Introduction to Matplotlib

Course Code: CPE 031	**Program**: Computer Engineering
Course Title: Visualization and Data Analysis	**Date Performed**: 22/10/2024
Section: CPE21S4	**Date Submitted**:22/10/2024
Name: Kenn Jie Valleser	**Instructor**: Engr. Ma. Rizette Sayo

Intended Learning Outcomes (ILO):

By the end of this laboratory session, learners will be able to:

1. Utilize Matplotlib's pyplot interface to create a variety of visualizations, including line plots, scatter plots, histograms, and box plots, demonstrating an understanding of the library's syntax and functionality.

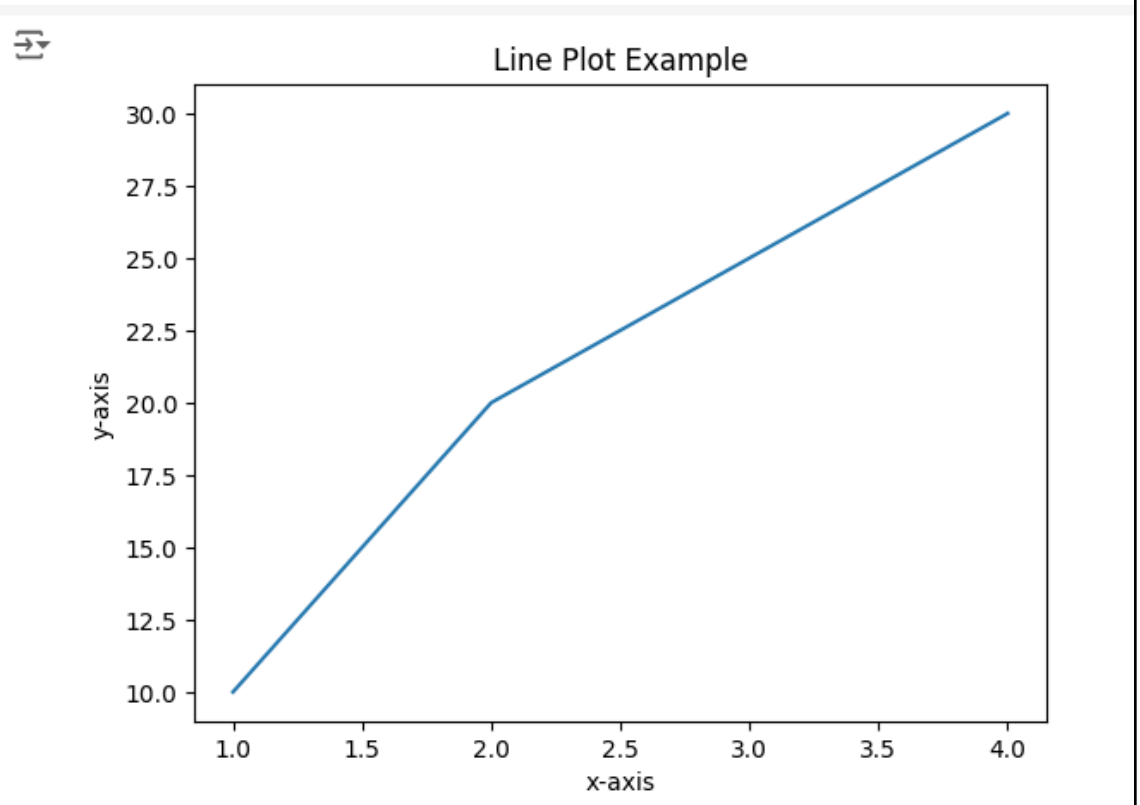
2. Customize visual elements such as titles, labels, and legends to enhance the clarity and aesthetics of their plots, applying best practices in data visualization.

3. Analyze and interpret visual data representations to extract meaningful insights, effectively communicating findings through well-structured graphical presentations.

Part 1: Perform the following codes, and understand the difference between line plot, scatter plot, histogram, bar chart, box plot, and pie chart using matplotlib’s pyplot sub-module. **(Provide a screenshot of your output.)**

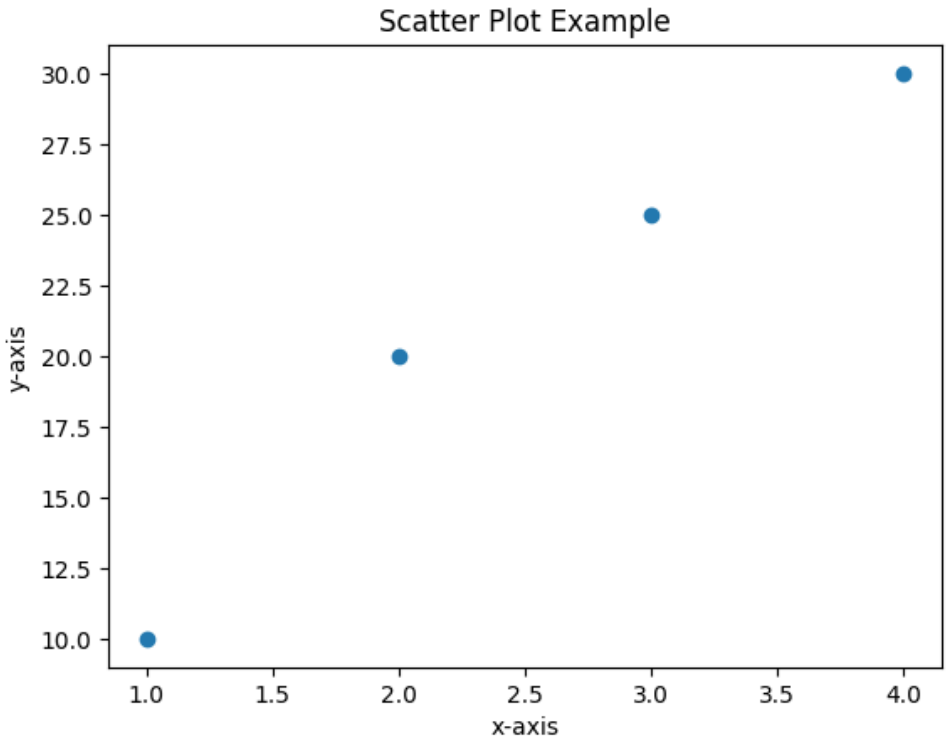
1. Line Plot

```python
import matplotlib.pyplot as plt

x = [1, 2, 3, 4]
y = [10, 20, 25, 30]
plt.plot(x, y)
plt.title("Line Plot Example")
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.show()
```



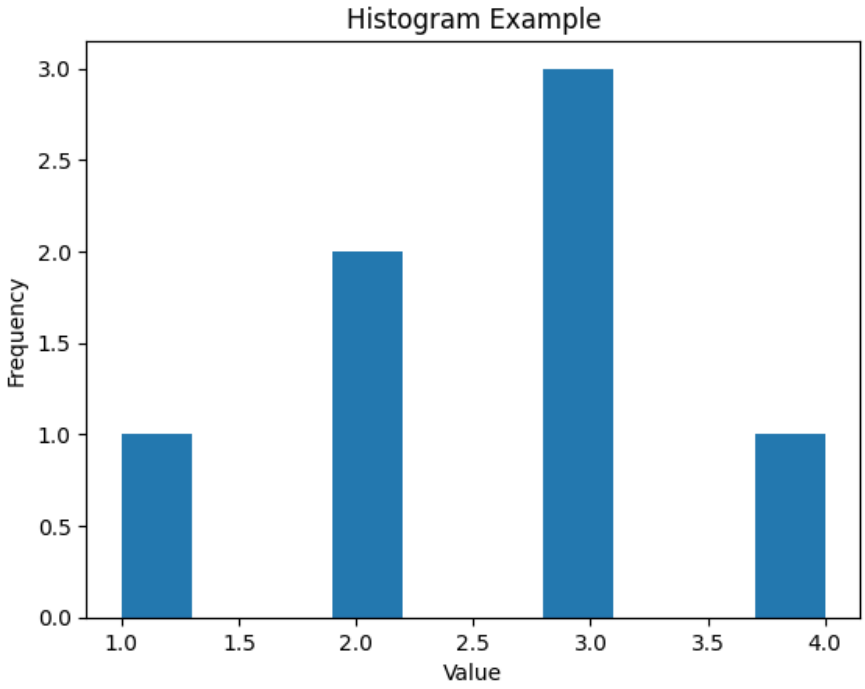
2. Scatter Plot

```python
import matplotlib.pyplot as plt

x = [1, 2, 3, 4]
y = [10, 20, 25, 30]
plt.scatter(x, y)
plt.title("Scatter Plot Example")
plt.xlabel("X-axis")
plt.ylabel("Y-axis")
plt.show()
```



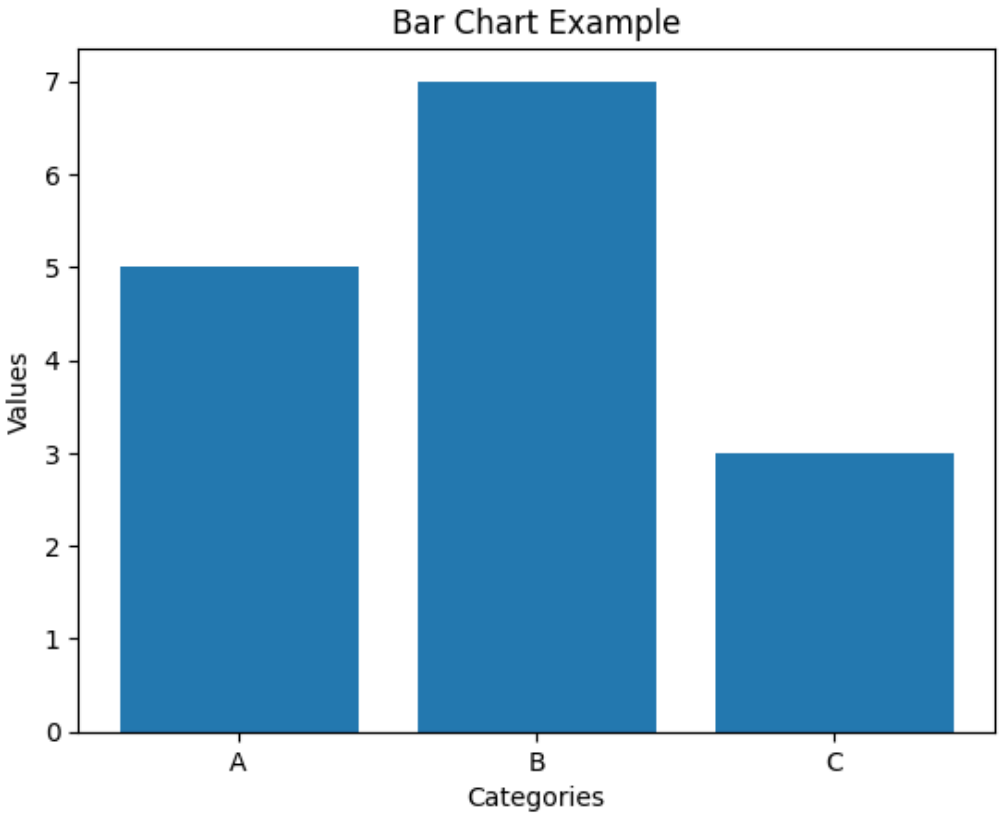
3. Histogram

```python
import matplotlib.pyplot as plt

data = [1, 2, 2, 3, 3, 3, 4]
plt.hist(data)
plt.title("Histogram Example")
plt.xlabel("Value")
plt.ylabel("Frequency")
plt.show()
```



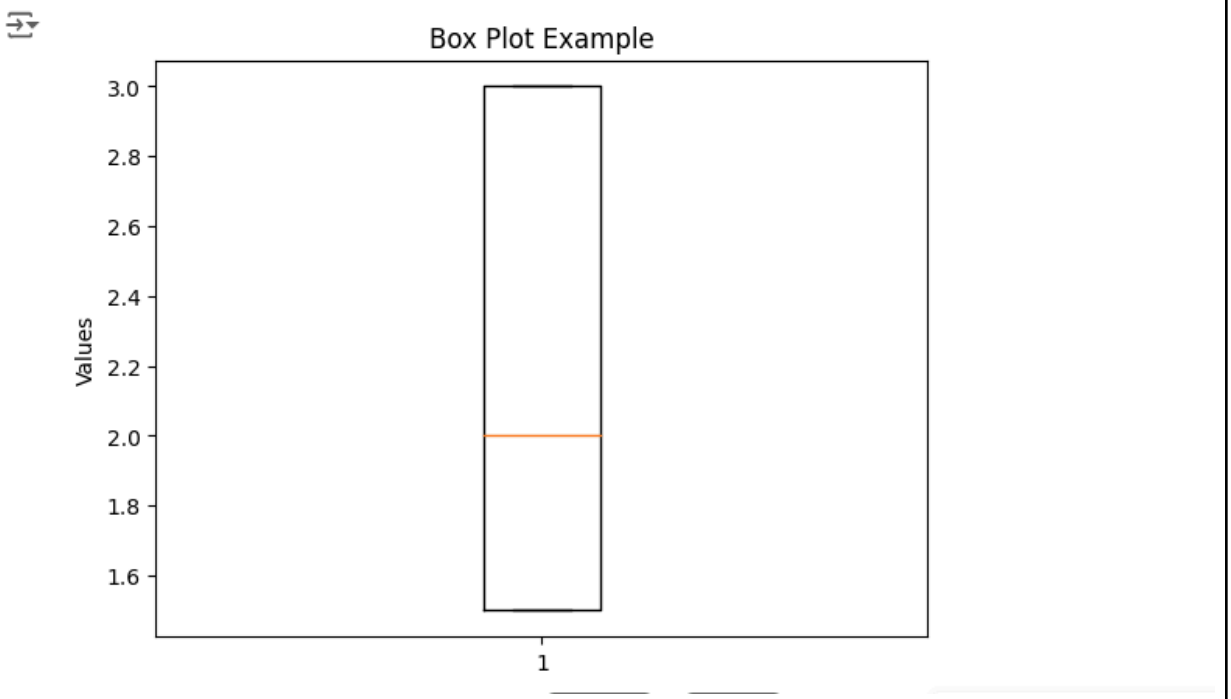
4. Bar Chart

```python
import matplotlib.pyplot as plt

categories = ['A', 'B', 'C']
values = [5, 7, 3]
plt.bar(categories, values)
plt.title("Bar Chart Example")
plt.xlabel("Categories")
plt.ylabel("Values")
plt.show()
```



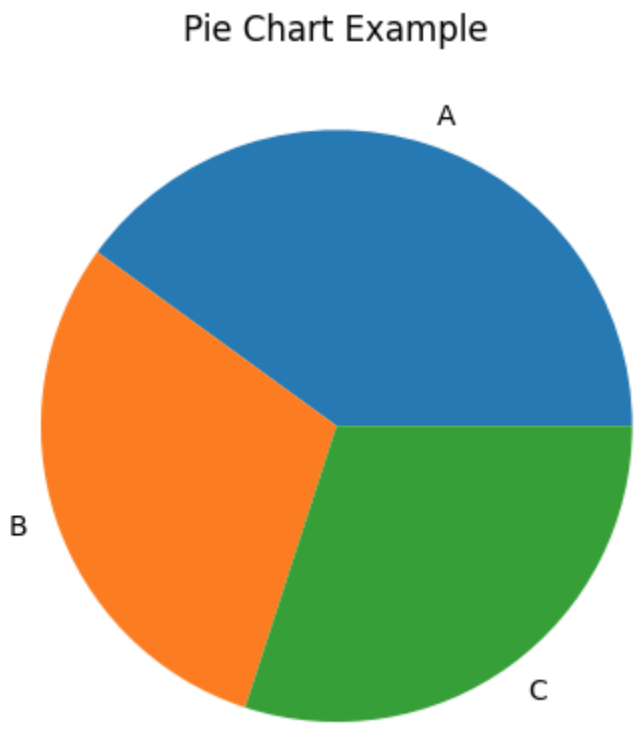
5. Box plot

```python
import matplotlib.pyplot as plt

data = [[1.5]*10 + [2]*10 + [3]*10]

plt.boxplot(data)

plt.title("Box Plot Example")
plt.ylabel("Values")
plt.show()
```



6. Pie chart

```python
import matplotlib.pyplot as plt

labels = ['A', 'B', 'C']
sizes = [40, 30, 30]

plt.pie(sizes, labels=labels)
plt.title("Pie Chart Example")
plt.show()
```



Part 2: Refer to the instructions below.

1. **Find a dataset for this activity**: Please visit Kaggle and look for a new dataset that would allow you to perform visualization and analysis using matplotlib.

2. **Creating a dataframe from your CSV file:** Once you have successfully loaded your dataset, you need to create a dataframe from your uploaded CSV file

	character	rarity	path	combat_type	wb_b-atk	wb_skill	wb_ult	er_b-atk	er_skill	er_ult	...	hp_40	atk_20a	def_20a	hp_20a	at
0	march_7th	4	preservation	ice	30	0	60	20	30	5	...	540	163.56	183.30	338	1
1	dan_heng	4	hunt	wind	30	60	90	20	30	5	...	450	174.84	126.90	282	1
2	himeko	5	erudition	fire	30	60	60	20	10	5	...	535	241.96	139.59	335	2
3	welt	5	nihility	imaginary	30	30	60	20	30	5	...	574	198.53	162.86	360	1
4	kafka	5	nihility	lightning	30	60	60	20	30	5	...	554	217.14	155.10	347	1

5 rows × 51 columns

3. **Import the matplotlib.pyplot**

4. **Based on your chosen dataset, you will develop three questions that you will answer using pyplot visualizations. This means that you will need to produce at least three pyplot visualizations. You are also required to make certain customizations on your data vizes.**

Who has the highest HP at lvl 80?

```
import pandas as pd
import matplotlib.pyplot as plt



df = pd.read_csv('hsr_character-data.csv')



plt.figure(figsize=(10, 6))
plt.bar(df['character'], df['hp_80'], alpha=0.7, color='salmon')
plt.title('Characters HP at Level 80')
plt.xlabel('Character Names')
plt.ylabel('hp at Level 80')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

Who has the highest DEF at lvl 80?

```
import pandas as pd
import matplotlib.pyplot as plt



df = pd.read_csv('hsr_character-data.csv')



plt.figure(figsize=(10, 6))
```

```
plt.bar(df['character'], df['def_80'], alpha=0.7,
color='lightgreen')
plt.title('Characters DEF at Level 80')
plt.xlabel('Character Names')
plt.ylabel('Def at Level 80')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

Who has the highest ATK at lvl 80?

```
import pandas as pd
import matplotlib.pyplot as plt



df = pd.read_csv('hsr_character-data.csv')



plt.figure(figsize=(10, 6))
plt.bar(df['character'], df['atk_80'], alpha=0.7, color='skyblue')
plt.title('Characters ATK at Level 80')
plt.xlabel('Character Names')
plt.ylabel('ATK at Level 80')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

5. Provide observations for each of your data viz, then **produce one insight not longer than five sentences given your three observations**. Your output shall follow this outline:
    a. Introduction (Describe your dataset)
    b. Questions
    c. Visualization and Observation
    d. Insight
    -The dataset that i used contains values for various attributes of the characters from the Honkai Star Rail Game. Each Attribute contributes to how well the characters work in the game. The Questions created were "**Who has the highest HP at lvl 80?", " Who has the highest ATK at lvl 80?", and "Who has the highest DEF at lvl 80?".** The results and observation for the visualization are that characters like Himeko and Kafka have the highest attack

stats, while March_7th and Welt lead in defense and HP, meaning they are more defensively oriented. My insight from the analysis is that characters like Himeko and Kafka are optimized for offensive roles with high attack, while March_7th and Welt are more suited for defensive roles, excelling in both defense and HP.

6. Your grade will depend on the quality of the question, difficulty/complexity of the visualization, and value-add of the insight that you will generate.