Data Science Internship

Project Level - Beginner

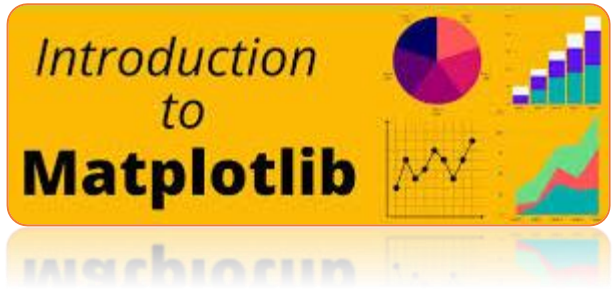
Guided By: Mr. Hari

Program Mentor From ShadowFox

Presented By: Sumit Sahu

Institute: Central University of Jammu

Objective: A documentation guide for popular Python visualization libraries



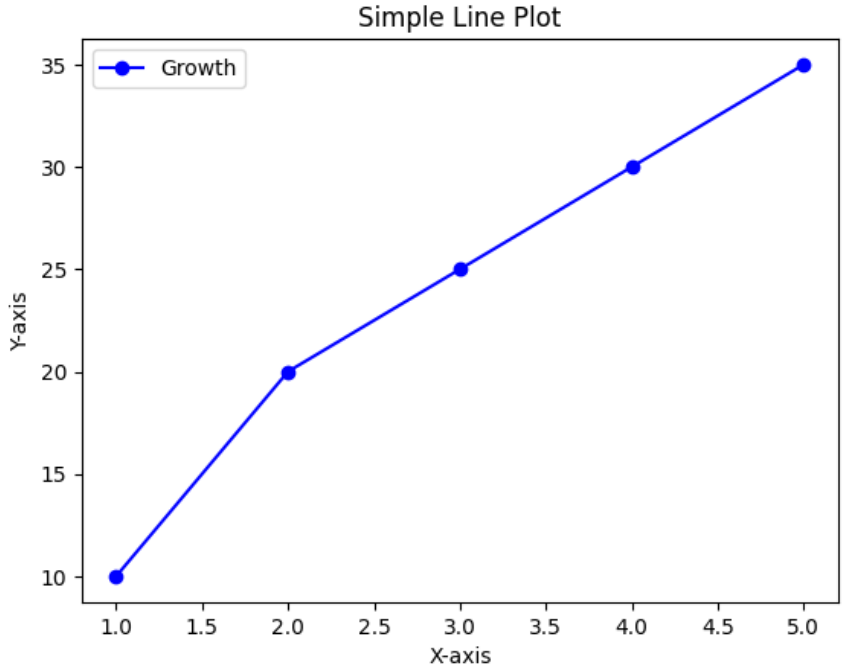
Matplotlib is a highly customizable plotting library that allows users to create a variety of static, animated, and interactive visualizations.

- **Common Graphs Matplotlib Can Create:**

1. **Line Plots** (ideal for trends over time)
2. **Bar Charts** (comparison between categories)
3. **Histograms** (distribution of data)
4. **Scatter Plots** (correlation between two variables)
5. **Pie Charts** (composition of a whole)

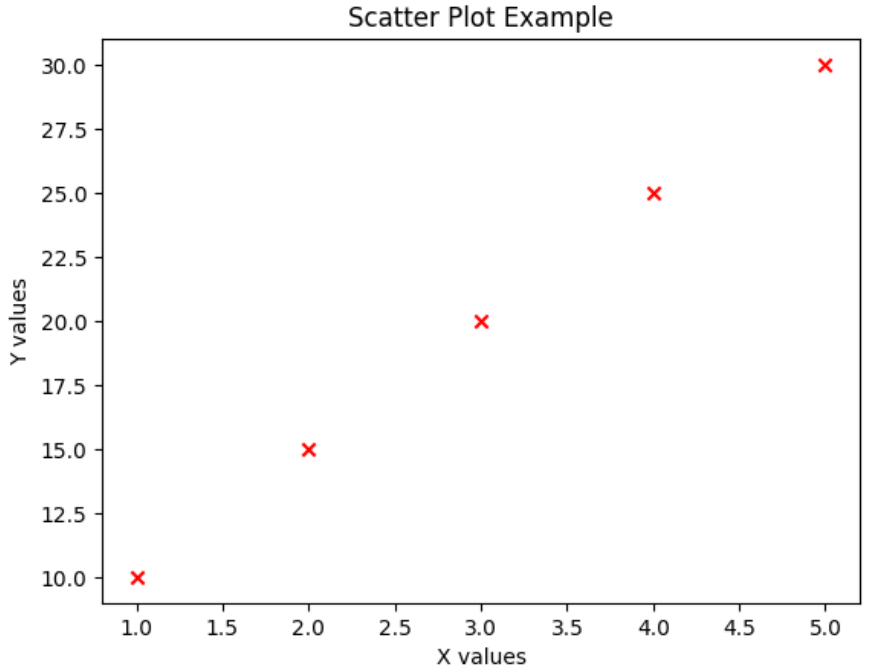
Example 01 : Line Plot

```
import matplotlib.pyplot as plt
x = [1, 2, 3, 4, 5]
y = [10, 20, 25, 30, 35]
plt.plot(x, y, marker='o', linestyle='-', color='b', label="Growth")
plt.xlabel('X-axis')
plt.ylabel('Y-axis')
plt.title('Simple Line Plot')
plt.legend()
plt.show()
```



Example 02 : Scatter Plot

```
import matplotlib.pyplot as plt
x = [1, 2, 3, 4, 5]
y = [10, 15, 20, 25, 30]
plt.scatter(x, y, color='red', marker='x')
plt.xlabel('X values')
plt.ylabel('Y values')
plt.title('Scatter Plot Example')
plt.show()
```



Pandas is a Python package designed to work with tabular data.

- **Two main data structures of Pandas are :**
  - **Series**, for one-dimensional data.
  - **DataFrame**, for two-dimensional data.

- **To use a module, we import it.**

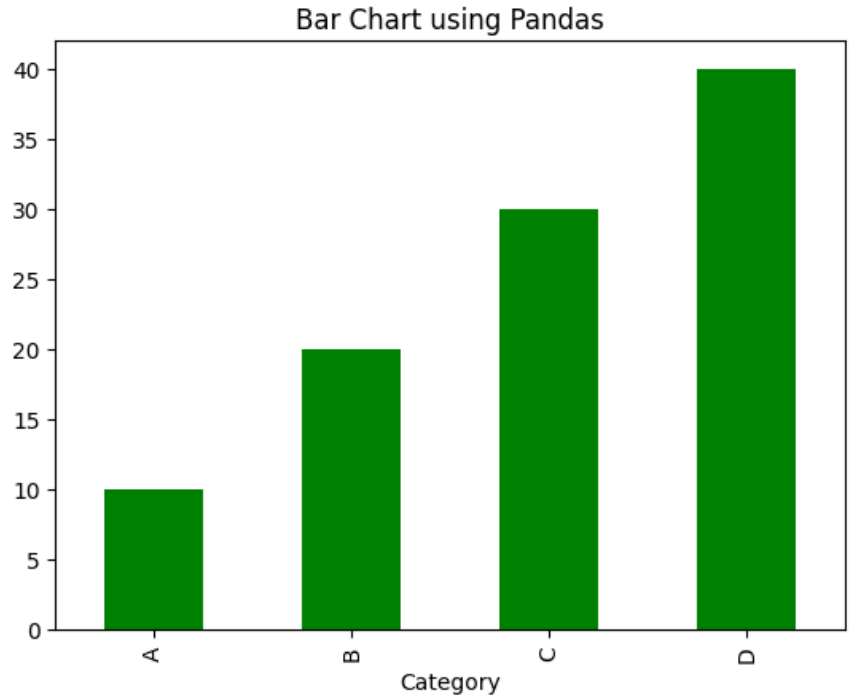
```
>>> import numpy as np
>>> import pandas as pd
>>> from matplotlib import pyplot as plt
```

Common Graphs Pandas Can Create:

1. **Line Plots** (quick visualization of Series/DataFrame)
2. **Bar Charts** (easy category visualization)
3. **Histograms** (automatic binning of data)
4. **Box Plots** (summary statistics and outliers)
5. **Scatter Plots** (simple point distribution)

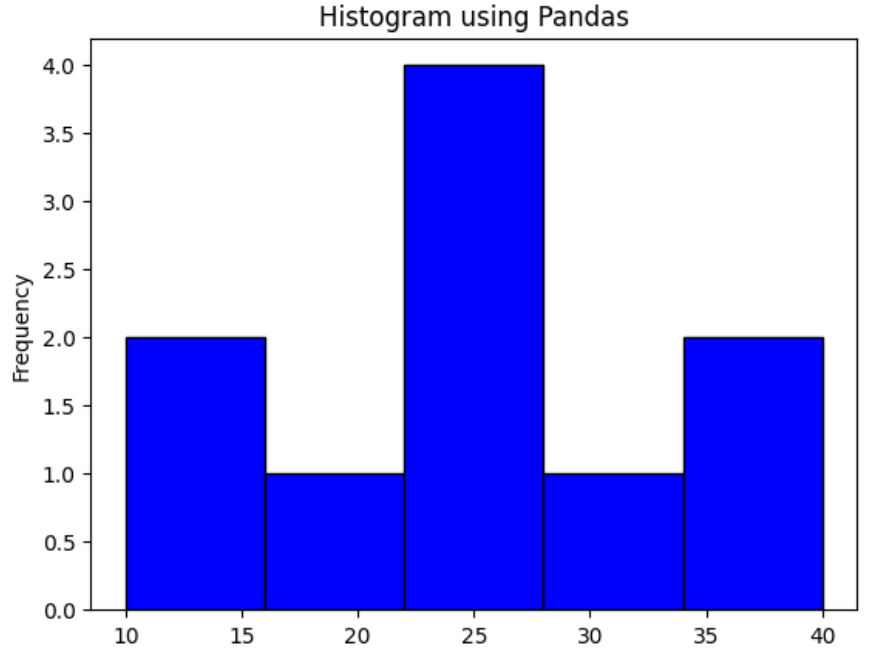
Example 01 : Bar Chart

```
import pandas as pd
import matplotlib.pyplot as plt
data = {'Category': ['A', 'B', 'C', 'D’], 'Values': [10, 20, 30, 40]}
df = pd.DataFrame(data)
df.plot(kind='bar', x='Category', y='Values', color='green',
legend=False)
plt.title('Bar Chart using Pandas’)
plt.show()
```



Example 02 : Histogram

```
import pandas as pd
import matplotlib.pyplot as plt
data = {'Values': [10, 15, 20, 22, 23, 24, 25, 30, 35, 40]}
df = pd.DataFrame(data)
df['Values'].plot(kind='hist', bins=5, color='blue’, edgecolor='black’)
plt.title('Histogram using Pandas’)
plt.show()
```



Comparison

Feature	Matplotlib	Pandas
Primary Purpose	Data Visualization	Data Analysis & Visualization
Customization	Highly customizable	Limited customization
Ease of Use	Requires more coding for complex plots	Quick, simple plots from DataFrames
Graph Types	Supports a wide variety	Limited to basic types
Integration	Works with Pandas, NumPy, etc.	Built-in integration with Pandas
Best Use Case	Complex, professional-grade plots	Quick exploratory data analysis

Thank You!

Learn, Create, Lead