

# Data Visualization

**Data visualization** is the presentation of data in graphical format. It helps people understand the significance of data by summarizing and presenting a huge amount of data in a simple and easy to understand format and helps communicate information clearly and effectively

## Matplotlib

It is an amazing visualization library in Python that used for 2D plots of arrays. It is a multi-platform data visualization library which build NumPy arrays. Matplotlib 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 shell, web application servers and various graphical user interface toolkits.

To get matplotlib up and running in our environment, we need to import it.

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

### Case Study Based MCQs

#### 1. Consider the following program and answer any four questions

from (i) to (v):

```
1. import _____ as plt
2. plt.bar ([2,3,4,8,1],[2,4,7,3,5], label= _____ )
3. plt.legend( )
4. plt.xlabel(_____)
5. plt.ylabel('Height')
6. plt._____ ('Vertical Bar Chart')
7. _____
```

(i) Which Module will be imported in Line 1 for above code?

- (a) matplotlib
- (b) matplotlib.pyplot
- (c) plotlib
- (d) None of these

**Ans: (b)**

(ii) Name the label that can be used to represent the bar chart in Line 2.

- (a) Data
- (b) Data Values
- (b) Values for X axis
- (d) All of these

**Ans: (d)**

(iii) Which message is best suited for xlabel?

- (a) X values                      (b) Y values  
(c) Legend                      (d) Vertical

**Ans: (a)**

(iv) Which method will take place at Line 6 for setting heading on the top of Chart?

- (a) Title ()                      (b) title()  
(c) Head()                      (d) All of these.

Ans: (b)

(v) Choose the statement to be placed at Line 7 of the above code.

- (a) plt.print()                      (b) plt.show()  
(c) plt.display()                      (d) plot.show()

**Ans: (b)**

**2. If you are given to plot a histogram using numpy array as per the code given below then answer any of four question from (i) to (v)**

```
from matplotlib import _____ as plt
import numpy as np
fig, ax = plt._____ (1, 1)
a= np.array([26,59,44,39,76,16,23,11,18,78])
ax.hist(a, bins=[0,10,20,30,40])
ax._____ ('Histogram')
ax.set_xticks ([0,10,20,30,40, ])
ax.set_xlabel('Percentage')
ax._____ ('Students')
_____
```

**(i)** Choose the correct option to import for given program:

- (a) matplotlib                      (b) matplot

- (c) pyplot (d) plot

**Ans: (c)**

(ii) Fill in the blank at Line 3

- (a) subplots (b) subplot  
(c) plot (d) subplt

**Ans: (a)**

(iii) Which statement is used to set title in Line 6?

- (a) title (b) set\_title  
(c) set (d) Title

**Ans: (b) set\_title**

(iv) How to set Y-axis label?

- (a) label () (b) set\_y  
(c) set\_ylab () (d) set\_ylabel

**Ans: (d) set\_ylabel**

(v) To fill in blank on Line 10 for showing histogram what can be used?

- (a) plt.show() (b) plt\_show()  
(c) plot\_show() (d) plt.show

**Ans: plt.show()**

**3. Consider the following case and answer the from (i) to (v)**

```
import ..... as pd
import matplotlib. _____ as plt
data= {'Name' : ['Karan', 'Adi', 'Abhinav', 'Kirti', 'Rahul' ],
'Height' : [60,61,63,65,61],
'Weight' : [47,89,52,58,50,47]}
df=pd. _____ (data)
```

df.\_\_\_\_\_ (Kind = 'hist', edgecolor = 'Green', linewidth = 2, linestyle = ':', fill = False)

---

(i) Fill in the blank in Line 1.

- (a) numpy                      (b) pandas  
(c) Python                      (d) matplotlib

Ans: (b) pandas

(ii) Fill in the blank in Line 2.

- (a) pyplot                      (b) plot  
(c) pyp                          (d) None of these

Ans: (a) pyplot

(iii) Which of the following is used in Line 6 to represent the data?

- (a) Series                      (b) Matplot  
(c) DataFrame                      (d) Plot

Ans: (c) DataFrame

(iv) For blank of Line 7 command used may be:

- (a) plt                          (b) pyplot  
(c) plot                          (d) figure

Ans: (c) plot

(v) To show the above graph which statement is used in Line 8?

- (a) plot.show                      (b) plot.show()  
(c) plt.show                      (d) plt.show()

Ans: (d) plt.show()

1. What is true about Data Visualization?

- A. Data Visualization is used to communicate information clearly and efficiently to users by the usage of information graphics such as tables and charts.
- B. Data Visualization helps users in analyzing a large amount of data in a simpler way.
- C. Data Visualization makes complex data more accessible, understandable, and usable.
- D. All of the above

Ans: D

Explanation: Data Visualization is used to communicate information clearly and efficiently to users by the usage of information graphics such as tables and charts. It helps users in analyzing a large amount of data in a simpler way. It makes complex data more accessible, understandable, and usable.

2. Which method is used to save the output of pyplot in the form of image file?

- A. `savefig('filename')`
- B. `save_fig('filename')`
- C. `save_figure('filename')`
- D. `save_img('filename')`

Answer: A

3. Data visualization is also an element of the broader \_\_\_\_\_.

- A. deliver presentation architecture
- B. data presentation architecture
- C. dataset presentation architecture
- D. data process architecture

Answer: B

4. The command used to give a heading to a graph is \_\_\_\_\_
- A. plt.show()
  - B. plt.plot()
  - C. plt.xlabel()
  - D. plt.title()

Answer: D

5. Common use cases for data visualization include?

- A. Politics
- B. Sales and marketing
- C. Healthcare
- D. All of the above

Answer: D

6. Which are pros of data visualization?

- A. It can be accessed quickly by a wider audience.
- B. It can misrepresent information
- C. It can be distracting
- D. None Of the above

Ans: A

Explanation: Pros of data visualization: it can be accessed quickly by a wider audience.

7. Which are cons of data visualization?

- A. It conveys a lot of information in a small space.
- B. It makes your report more visually appealing.
- C. visual data is distorted or excessively used.
- D. None Of the above

Ans: C

Explanation: It can be distracting: if the visual data is distorted or excessively used.

8. Which one of the following is most basic and commonly used techniques?

- A. Line charts
- B. Scatter plots
- C. Population pyramids
- D. Area charts

Answer: A

Explanation: Line charts. This is one of the most basic and common techniques used. Line charts display how variables can change over time.

9. Which is a python package used for 2D graphics?

- A. matplotlib.pyplot
- B. matplotlib.pip
- C. matplotlib.numpy
- D. matplotlib.plt

ANS: A

10. Observe the output figure. Identify the coding for obtaining this output.

A. import matplotlib.pyplot as plt

```
plt.plot([1,2,3],[4,5,1])
```

```
plt.show()
```

B. import matplotlib.pyplot as plt

```
plt.plot([1,2],[4,5])
```

```
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()
```

ANS: A

12. Identify the right type of chart using the following hints.

Hint 1: This chart is often used to visualize a trend in data over intervals of time.

Hint 2: The line in this type of chart is often drawn chronologically.

- A. Line chart
- B. Bar chart
- C. Pie chart
- D. Scatter plot

ANS: A

13. Which of the following method will be add inside the file to display plot?

- A. show()
- B. display()
- C. plot()
- D. execute()

ANS: show()

14. Which method is used to plot horizontal bar graph in pyplot?

- A. horizontal\_bar()
- B. barh()
- C. hbar()
- D. bar()

ANS: C

15. Which method is used to plot histogram in pyplot?

- A. his()
- B. hist()
- C. Hist()
- D. histogram()

ANS: B