SKILLS

Assignment

Data Visualization

MATPLOTLIB ASSIGNMENT:

(Use Matplotlib for the visualization of the given questions)

1. Create a scatter plot using Matplotlib to visualize the relationship between two arrays, x and y for the given data.

```
x = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
y = [2, 4, 5, 7, 6, 8, 9, 10, 12, 13]
```

2. Generate a line plot to visualize the trend of values for the given data.

```
data = np.array([3, 7, 9, 15, 22, 29, 35])
```

3. Display a bar chart to represent the frequency of each item in the given array categories.

```
categories = ['A', 'B', 'C', 'D', 'E']
values = [25, 40, 30, 35, 20]
```

4. Create a histogram to visualize the distribution of values in the array data.

```
data = np.random.normal(0, 1, 1000)
```

5. Show a pie chart to represent the percentage distribution of different sections in the array `sections`.

```
sections = ['Section A', 'Section B', 'Section C', 'Section D']
sizes = [25, 30, 15, 30]
```

SEABORN ASSIGNMENT:

(Use Seaborn for the visualization of the given questions)

1. Create a scatter plot to visualize the relationship between two variables, by generating a synthetic dataset.

2. Generate a dataset of random numbers. Visualize the distribution of a numerical variable.

3. Create a dataset representing categories and their corresponding values. Compare different categories based on numerical values.

4. Generate a dataset with categories and numerical values. Visualize the distribution of a numerical variable across different categories.

5. Generate a synthetic dataset with correlated features. Visualize the correlation matrix of a dataset using a heatmap.

PLOTLY ASSIGNMENT:

(Use Plotly for the visualization of the given questions)

1. Using the given dataset, to generate a 3D scatter plot to visualize the distribution of data points in a three-dimensional space.

```
np.random.seed(30)
data = {
    'X': np.random.uniform(-10, 10, 300),
    'Y': np.random.uniform(-10, 10, 300),
    'Z': np.random.uniform(-10, 10, 300)
}
df = pd.DataFrame(data)
```

2. Using the Student Grades, create a violin plot to display the distribution of scores across different grade categories.

```
np.random.seed(15)
data = {
    'Grade': np.random.choice(['A', 'B', 'C', 'D', 'F'], 200),
    'Score': np.random.randint(50, 100, 200)
}
df = pd.DataFrame(data)

1. Using the sales data, generate a heatmap to visualize the variation in sales across
   different months and days.
np.random.seed(20)
data = {
    'Month': np.random.choice(['Jan', 'Feb', 'Mar', 'Apr', 'May'], 100),
    'Day': np.random.choice(range(1, 31), 100),
    'Sales': np.random.randint(1000, 5000, 100)
}
df = pd.DataFrame(data)
```

3. Using the sales data, generate a heatmap to visualize the variation in sales across different months and days.

```
np.random.seed(20)
data = {
    'Month': np.random.choice(['Jan', 'Feb', 'Mar', 'Apr', 'May'], 100),
    'Day': np.random.choice(range(1, 31), 100),
    'Sales': np.random.randint(1000, 5000, 100)
}
df = pd.DataFrame(data)
```

4. Using the given x and y data, generate a 3D surface plot to visualize the function $z = sin(\sqrt{x^2 + y^2})$

```
x = np.linspace(-5, 5, 100)
y = np.linspace(-5, 5, 100)
x, y = np.meshgrid(x, y)
z = np.sin(np.sqrt(x**2 + y**2))

data = {
    'X': x.flatten(),
    'Y': y.flatten(),
    'Z': z.flatten()
}
df = pd.DataFrame(data)
```

5. Using the given dataset, create a bubble chart to represent each country's population (y-axis), GDP (x-axis), and bubble size proportional to the population.

```
np.random.seed(25)
data = {
    'Country': ['USA', 'Canada', 'UK',
'Germany', 'France'],
    'Population':
np.random.randint(100, 1000, 5),
    'GDP': np.random.randint(500, 2000,
5)
}
df = pd.DataFrame(data)
```

BOKEH ASSIGNMENT:

(Use Bokeh for the visualization of the given questions)

1.Create a Bokeh plot displaying a sine wave. Set x-values from 0 to 10 and y-values as the sine of x.

2.Create a Bokeh scatter plot using randomly generated x and y values. Use different sizes and colors for the markers based on the 'sizes' and 'colors' columns.

3. Generate a Bokeh bar chart representing the counts of different fruits using the following dataset.

```
fruits = ['Apples', 'Oranges', 'Bananas', 'Pears']
counts = [20, 25, 30, 35]
```

4. Create a Bokeh histogram to visualize the distribution of the given data.

```
data_hist = np.random.randn(1000)
hist, edges = np.histogram(data_hist, bins=30)
```

5. Create a Bokeh heatmap using the provided dataset.

```
data_heatmap = np.random.rand(10, 10)
x = np.linspace(0, 1, 10)
y = np.linspace(0, 1, 10)
xx, yy = np.meshgrid(x, y)
```