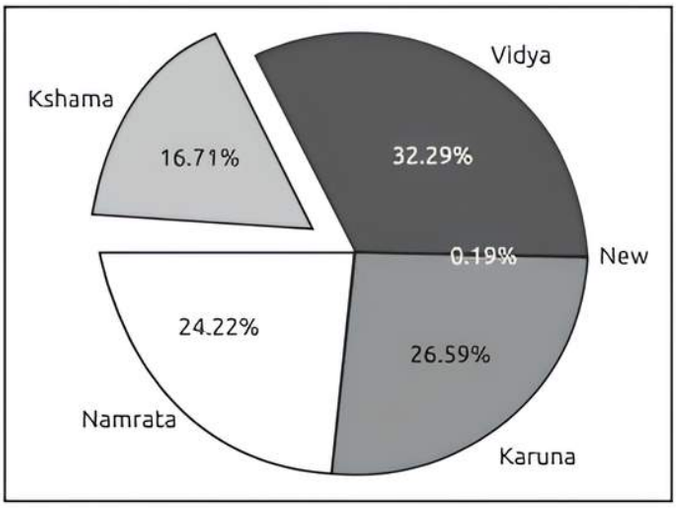
A line chart or line graph is a type of chart which displays information as a series of data points called **'markers'** connected by straight line segments.



A bar chart is a chart that presents categorical data with rectangular bars with heights proportional to the values that they represent.



A scatter plot is a two-dimensional data visualisation that uses dots to represent the values obtained for two different variables — one plotted along the x-axis and the other plotted along the y-axis.



A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion.

▶ **List of pyplot functions to plot different charts:**

`plot(\*args[, scalex, scaley, data])`	Plot x versus y as lines and/or markers.
`bar(x, height[, width, bottom, align, data])`	Make a bar plot.
`boxplot(x[, notch, sym, vert, whis, ...])`	Make a box and whisker plot.
`hist(x[, bins, range, density, weights, ...])`	Plot a histogram.
`pie(x[, explode, labels, colors, autopct, ...])`	Plot a pie chart.
`scatter(x, y[, s, c, marker, cmap, norm, ...])`	A scatter plot of x versus y.

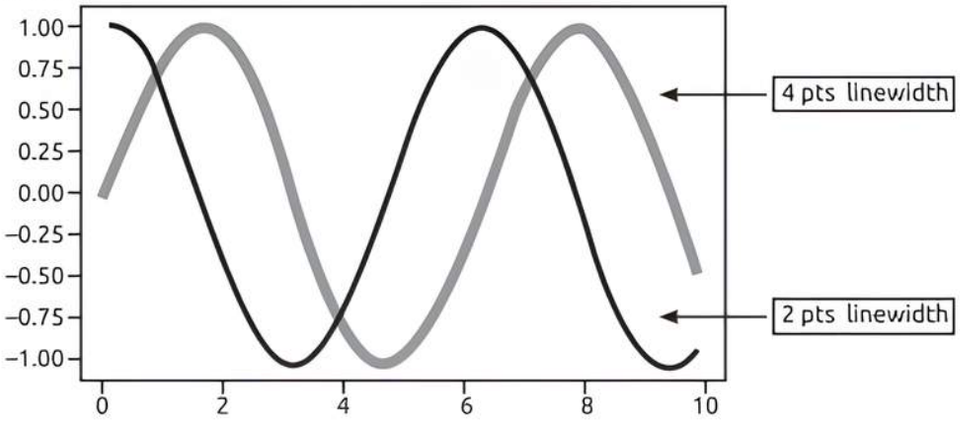
▶ **Marker:** A marker is any symbol that represents a data value in a line chart or a scatter plot. In other words, the data points being plotted on a graph/chart are called markers.

marker	description	marker	description
'.'	point marker	'H'	hexagon2 marker
','	pixel marker	'1'	tri_down marker
'o'	circle marker	'2'	tri_up marker
'+'	plus marker	'3'	tri_left marker
'x'	x marker	'4'	tri_right marker
'D'	diamond marker	'v'	triangle_down marker
'd'	thin_diamond marker	'^'	triangle_up marker
's'	square marker	'<'	triangle_left marker
'p'	pentagon marker	'>'	triangle_right marker
'*'	star marker	'|', '_'	vline, hline markers
'h'	hexagon1 marker		

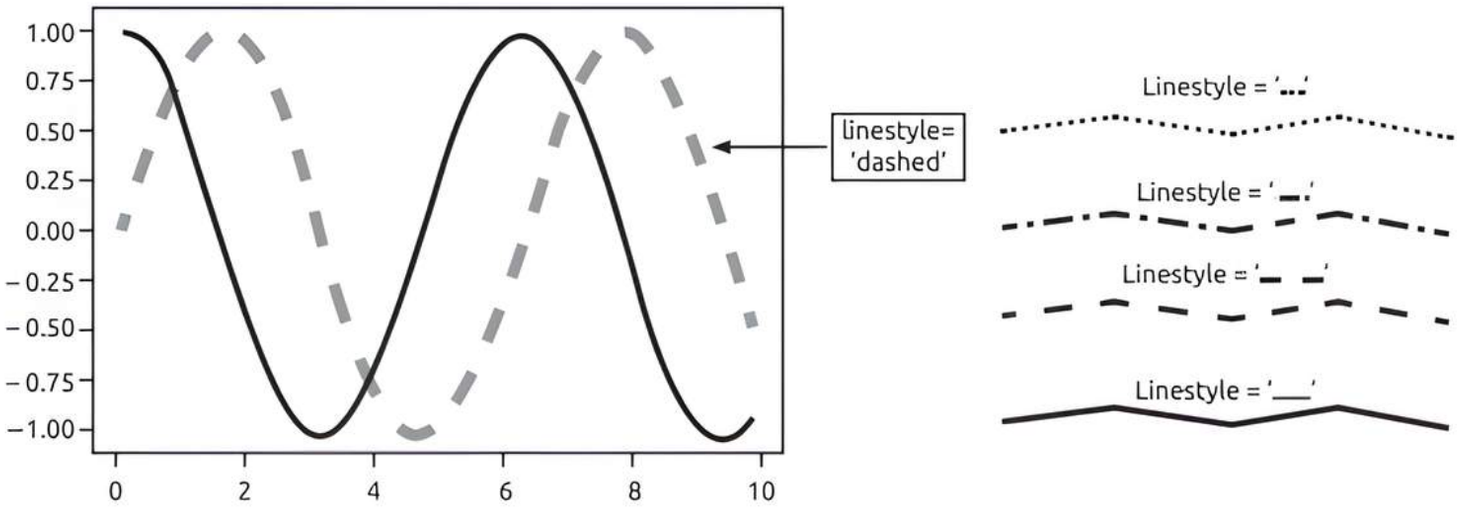
▶ **Colour:** It is also possible to format the plot further by changing the colour of the plotted data. We can either use character codes or the colour names as values to the parameter colour in the plot().

character	colour	character	colour
'b'	blue	'y'	yellow
'g'	green	'k'	black
'r'	red	'c'	cyan
'm'	magenta	'w'	white

▶ **Linewidth and Linestyle:** The linewidth and linestyle property can be used to change the width and the style of the line chart. Linewidth is specified in pixels. The default line width is 1 pixel showing a thin line. Thus, a number greater than 1 will output a thicker line depending on the value provided.



▶ We can also set the linestyle of a line chart using the linestyle parameter. It can take a string such as "solid", "dotted", "dashed" or "dashdot".



▶ **The Pandas Plot Function:** The plot( ) method of Pandas accepts a considerable number of arguments that can be used to plot a variety of graphs. It allows customising different plot types by supplying the kind keyword arguments.

Syntax: plt.plot(kind),where **kind** accepts a string indicating the type of **.plot**

Kind =	Plot Type	Kind =	Plot Type
line	Line plot (default)	box	Boxplot
bar	Vertical bar plot	Area	Area plot
barh	Horizontal bar plot	Ple	Ple plot
Hist	Histogram	scatter	Scatter plot

Practice Exercise

Multiple Choice Questions

Q 1. Pyplot is an interface of Python's ..... library.
a. seaborn b. plotly
c. ggplot d. matplotlib

Q 2. For 2D plotting using a Python library, which library interface is often used?
a. Seaborn b. Plotly
c. Matplotlib d. Matplotlib.pyplot

Q 3. Which of the following is not a valid chart type?
a. Histogram b. Statistical
c. Box d. Either b. or c.

Q 4. Which of the following is not a valid plotting function of pyplot?
a. plot( ) b. bar( )
c. line( ) d. pie( )

Q 5. Which of the following plotting functions does not plot multiple data series?
a. plot( ) b. bar( )
c. pie( ) d. barh( )

Q 6. The plot which tells the trend between two graphed variables is the ..... graph/chart.
a. line b. scatter
c. bar d. pie

Q 7. The plot which tells the correlation between two variables which may not be directly related is ..... graph/chart.
a. line b. scatter
c. bar d. pie

Q 8. A ..... is a summarisation tool for discrete or continuous data.
a. quartile b. histogram
c. mean d. median

Q 9. A visual representation of the statistical five number summary of a given dataset is known as .....
a. histogram b. frequency distribution
c. boxplot d. frequency polygon

Q 10. Which of the following function is used to create a line chart?
a. line( ) b. plot( )
c. chart( ) d. plotline( )

Q 11. Which of the following function will produce a bar chart?
a. plot( ) b. bar( )
c. barh( ) d. Either b. or c.

Q 12. Which of the following will create a vertical bar chart?
a. plot( ) b. bar( )
c. plotbar( ) d. barh( )

Q 13. Which of the following function will create a horizontal bar chart?
a. plot( ) b. bar( ) c. plotbar( ) d. barh( )

Q 14. To specify the style of line as dashed, which argument of plot( ) needs to be set?
a. Line b. Width c. Style d. Linestyle

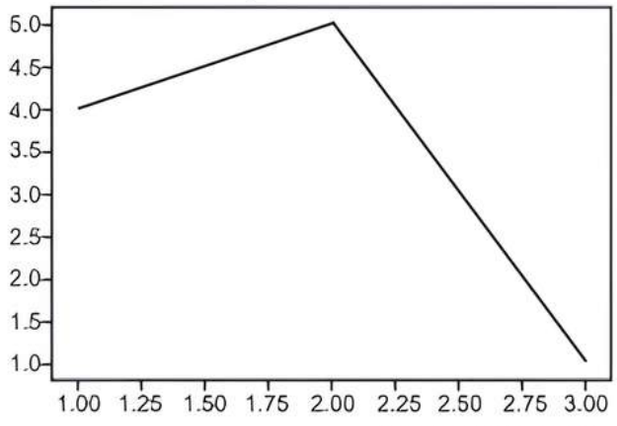
Q 15. The datapoints plotted on a graph are called .....
a. points b. pointers c. marks d. markers

Q 16. Out of the following, which function cannot be used for customisation of charts in Python? *[CBSE SQP 2021 Term-1]*
a. xlabel() b. colour()
c. title() d. xticks()

Q 17. What is the minimum number of arguments required for plot () function in matplotlib? *[CBSE SQP 2021 Term-1]*
a. 1 b. 3 c. 2 d. 4

Q 18. ..... is the function to save the graph. *[CBSE SQP 2021 Term-1]*
a. Savefig () b. Savegraph ()
c. Savefigure () d. Savechart ()

Q 19. Observe the following figure. Identify the coding for obtaining this as output.



[CBSE SQP 2021 Term-1]

a.
```
import matplotlib.pyplot as plt
plt.plot([1,2],[4,5])
plt.show()
```

b.
```
import matplotlib.pyplot as plt
plt.plot([1,2,3],[4,5,1])
plt.show()
```

c.
```
import matplotlib.pyplot as plt
plt.plot([2,3],[5,1])
plt.show()
```

d.
```
import matplotlib.pyplot as plt
plt.plot([1,3],[4,1])
plt.show()
```

Q 20. Read the statements given below and identify the right option to draw a histogram.

Statement 1: To make a Histogram with Matplotlib, we can use the plt.hist() function.

Statement 2: The bin parameter is compulsory to create histogram. *[CBSE SQP 2021 Term-1]*
a. Statement 1 is correct.
b. Statement 2 is correct.
c. Statement 1 is correct, but Statement 2 is incorrect.
d. Statement 1 is incorrect, but Statement 2 is correct.

Q 21. Which graph should be used where each column represents a range of values and the height of a column corresponds to how many values are in that range? *[CBSE SQP 2021 Term-1]*
a. plot b. line
c. bar d. histogram

Q 22. To compare data values of commission earned by salesmen over a year, which of the following type of graph should preferably be used? *[CBSE 2023]*
a. line b. area
c. bar d. scatter

Q 23. To create scatter charts using plot( ), which argument is skipped?
a. Marker b. Linestyle
c. Markeredgecolor d. Linewidth

Q 24. In scatter( ), which argument is used to specify the size of datapoints?
a. size b. s
c. marker d. markersize

Q 25. Which argument of bar( ) lets you set the thickness of bar?
a. thick b. thickness
c. width d. barwidth

Q 26. Which function lets you set the title of the plot?
a. title( ) b. plottitle( )
c. graphtitle( ) d. All of these

Q 27. The command used to give a heading to a graph is .....
a. plt.show( ) b. plt.plot( )
c. plt.xlabel( ) d. plt.title( )

Q 28. Which function would you use to set the limits for x-axis of the plot?
a. limits( ) b. xlimits( )
c. xlim( ) d. lim( )

Q 29. Which function is used to show legends?
a. display( ) b. show( )
c. legend( ) d. legends( )

Q 30. Which argument must be set with plotting functions for legend( ) to display the legends?
a. Data b. Label
c. Name d. Sequence

Q 31. Which function is used to create a histogram?
a. histo( ) b. histogram( )
c. hist( ) d. histtype

Q 32. Which argument in hist( ) is used to create a stacked bar type histogram?
a. histt b. histtype
c. type d. barstacked

Q 33. Which of the following functions can plot only one data series?
a. plot( ) b. bar( )
c. boxplot( ) d. pie( )

Q 34. Which argument must be provided to create wedges out of a pie chart?
a. label b. autopct
c. explode d. wedge

Q 35. Which argument should be set to display percentage share of each pie on a pie chart?
a. Label b. Autopct
c. Explode d. Wedge

Q 36. Which function creates a boxplot?

a. box( ) b. plot( )
c. boxplot( ) d. showbox( )

Q 37. To change the width of bars in a bar chart, which of the following arguments with a float value is used? *[CBSE 2021 Term-1]*

a. hwidth b. width
c. breath d. barwidth

Q 38. Which of the following command is used to import matplotlib for coding? *[CBSE 2021 Term-1]*

a. import matplotlib. pyplot as plt
b. import plt.matplotlib as plt
c. import by.matplotlib as plt
d. import pyplot.matplotlib as plt

Q 39. Consider the following statements with reference to Line charts.

Statement 1: line graphs is a tool for comparison and is created by plotting a series of several points and connecting them with a straight line.

Statement 2: You should never use line chart when the chart is in a continuous data set.

[CBSE 2021 Term-1]

a. Statement 1 is correct.
b. Statement 2 is correct.
c. Statement 1 is correct, but Statement B is incorrect.
d. Statement 1 is incorrect, but Statement B is correct.

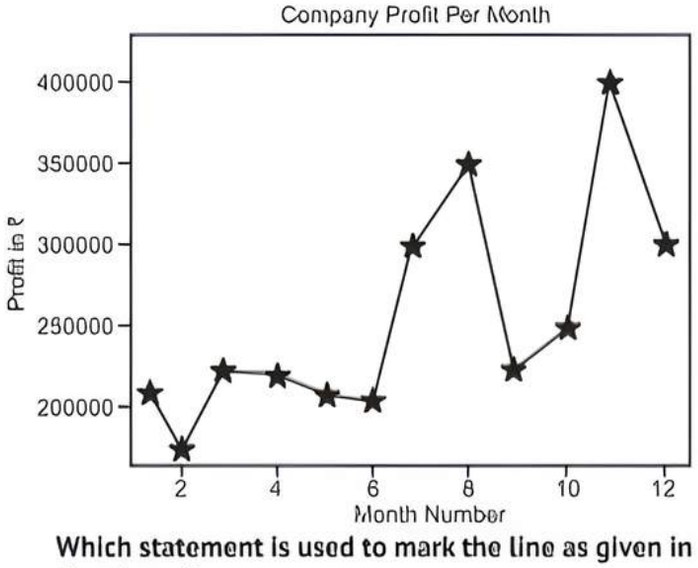
Q 40. What is not true about Data Visualisation?

[CBSE 2021 Term-1]

a. Graphical representation of information and data.
b. Helps users in analysing a large amount of data in a simpler way.
c. Data Visualisation makes complex data more accessible, understandable and usable.
d. No library needs to be imported to create charts in Python language.

Q 41. Ms. Kalpana is working with an IT company and she wants to create charts from the data provided to her. She generates the following graph.

[CBSE 2021 Term-1]



Which statement is used to mark the line as given in the above fig:

a. `plt.plot(x,y,marker = '0',markersize = 10,color='red',linestyle='dashdot')`
b. `plt.plot(x,y,marker='star',markersize =10,color='red')`
c. `plt. plot(x,y, marker ='@', markersize = 10, color = 'red', linestyle ='dashdot')`
d. `plt.plot(x,y,marker='*',markersize=10, color='red')`

Fill in the Blanks Type Questions

Q 42. A ..... is a plot that shows the underlying frequency distribution of a set of continuous data.

Q 43. Pyplot interface is a collection of methods within ..... library of Python.

Q 44. Pyplot's ..... function is used to create line charts.

Q 45. Pyplot's ..... function is used to create horizontal bar charts.

Q 46. Pyplot's ..... functions is used to create scatter charts.

Q 47. Pyplot's ..... function is used to create histogram.

Q 48. The datapoints plotted on a graph are called .....

Q 49. The ..... argument of plot ( ) specifies the width for the line.

Assertion & Reason Type Questions

Directions (Q.Nos. 50-57): *In the questions given below, there are two statements marked as Assertion (A) and Reason (R). Read the statements and choose the correct option.*

a. Both Assertion (A) and Reason (R) are true and Reason (R) is the correct explanation of Assertion (A).
b. Both Assertion (A) and Reason (R) are true, but Reason (R) is not the correct explanation of Assertion (A).
c. Assertion (A) is true, but Reason (R) is false.
d. Assertion (A) is false, but Reason (R) is true.

Q 50. Assertion (A): A histogram is a plot that shows the underlying frequency distribution of a set of continuous data.
Reason (R): Pyplot interface is a collection of methods within matplotlib library of Python.

Q 51. Assertion (A): Pyplot's plot( ) function is used to create line charts.
Reason (R): Pyplot's barh( ) function is used to create horizontal bar charts.

Q 52. Assertion (A): Pyplot's scatter( ) function is used to create scatter charts.
Reason (R): Pyplot's hist( ) function is used to create histogram.

Q 53. Assertion (A): The datapoints plotted on a graph are called markers.
Reason (R): The width argument of plot( ) specifies the width for the line.

Q 54. Assertion (A): The linestyle argument of plot ( ) specifies the style of the line.
Reason (R): The line argument of bar( ) specifies the bar width.

Q 55. Assertion (A): The xticks( ) function is used to specify ticks for x-axis.
Reason (R): To save a plot, savefig( ) function is used.

Q 56. Assertion (A): The orientation argument of hist( ) is set to create a horizontal histogram.
Reason (R): The showmeans argument shows the arithmetic mean on a boxplot.

Q 57. Assertion (A): The notch argument in a boxplot( ) creates a notched boxplot.
Reason (R): The loc argument of legend( ) provides the location of legend.

Answers

1. (d)	2. (d)	3. (d)	4. (c)	5. (c)
6. (a)	7. (b)	8. (b)	9. (c)	10. (b)
11. (d)	12. (b)	13. (d)	14. (d)	15. (d)
16. (b)	17. (a)	18. (a)	19. (b)	20. (c)
21. (d)	22. (c)	23. (b)	24. (b)	25. (c)
26. (a)	27. (d)	28. (c)	29. (c)	30. (b)
31. (c)	32. (b)	33. (d)	34. (c)	35. (b)
36. (c)	37. (b)	38. (a)	39. (c)	40. (d)

41. (d)

42. histogram
43. matplotlib
44. plot( )
45. barh( )
46. scatter( )
47. hist( )
48. markers
49. linewidth

50. (b) 51. (b) 52. (b) 53. (c)
54. (c) 55. (b) 56. (b) 57. (b)

Case Study Based Questions

Case Study 1

Customising the Plot: Data visualisation demand much more from a graph/plot. The graph or plot should have a proper title, X and Y limits defined, labels, legends, etc. All this make understanding the plot and taking the decision easier. Any graph or chart that you create using matplotlib's PyPlot interface is created as per a specific structure of a plot or shall we say a specific anatomy. Pyplot charts have hierarchical structures or in simple words they are actually like containers containing multiple items/things inside it.

Q 1. ..... refers to study of bodily structures (or parts) of something.
a. Anatomy b. Figure
c. Axes d. None of these

Q 2. Pyplot, by default, plots every chart into an area called .....
a. figsize b. figure
c. title d. axes

Q 3. The ..... are individual points marked on *x*-axis and *y*-axis.
a. Limits b. Axis Label
c. Tick_Marks d. Legends

Q 4. ..... are the different colours that identify different sets of data plotted on the plot.
a. Legends b. Axes
c. Figure d. None of these

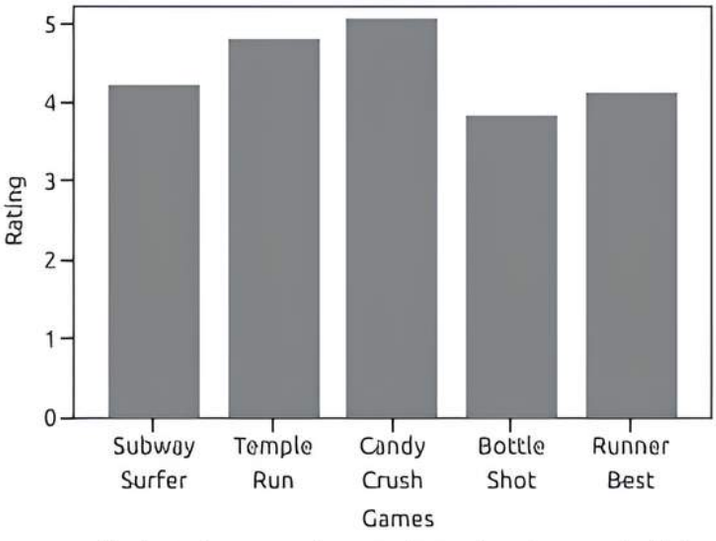
Q 5. ..... defines the range of values and number of values marked on *x*-axis and *y*-axis.
a. Tick_Marks b. Axes
c. Axis Label d. Limits

Answers

1. (a) 2. (b) 3. (c) 4. (a) 5. (d)

Case Study 2

Mr. Sharma is working in a game development industry and he was comparing the given chart on the basis of the rating of the various games available on the play store.



He is trying to write a code to plot the graph. Help Mr. Sharma to fill in the blanks of the code and get the desired output.

```
import _____
#Statement 1
Games=["Subway Surfer","Temple
Run","Candy Crush","Bottle Shot","Runner
Best"]
Rating=[4.2,4.8,5.0,3.8,4.1]
plt._____(Games,Rating)
#Statement 2
plt.xlabel("Games")
plt._____("Rating")
#Statement 3
plt._____
#Statement 4
```

Q 1. Choose the right code from the following for statement 1.

a. matplotlib as plt
b. pyplot as plt
c. matplotlib.pyplot as plt
d. matplotlib.plt as pyplot

Q 2. Identify the name of the function that should be used in statement 2 to plot the above graph.

a. line( ) b. bar( )
c. hist( ) d. barh( )

Q 3. Choose the correct option for the statement 3.

a. title("Rating") b. ytitle("Rating")
c. ylabel("Rating") d. yaxis("Rating")

Q 4. Choose the right function/method from the following for the statement 4.

a. display() b. print()
c. bar() d. show()

Q 5. In case, Mr. Sharma wants to change the above plot to any other shape, which statement should be change?

a. Statement 1 b. Statement 2
c. Statement 3 d. Statement 4

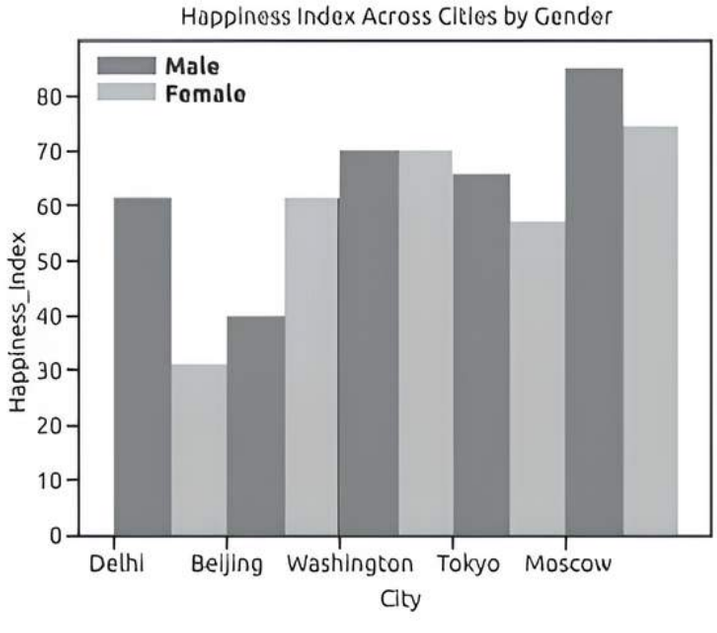
Answers

1. (c) **2.** (b) **3.** (c) **4.** (d) **5.** (b)

Case Study 3

Gaurav has written a Python Code to create a bar plot as given below using the following data:

City	Happiness_Index Male	Happiness_Index Female
Delhi	60	30
Beijing	40	60
Washington	70	70
Tokyo	65	55
Moscow	85	75




```
import _____ as _____
City=['Delhi','Beijing','Washington','Tokyo','Moscow']
Gender=['Male','Female'] Happiness_Index_
Male=[60,40,70,65,85]
Happiness_Index_Female=[30,60,70,55,75]
plt.bar([0.25,1.25,2.25,3.25,4.25],Happiness_Index_
Male,colour='blue',label=" Male",width=.5)
plt .....([.75,1.75,2.75,3.75,4.75],Happiness_Index_
Female,colour='Green', width=.5,label="Female")
pos=range(len(City)) print(pos)
plt.xticks(pos,City,fontsize=10)
plt.xlabel('City', fontsize=16)
plt.ylabel('Happiness_Index', fontsize=16)
```

Q 1. Identify the suitable code to be used in the blank space in line marked as Statement 1.

a. matplotlib as plt b. numpy as np
c. pandas as pd d. matplotlib.pyplot as plt

Q 2. What is the name of the function to plot the required bar graph in the line marked as Statement 2?

a. hist() b. pie()
c. bar() d. scatter()

Q 3. Fill in the blank in statement 3 to set Chart Title as "Happiness Index across cities by gender "and font size as 18.

a. plt.xtitle("Happiness Index across cities by gender",fontsize=18)
b. plt.title("Happiness Index across cities by gender",fontsize=18)
c. plt.ytitle("Happiness Index across cities by gender",fontsize=18)
d. plt.show("Happiness Index across cities by gender",fontsize=18)

Q 4. Identify the suitable code for line marked as Statement 4 to display the legends as shown in the plot.

a. plt.showlegend() b. plt.legend()
c. plt.display() d. plt.show()

Q 5. Fill in the blank marked in Statement 5 to display the plot.

a. plt.plot() b. plt.showplot()
c. plt.display() d. plt.show()

Answers

1. (d) **2.** (c) **3.** (b) **4.** (b) **5.** (d)

Case Study 4

Pyplot is a Matplotlib module that provides a MATLAB-like interface. Matplotlib is designed to be as usable as MATLAB, with the ability to use Python and the advantage of being free and open source. Each Pyplot function makes some change to a figure: for example, creates a figure, creates a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc. The various plots we can utilise using Pyplot are line plot, histogram, scatter, 3D plot, image, contour and polar.

Q 1. What is the role of legends in a graph/chart?

Q 2. What will happen if you use legend() without providing any label for the data series being plotted?

Q 3. A list temp contains average temperatures for seven days of last week. You want to see how the temperature changes L last seven days. Which chart type will you plot for the same and why?

Q 4. What do you mean by Linewidth and Linestyle Property?

Q 5. Define marker.

Answers

1. When we plot multiple ranges on a single plot, it becomes necessary that legends are specified. Legends are colour or mark linked to a specific data range plotted. legend() function is used to add legends to a plot.
2. If the legend() function is used without specifying any labels then the warning 'No handles with labels found to put in legend' will be raised.
3. A bar chart should be used to plot such data because the data is categorised in seven week days. The bars will be adequate to show the average temperature on these week days.
4. The linewidth and linestyle property can be used to change the width and the style of the line chart.
5. A marker is any symbol that represents a data value in a line chart or a scatter plot.

Case Study 5

"Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter notebook, web application servers and four graphical user interface toolkits." Matplotlib is a low-level library of Python which is used for data visualisation. It is easy to use and emulates MATLAB like graphs and visualisation. This library is built on the top of NumPy arrays and consists of several plots like line chart, bar chart, histogram, etc. It provides a lot of flexibility but at the cost of writing more code.

Q 1. What do you mean by data visualisation technique?

Q 2. How data visualisation can help in decision making?

Q 3. Name the library and interface used to plot a chart in Python.

Q 4. What are the ways of importing matplotlib?

Q 5. What are the basic elements/components of the chart?

Answers

1. The data visualisation technique refers to the graphical or pictorial or visual representation of data. This can be achieved by charts, graphs, diagrams or maps.
2. Mostly data visualisation technique provides data in form of charts or graphs. These charts majorly used in producing various reports like trends, outliers, and other diagrams. By observing these all user can understand the data easily and he/she can take his decision.
3. Library–matplotlib
   interface–pyplot
4. We can import in following two ways:
   (i) Using alias name: import matplotlib.pyplot as pp
   (ii) Without alias name: import matplotlib.pyplot
5. The chart has following elements/components:
   (i) Chart area or figure (ii) Axis
   (iii) Artist (iv) Titles
   (v) Legends

Very Short Answer Type Questions

Q 1. Name the library of which the Pyplot is an interface.

Ans. Pyplot is an interface of Matplotlib.

Q 2. Write the statement to install matplotlib.

Ans. Matplotlib can be installed using the following pip command from the command prompt: **pip install matplotlib.**

Q 3. Write the statement to import Pyplot in your script.

Ans. import matplotlib.pyplot

Q 4. Name the functions to create the following:
(i) line chart **(ii) bar chart**

Ans. (i) plot( ) (ii) bar() or barh()

Q 5. What do you mean by Pyplot?

Ans. Pyplot is a collection of methods within matplotlib which allows user to construct 2D plots easily and interactively.

Q 6. Define figure.

Ans. A figure is the overall window where the outputs of Pyplot functions are plotted. A figure contains a notes plotting area, legend, axis labels, ticks, title, etc.

Q 7. What is a line chart?

Ans. Chart which display information as a series of data points connected by straight line is called a line chart.

Q 8. What is a scatter plot?

Ans. It is similar to a line chart, the major difference is that while line graph generates the data points with a line, scatter chart plots the data points to show the trend in the data.

Q 9. What is a pie chart?

Ans. A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion.

Q 10. What is a bar chart?

Ans. A bar chart or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent.

Q 11. What is a histogram?

Ans. Histogram is a type of graph that provides a visual interpretation of numerical data by indicating the number of data points that lie within a range of value.

Q 12. What is a boxplot?

Ans. A boxplot is the visual representation of a statistical 5 number summary of a given data set.

Q 13. What do understand by the term marker?

Ans. A marker is any symbol that represents a data value in a line chart or a scatter plot. In other words, the data points being plotted on a graph/chart are called markers.

Q 14. Write the use of linewidth and linestyle.

Ans. The linewidth and linestyle property can be used to change the width and the style of the line chart.

Q 15. What is a frequency polygon?

Ans. Frequency polygon is a type of frequency distribution graph.

Q 16. Name the function to label axes.

Ans. xlabel( ), ylabel( )

Q 17. Answer the following questions:

(i) Name the function to give title to a plot.

(ii) Name the function to set figure size of a plot.

Ans. (i) title( )

(ii) figure( )

Q 18. Answer the following questions:

(i) Name the function to set limits for the axes.

(ii) Name the function to show legends on a plot.

Ans. (i) xlim( )

(ii) legend( )

Q 19. Name the function to add ticks on axes.

Ans. xticks( ), yticks( )

Q 20. Define axes.

Ans. The axes define the area (mostly rectangular in shape for simple plots) on which actual plot (line or bar or graph etc.) will appear. Axis have properties like label, limits and tick marks on them.

Short Answer Type-I Questions

Q 1. What is the significance of data visualisation?

Ans. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognised easier with data visualisation techniques or tools such as line chart, bar chart, pie chart, histogram, scatter chart, etc. Thus, with data visualisation tools, information can be processed in efficient manner and hence better decisions can be made.

Q 2. How does Python support data visualisation?

Ans. Python supports data visualisation by providing some useful libraries for visualisation. Most commonly used data visualisation library is matplotlib.

Q 3. What is the use of Matplotlib and Pyplot?

Ans. Matplotlib is a Python library, also sometimes known as the plotting library. The matplotlib library offers very extensive range of 2D plot types and output formats.

It can also be used for animations as well. There are many other libraries of Python that can be used for data visualisation but matplotlib is very popular for 2D plotting.

Pyplot is a collection of methods within matplotlib library (of Python) which allows user to construct 2D plots easily and interactively.

Knowledge BOOSTER

It offers complete 2D support along with limited 3D graphic support. It is useful in producing publication quality figures in interactive environment across platforms.

Q 4. What are the popular ways of plotting data?

Ans. The popular ways of plotting data are:

(i) Line Chart (ii) Bar Chart
(iii) Scatter Plot (iv) Pie Chart
(v) Histogram Plot (vi) Boxplot Chart

Q 5. Name the functions to create the following:

(a) Horizontal bar chart **(b) Histogram**
(c) Scatter chart **(d) Boxplot**

Ans. (a) barh( ) (b) hist( )
(c) scatter( ) (d) boxplot()

Q 6. Compare bar( ) and barh( ) functions.

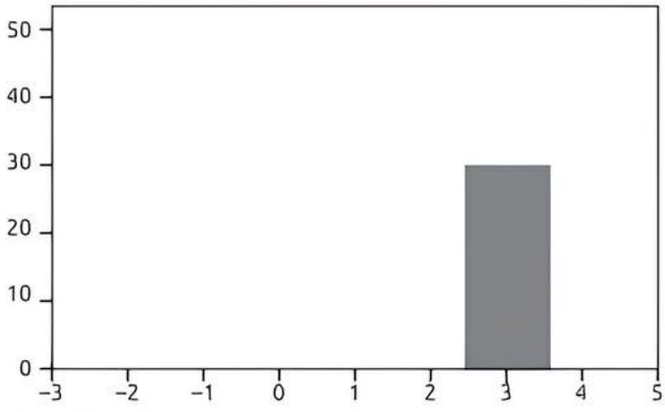
Ans. bar( ) function is used plot vertical bar graph while barh( ) function is used to plot horizontal bar graph.

Q 7. What is the role of legends in a graph/chart?

Ans. When we plot multiple ranges on a single plot, it becomes necessary that legends are specified. Recall that a legend is a colour or mark linked to a specific data range plotted.

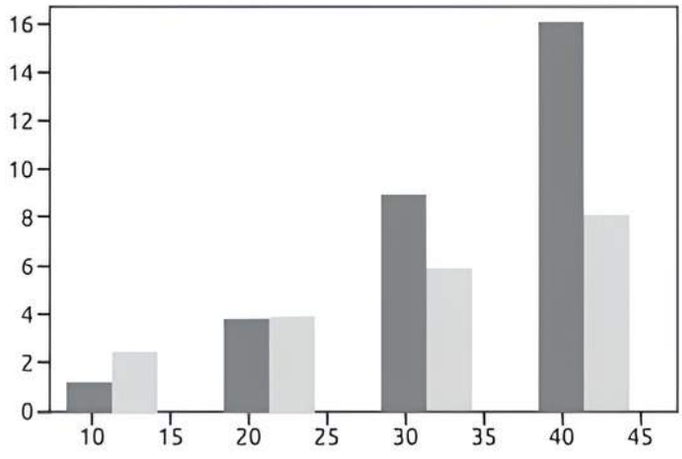
Q 8. Consider the code given below (all required libraries are imported) and the output produced by it. Why is the chart showing one bar only while we are plotting four values on the chart?

```
a = [3, 6, 9, 12]
b = [30, 48, 54, 48]
plt.xlim (-3, 5)
plt.bar (a, b)
plt. show ( )
```



Ans. The given chart is showing a single bar as the limits of *x*-axis have been set as –3 to 5. On this range, only one value from the data range being plotted falls *i.e.*, only a(0) and b(0) fall on this range. Thus, only a single value b(0) *i.e.*, 30 is plotted against a(0) *i.e.*, 3.

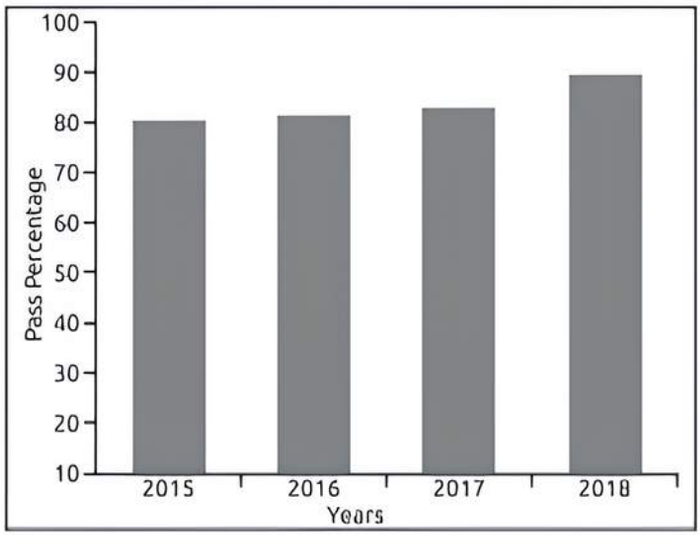
Q 9. Given an ndarray p as ([1, 2, 3, 4]). Write code to plot a bar chart having bars for p and p2(with dark colour) and another bar for p vs p*2 (with light colour). (Assume that libraries have been imported)**



Ans. plt.bar (p, p** 2, colour = 'r', width = 0.3)
plt.bar (p + 0.3, p * 2, colour = 'b', width = 0.3)

Q 10. Write a code to plot a bar chart to depict the pass percentage of students in CBSE exams for the years 2015 to 2018 as shown below.

[CBSE SQP 2019-20]



Ans.
```
import matplotlib.pyplot as plt
import numpy as np
objects=('2015', '2016', '2017', '2018')
y_pos=np.arrange(len(objects))
percentage=(82.83.85.90)
plt.bar(y_pos, percentage. align='Centre'. colour='Blue')
plt.xticks(y_pos.objects)
plt.ylabel("Pass Percentage")
plt.xlabel('Years')
plt.show( )
```

Q 11. What will happen if you use legend( ) without providing any label for the data series being plotted?

Ans. Then Error will come as 'No handles with labels found to put in legend."

Q 12. What do you understand by xlimit and ylimit? How are these linked to data being plotted?

Ans. If we need to have own limits specified for *X* and *Y* axes. For this, you can use xlim( ) and ylim( ) functions to set limits for *x*-axis and *y*-axis respectively.

Both xlim( ) and ylim( ) are used as per following format:

<matplotlib.pyplot>.xlim(<xmin>, <xmax>)

<matplotlib.

pyplot>.ylim(<ymin>, <ymax>)

Q 13. A list namely temp contains average temperatures for seven days of last week. You want to see how the temperature changed in last seven days. Which chart type will you plot for the same and why?

Ans. Bar chart is best one for such a situation. Because using bar graph we can easily plot and show data of seven days of last week.

Q 14. What is histogram? How do you create histograms in Python?

Ans. A histogram is a statistical tool used to summarise discrete or continuous data. It provides a visual interpretation of numerical data by showing the number of data points that fall within a specified range of values (called "bins").

The syntax for using hist( ) function of pyplot is:

matplotlib.pyplot.hist(x, bins = None, cumulative = False, histtype = 'bar', align='mid', orientation = 'vertical',)

COMMON ERROR

Some students get confused between bar graph and histogram.

Q 15. What are various types of histograms that can be created through hist( ) function?

Ans. The various types of histograms that can be created through hist( ) function are:

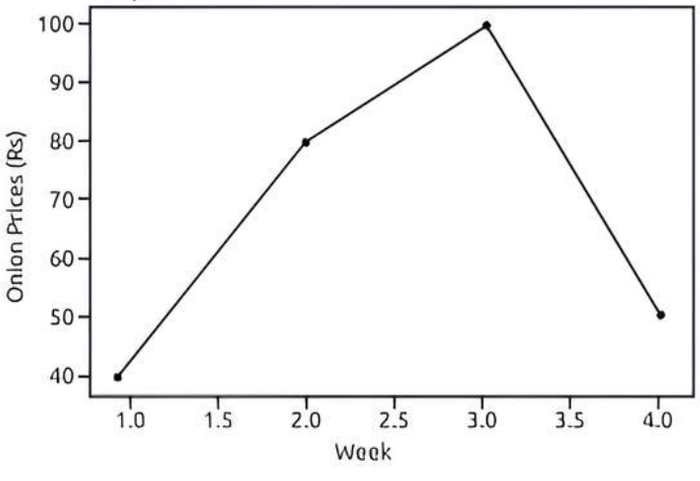
(i) Step type histogram
(ii) Regular histogram
(iii) Cumulative histogram
(iv) Horizontal histogram
(v) Vertical histogram

Short Answer Type-II Questions

Q 1. Write a program to plot a line chart to depict the changing weekly onion prices for four weeks. Give appropriate axes labels.

Ans.
```
import matplotlib.pyplot as plt
week = [1, 2, 3, 4]
Prices = [40, 80, 100, 50]          # onion prices
# Plotting line graph
plt.plot(week, Prices)
# Set the x-axis and y-axis labels
plt.xlabel ('Week')
plt.ylabel ('Onion Prices (Rs.)')
plt.show( )
```

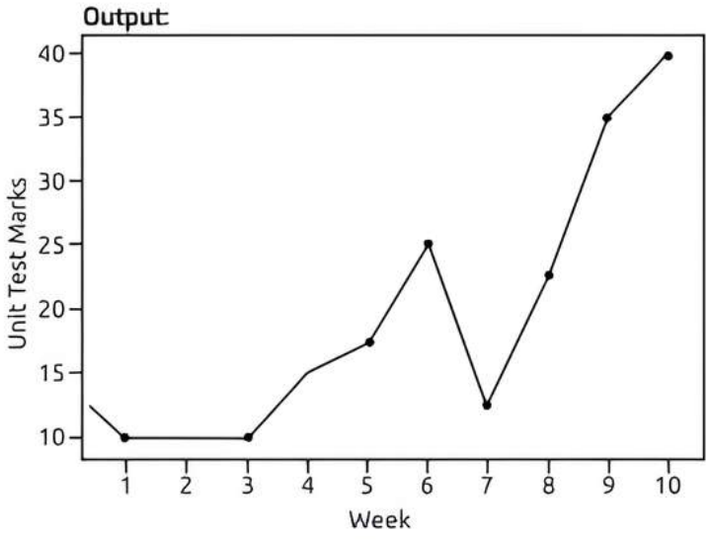
Output:



Q 2. Marks is a list that stores marks of a student in 10 unit tests. Write a program to plot the student's performance in these 10 unit tests.

Ans.
```
import matplotlib.pyplot as plt
week = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Marks = [12, 10, 10, 15, 17, 25, 12, 22, 35, 40]
# Plotting line graph
plt.plot (week, Marks)
# Set the x-axis and y-axis labels.
plt.xlabel ('Week')
plt.ylabel ('Unit Test marks')
plt.show( )
```

Output:



Q 3. Why is following code not producing any result? Why is it giving errors?
(Note: All required libraries have been imported and are available)
a = range (10, 50, 12)
b = range (90, 200, 20)
matplotlib.pyplot.plot (a, b)

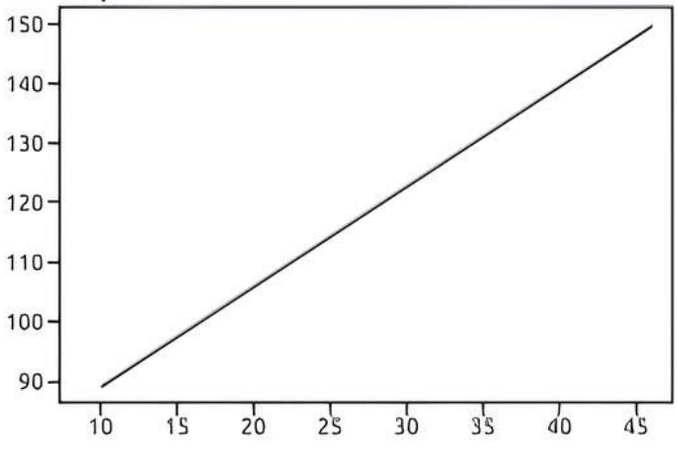
Ans. The above code is producing errors because the two sequences being plotted *i.e.*, **a** and **b** do not match in shape. While sequence **'a'** contains 4 elements, sequence **'b'** contains 6 elements. For plotting, it is necessary that the two sequences being plotted must match in their shape.

Q 4. What changes will you recommend to rectify the error in previous question's code?

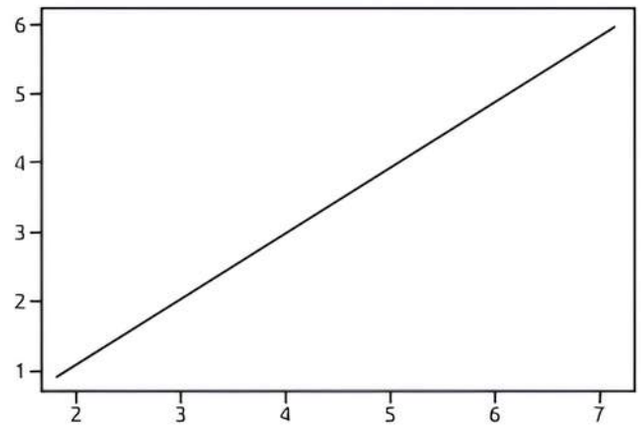
Ans. Since both the sequences being plotted must match in their shape, we can achieve this either by adding two elements to sequence **a** so that it has the same shape as sequence **b** (*i.e.*, 6 elements) or by removing two elements from sequence **b** so that it matches the shape of sequence **a** (*i.e.*, 4 elements).
For example,

```
a = range (10, 50, 12)
b = range (90, 160, 20)
matplotlib.pyplot.plot (a, b)
```

Output:



Q 5. Consider the following graph. Write the code to plot it. *[CBSE SQP 2020-21]*



Ans. import matplotlib.pyplot as plt
plt.plot ([2, 7], [1, 6])
plt. show ( )

Alternative answer:
import matplotlib.pyplot as plt
a = [1, 2, 3, 4, 5, 6]
b = [2, 3, 4, 5, 6, 7]
plt.plot (a, b)

Q 6. What is scatter chart? How is it different from line chart?

Ans. The scatter chart is a graph of plotted points that show the relationship between two sets of data. With a scatter plot, a mark or marker (usually a dot or small circle), represents a single data point. With one mark (point) for every data point a visual distribution of the data can be seen. Depending on how tightly the points cluster together, we may be able to discern a clear trend in data.

The difference is that with a scatter plot, the decision is made from the data points such that the individual points should not be connected directly together with a line but, instead express a trend.

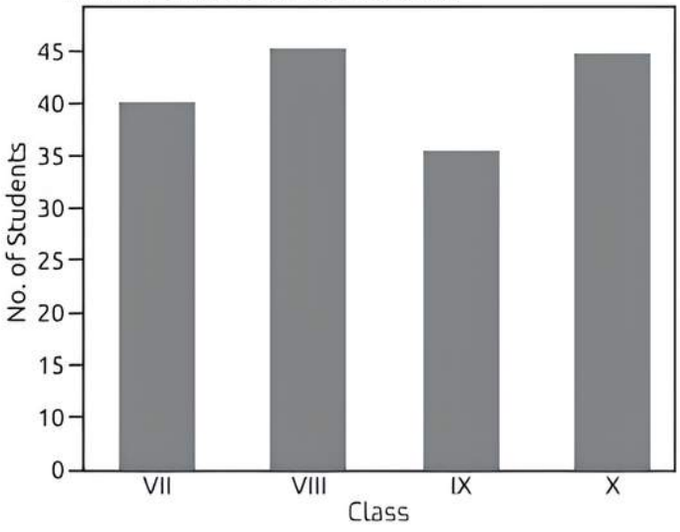
Q 7. The table below shows the marks of two students for the four unit tests for academic session 2019-2020. Fill in the blanks to draw a line graph with Test Names on the *x*-axis and marks on the *y*-axis. *[CBSE 2020]*

Tests	Marks	
	Rohit	Suman
Unit 1	85	97
Unit 2	88	99
Unit 3	89	90
Unit 4	87	92

import matplotlib.pyplot as plt
Tests = ... # Assign Test Names
Rohit = ... # Assign Marks of Rohit
Suman = ... # Assign Marks of Suman
plt.plot (Tests, Rohit, Suman)
..... # Label *y*-axis as Marks
..... # Add legends "Rohit", "Suman" for the lines plt.show ( )

Ans. ['Unit1', 'Unit2', 'Unit3', 'Unit4']
[85, 88, 89, 87]
[97, 99, 90, 92]
plt.ylabel ('Marks')
plt.legend (['Rohit', 'Suman'])

Q 8. Draw the following bar graph representing the number of students in each class.



[CBSE SQP 2020-21]

Ans. import matplotlib.pyplot as plt
Classes = ['VII', 'VIII', 'IX', 'X']
Students = [40, 45, 35, 44]
plt.bar (classes, students)
plt.show( )

Q 9. Following code is plotting the desired graph but legends are not showing despite giving the legend() of pyplot. What could be the reason? Suggest a solution for the problem.
plt.plot (*x*, *y*)
plt.plot (*x*, *z*)
plt.legend (loc = "upper left")

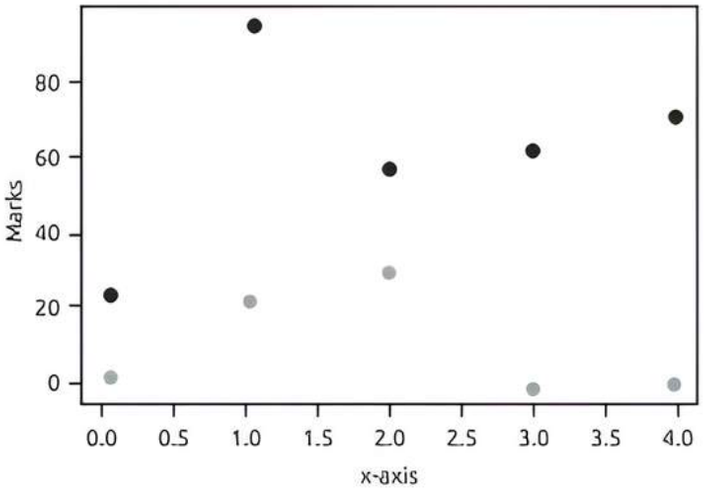
Ans. The above code won't print the legends because with the plot ( ), the labels are missing. The legend( ) will work only when we specify label for data series being plotted in the plot( ).
The solution for above problem will be:
plt. plot (x, y, label = "Y data")
plt. plot (x, z, label = "Z data")
plt. legend (loc = 'upper left')

Q 10. Consider a DataFrame mksdf as shown below:

	Name	Age	Pre-board Marks	Board Marks
0	Karan	17	4	25
1	Alex	19	24	94
2	Ani	18	31	57
3	Javed	18	2	62
4	Amrit	17	3	70

Write a program to plot pre-board marks and board marks from above DataFrame on the same scatter chart.

Ans.
```
import pandas as pd
import matplotlib.pyplot as plt
: #mksdf created or loaded
plt.scatter (x = mksdf.index, y = mksdf.
PreBoardMarks, c = 'g', s= mksdf. Age)
plt. scatter (x = mksdf.index, y = mksdf. BoardMarks,
c = 'r', s = mksdf. Age)
plt. xlabel ("X-axis")
plt.ylabel ("Marks")
plt.show ( )
```
Output:



Long Answer Type Questions

Q 1. Tanushree is doing some research. She has a stored line of Pascal's triangle numbers as ar2 as shown below:

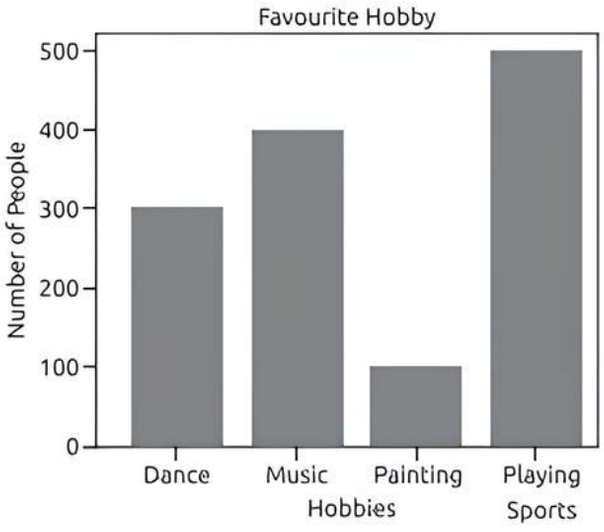
ar2 = [1, 7, 21, 35, 35, 21, 7, 1]

- **She wants to plot the sine (numpy. sin( )), cosine (numpy.cos( )) and tangent values (numpy.tan( )) for the same array (ar2).**
- **She wants cyan colour for sine plot line, red colour for cosine plot line and the black colour for tangent plot line.**
- **Also, the tangent line should be dashed.**

Write a program to accomplish all this.

Ans.
```
import matplotlib.pyplot as plt
import numpy as np
ar2 = [1, 7, 21, 35, 35, 21, 7, 1]
# calculating sin( ), cos( ) and tan( ) values
s2 = np.sin (ar2)
c2 = np.cos (ar2)
t2 = np.tan (ar2)
# plotting line chart
plt.figure (figsize = (15, 7))          # figure size specified
plt.plot (ar2, s2, 'c')                 # sine line with cyan
plt.plot (ar2, c2, 'r')                 # cosine line with red
plt.plot (ar2, t2, 'k', linestyle = 'dashed')   # tan line with black
# Set the x-axis label of the current axis.
plt.xlabel ('Array values')
# Set the y-axis label of the current axis.
plt.ylabel ('Sine, Cosine and Tangent Values')
plt. show ( )
```

Q 2. Write suitable Python code to create 'Favourite Hobby' Bar Chart as shown below:



Also give suitable python statement to save this chart. *[CBSE SQP 2023-24]*

Ans.
```
import matplotlib . pyplot as plt
hobby = ('Dance', 'Music', 'Painting', 'Playing Sports')
users = (300, 400, 100, 500)
plt . bar (hobby, users)
plt . title ("Favourite Hobby")
plt . ylabel ("Number of people")
plt . xlabel ("Hobbies")
plt . show ( )
plt . savefig ("hobbies . jpg")
```

Q 3. The heights of 10 students of eighth grade are given below:

Height_cms=[145, 141, 142, 142, 143, 144, 141, 140, 143, 144]

Write suitable Python code to generate a histogram based on the given data, along with an appropriate chart title and both axis labels.

Also give suitable python statement to save this chart. *[CBSE SQP 2023-24]*

Ans.
```
import matplotlib . pyplot as plt
Height_cms=[145, 141, 142, 142, 143, 143, 141, 140, 143,
144]
plt . hist (Height_cms)
plt . title ("Height Chart")
plt . xlabel ("Height in cms")
plt . ylabel ("Number of people")
plt . show ( )
plt . savefig ("heights . jpg")
```

Q 4. First 10 terms of Fibonacci series are stored in a list namely fib:

fib = [0, 1, 1, 2, 3, 5, 8, 13, 21, 34]

Write a program to plot Fibonacci terms and their square-roots with two separate lines on the same plot.

- **The Fibonacci series should be plotted as a cyan line with 'o' markers having size as 5 and edge-colour as red.**
- **The square-root series should be plotted as a black line with '+' markers having size as 7 and edge-colour as red.**

Ans.
```
import matplotlib.pyplot as plt
import numpy as np
fib = [0, 1, 1, 2, 3, 5, 8, 13, 21, 34]
sqfib = np.sqrt (fib)
fib = [0, 1, 1, 2, 3, 5, 8, 13, 21, 34]
sqfib = np.sqrt (fib)
plt. figure (figsize = (10, 7))
plt.plot (range (1, 11), fib, 'co', markersize = 5, linestyle = "solid", markeredgecolour = 'r')
plt.plot (range (1, 11), sqfib, 'k+', markersize = 7, linestyle = "solid", markeredgecolour = 'r')
plt. show ( )
```

Q 5. Given a series *nfib* that contains reversed Fibonacci numbers with Fibonacci numbers as shown below:
[0, –1, –1, –2, –3, –5, –8, –13, –21, –34, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34]
Write a program to plot nfib with following specifications:
- **The line colour should be magenta.**
- **The marker edge-colour should be black with size 5.**
- **Grid should be displayed.**

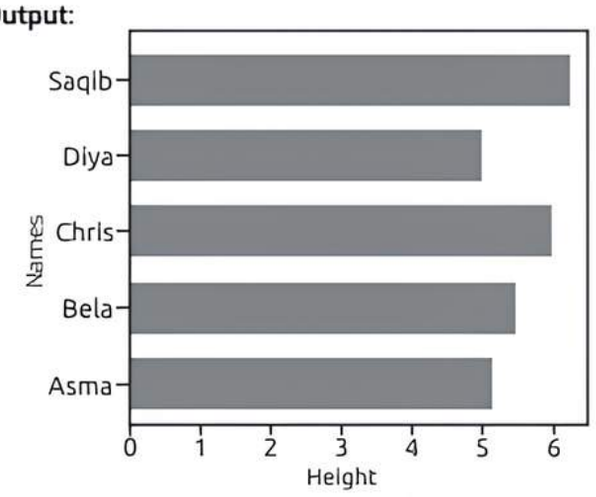
Ans.
```
Import matplotlib.pyplot as plt
import numpy as np
nfib = [0, –1, –1, –2, –3, –5, –8, –13, –21, –34, 0, 1, 1, 2, 3,
        5, 8, 13, 21, 34 ]
plt.plot (range (–10, 10), nfib, 'mo', markersize = 5,
markeredgecolour = 'k', linestyle = 'solid')
plt.grid (True)
plt.show( )
```

TiP
Do not skip any specification while writing the program and write correct syntax.

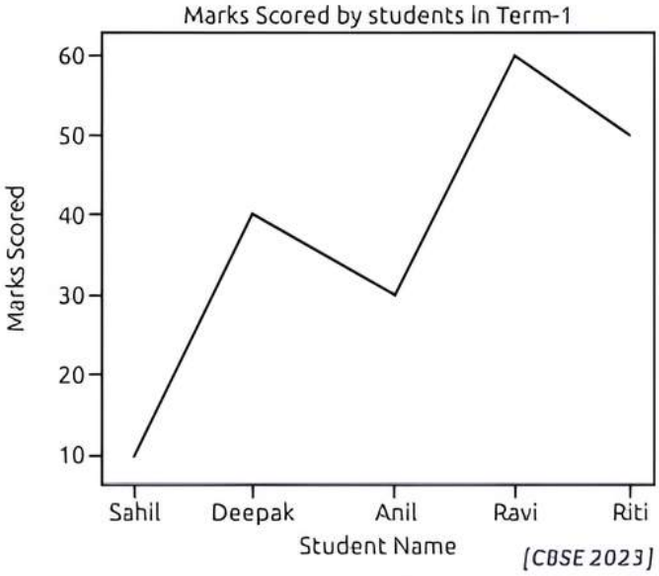
Q 6. Write a program to plot a horizontal bar chart from the heights of some students.

Ans.
```
import matplotlib.pyplot as plt
height = [5.1, 5.5, 6.0, 5.0, 6.3]
Names = ('Asma', 'Bela', 'Chris', 'Diya', 'Saqib')
# Create horizontal bars
plt.barh(Names, height)
plt.xlabel ("Height")
plt.ylabel ("Names")
plt.show( )
```

Output:



Q 7. Consider the following graph. Write the Python code to plot it. Also add the Title label for *X* and *Y*-axis.
Use the following data for plotting the graph
smarks=[10, 40, 30, 60, 55]
sname=["Sahil", "Deepak", "Anil", "Ravi", "Riti"]



[CBSE 2023]

Ans.
```
import matplotlib . pyplot as plt
sname=['Sahil', 'Deepak', 'Anil', 'Ravi', 'Riti']
smarks=[10, 40, 30, 60, 55]
plt . plot (sname, smarks)
plt . title ("Marks secured by students in Term-I")
plt . xlabel ('Student Name')
plt . ylabel ("Marks scored")
plt . show ( )
```

Q 8. Create multiple line charts on common plot where three data ranges are plotted on same chart. The data range (s) to be plotted is/are:
Data = [[5., 25., 45., 20.], [8., 13., 29., 27.], [9., 29., 27., 39.]]

Ans.
```
import numpy as np
import matplotlib.pyplot as plt
Data = [[5., 25., 45., 20.], [8., 13., 29., 27.], [9., 29., 27.,
39.]
X = np.arange (4)
plt.plot (X, Data [0], color = 'b', label = 'range1')
plt.plot (X, Data [1], colour = 'g', label = 'range2')
```

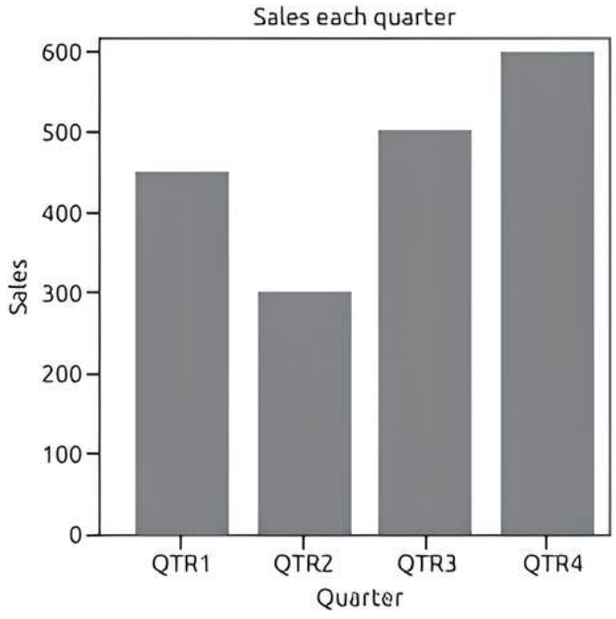
```
plt.plot (X, Data [2], colour = 'r', label = 'range3')
plt.legend (loc = 'upper left')
plt.title ("MultiRange Line chart")
plt.xlabel ('X')
plt.ylabel ('Y')
plt.show ( )
```

Q 9. Write Python code to draw the following bar graph representing the total sales in each quarter. Add the Title, Label for *X*-axis and *Y*-axis.

Use the following data for plotting the graph:

sales=[450, 300, 500, 650]

qtr=["QTR1", "QTR2", "QTR3", "QTR4"] *[CBSE 2023]*

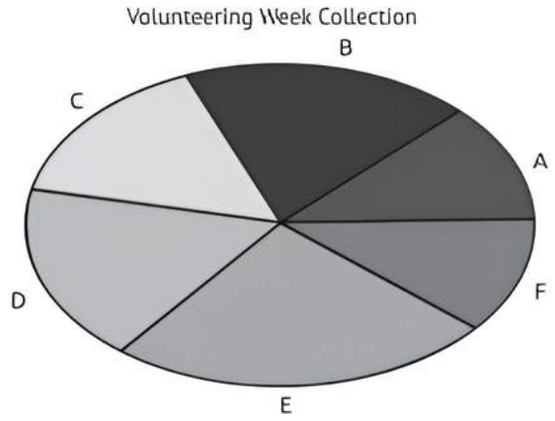


Ans.
```
import matplotlib . pyplot as plt
qtr=['QTR1', 'QTR2', 'QTR3', 'QTR4']
sales=[450, 300, 500, 650]
plt . plot (qtr, sales)
plt . title ("Sales each quarter")
plt . xlabel ("Quarter")
plt . ylabel ("Sales")
plt . show ( )
```

Q 10. TSS school celebrated volunteering week where each section of class XI dedicated a day for collecting amount for charity being supported by the school. Section A volunteered on Monday, B on Tuesday, C on Wednesday and so on. There are six sections in class XI. Amounts collected by sections A to F are 8000, 12000, 9800, 11200, 15500, 7300. Write a program to create a pie chart showing collection amount section wise.

Ans.
```
import matplotlib.pyplot as plt
Col = [8000, 12000, 9800, 11200, 15500, 7300]
Section = ['A', 'B', 'C', 'D', 'E', 'F']
plt.title ("Volunteering Week Collection")
plt.pie(Col, labels = Section)
plt.show( )
```

Output:



Q 11. Consider the following dataframe *prodf*:

	Fruits	Pulses	Rice	Wheat
Andhra P.	7830.0	931.0	7452.4	NaN
Gujarat	11950.0	818.0	1930.0	2737.0
Kerala	113.1	1.7	2604.8	NaN
Punjab	7152.0	33.0	11586.2	16440.5
Tripura	44.1	23.2	814.6	0.5
Uttar P.	24169.2	2184.4	13754.0	30056.0

Write a program to plot a scatter chart with the columns *Pulses*.

Ans. Since the given dataframe **prodf** does not have number index, we cannot use it for plotting because x has the numeric for **scatter( )**.

We can create numeric values for *x*-axis as:

```
x = range (0, len (prodf) )        # will generate [0, 1, 2, 3, 4, 5]
```

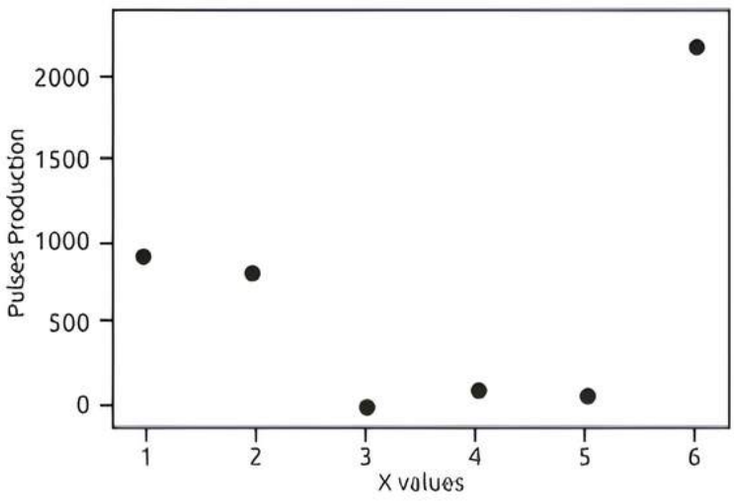
Or

```
x = range (1, len (prodf) + 1)     # will generate [1, 2, 3, 4, 5, 6]
```

Now the program code for above given problem will be:

```
import matplotlib.pyplot as plt
: # prodf created or loaded
plt. scatter (x = range (1, len (prodf) + 1), y = prodf. Pulses)
plt. xlabel ("X values")
plt. ylabel ("Pulses production")
plt. show ( )
```

Output:



Chapter Test

Multiple Choice Questions

Q 1. Which argument of boxplot( ) is used to create a filled boxplot?
a. fill b. box
c. patch_artish d. patch

Q 2. A..... is a plot that shows the underlying frequency distribution of a set of continuous data.
a. Histogram b. Pyplot
c. Bar chart d. Pie chart

Q 3. Pyplot interface is a collection of methods within ..... library of Python.
a. Matplotlib b. Numpy
c. Pandas d. None of these

Q 4. Pyplot's ..... function is used to create line charts.
a. plot( ) b. barh( )
c. scatter( ) d. hist( )

Q 5. Pyplot's ..... function is used to create horizontal bar charts.
a. plot( ) b. barh( )
c. scatter( ) d. hist( )

Fill in the Blanks

Q 6. The ..... argument of plot( ) specifies the style of the line.

Q 7. The ..... argument of bar( ) specifies the bar width.

Q 8. The ..... function is used to specify ticks for x-axis.

Assertion & Reason Type Questions

Directions (Q. Nos. 9-11): *In the questions given below, there are two statements marked as Assertion (A) and Reason (R). Read the statements and choose the correct option.*

a. Both Assertion (A) and Reason (R) are true and Reason (R) is the correct explanation of Assertion (A).
b. Both Assertion (A) and Reason (R) are true, but Reason (R) is not the correct explanation of Assertion (A).
c. Assertion (A) is true, but Reason (R) is false.
d. Assertion (A) is false, but Reason (R) is true.

Q 9. Assertion (A): Using Python Matplotlib, histogram can be used to count how many values fall into each interval.
Reason (R): Pyplot's barh( ) function is used to create line chart.

Q 10. Assertion (A): Statement import pyplot.matplotlib is a valid statement for working on pyplot functions.
Reason (R): By default, pie chart is printed in elliptical or oval shape.

Q 11. Assertion (A): The axes define the area (mostly rectangular in shape for simple plots) on which actual plot (line or bar or graph etc.) will appear. Axis have properties like label, limits and tick marks on them.
Reason (R): A Boxplot is the visual representation of a statistical 5 number summary of a given data set.

Case Study Based Questions

Q 12. Data Visualisation is the process of presenting data in the form of graphs or charts. It helps to understand large and complex amounts of data very easily. It allows the decision-makers to make decisions very efficiently and also allows them in identifying new trends and patterns very easily. It is also used in high-level data analysis for Machine Learning and Exploratory Data Analysis (EDA). Data visualisation can be done with various tools like Tableau, Power BI and Python.

(i) The value of ..... chart is calculated in terms of percentage.
a. Histogram b. Boxplot
c. Bar d. Pie

(ii) By default Plot() function plots a:
a. Bar chart b. Pie chart
c. Line chart d. Horizontal bar chart

(iii) Symbol which represents single data value(points) in a chart is called:
a. ticks b. marker
c. title d. data labels

(iv) Raju want to create scatter chart for the given data:
X=[1,2,3,4]
Y=[12,23,16,28]
Help him to write correct code (important libraries are imported)
a. Plt.plot(x,y,'o')
b. Plt.scatter(x,y)
c. Df=pd.DataFrame({'c1':x,'c2':y})
Df.plot(kind='scatter',x='c1',y='c2')
d. All of the above

(v) Which of the following argument cannot used with hist()?
a. Bin b. Width
c. Histtype d. Cumulative

Q 13. Data Visualisation refers to the graphical or visual representation of information and data using visual elements like charts, graphs and maps, etc. The purpose of plotting data is to visualise or show relationships between variables. The matplotlib is a Python library that provides many interfaces and functionality for 2D– graphics similar to MATLAB's in various forms. It provides both a very quick way to visualise data from Python and publication – quality figures in many formats. Pyplot: It is a collection of methods within matplotlib which allows user to construct 2D plots easily and interactively.

Matplotlib can be installed using the following pip command from the command prompt:

pip install matplotlib

For plotting using Matplotlib, we need to import its pyplot module using the following command:

import matplotlib.pyplot as plt

(i) Define Boxplot().

(ii) What do you mean by the term outlier?

(iii) What are the five important numbers of a data range used by box plot?

(iv) Define loc arguments of legends().

(v) Define markers.

Very Short Answer Type Questions

Q 14. What do you mean by tick marks?

Q 15. Write the correct syntax for saving a figure.

Q 16. Fill in the blank with the correct statement to plot a bar graph using a matplotlib method, so that company ABC can see the graphical presentation of its profit figures for the 2nd quarter of the financial year 2019 (*i.e.* August, September, October, November).

```
import matplotlib.pyplot as mtp
Months = ['AUG','SEP','OCT','NOV'] #x-axis
Profits = [125, 220, 230, 175] # y-axis
mtp. show()
```

Short Answer Type-I Questions

Q 17. What is frequency polygon?

Q 18. Write all the steps to draw frequency polygon manually.

Short Answer Type-II Questions

Q 19. Three days' prices are available in three lists as shown below:

Day1 = [74.25, 76.06, 69.5, 72.55]

Day2 = [56.03, 68.71, 62.89, 56.42]

Day3 = [59.3, 72.07, 77.65, 66.46]

Write a program to create filled boxplot from this data with their unique labels.

Q 20. A survey gathers height and weight of 100 participants and recorded the participants' ages as:

ages = [1, 1, 2, 3, 5, 7, 8, 9, 10, 10, 11, 13, 13, 15, 16, 17, 18, 19, 20, 21, 21, 23, 24, 24, 24, 25, 25, 25, 25, 26, 26, 26, 27, 27, 27, 27, 27, 29, 30, 30, 30, 30, 31, 33, 34, 34, 34, 35, 36, 36, 37, 37, 37, 38, 38, 39, 40, 40, 41, 41, 42, 43, 45, 45, 46, 46, 46, 47, 48, 48, 49, 50, 51, 51, 52, 52, 53, 54, 55, 56, 57, 58, 60, 61, 63, 65, 66, 68, 70, 72, 74, 75, 77, 81, 83, 84, 87, 89, 90, 91]

Write a program to plot a histogram from above with 20 bins.

Long Answer Type Questions

Q 21. Answer the following questions:

(i) Is it possible to show multiple plots in the same visualisation? Show with an example.

(ii) Write a python program to show sine chart with line colour as red, line width as "4 pixels" and dashed line style.

(iii) Execute the following code and find out what happens.

```
A=np.arange(2,20,2)
B=np.log(A)
plt.pie(A,B)
```

will code produce any error? Why?

Q 22. Give four sequences as given below:

```
X=[1,2,3,4]
Y=[10,20,25,30]
A=[0.3,3.8,1.2,2.5]
B=[11,25,9,26]
```

Write a program to plot in the same chart as:

- A line graph plotted with *X* and *Y* with blue colour and having line width as 3.
- A scatter graph plotted with *A* and *B* with triangular marker of magenta colour.