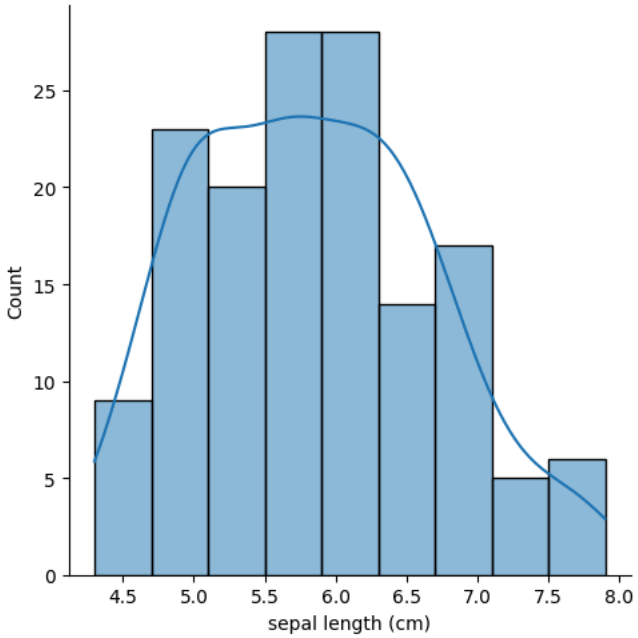
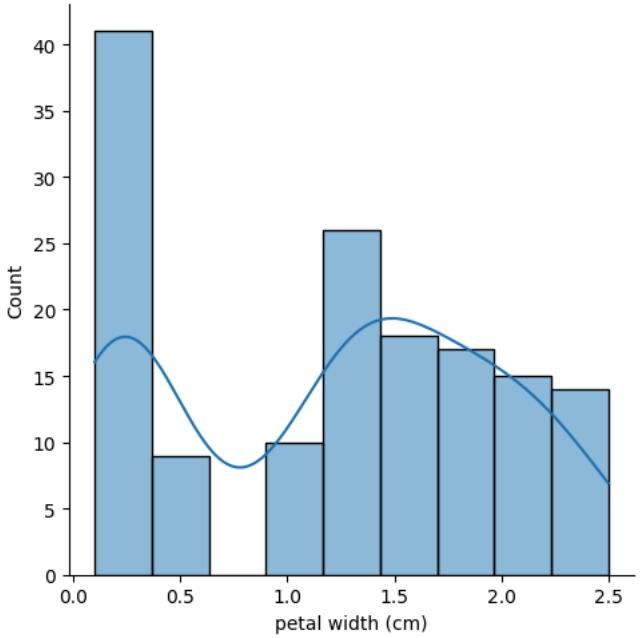
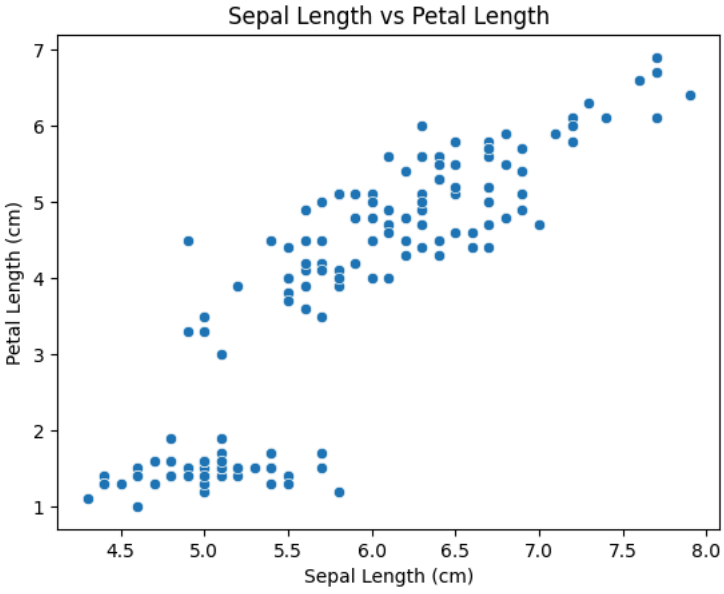
```
import pandas as pd
import seaborn as sns
from scipy import stats
from sklearn.datasets import load_iris
import matplotlib.pyplot as plt

# Load the iris dataset
iris = load_iris()
iris_df = pd.DataFrame(data=iris.data, columns=iris.feature_names)
iris_df['species'] = iris.target

# 1. Find the mean and median of the 'sepal_length' column.
sepal_length_mean = iris_df['sepal length (cm)'].mean()
sepal_length_median = iris_df['sepal length (cm)'].median()
print("Mean sepal length:", sepal_length_mean)
print("Median sepal length:", sepal_length_median)

# 2. Find the most occurring species in the dataset.
most_occuring_species = iris_df['species'].mode()[0]
print("Most occurring species:", most_occuring_species)

# 3. Plot the distributions of 'sepal_length' and 'petal_width' using seaborn.
sns.displot(data=iris_df, x='sepal length (cm)', kde=True)
sns.displot(data=iris_df, x='petal width (cm)', kde=True)
plt.show()

# 4. Find the range and standard deviation of the 'petal_length' column.
petal_length_range = iris_df['petal length (cm)'].max() - iris_df['petal length (cm)'].min()
petal_length_std = iris_df['petal length (cm)'].std()
print("Petal length range:", petal_length_range)
print("Petal length standard deviation:", petal_length_std)

# 5. Find skewness and kurtosis of the 'sepal_width' column using the stats module from scipy.
sepal_width_skewness = stats.skew(iris_df['sepal width (cm)'])
sepal_width_kurtosis = stats.kurtosis(iris_df['sepal width (cm)'])
print("Sepal width skewness:", sepal_width_skewness)
print("Sepal width kurtosis:", sepal_width_kurtosis)

# 6. Plot a scatter plot with 'sepal_length' against 'petal_length'.
sns.scatterplot(data=iris_df, x='sepal length (cm)', y='petal length (cm)')
plt.xlabel('Sepal Length (cm)')
plt.ylabel('Petal Length (cm)')
plt.title('Sepal Length vs Petal Length')
plt.show()

# 7. Find the quartiles of the 'petal_width' column.
petal_width_quartiles = iris_df['petal width (cm)'].quantile([0.25, 0.5, 0.75])
print("Petal width quartiles:")
print(petal_width_quartiles)

# 8. Find the 30th percentile of the 'sepal_width' column.
sepal_width_30th_percentile = iris_df['sepal width (cm)'].quantile(0.3)
print("30th percentile of sepal width:", sepal_width_30th_percentile)

# 9. Use the iris dataset and perform a t-test on two different species to see if the data for 'petal_length' is significantly different
# For example, comparing species 0 and species 1:
species_0_petal_length = iris_df[iris_df['species'] == 0]['petal length (cm)']
species_1_petal_length = iris_df[iris_df['species'] == 1]['petal length (cm)']
t_statistic, p_value = stats.ttest_ind(species_0_petal_length, species_1_petal_length)
print("T-statistic:", t_statistic)
print("P-value:", p_value)
```

```
Mean sepal length: 5.843333333333334
Median sepal length: 5.8
Most occurring species: 0
```






```
Petal length range: 5.9
Petal length standard deviation: 1.7652982332594662
Sepal width skewness: 0.31576710633893473
Sepal width kurtosis: 0.18097631752246768
```

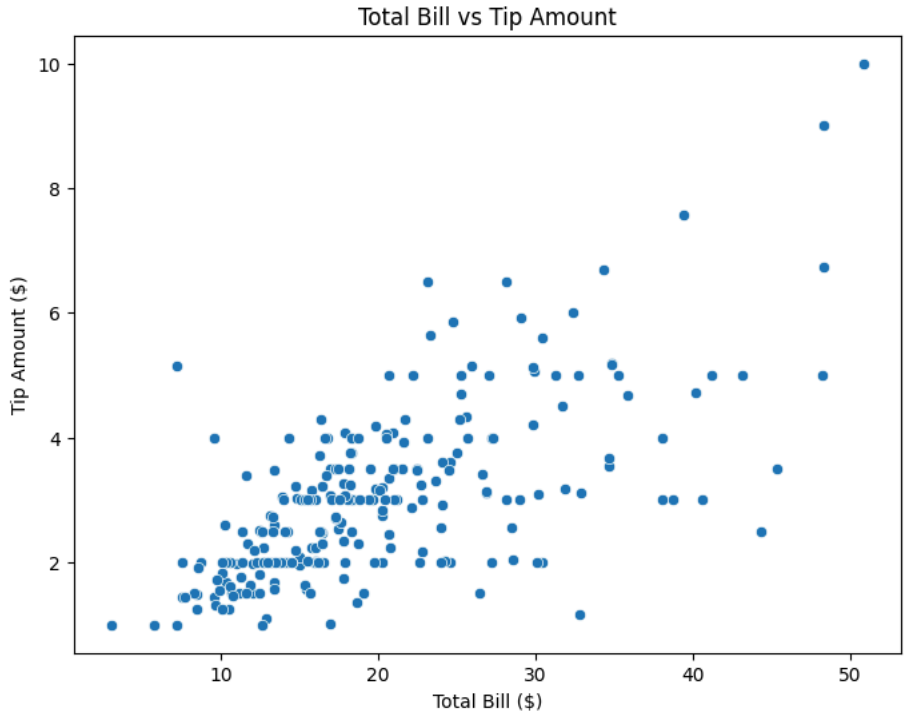


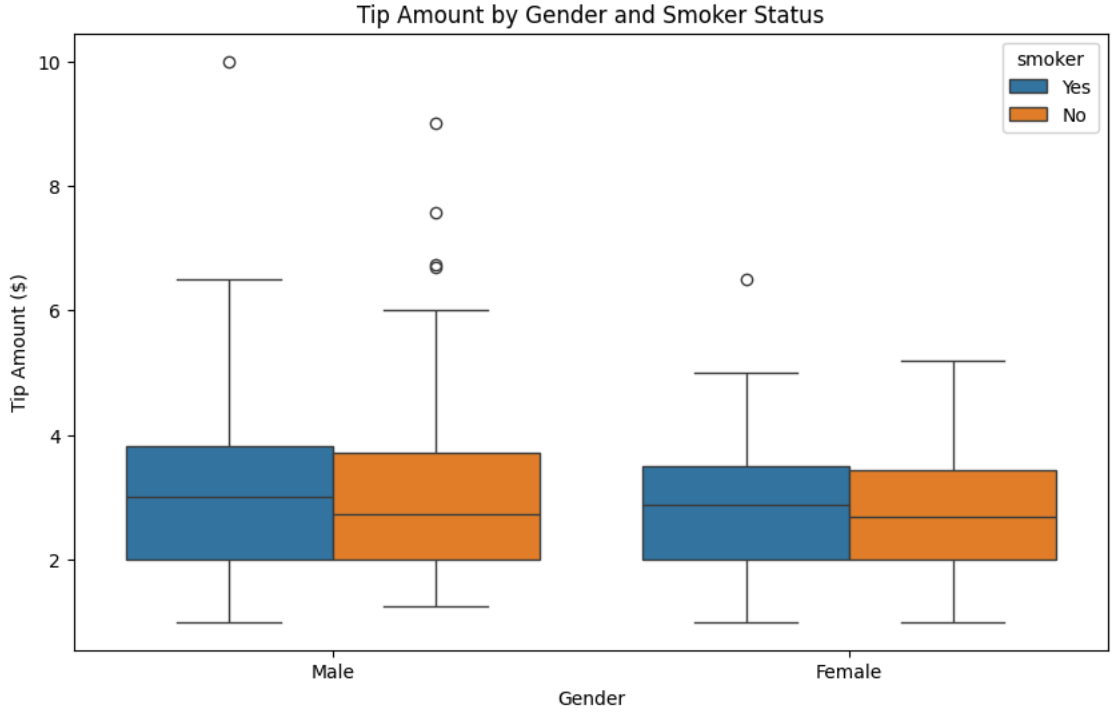

```
Petal width quantiles:
```

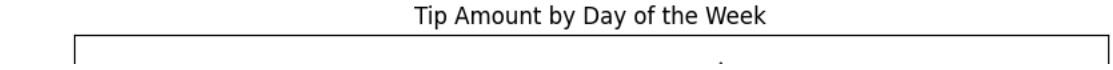
```
Petal width quartiles:
0.25    0.3
0.50    1.3
0.75    1.8
Name: petal width (cm), dtype: float64
30th percentile of sepal width: 2.8
T-statistic: -39.492719391538095
P-value: 5.404910513441677e-62
```

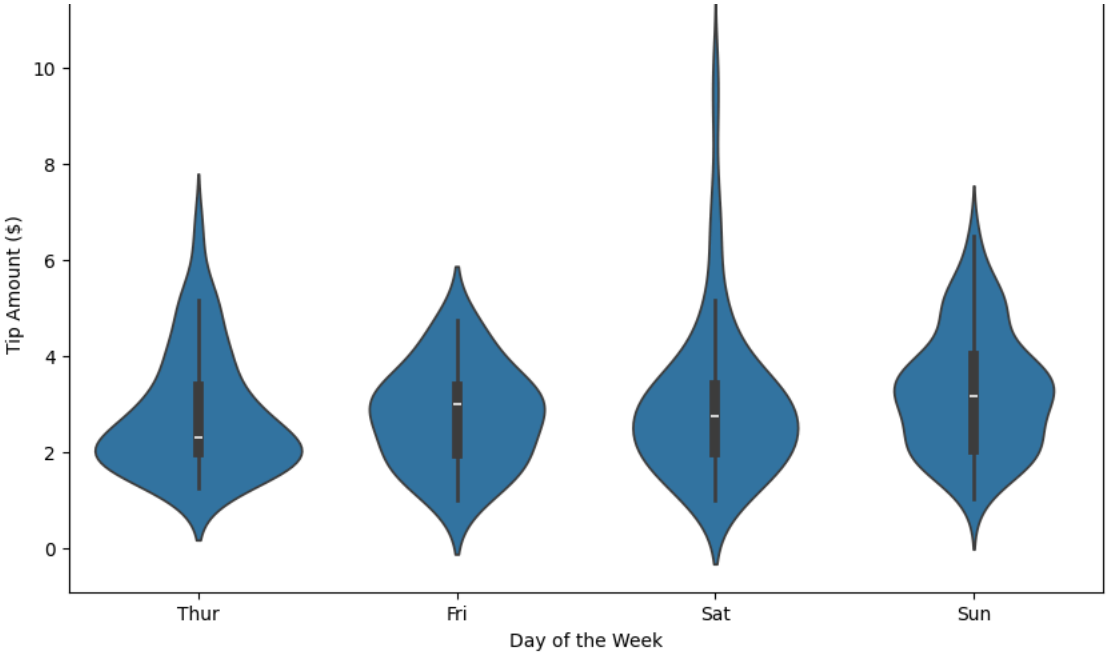
```
# Step 1: Load the "tips" dataset from the Seaborn library into a Pandas DataFrame
import seaborn as sns
import pandas as pd

# Load the tips dataset
tips = sns.load_dataset("tips")

# Step 2: Perform exploratory data analysis (EDA) to understand the distribution of the total bill amount and the tip amount
print(tips.info())
print(tips.describe())

# Step 3: Visualize the relationship between the total bill amount and the tip amount using a scatter plot
import matplotlib.pyplot as plt

plt.figure(figsize=(8, 6))
sns.scatterplot(x='total_bill', y='tip', data=tips)
plt.title('Total Bill vs Tip Amount')
plt.xlabel('Total Bill ($)')
plt.ylabel('Tip Amount ($)')
plt.show()

# Step 4: Analyze whether there is a difference in tipping behavior between genders and between smokers and non-smokers. Visualize these
plt.figure(figsize=(10, 6))
sns.boxplot(x='sex', y='tip', hue='smoker', data=tips)
plt.title('Tip Amount by Gender and Smoker Status')
plt.xlabel('Gender')
plt.ylabel('Tip Amount ($)')
plt.show()

# Step 5: Investigate whether the day of the week influences tipping behavior. Visualize the distribution of tips across different days o
plt.figure(figsize=(10, 6))
sns.violinplot(x='day', y='tip', data=tips, order=['Thur', 'Fri', 'Sat', 'Sun'])
plt.title('Tip Amount by Day of the Week')
plt.xlabel('Day of the Week')
plt.ylabel('Tip Amount ($)')
plt.show()

# Step 6: Optionally, explore how the time of the meal (Lunch or Dinner) and the size of the dining party affect tipping behavior
plt.figure(figsize=(12, 6))
sns.boxplot(x='time', y='tip', hue='size', data=tips)
plt.title('Tip Amount by Time of Meal and Party Size')
plt.xlabel('Time of Meal')
plt.ylabel('Tip Amount ($)')
plt.show()
```

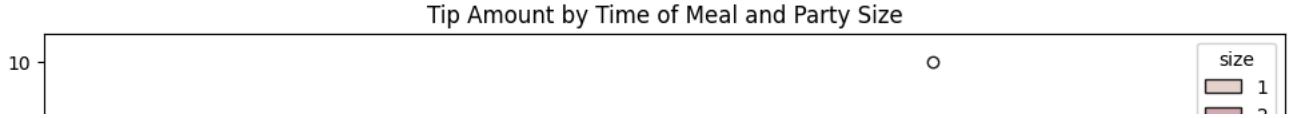
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 244 entries, 0 to 243
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype   
---  ------      --------------  -----   
 0   total_bill  244 non-null    float64 
 1   tip         244 non-null    float64 
 2   sex         244 non-null    category
 3   smoker      244 non-null    category
 4   day         244 non-null    category
 5   time        244 non-null    category
 6   size        244 non-null    int64   
dtypes: category(4), float64(2), int64(1)
memory usage: 7.4 KB
None
       total_bill         tip        size
count  244.000000  244.000000  244.000000
mean    19.785943    2.998279    2.569672
std      8.902412    1.383638    0.951100
min      3.070000    1.000000    1.000000
25%     13.347500    2.000000    2.000000
50%     17.795000    2.900000    2.000000
75%     24.127500    3.562500    3.000000
max     50.810000   10.000000    6.000000
```








10
8
6
4
2
0
Tip Amount ($)
Thur
Fri
Sat
Sun
Day of the Week


Tip Amount by Time of Meal and Party Size
10
size
1