EX.No:10	Generating and interpreting statistical visualizations using the Seaborn package.
DATE:	

AIM:

To write a program that performs data visualization using Bar chart, Pie chart, Scatterplot, histogram

ALGORITHM:

1. Import Libraries: Import Pandas, Matplotlib, and Seaborn for data manipulation and visualization.

   Instruction: In Spyder, go to the IPython Console tab, open the console's Terminal/Command Prompt, and run:
   conda install seaborn
   This will install the Seaborn package in your active Anaconda environment so Spyder can use it.
2. Generate Data: Create a hypothetical dataset with categories and numerical values.
3. Bar Chart: Use Seaborn's barplot to visualize values across categories.
4. Pie Chart: Utilize Matplotlib's pie to represent the distribution of a subset of values as percentages.
5. Scatter Plot: Create a scatter plot using Matplotlib to show the relationship between two numerical variables.
6. Histogram: Plot a histogram with Matplotlib to display the frequency distribution of a single variable.
7. Adjustments: Customize titles, labels, and styling for each plot.
8. Display Plots: Use plt.show() to display the individual plots.
9. Install Dependencies: Ensure Matplotlib and Seaborn are installed (pip install matplotlib seaborn).
10. Run Program: Execute the script, visualizing the data with bar charts, pie charts, scatter plots, and histograms.

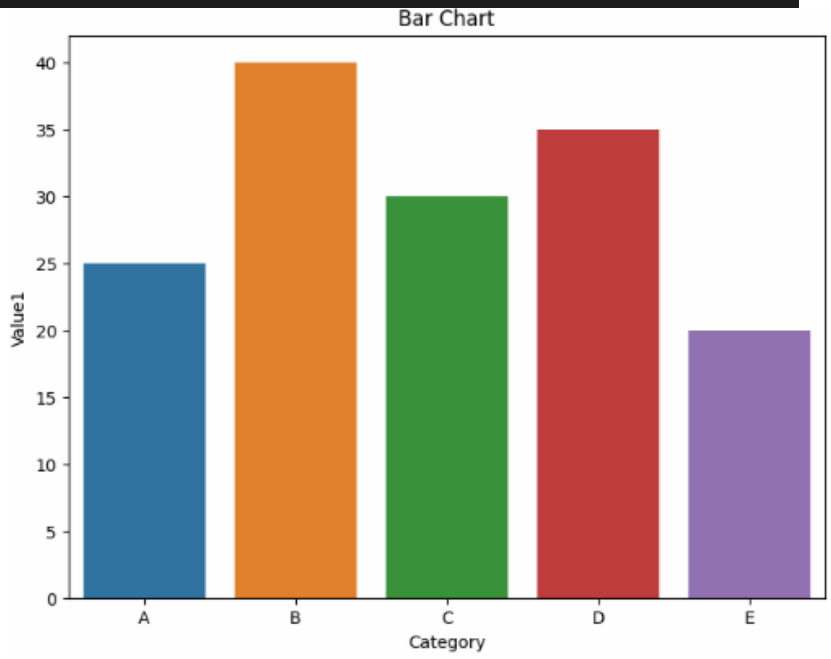
PROGRAM:

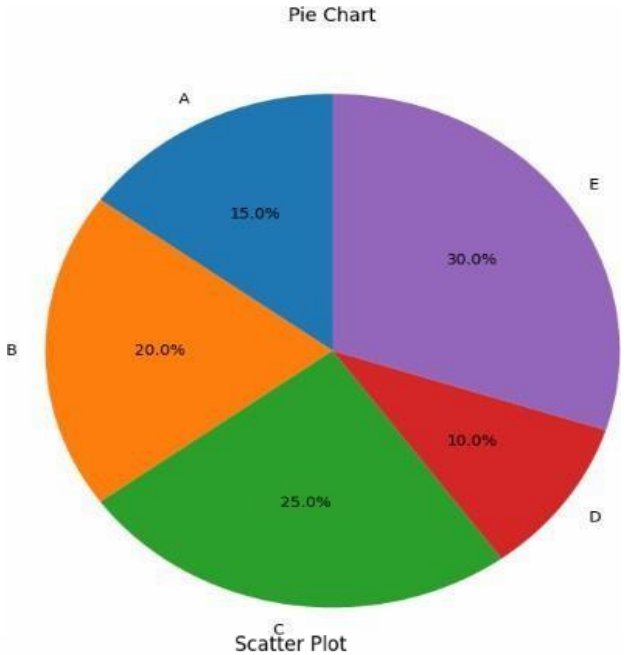
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
# Generate a hypothetical
dataset data = {
   'Category': ['A', 'B', 'C', 'D', 'E'],
   'Value1': [25, 40, 30, 35, 20],
   'Value2': [15, 20, 25, 10, 30]
}
```

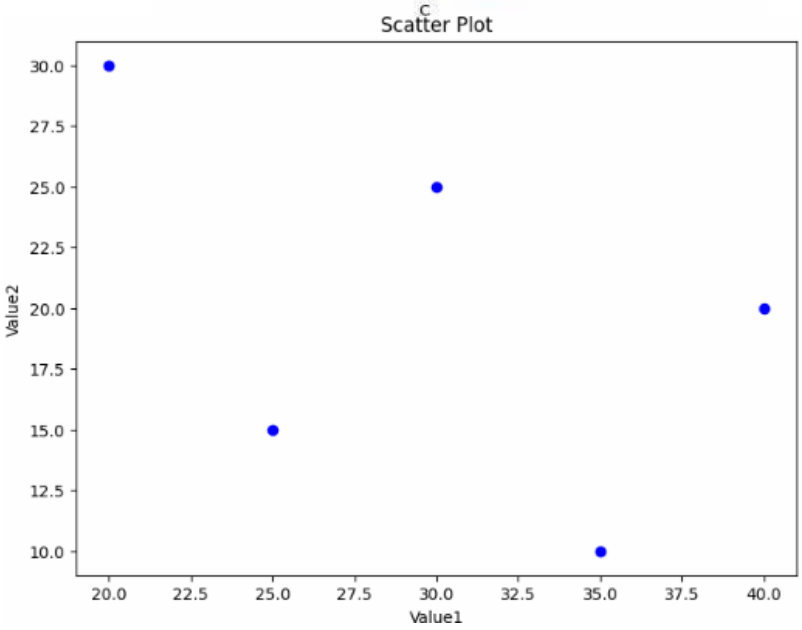
```
df = pd.DataFrame(data)

# Bar chart
plt.figure(figsize=(8, 6))
sns.barplot(x='Category', y='Value1', data=df)
plt.title('Bar Chart')
plt.show()

# Pie chart
plt.figure(figsize=(8, 8))
plt.pie(df['Value2'], labels=df['Category'], autopct='%1.1f%%', startangle=90)
plt.title('Pie Chart')
plt.show()

# Scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(df['Value1'], df['Value2'], c='blue',
marker='o') plt.title('Scatter Plot')
plt.xlabel('Value1')
plt.ylabel('Value
2') plt.show()

# Histogram
plt.figure(figsize=(8, 6))
plt.hist(df['Value1'], bins=np.arange(15, 45, 5), edgecolor='black')
plt.title('Histogram')
plt.xlabel('Value1')
plt.ylabel('Frequency')
plt.show()
```

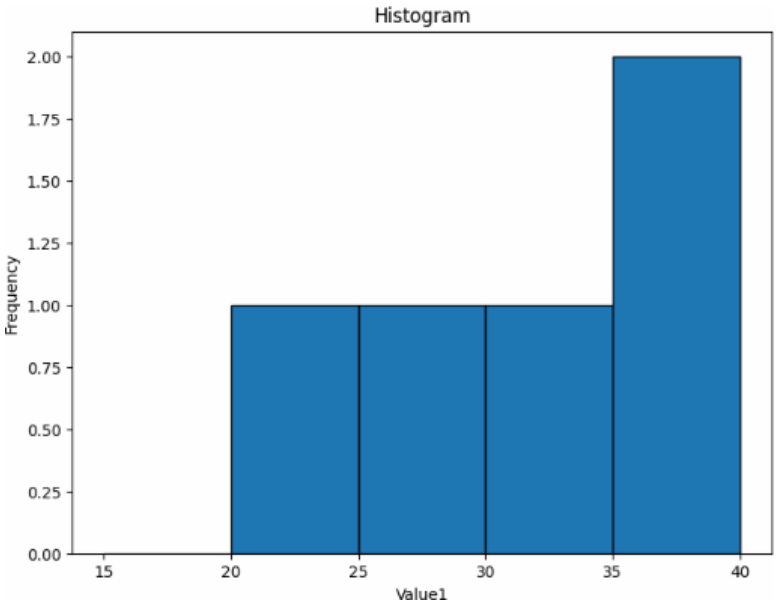
OUTPUT:

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np

# Generate a hypothetical dataset
data = {
    'Category': ['A', 'B', 'C', 'D', 'E'],
    'Value1': [25, 40, 30, 35, 20],
    'Value2': [15, 20, 25, 10, 30]
}

df = pd.DataFrame(data)

# Bar chart
plt.figure(figsize=(8, 6))
sns.barplot(x='Category', y='Value1', data=df)
plt.title('Bar Chart')
plt.show()

# Pie chart
plt.figure(figsize=(8, 8))
plt.pie(df['Value2'], labels=df['Category'], autopct='%1.1f%%', startangle=90)
plt.title('Pie Chart')
plt.show()

# Scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(df['Value1'], df['Value2'], c='blue', marker='o')
plt.title('Scatter Plot')
plt.xlabel('Value1')
plt.ylabel('Value2')
plt.show()

# Histogram
plt.figure(figsize=(8, 6))
plt.hist(df['Value1'], bins=np.arange(15, 45, 5), edgecolor='black')
plt.title('Histogram')
plt.xlabel('Value1')
plt.ylabel('Frequency')
plt.show()
```




Pie Chart
A
B
C
D
E
15.0%
20.0%
25.0%
10.0%
30.0%


Scatter Plot
Value2
Value1
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
10.0
20.0
22.5
25.0
27.5
30.0
32.5
35.0
37.5
40.0



RESULT:

Thus, python program is successfully implemented that performs data visualization using Bar chart,Pie chart,Scatter plot,histogram

Exercise :

1. **Create a Histogram and KDE Plot**
   Using the tips dataset from Seaborn, create a histogram and kernel density estimate (KDE) plot to visualize the distribution of the total bill amount. Interpret what the shape of the distribution tells you about the data.
2. **Boxplot for Comparing Groups**
   Use the tips dataset to create a boxplot comparing the distribution of total bills across

different days of the week. What can you infer about the variability and median total bill on different days?

3. **Scatterplot with Regression Line**
   Generate a scatterplot with a regression line (lmplot) to explore the relationship between total bill and tip amount in the tips dataset. Describe the trend and strength of the relationship based on the plot.
4. **Heatmap of Correlation Matrix**
   Calculate the correlation matrix for the numerical variables in the iris dataset. Visualize this correlation matrix using a heatmap. Which pairs of variables show the strongest positive and negative correlations?
5. **Countplot with Hue**
   Using the titanic dataset, create a countplot showing the count of passengers by class (Pclass) and use the hue parameter to break down the counts by survival status (Survived). What insights can you gain about survival rates across passenger classes?