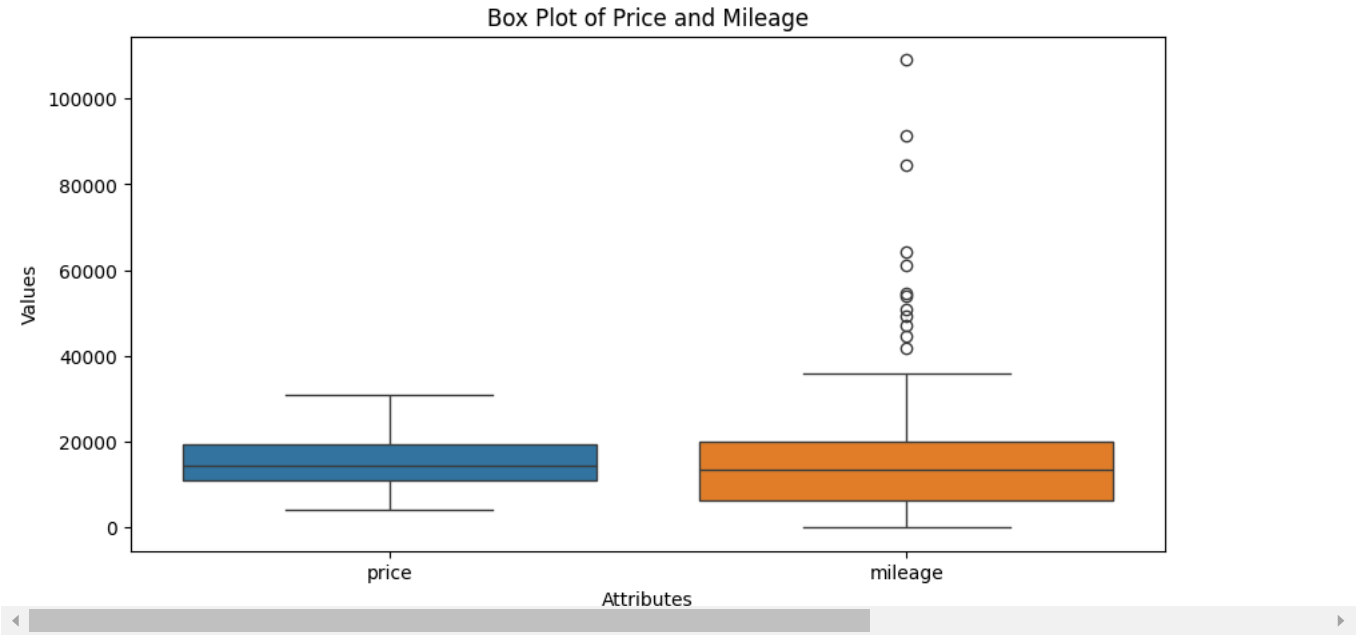
Box-Plot

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

file_path = "dataset.xlsx"
df = pd.read_excel(file_path)

selected_attributes = ['price', 'mileage']

for attr in selected_attributes:
    mean_value = df[attr].mean()
    median_value = df[attr].median()
    mode_value = df[attr].mode()[0]

    print(f"{attr.capitalize()} Statistics:")
    print(f"Mean: {mean_value}")
    print(f"Median: {median_value}")
    print(f"Mode: {mode_value}")
    print("-" * 30)

plt.figure(figsize=(10, 5))
sns.boxplot(data=df[selected_attributes])

plt.xlabel("Attributes")
plt.ylabel("Values")
plt.title("Box Plot of Price and Mileage")

plt.show()
```

```
Price Statistics:
Mean: 15734.47
Median: 14542.5
Mode: 11000
------------------------------
Mileage Statistics:
Mean: 18489.99
Median: 13471.5
Mode: 4990
------------------------------
```



Double-click (or enter) to edit

Double-click (or enter) to edit

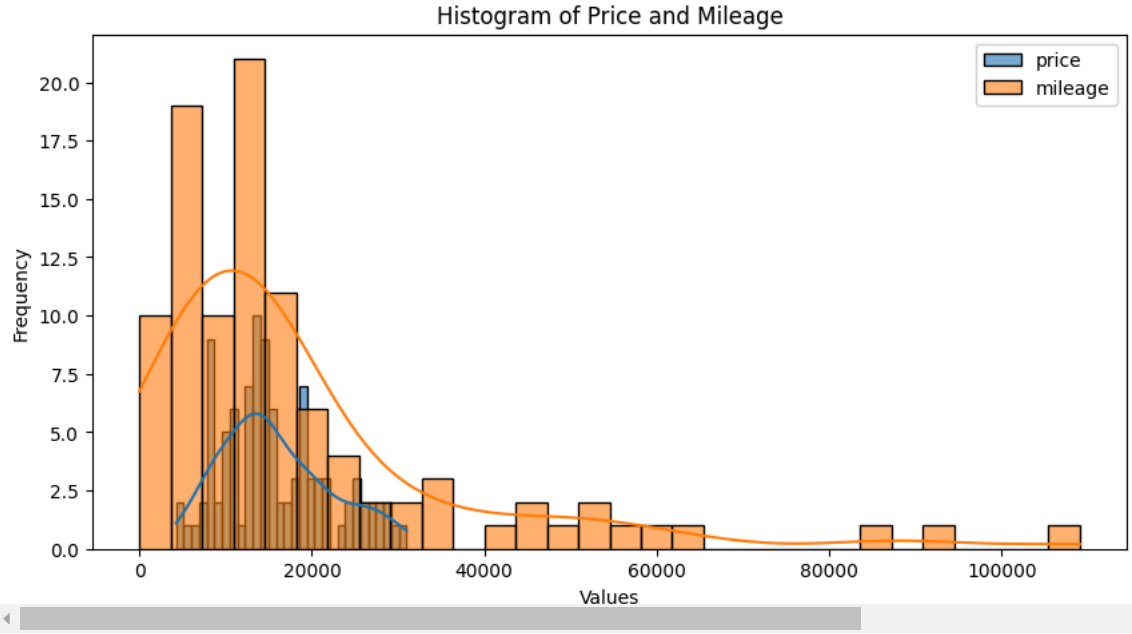
Double-click (or enter) to edit

Double-click (or enter) to edit

Histogram

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

file_path = "dataset.xlsx"
df = pd.read_excel(file_path)

selected_attributes = ['price', 'mileage']

for attr in selected_attributes:
    mean_value = df[attr].mean()
    median_value = df[attr].median()
    mode_value = df[attr].mode()[0]

    print(f"{attr.capitalize()} Statistics:")
    print(f"Mean: {mean_value}")
    print(f"Median: {median_value}")
    print(f"Mode: {mode_value}")
    print("-" * 30)

plt.figure(figsize=(10, 5))

for attr in selected_attributes:
    sns.histplot(df[attr], kde=True, bins=30, label=attr, alpha=0.6)

plt.xlabel("Values")
plt.ylabel("Frequency")
plt.title("Histogram of Price and Mileage")
plt.legend()

plt.show()
```

```
Price Statistics:
Mean: 15734.47
Median: 14542.5
Mode: 11000
------------------------------
Mileage Statistics:
Mean: 18489.99
Median: 13471.5
Mode: 4990
------------------------------
```



Double-click (or enter) to edit

Double-click (or enter) to edit

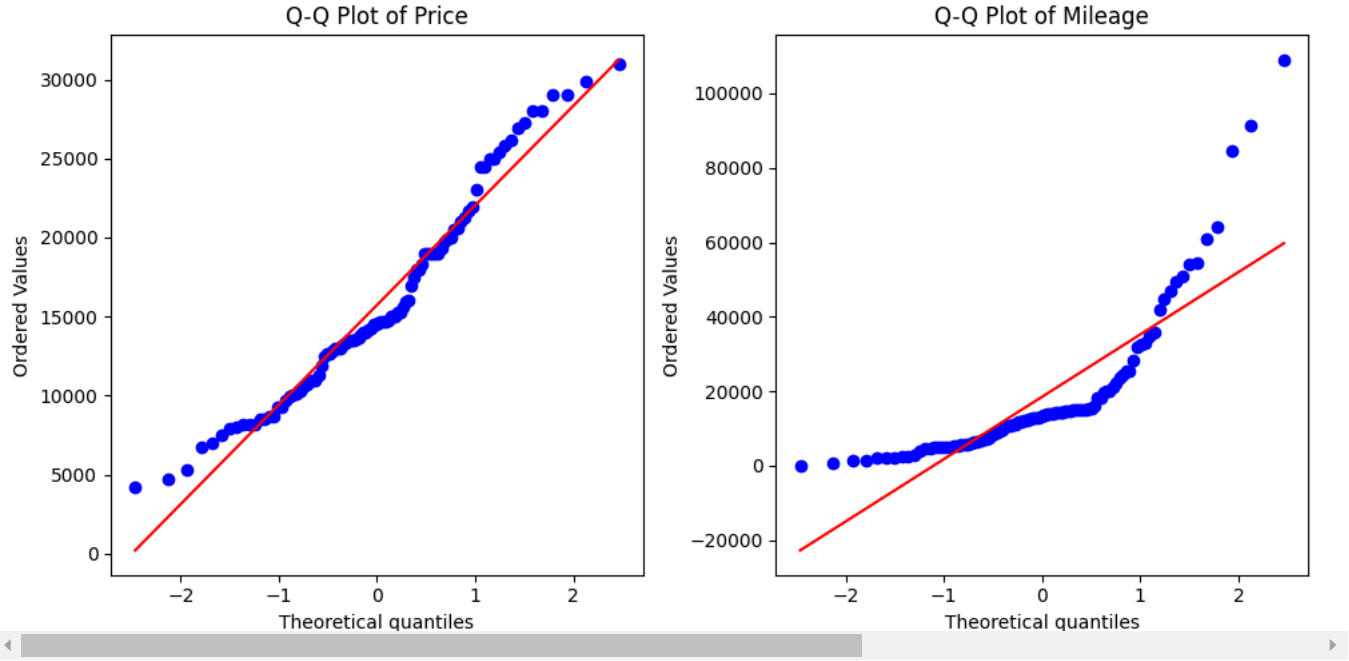
Double-click (or enter) to edit

Double-click (or enter) to edit

Quantile Plot

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

file_path = "dataset.xlsx"
df = pd.read_excel(file_path)

selected_attributes = ['price', 'mileage']

for attr in selected_attributes:
    mean_value = df[attr].mean()
    median_value = df[attr].median()
    mode_value = df[attr].mode()[0]

    print(f"{attr.capitalize()} Statistics:")
    print(f"Mean: {mean_value}")
    print(f"Median: {median_value}")
    print(f"Mode: {mode_value}")
    print("-" * 30)

plt.figure(figsize=(10, 5))

for i, attr in enumerate(selected_attributes, 1):
    plt.subplot(1, 2, i)
    stats.probplot(df[attr], dist="norm", plot=plt)
    plt.title(f"Q-Q Plot of {attr.capitalize()}")

plt.tight_layout()
plt.show()
```

```
Price Statistics:
Mean: 15734.47
Median: 14542.5
Mode: 11000
------------------------------
Mileage Statistics:
Mean: 18489.99
Median: 13471.5
Mode: 4990
------------------------------
```



Double-click (or enter) to edit

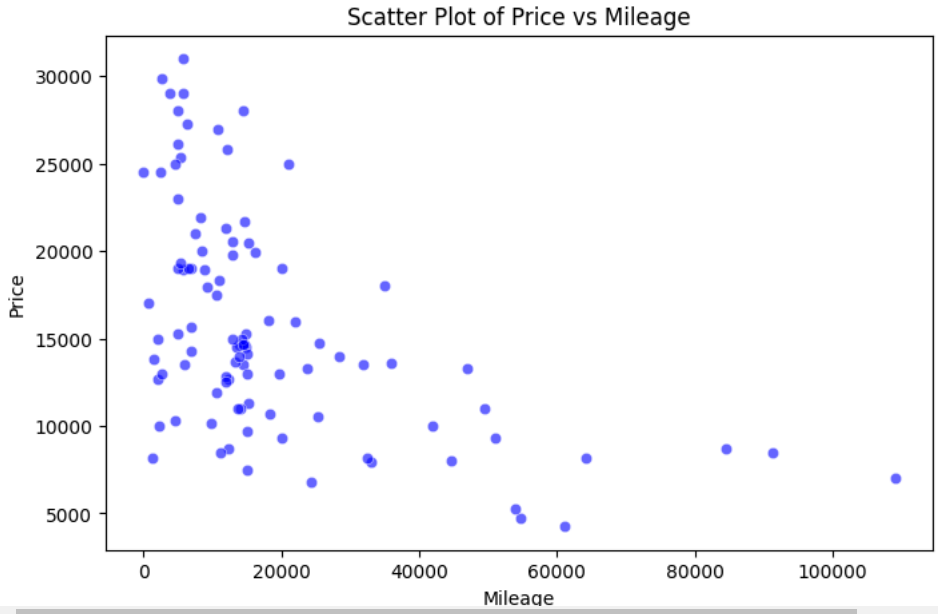
Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Scatter-Plot

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

file_path = "dataset.xlsx"
df = pd.read_excel(file_path)

plt.figure(figsize=(8, 5))
sns.scatterplot(x=df["mileage"], y=df["price"], color="blue", alpha=0.6)

plt.xlabel("Mileage")
plt.ylabel("Price")
plt.title("Scatter Plot of Price vs Mileage")
plt.show()
```



Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Double-click (or enter) to edit

Skewness

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

file_path = "dataset.xlsx"
df = pd.read_excel(file_path)

selected_attributes = ['price', 'mileage']

plt.figure(figsize=(10, 5))

for i, attr in enumerate(selected_attributes, 1):
    plt.subplot(1, 2, i)
    sns.histplot(df[attr], kde=True, bins=30, color="blue", alpha=0.6)

    mean_value = df[attr].mean()
    median_value = df[attr].median()
    mode_value = df[attr].mode()[0]
    skewness = df[attr].skew()

    plt.axvline(mean_value, color='red', linestyle='dashed', label="Mean")
    plt.axvline(median_value, color='green', linestyle='dashed', label="Median")
    plt.axvline(mode_value, color='purple', linestyle='dashed', label="Mode")

    if skewness > 0:
        skew_text = "Positively Skewed"
    elif skewness < 0:
        skew_text = "Negatively Skewed"
    else:
        skew_text = "Symmetric"

    print(f"{attr.capitalize()} Skewness: {skewness:.3f} ({skew_text})")


    plt.text(x=mean_value, y=plt.ylim()[1] * 0.9, s=skew_text, fontsize=12, color="black", ha="center")

    plt.title(f"Distribution of {attr.capitalize()}")
    plt.legend()

plt.tight_layout()
plt.show()
```

```
Price Skewness: 0.549 (Positively Skewed)
Mileage Skewness: 2.414 (Positively Skewed)
```