

Practical – 3

Aim: 1. Construction of Discrete and continuous frequency distribution 2. Construction of Bar Diagram, Histogram, Pie Diagram, Frequency curve and Frequency polygon

- Code for Discrete Frequency:

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

data_set = pd.read_csv("/content/drive/MyDrive/temp/dsc_data_prac_3.csv", nrows=11)
data_set.columns = ["Marks (out of 10)", "Number of Students (Frequency)"]

# it is necessary coz data is not properly formatted then this will not work, plot function will crash
marks_numeric = data_set["Marks (out of 10)"]

fig, axes = plt.subplots(2, 3, figsize=(18, 10))

# ! bar chart
axes[0, 0].bar(data_set["Marks (out of 10)"], data_set["Number of Students (Frequency)"],
color='blue', edgecolor='black')
axes[0, 0].set_xlabel("Marks (out of 10)")
axes[0, 0].set_ylabel("Number of Students (Frequency)")
axes[0, 0].set_title("Bar Chart of Marks Distribution")

# ! histogram
axes[0, 1].hist(data_set["Number of Students (Frequency)"], bins=6, color='purple',
edgecolor='black')
axes[0, 1].set_xlabel("Number of Students (Frequency)")
axes[0, 1].set_ylabel("Frequency Count")
axes[0, 1].set_title("Histogram of Student Count")

# ! pie chart
axes[0, 2].pie(data_set["Number of Students (Frequency)"], labels=data_set["Marks (out of 10)"],
autopct='%1.1f%%', colors=plt.cm.Set3.colors, startangle=140)
axes[0, 2].set_title("Pie Chart of Marks Distribution")

# ! curve chart
axes[1, 0].plot(marks_numeric, data_set["Number of Students (Frequency)"], marker='o',
linestyle='-', color='red', label="Frequency Curve")
axes[1, 0].set_xlabel("Marks (out of 10)")
axes[1, 0].set_ylabel("Number of Students (Frequency)")
axes[1, 0].set_title("Frequency Curve")
```

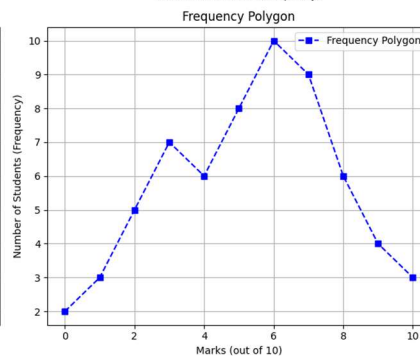
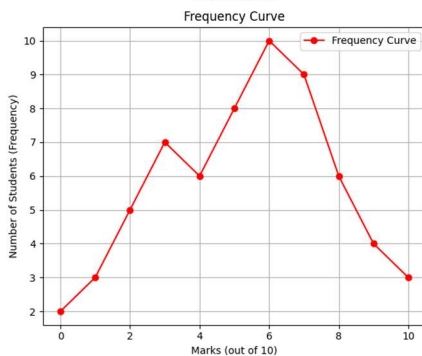
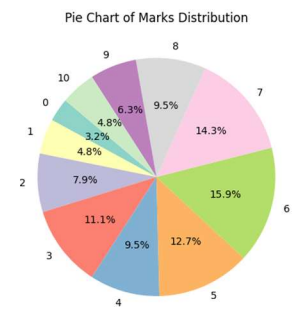
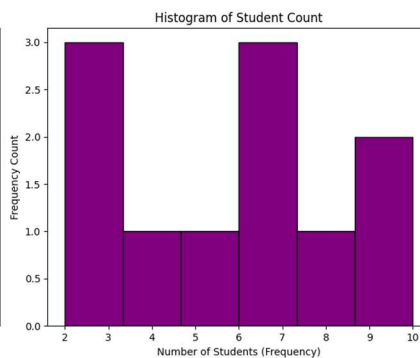
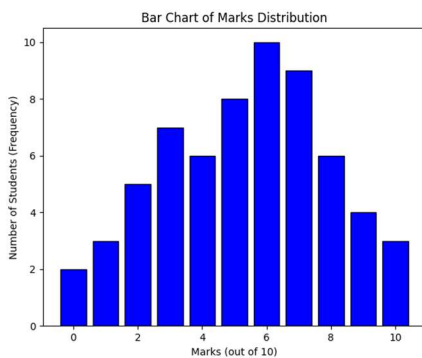
```
axes[1, 0].grid(True)
axes[1, 0].legend()
```

```
# ! polygon
axes[1, 1].plot(marks_numeric, data_set["Number of Students (Frequency)"], marker='s',
linestyle='--', color='blue', label="Frequency Polygon")
axes[1, 1].set_xlabel("Marks (out of 10)")
axes[1, 1].set_ylabel("Number of Students (Frequency)")
axes[1, 1].set_title("Frequency Polygon")
axes[1, 1].grid(True)
axes[1, 1].legend()
```

```
# Remove the last empty subplot (2,2)
fig.delaxes(axes[1, 2])
```

```
# ! removing the empty chart coz we made array of 3x3 and one was empty chart
plt.tight_layout()
plt.show()
```

- Output of Discrete Frequency



- Code for continuous frequency distribution:

```
data_set = pd.read_csv("/content/drive/MyDrive/temp/dsc_data_prac_3.csv", skiprows=14)
data_set.columns = ["Hour of the Day", "Number of Vehicles (Frequency)"]

# ! we need int data coz hour col is in alphanumeric so histograms, curves, and polygons need data
in int form
hours_numeric = np.arange(len(data_set))

# ! creating array for the plots
fig, axes = plt.subplots(2, 3, figsize=(18, 10))

# ! barchart
axes[0, 0].bar(data_set["Hour of the Day"], data_set["Number of Vehicles (Frequency)"],
color='blue', edgecolor='black')
axes[0, 0].set_xlabel("Hour of the Day")
axes[0, 0].set_ylabel("Number of Vehicles (Frequency)")
axes[0, 0].set_title("Bar Chart of Vehicles per Hour")
axes[0, 0].tick_params(axis='x', rotation=45)

# ! histogram
axes[0, 1].hist(data_set["Number of Vehicles (Frequency)"], bins=6, color='purple',
edgecolor='black')
axes[0, 1].set_xlabel("Number of Vehicles (Frequency)")
axes[0, 1].set_ylabel("Frequency Count")
axes[0, 1].set_title("Histogram of Vehicle Count")

# ! pie chart
axes[0, 2].pie(data_set["Number of Vehicles (Frequency)"], labels=data_set["Hour of the Day"],
autopct='%1.1f%%', colors=plt.cm.Set3.colors, startangle=140)
axes[0, 2].set_title("Pie Chart of Vehicles per Hour")

# ! curve chart
axes[1, 0].plot(hours_numeric, data_set["Number of Vehicles (Frequency)"], marker='o',
linestyle='-', color='red', label="Frequency Curve")
axes[1, 0].set_xticks(hours_numeric)
axes[1, 0].set_xticklabels(data_set["Hour of the Day"], rotation=45)
axes[1, 0].set_xlabel("Hour of the Day")
axes[1, 0].set_ylabel("Number of Vehicles (Frequency)")
axes[1, 0].set_title("Frequency Curve")
axes[1, 0].grid(True)
axes[1, 0].legend()

# ! polygon chart
axes[1, 1].plot(hours_numeric, data_set["Number of Vehicles (Frequency)"], marker='s',
linestyle='-', color='blue', label="Frequency Polygon")
axes[1, 1].set_xticks(hours_numeric)
```

```

axes[1, 1].set_xticklabels(data_set["Hour of the Day"], rotation=45)
axes[1, 1].set_xlabel("Hour of the Day")
axes[1, 1].set_ylabel("Number of Vehicles (Frequency)")
axes[1, 1].set_title("Frequency Polygon")
axes[1, 1].grid(True)
axes[1, 1].legend()

```

```

# ! removing the empty chart coz we made array of 3x3 and one was empty chart
fig.delaxes(axes[1, 2])

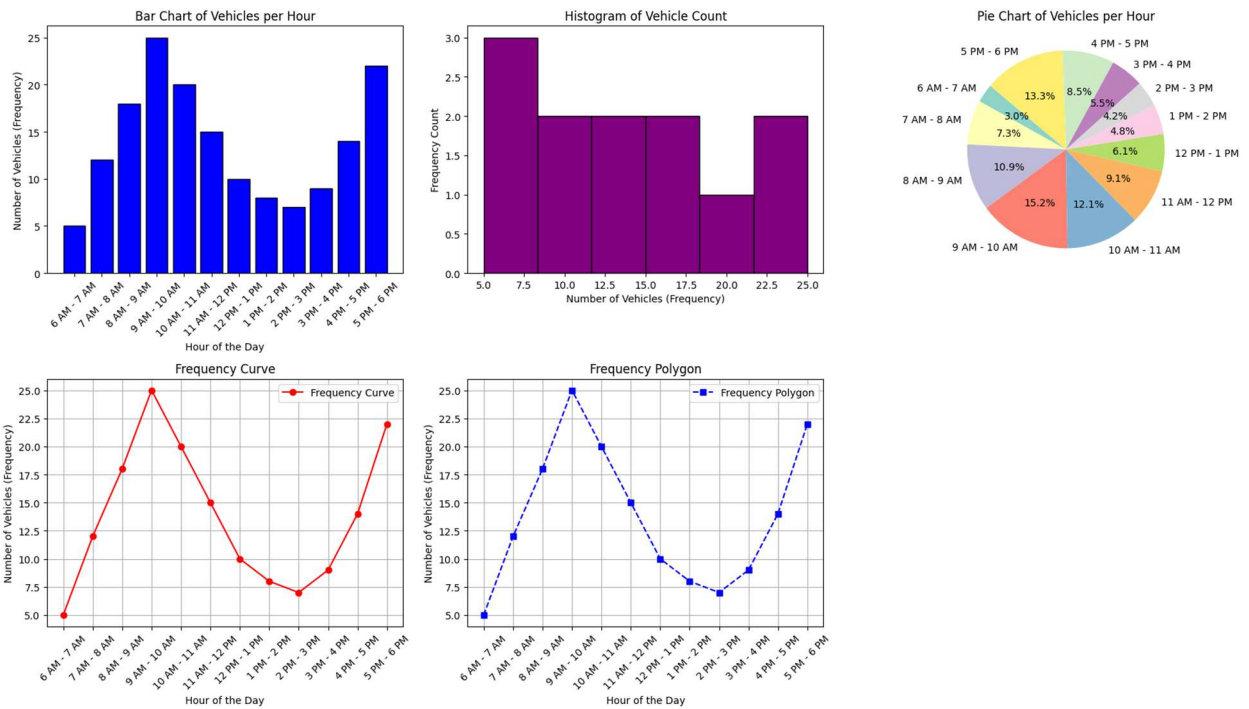
```

```

plt.tight_layout()
plt.show()

```

- Output of continuous frequency distribution:



Faculty Signature: _____

Date: _____