CHAPTER 16

۲

DATA VISUALIZATION USING PYPLOT: LINE CHART, PIE CHART AND BAR CHART

Q Learning Objectives

Unit

After learning this chapter, the learners will be able to

- Define the term Data Visualization.
- List the types of Data Visualization.
- List the uses of Data Visualization.
- List the types of Visualizations in Matplotlib.
- Practice installing Matplotlib.
- Explore importing Matplotlib.
- Classify the types of Data Visualization plots.
- Practice creating various types of plots using Matplotlib.

16.1 Data Visualization Definition ⊢

Data Visualization is the graphical representation of information and data. The objective of Data Visualization is to communicate information visually to users. For this, data visualization uses statistical graphics. Numerical data may be encoded using dots, lines, or bars, to visually communicate a quantitative message.

General types of Data Visualization

• Charts

۲

- Tables
- Graphs
- Maps
- Infographics
- Dashboards

Data visualization - Uses

• Data Visualization help users to analyze and interpret the data easily.



۲

- It makes complex data understandable and usable.
- Various Charts in Data Visualization helps to show relationship in the data for one or more variables.

Infographics \rightarrow An infographic (information graphic) is the representation of information in a graphic format.

Dashboard \rightarrow A dashboard is a collection of resources assembled to create a single unified visual display. Data visualizations and dashboards translate complex ideas and concepts into a simple visual format. Patterns and relationships that are undetectable in text are detectable at a glance using dashboard.

Introduction to Matplotlib — Data Visualization in Python

Matplotlib is the most popular data visualization library in Python. It allows you to create charts in few lines of code.

Types of Visualizations in Matplotlib

There are many types of Visualizations under Matplotlib. Some of them are:

• Line plot

۲

- Scatter plot
- Histogram
- Box plot
- Bar chart and
- Pie chart

Scatter plot: A scatter plot is a type of plot that shows the data as a collection of points. The position of a point depends on its two-dimensional value, where each value is a position on either the horizontal or vertical dimension.

Box plot: The box plot is a standardized way of displaying the distribution of data based on the five number summary: minimum, first quartile, median, third quartile, and maximum.

Installing Matplotlib

You can install matplotlib using pip. Pip is a management software for installing python packages.



<mark>16.2</mark> Getting Started ⊢

After installing Matplotlib, we will begin coding by importing Matplotlib using the command:

import matplotlib.pyplot as plt

Now you have imported Matplotlib in your workspace. You need to display the plots. Using Matplotlib from within a Python script, you have to add plt.show() function inside the file to display your plot.

Example

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

Output

۲

This window is a matplotlib window, which allows you to see your graph. You can hover the graph and see the coordinates in the bottom right.



You may be wondering why the x-axis ranges from 0-3 and the y-axis from 1-4. If you provide a single list or array to the plot () command, matplotlib assumes it is a sequence of y values, and automatically generates the x values for you. Since python ranges start with 0, the default x vector has the same length as y but starts with 0. Hence the x data are [0, 1, 2, 3].

۲

plot() is a versatile command, and will take an arbitrary number of arguments.

Program

For example, to plot x versus y, you can issue the command: import matplotlib.pyplot as plt plt.plot([1,2,3,4], [1,4,9,16]) plt.show()

This .plot takes many parameters, but the first two here are 'x' and 'y' coordinates. This means, you have 4 co-ordinates according to these lists: (1,1), (2,4), (3,9) and (4,16).



Figure 16.2

323

۲

۲

Plotting Two Lines

To plot two lines, use the following code: import matplotlib.pyplot as plt

```
x = [1,2,3]
y = [5,7,4]
x2 = [1,2,3]
y2 = [10,14,12]
plt.plot(x, y, label='Line 1')
plt.plot(x2, y2, label='Line 2')
plt.xlabel('X-Axis')
plt.ylabel('Y-Axis')
plt.title('LINE GRAPH')
plt.legend()
plt.show()
```

Output

۲

With plt.xlabel and plt.ylabel, you can assign labels to those respective axis. Next, you can assign the plot's title with plt.title, and then you can invoke the default legend with plt. legend().

۲



۲

Buttons in the output

In the output figure, you can see few buttons at the bottom left corner. Let us see the use of these buttons.

۲



Figure 16.4

Home Button \Rightarrow The Home Button will help once you have begun navigating your chart. If you ever want to return back to the original view, you can click on this.

Forward/Back buttons \rightarrow These buttons can be used like the Forward and Back buttons in your browser. You can click these to move back to the previous point you were at, or forward again.

Pan Axis \Rightarrow This cross-looking button allows you to click it, and then click and drag your graph around.

Zoom \rightarrow The Zoom button lets you click on it, then click and drag a square that you would like to zoom into specifically. Zooming in will require a left click and drag. You can alternatively zoom out with a right click and drag.

Configure Subplots \rightarrow This button allows you to configure various spacing options with your figure and plot.

Save Figure \rightarrow This button will allow you to save your figure in various forms.

<mark>- 16.3</mark> Special Plot Types

Matplotlib allows you to create different kinds of plots ranging from histograms and scatter plots to bar graphs and bar charts.

Line Chart

۲

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 Line Chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically.

325

۲

XII Std - CS EM Chapter-16.indd 325

Example: Line plot

import matplotlib.pyplot as plt

years = [2014, 2015, 2016, 2017, 2018]

total_populations = [8939007, 8954518, 8960387, 8956741, 8943721]

۲

plt.plot (years, total_populations)

plt.title ("Year vs Population in India")

plt.xlabel ("Year")

plt.ylabel ("Total Population")

plt.show()

In this program,

 $Plt.title() \rightarrow specifies title to the graph$

 $Plt.xlabel() \rightarrow specifies label for X-axis$

 $Plt.ylabel() \rightarrow specifies label for Y-axis$

Output

۲



Figure 16.5

Bar Chart

A BarPlot (or BarChart) is one of the most common type of plot. It shows the relationship between a numerical variable and a categorical variable.

Bar chart represents categorical data with rectangular bars. Each bar has a height corresponds to the value it represents. The bars can be plotted vertically or horizontally. It's useful when we want to compare a given numeric value on different categories. To make a bar chart with Matplotlib, we can use the plt.bar() function.

XII Std Computer Science

326

Example

import matplotlib.pyplot as plt

Our data

labels = ["TAMIL", "ENGLISH", "MATHS", "PHYSICS", "CHEMISTRY", "CS"] usage = [79.8, 67.3, 77.8, 68.4, 70.2, 88.5]

۲

Generating the y positions. Later, we'll use them to replace them with labels.

y_positions = range (len(labels))

Creating our bar plot

plt.bar (y_positions, usage)

plt.xticks (y_positions, labels)

plt.ylabel ("RANGE")

plt.title ("MARKS")

plt.show()

Output

۲





The above code represents the following:

Labels \Rightarrow Specifies labels for the bars.

Usgae \rightarrow Assign values to the labels specified.

Xticks \rightarrow Display the tick marks along the x-axis at the values represented. Then specify the label for each tick mark.

Range \rightarrow Create sequence of numbers.

327

۲

Data Visualization using Pyplot

XII Std - CS EM Chapter-16.indd 327

08-12-2021 19:14:50

Bar Graph and Histogram are the two ways to display data in the form of a diagram.

Key Differences Between Histogram and Bar Graph

The differences between Histogram and bar graph are as follows

- 1. Histogram refers to a graphical representation; that displays data by way of bars to show the frequency of numerical data. A bar graph is a pictorial representation of data that uses bars to compare different categories of data.
- 2. A histogram represents the frequency distribution of continuous variables. Conversely, a bar graph is a diagrammatic comparison of discrete variables.
- 3. Histogram presents numerical data whereas bar graph shows categorical data.
- 4. The histogram is drawn in such a way that there is no gap between the bars. On the ot her hand, there is proper spacing between bars in a bar graph that indicates discontinuity.
- 5. Items of the histogram are numbers, which are categorised together, to represent ranges of data. As opposed to the bar graph, items are considered as individual entities.
- 6. In the case of a bar graph, it is quite common to rearrange the blocks, from highest to lowest. But with histogram, this cannot be done, as they are shown in the sequence of classes.
- 7. The width of rectangular blocks in a histogram may or may not be same while the width of the bars in a bar graph is always same.

Pie Chart

 (\bullet)

Pie Chart is probably one of the most common type of chart. It is a circular graphic which is divided into slices to illustrate numerical proportion. The point of a pie chart is to show the relationship of parts out of a whole.

To make a Pie Chart with Matplotlib, we can use the *plt.pie()* function. The autopct parameter allows us to display the percentage value using the Python string formatting.

Example

```
import matplotlib.pyplot as plt
sizes = [89, 80, 90, 100, 75]
labels = ["Tamil", "English", "Maths", "Science", "Social"]
plt.pie (sizes, labels = labels, autopct = "%.2f")
plt.axes().set_aspect ("equal")
plt.show()
```

XII Std Computer Science

۲

Output



Figure 16.7



- 1. Create a plot. Set the title, the *x* and *y* labels for both axes.
- 2. Plot a pie chart for your marks in the recent examination.
- 3. Plot a line chart on the academic performance of Class 12 students in Computer Science for the past 10 years.
- 4. Plot a bar chart for the number of computer science periods in a week.

XII Std - CS EM Chapter-16.indd 329

۲

08-12-2021 19:14:51

۲

	Evaluation
	Part - I
. Which is a python package use	d for 2D graphics?
a. matplotlib.pyplot	b. matplotlib.pip
c. matplotlib.numpy	d. matplotlib.plt
2. Identify the package manager f	for Python packages, or modules.
a. Matplotlib b	b. PIP c. plt.show() d. python package
8. Read the following code: Id option from the following	entify the purpose of this code and choose the right
C:\Users\YourName\AppDat version	a\Local\Programs\Python\Python36-32\Scripts>pip -
a. Check if PIP is Installed	b. Install PIP
c. Download a Package	d. Check PIP version
Read the following code: Id option from the following	entify the purpose of this code and choose the right
C:\Users\Your Name\AppI list	Data\Local\Programs\Python\Python36-32\Scripts>pip
a. List installed packages	b. list command
c. Install PIP	d. packages installed
5. To install matplotlib, the follo What does "-U"represents?	wing command will be typed in your command prompt.
Python –m pip install –U ma	ıtplotlib
a. downloading pip to the la	test version
b. upgrading pip to the lates	t version
c. removing pip	
d. upgrading matplotlib to t	he latest version
5. Observe the output figure. Iden	ntify 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()

۲

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

plt.show()

7. Read the code:

- a. import matplotlib.pyplot as plt
- b. plt.plot(3,2)
- c. plt.show()

Identify the output for the above coding.

۲

۲



- 8. Which key is used to run the module?
 - a. F6 b. F4 c. F3 d. F5
- 9. 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
- 10. Read the statements given below. Identify the right option from the following for pie chart.

Statement A: To make a pie chart with Matplotlib, we can use the plt.pie() function.

Statement B: The autopct parameter allows us to display the percentage value using the Python string formatting.

- a. Statement A is correct b. Statement B is correct
 - d. Both the statements are wrong
- c. Both the statements are correct

Part - II

- 1. Define: Data Visualization.
- 2. List the general types of data visualization.
- 3. List the types of Visualizations in Matplotlib.
- 4. How will you install Matplotlib?

XII Std Computer Science

۲

۲

5. Write the difference between the following functions: plt.plot([1,2,3,4]), plt. plot([1,2,3,4], [1,4,9,16]).

۲

Part - III

1. Draw the output for the following data visualization plot. import matplotlib.pyplot as plt plt.bar([1,3,5,7,9],[5,2,7,8,2], label="Example one") plt.bar([2,4,6,8,10],[8,6,2,5,6], label="Example two", color='g') plt.legend() plt.slabel('bar number') plt.ylabel('bar number') plt.ylabel('bar height') plt.title('Epic Graph\nAnother Line! Whoa') plt.show()

- 2. Write any three uses of data visualization.
- 3. Write the coding for the following:
 - a. To Install PIP in your PC.
 - b. To Check the version of PIP installed in your PC.
 - c. To list the packages in your system.
- 4. Write the plot for the following pie chart output.



Part - IV

- 1. Explain in detail the types of pyplots using Matplotlib.
- 2. Explain the various buttons in a matplotlib window.
- 3. Explain the purpose of the following functions:
 - a. plt.xlabel

۲

- b. plt.ylabel
- c. plt.title
- d. plt.legend()
- e. plt.show()

Reference

۲

1. https://towards datascience.com / data - science - with - python - intro -to- data -visualization-and-matplotlib-5f799b7c6d82.

۲

- 2. https://heartbeat.fritz.ai/introduction-to-matplotlib-data-visualization-in-python-d9143287ae39.
- 3. https://python programming.net / legends titles labels matplotlib tutorial/? completed=/matplotlib-intro-tutorial/.
- 4. https://keydifferences.com/difference-between-histogram-and-bar-graph.html.

۲