



```
In [1]: # Import libraries
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [2]: # Load dataset
tips = sns.load_dataset("tips")
```

## Task 1: Dataset Loading and Inspection

```
In [3]: # First five rows
tips.head()
```

```
Out[3]:   total_bill    tip      sex smoker  day    time  size
0        16.99  1.01  Female     No  Sun  Dinner     2
1        10.34  1.66    Male     No  Sun  Dinner     3
2        21.01  3.50    Male     No  Sun  Dinner     3
3        23.68  3.31    Male     No  Sun  Dinner     2
4        24.59  3.61  Female     No  Sun  Dinner     4
```

```
In [4]: # Rows and columns
tips.shape
```

```
Out[4]: (244, 7)
```

```
In [5]: # Data types
tips.info()
```

```
<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
```

```
In [6]: # Numerical and categorical columns
tips.describe()
```

Out[6]:

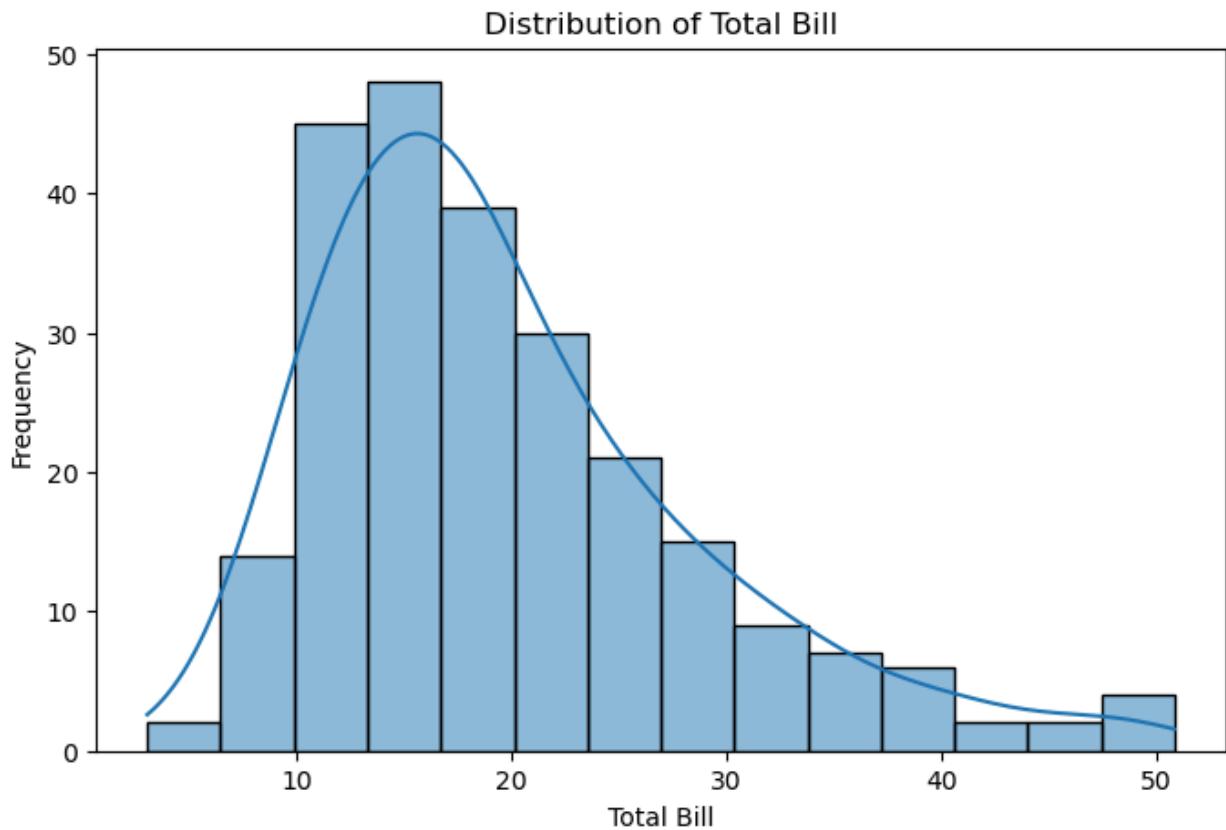
	total_bill	tip	size
<b>count</b>	244.000000	244.000000	244.000000
<b>mean</b>	19.785943	2.998279	2.569672
<b>std</b>	8.902412	1.383638	0.951100
<b>min</b>	3.070000	1.000000	1.000000
<b>25%</b>	13.347500	2.000000	2.000000
<b>50%</b>	17.795000	2.900000	2.000000
<b>75%</b>	24.127500	3.562500	3.000000
<b>max</b>	50.810000	10.000000	6.000000

In [7]: `tips.columns`

Out[7]: `Index(['total_bill', 'tip', 'sex', 'smoker', 'day', 'time', 'size'], dtype='object')`

## Task 2: Distribution of Total Bill

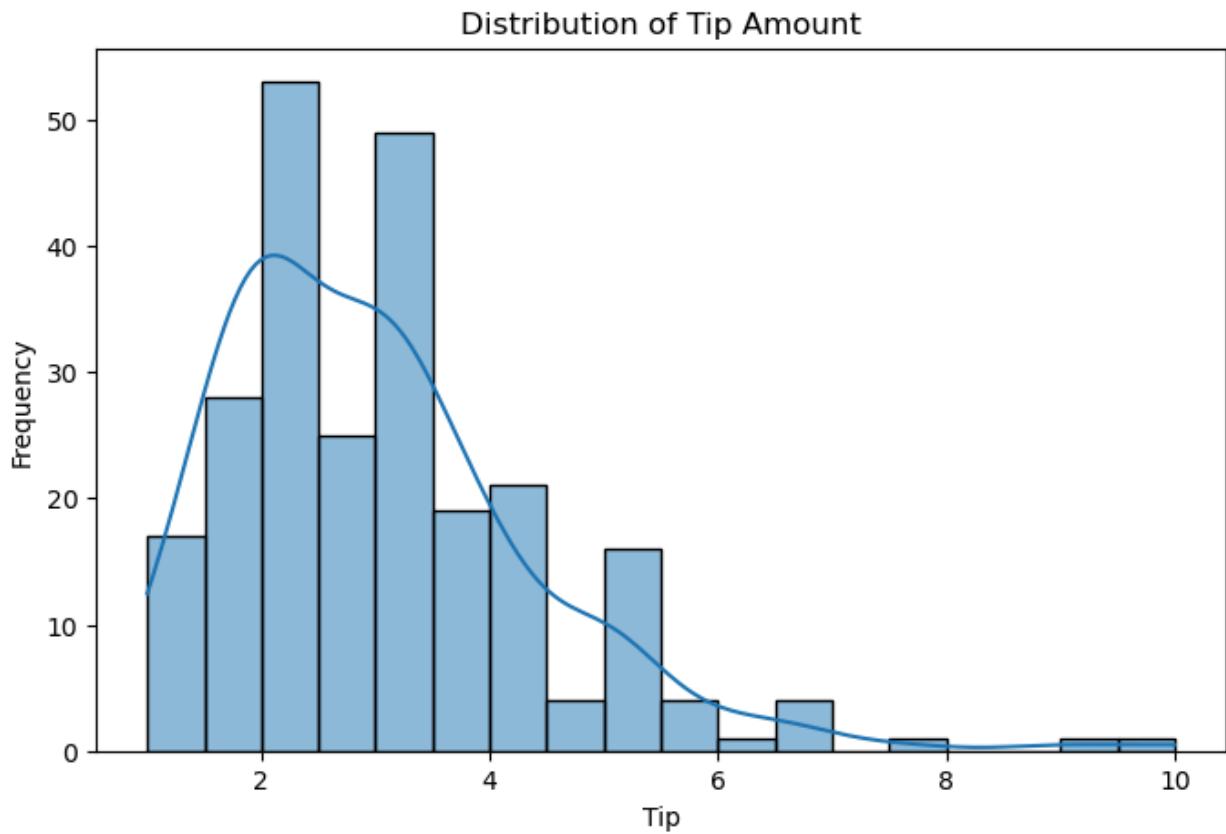
In [8]: `plt.figure(figsize=(8,5))  
sns.histplot(data=tips, x='total_bill', kde=True)  
plt.title("Distribution of Total Bill")  
plt.xlabel("Total Bill")  
plt.ylabel("Frequency")  
plt.show()`



Insight: Most customers spend moderate amounts; distribution is right skewed.

### Task 3: Distribution of Tip Amount

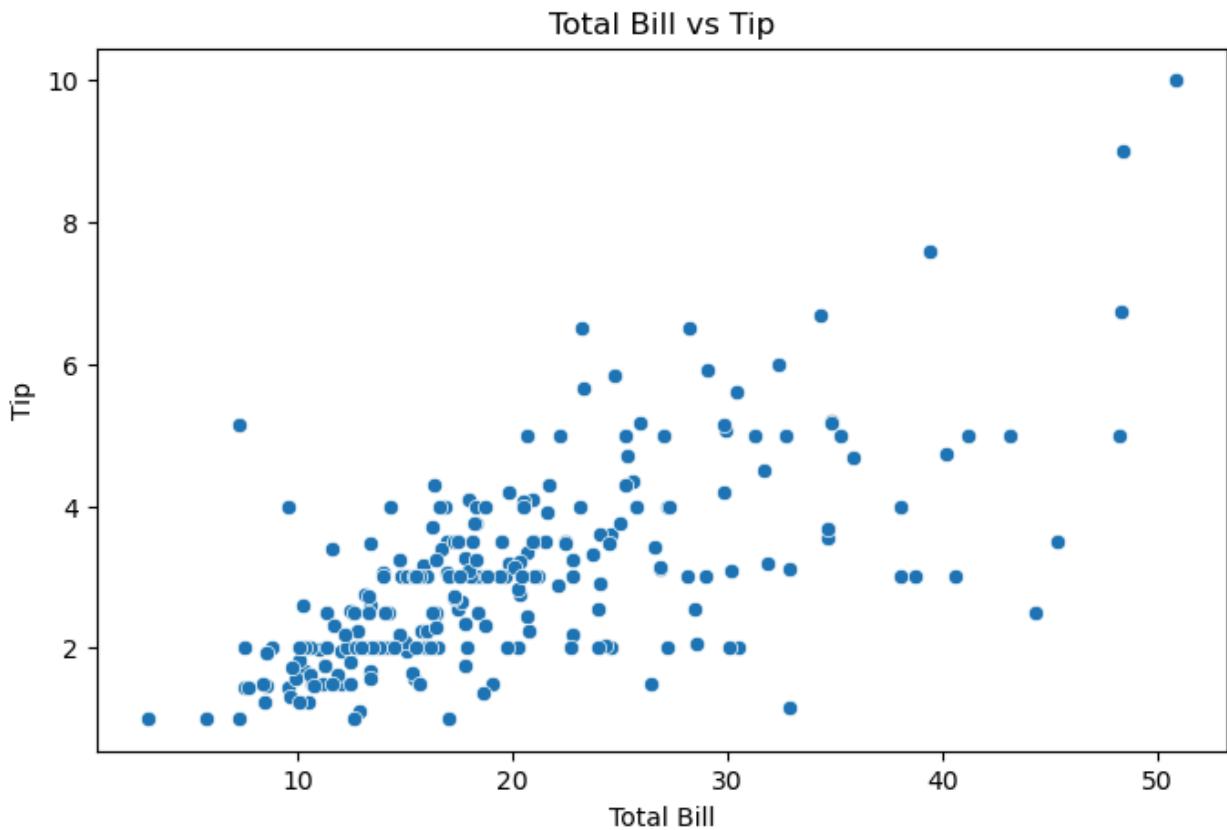
```
In [9]: plt.figure(figsize=(8,5))
sns.histplot(data=tips, x='tip', kde=True)
plt.title("Distribution of Tip Amount")
plt.xlabel("Tip")
plt.ylabel("Frequency")
plt.show()
```



**Insight:** Tips are usually small with few high values.

#### Task 4: Relationship Between Total Bill and Tip

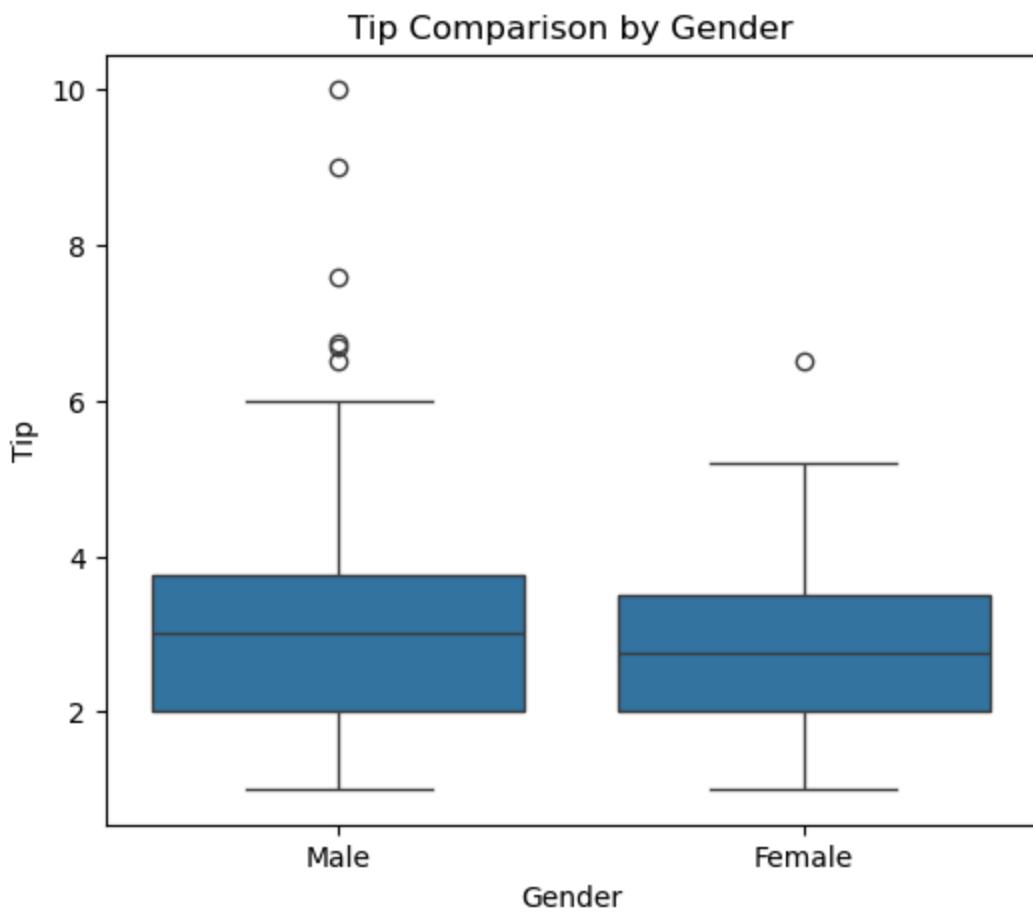
```
In [10]: plt.figure(figsize=(8,5))
sns.scatterplot(data=tips, x='total_bill', y='tip')
plt.title("Total Bill vs Tip")
plt.xlabel("Total Bill")
plt.ylabel("Tip")
plt.show()
```



Insight: Higher bills generally receive higher tips.

### Task 5: Gender-wise Tip Comparison

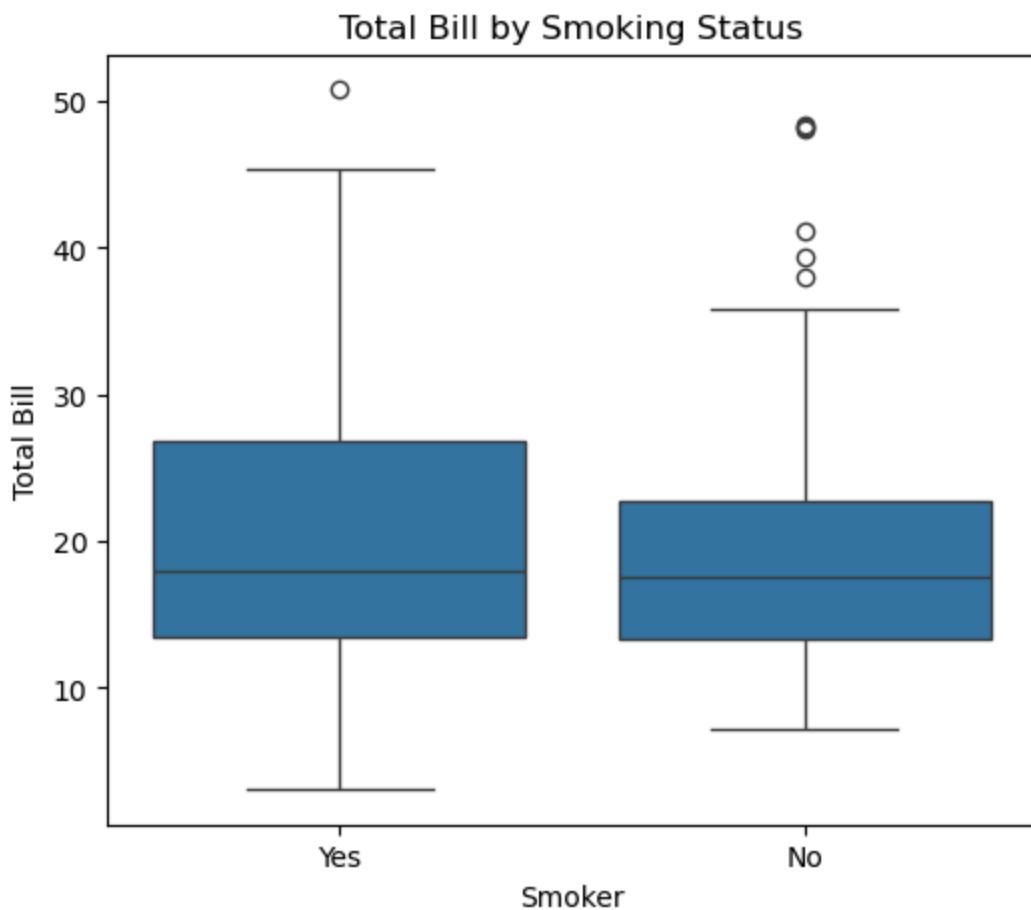
```
In [11]: plt.figure(figsize=(6,5))
sns.boxplot(data=tips, x='sex', y='tip')
plt.title("Tip Comparison by Gender")
plt.xlabel("Gender")
plt.ylabel("Tip")
plt.show()
```



Insight: Tipping behavior is similar for both genders.

### Task 6: Smoker vs Non-Smoker Spending

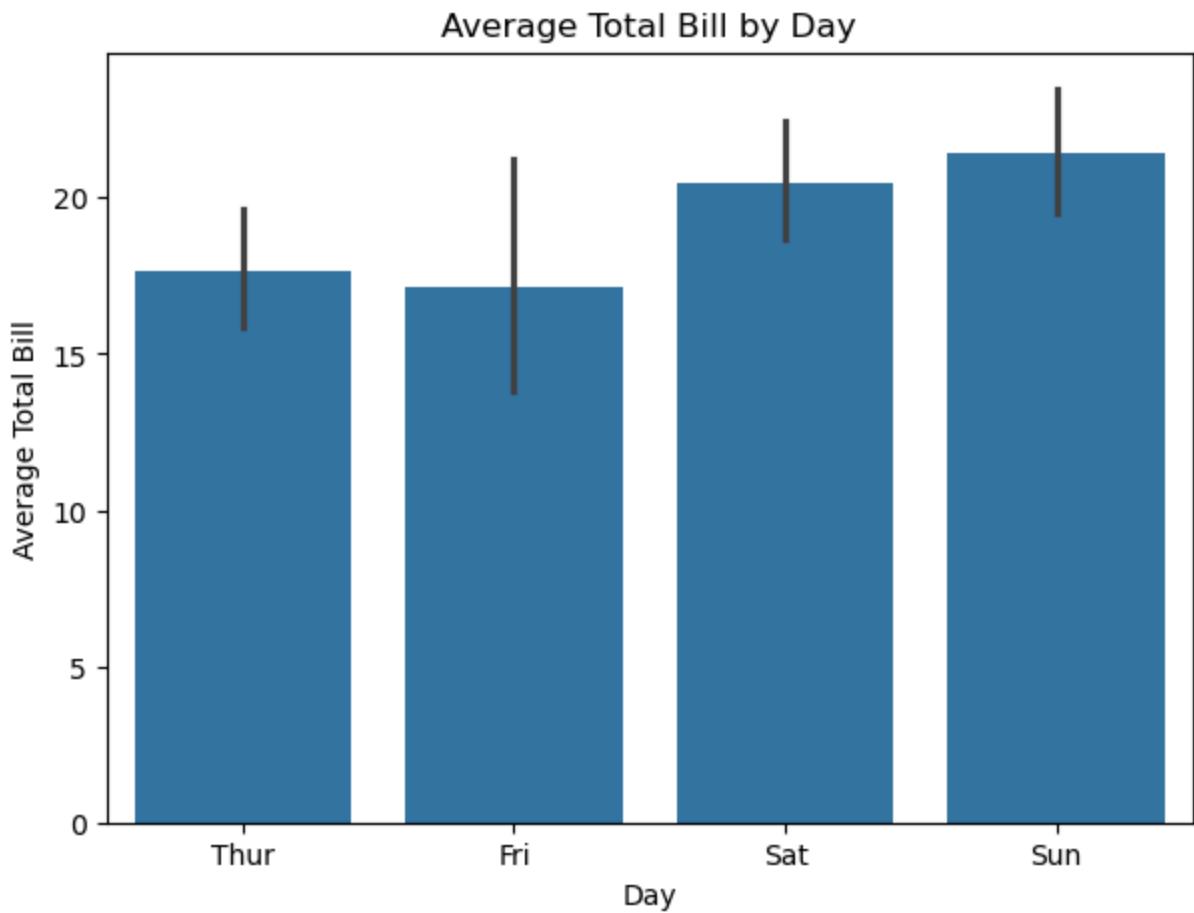
```
In [12]: plt.figure(figsize=(6,5))
sns.boxplot(data=tips, x='smoker', y='total_bill')
plt.title("Total Bill by Smoking Status")
plt.xlabel("Smoker")
plt.ylabel("Total Bill")
plt.show()
```



Insight: Smokers show slightly higher spending in some cases.

### Task 7: Day-wise Average Total Bill

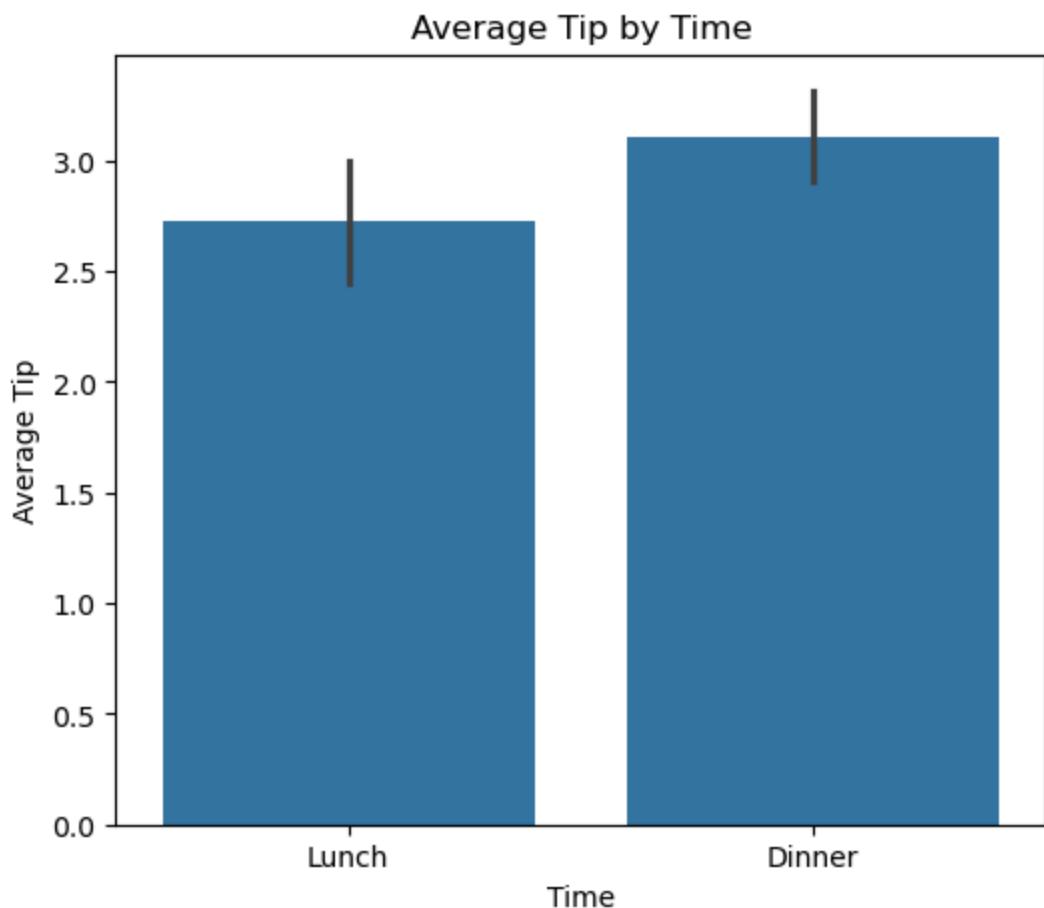
```
In [13]: plt.figure(figsize=(7,5))
sns.barplot(data=tips, x='day', y='total_bill')
plt.title("Average Total Bill by Day")
plt.xlabel("Day")
plt.ylabel("Average Total Bill")
plt.show()
```



Insight: Weekend days show higher average spending.

### Task 8: Time-based Tip Analysis

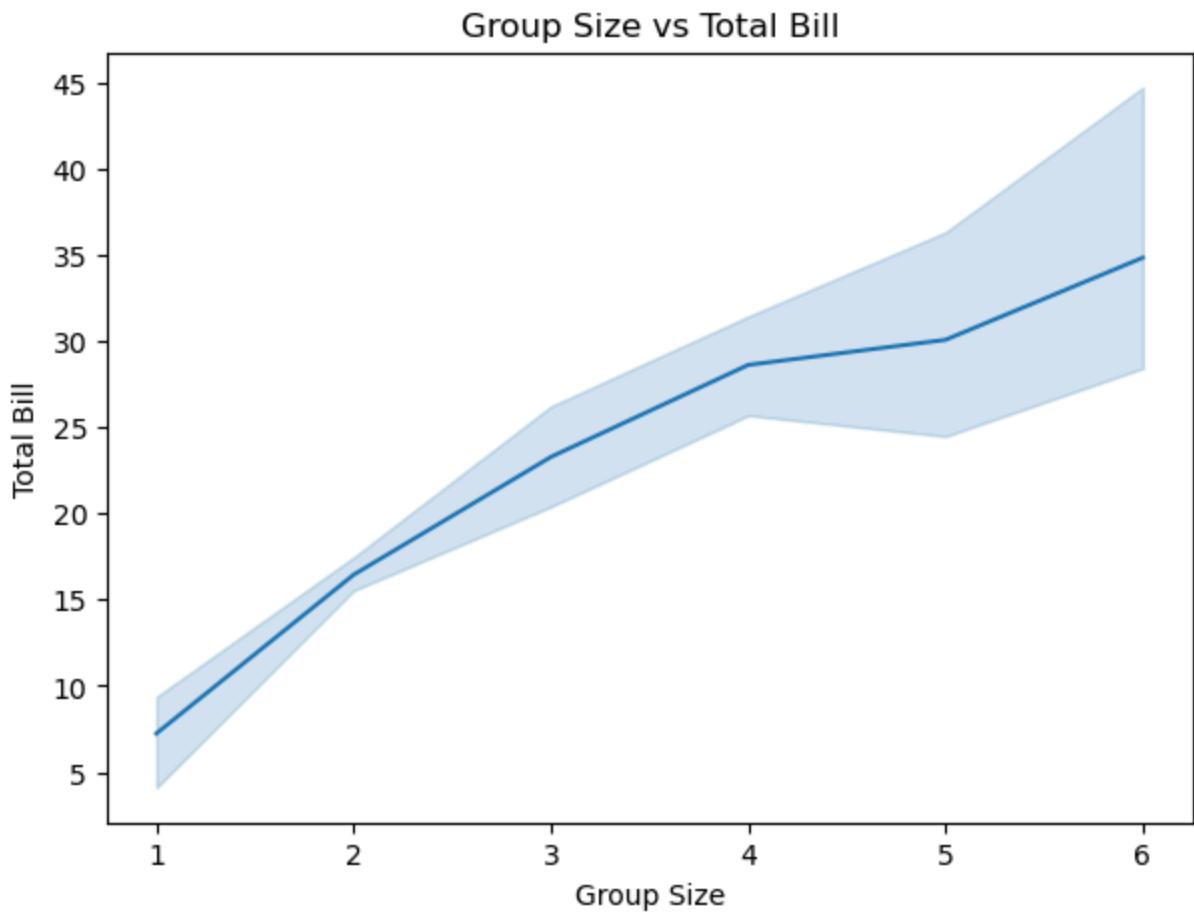
```
In [14]: plt.figure(figsize=(6,5))
sns.barplot(data=tips, x='time', y='tip')
plt.title("Average Tip by Time")
plt.xlabel("Time")
plt.ylabel("Average Tip")
plt.show()
```



Insight: Dinner time receives higher tips than lunch.

### Task 9: Group Size vs Total Bill

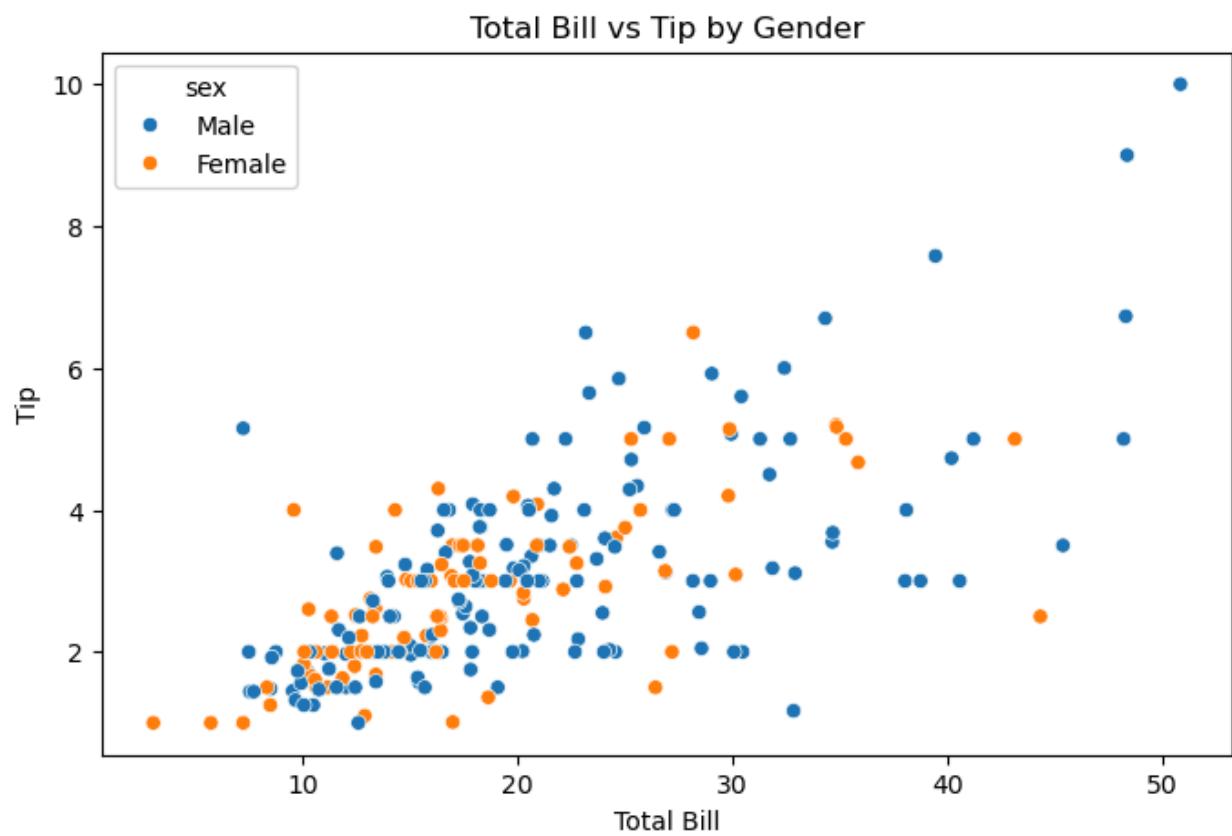
```
In [16]: plt.figure(figsize=(7,5))
sns.lineplot(data=tips, x='size', y='total_bill')
plt.title("Group Size vs Total Bill")
plt.xlabel("Group Size")
plt.ylabel("Total Bill")
plt.show()
```



Insight: Total bill increases as group size increases.

## Task 10: Combined Analysis Using Hue

```
In [17]: plt.figure(figsize=(8,5))
sns.scatterplot(data=tips, x='total_bill', y='tip', hue='sex')
plt.title("Total Bill vs Tip by Gender")
plt.xlabel("Total Bill")
plt.ylabel("Tip")
plt.show()
```



**Insight: For similar bills, tipping is almost same across genders.**

## Analytical Answers

1. Which factor appears to influence tip amount the most?

The total bill amount influences the tip the most. As the total bill increases, the tip amount generally increases.

2. What customer segment should the restaurant focus on to maximize revenue?

The restaurant should focus on customers who come in larger groups, especially during dinner time on weekends, as they generate higher total bills.

3. Are tipping patterns consistent across days and time periods?

No, tipping patterns are not consistent. Tips are usually higher during dinner time and on weekends compared to lunch and weekdays.

In [ ]: