Ex3

February 25, 2024

DATA VISUALIZATION THROUGH PYTHON 08/01/2024

URK22AI1085

AIM: Write a python program to demonstrate the data visualization for data science application.

DESCRIPTION: Data visualization is the graphical representation of data to uncove patterns, trends, and insights that may not be apparent in raw, tabular formats. Line Plot: A line plot is a simple and effective way to visualize tr ds␣ over a continuous variable, typically used for time-series data. Syntax: df.plot(x='Date', y='Value', kind='line') Bar Plot: Bar plots are useful for comparing categorical dat They␣ represent data using rectangular bars, with the length of each bar corresponding to the value of the data. Syntax: df.plot(x='Category', y='Value', kind='bar') Scatter Plot: Scatter plots display individual data ints in a␣ two-dimensional space, allowing you to observe relationships between two continuous variables. Syntax: df.plot(x='X', y='Y', kind='scatter') Pie Chart: A pie chart is a circular statistical graphic at is divided␣ into slices to illustrate numerical proportions. Syntax: plt.pie(sizes, labels=labels, autopct='%1.1f%% startangle=90,␣ counterclock=False) Box plot: A box plot, also known as a box-and-whisker plot, is a␣ statistical visualization that provides a summary of the distribution of a dataset. Syntax: df.boxplot(by='Category', column='Value', grid=False) Histogram: A histogram is a graphical representation f the distribution of␣ a dataset. It is commonly used to visualize the underlying frequency distribution of a continuous variable. 1 Syntax: plt.hist(data, bins=10, color='skyblue', edgecolor='black')

```python
[2]: import pandas as pd
     import matplotlib.pyplot as plt

     df = pd.read_csv("Toyota.csv")
```

```python
[7]: #1.Draw a bar chart with FuelType and its count

     df['FuelType'].value_counts().plot(kind='bar')
     plt.xlabel('FuelType')
     plt.ylabel('Count')
     plt.title('FuelType Count')
     plt.show()
```
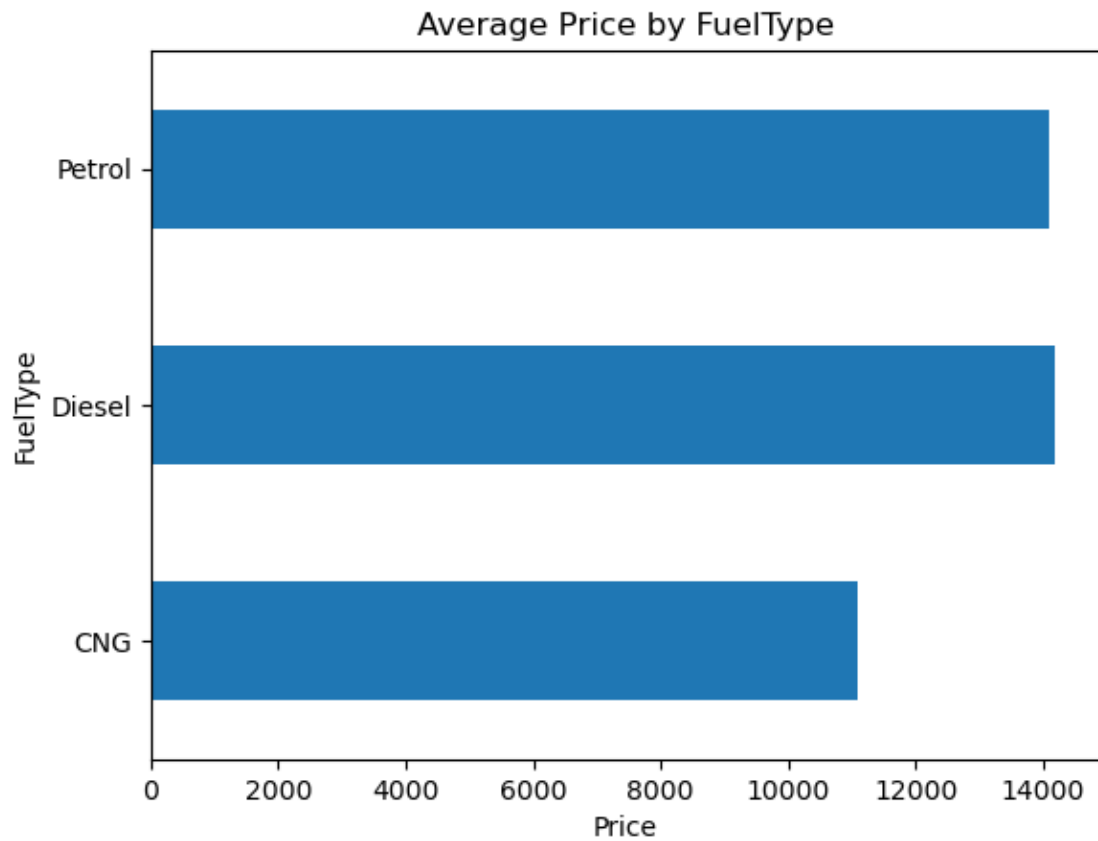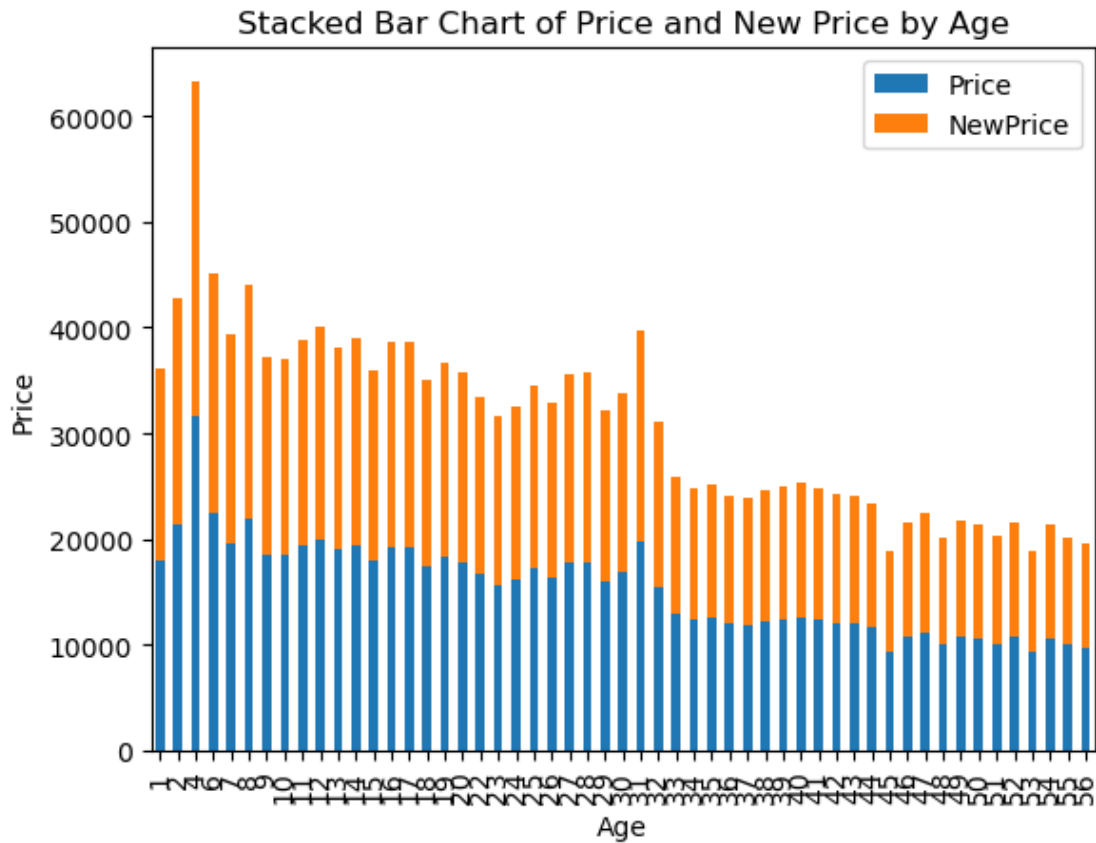
FuelType Count
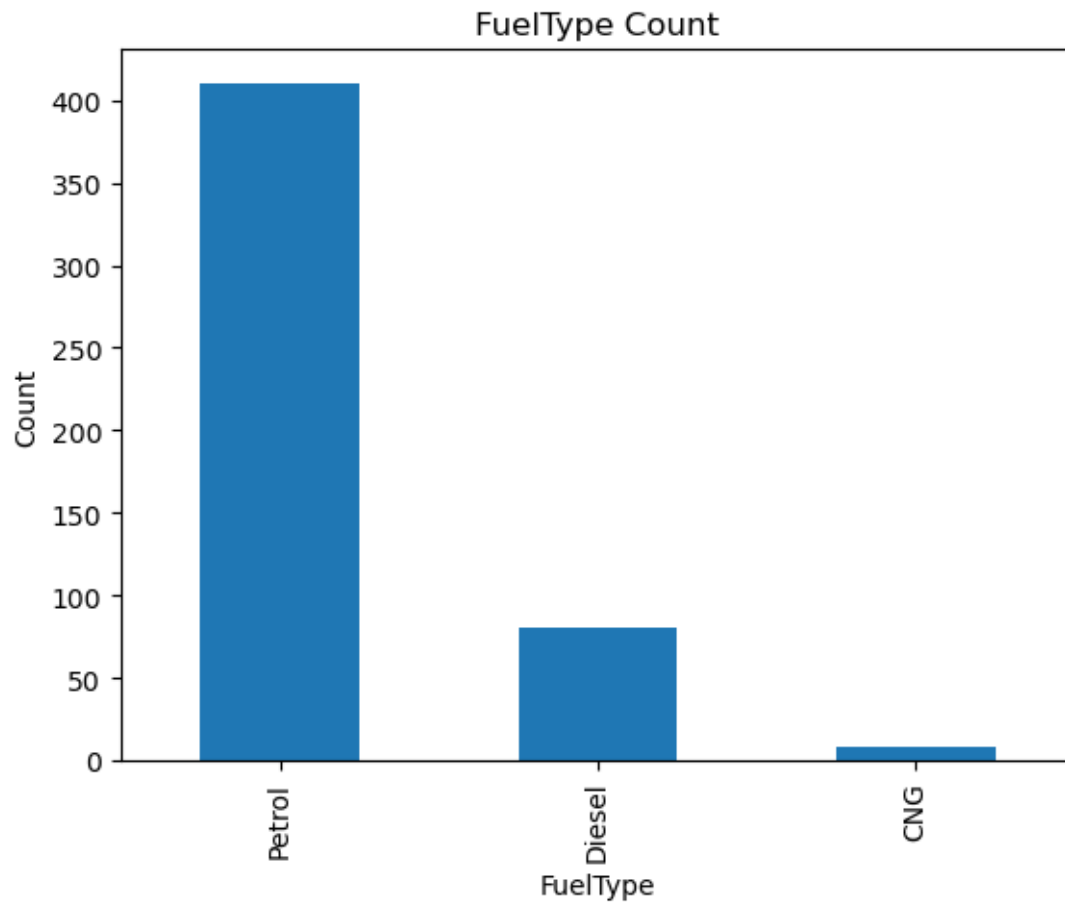

```
[4]: df['NewPrice'] = df['Price'] + 100
```

```
[6]: #2. Horizontal bar chart for FuelType and Price
     df.groupby('FuelType')['Price'].mean().plot(kind='barh')
     plt.xlabel('Price')
     plt.ylabel('FuelType')
     plt.title('Average Price by FuelType')
     plt.show()
```
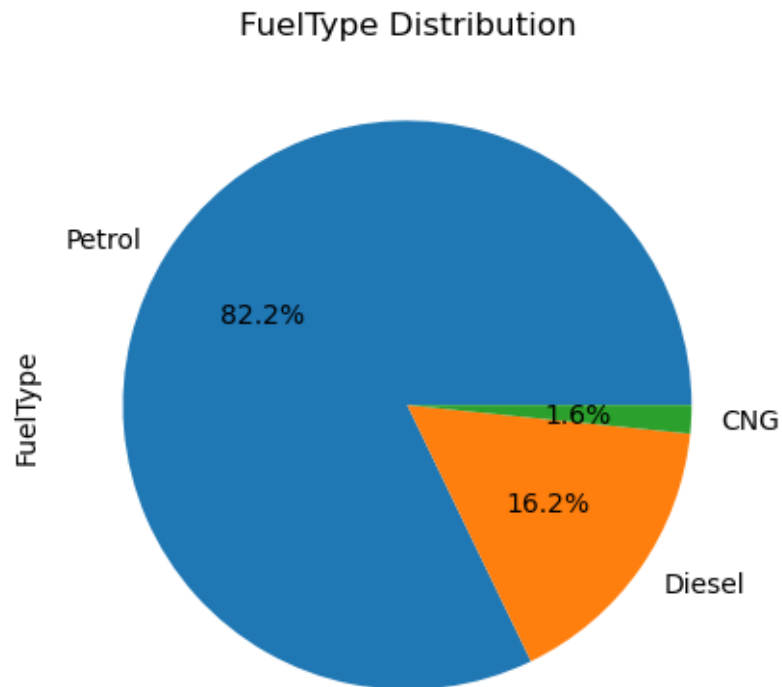
Average Price by FuelType

[7]:
```python
# 3. Stacked bar chart for price and new_price against the Age
df.groupby('Age')[['Price', 'NewPrice']].mean().plot(kind='bar', stacked=True)
plt.xlabel('Age')
plt.ylabel('Price')
plt.title('Stacked Bar Chart of Price and New Price by Age')
plt.show()
```
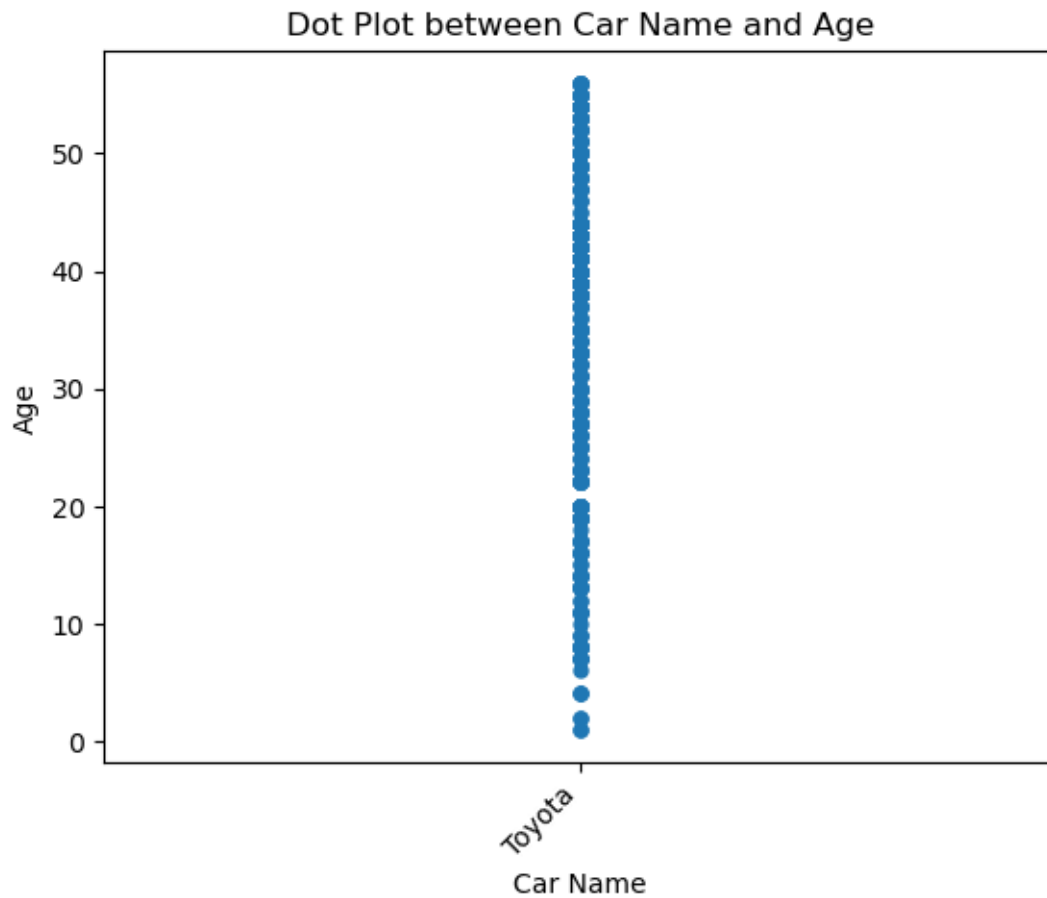
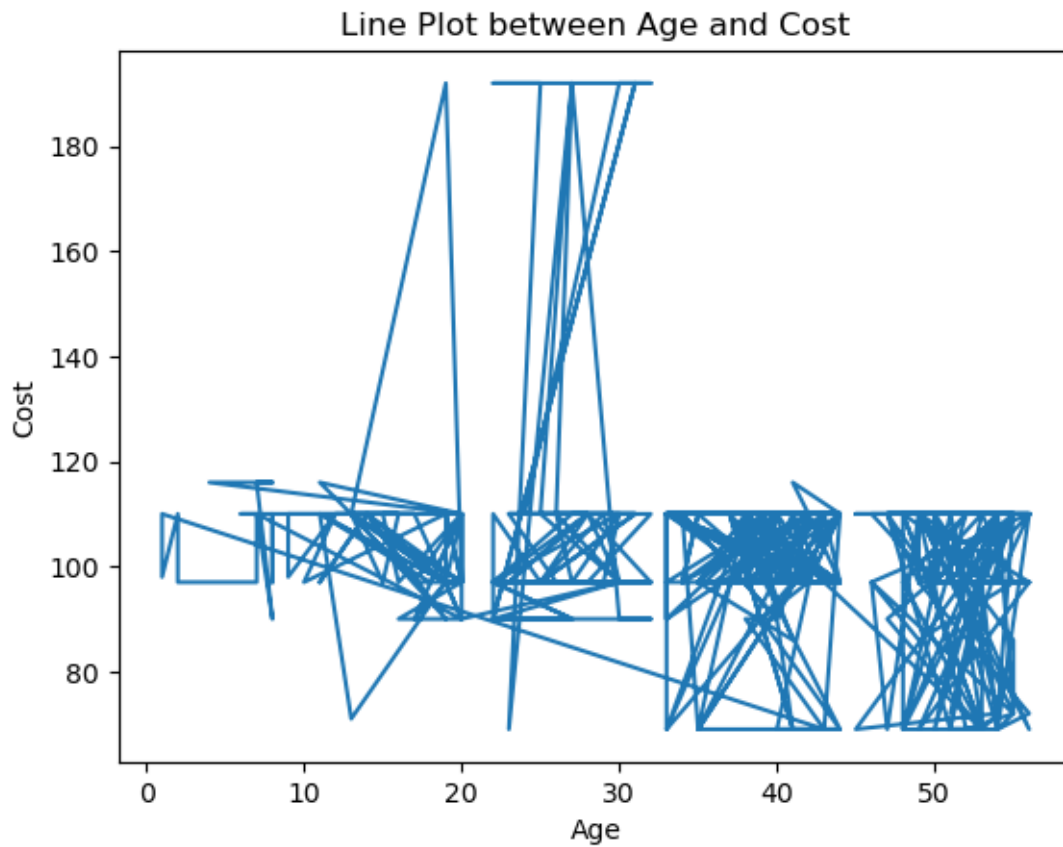Stacked Bar Chart of Price and New Price by Age

[8]: ```python
4. Pie chart with FuelType and its count

df['FuelType'].value_counts().plot(kind='pie', autopct='%1.1f%%')
plt.title('FuelType Distribution')
plt.show()
```

FuelType Distribution


```
[12]: # 5. Dot plot between car name and age
      df['CarName'] = 'Toyota'
      plt.plot(df['CarName'], df['Age'], 'o', markersize=5)
      plt.xlabel('Car Name')
      plt.ylabel('Age')
      plt.title('Dot Plot between Car Name and Age')
      plt.xticks(rotation=45, ha='right')
      plt.show()
```

Dot Plot between Car Name and Age


```
[13]:  # 6. Line plot between age and cost
       plt.plot(df['Age'], df['HP'])
       plt.xlabel('Age')
       plt.ylabel('Cost')
       plt.title('Line Plot between Age and Cost')
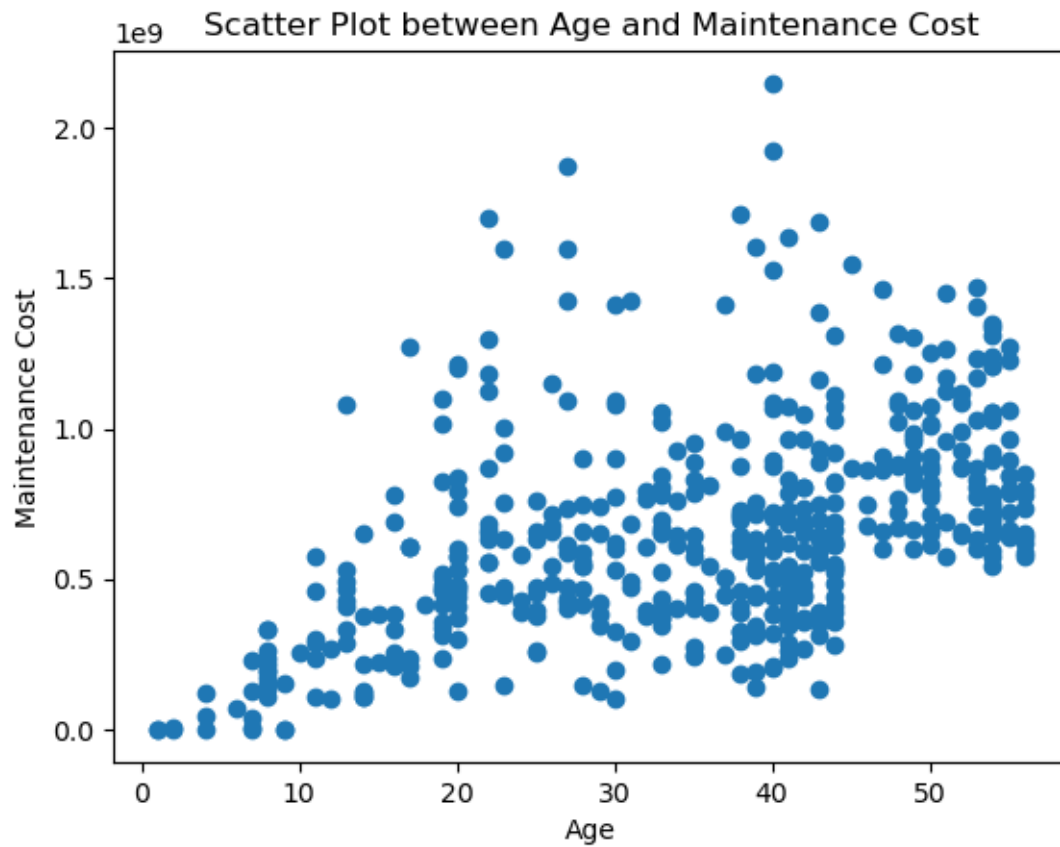       plt.show()
```

Line Plot between Age and Cost

[14]:
```python
# 7. Scatter plot between age and price
plt.scatter(df['Age'], df['Price'])
plt.xlabel('Age')
plt.ylabel('Price')
plt.title('Scatter Plot between Age and Price')
plt.show()
```

Scatter Plot between Age and Price
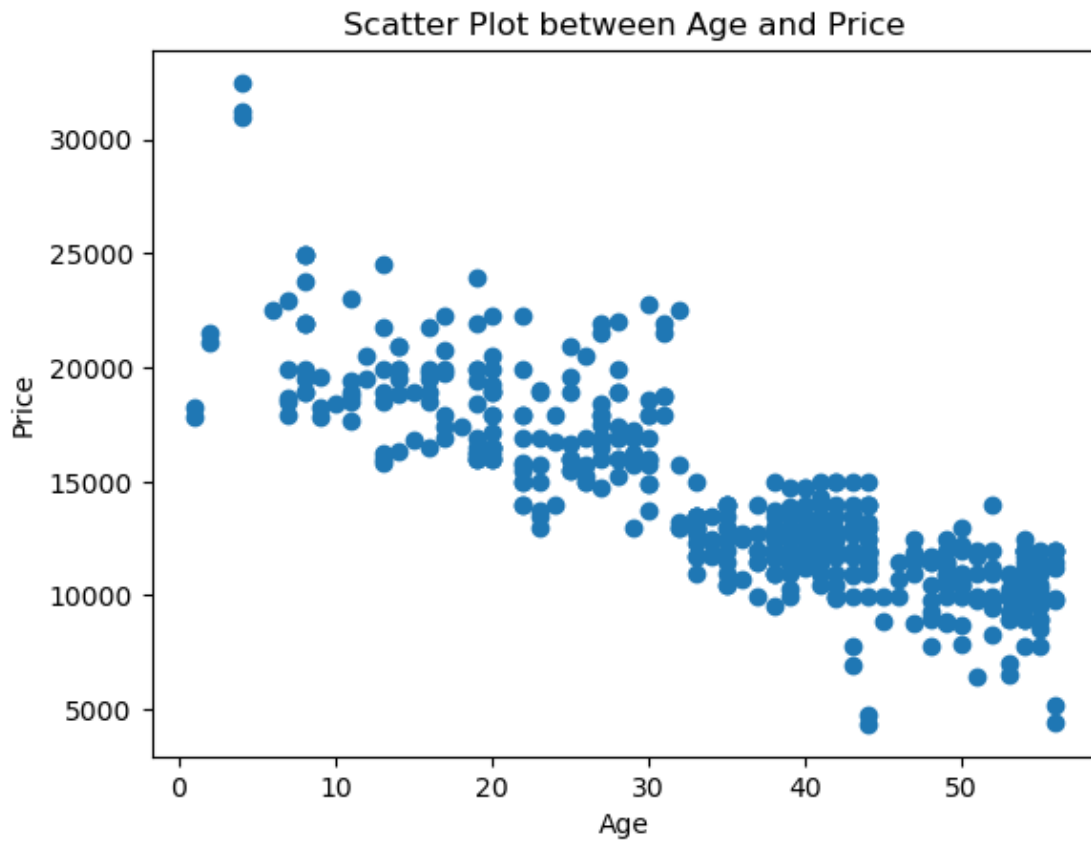
[15]: ```python
8. Scatter plot between age and maintenance_cost

df['MaintenanceCost'] = df['Price'] * df['KM']
plt.scatter(df['Age'], df['MaintenanceCost'])
plt.xlabel('Age')
plt.ylabel('Maintenance Cost')
plt.title('Scatter Plot between Age and Maintenance Cost')
plt.show()
```
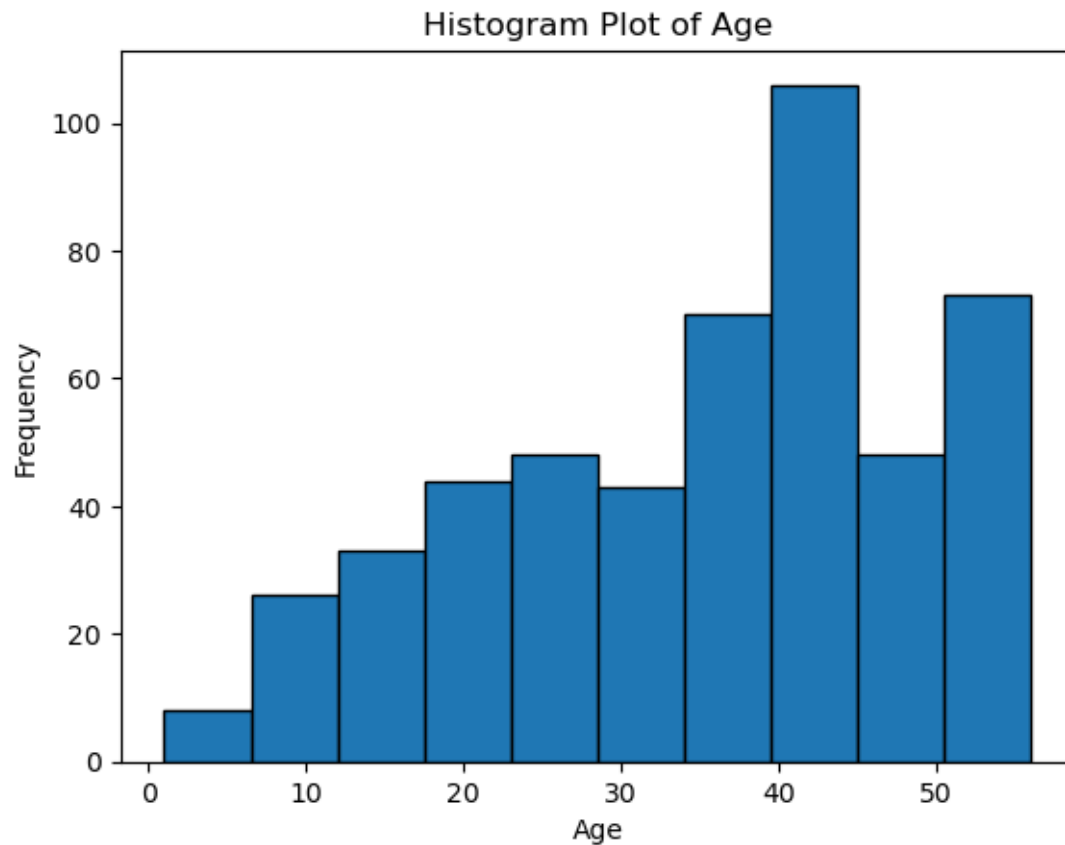
Scatter Plot between Age and Maintenance Cost

[16]:
```python
# 9. Box plot to show the statistical summary of Age column

plt.boxplot(df['Age'])
plt.xlabel('Age')
plt.title('Box Plot of Age')
plt.show()
```
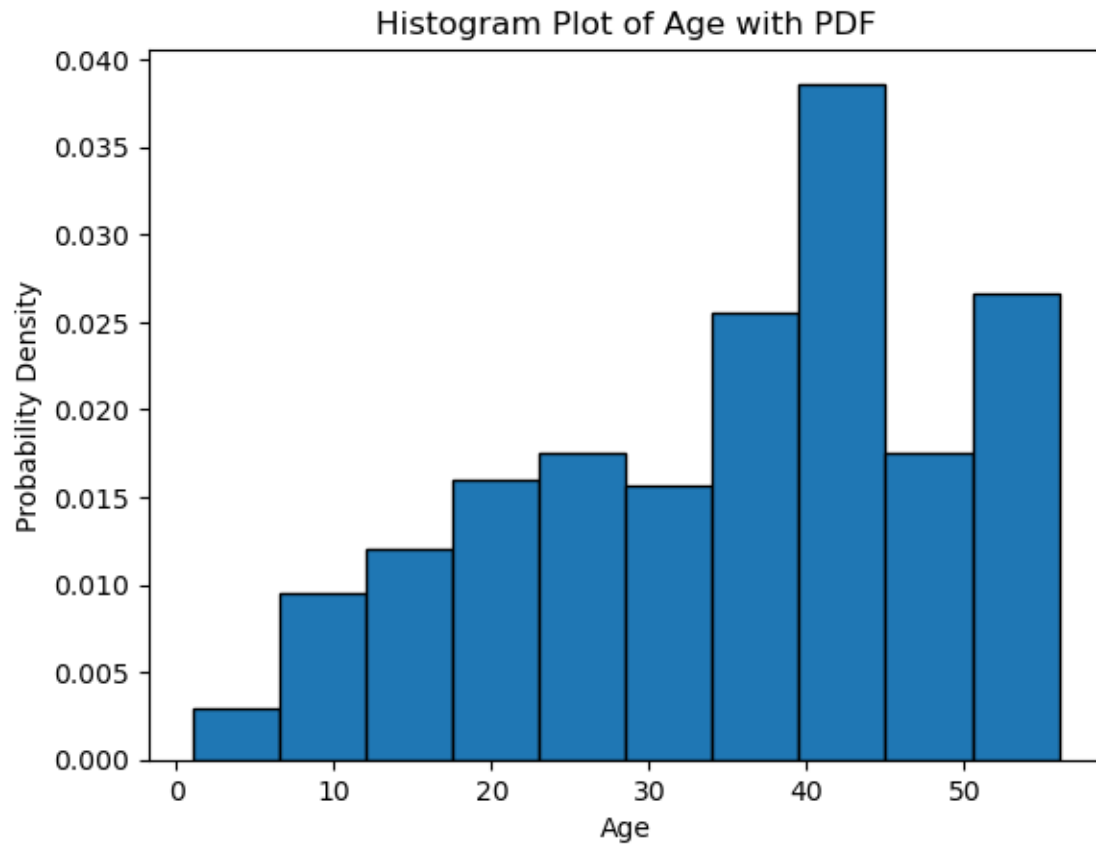
Box Plot of Age
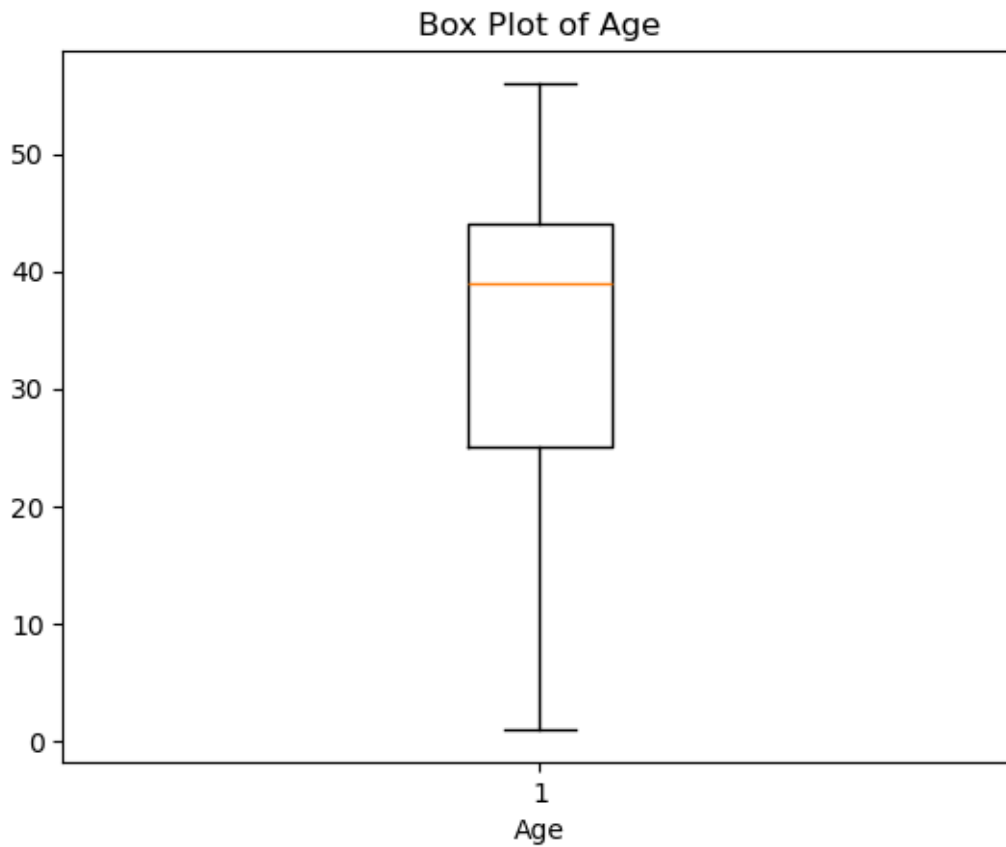
[17]: ```python
10. Histogram plot for Age column

plt.hist(df['Age'], bins=10, edgecolor='black')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.title('Histogram Plot of Age')
plt.show()
```

Histogram Plot of Age

```
[18]: # 11. Histogram plot for Age column with bin value and PDF

plt.hist(df['Age'], bins=10, density=True, edgecolor='black')
plt.xlabel('Age')
plt.ylabel('Probability Density')
plt.title('Histogram Plot of Age with PDF')
plt.show()
```

Histogram Plot of Age with PDF

RESULT: Python programs to demonstrate data visualization for data science application was successfully executed.