

# FY B.tech 2022-23

Name: Shivam Nivrutti Pokharkar

Roll: 546

PRN: 202201040078

Div: E3

Sub: Assignment 5

---

Code:

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

# Load the dataset into a Pandas DataFrame
df = pd.read_csv("D:\python progs\LAB\Assig 5\Assig 5\k_tips.csv")

# 1. How does the total bill vary with the tip amount?
plt.figure(figsize=(8, 6))
plt.scatter(df['total_bill'], df['tip'])
plt.xlabel('Total Bill')
plt.ylabel('Tip')
plt.title('Total Bill vs Tip Amount')

# 2. What is the distribution of total bills?
plt.figure(figsize=(8, 6))
plt.hist(df['total_bill'], bins=10)
plt.xlabel('Total Bill')
plt.ylabel('Frequency')
plt.title('Distribution of Total Bills')

# 3. How does the tip amount differ between males and females?
plt.figure(figsize=(8, 6))
plt.boxplot([df[df['sex'] == 'Male']['tip'], df[df['sex'] == 'Female']['tip']], labels=['Male', 'Female'])
plt.xlabel('Sex')
plt.ylabel('Tip Amount')
plt.title('Tip Amount by Gender')

# 4. How does the time of day (lunch/dinner) affect the total bill?
plt.figure(figsize=(8, 6))
df.groupby('time')['total_bill'].mean().plot(kind='bar')
plt.xlabel('Time of Day')
plt.ylabel('Average Total Bill')
plt.title('Average Total Bill by Time of Day')

# 5. Is there a difference in the tip amount between smokers and non-smokers?
plt.figure(figsize=(8, 6))
```

```

df.groupby('smoker')['tip'].mean().plot(kind='bar')
plt.xlabel('Smoker')
plt.ylabel('Average Tip')
plt.title('Average Tip Amount by Smoker')

# 6. How does the size of the group affect the total bill?
plt.figure(figsize=(8, 6))
df.groupby('size')['total_bill'].mean().plot(kind='bar')
plt.xlabel('Group Size')
plt.ylabel('Average Total Bill')
plt.title('Average Total Bill by Group Size')

# 7. Which day of the week has the highest total bill?
plt.figure(figsize=(8, 6))
df.groupby('day')['total_bill'].sum().plot(kind='bar')
plt.xlabel('Day of the Week')
plt.ylabel('Total Bill')
plt.title('Total Bill by Day of the Week')

# 8. What is the distribution of tips based on gender?
plt.figure(figsize=(8, 6))
plt.hist([df[df['sex'] == 'Male']['tip'], df[df['sex'] == 'Female']['tip']],
bins=10, label=['Male', 'Female'])
plt.xlabel('Tip Amount')
plt.ylabel('Frequency')
plt.title('Distribution of Tips by Gender')
plt.legend()

# 9. How does the tip amount vary with the total bill, categorized by day?
plt.figure(figsize=(8, 6))
days = df['day'].unique()
for day in days:
    plt.scatter(df[df['day'] == day]['total_bill'], df[df['day'] ==
day]['tip'], label=day)
plt.xlabel('Total Bill')
plt.ylabel('Tip Amount')
plt.title('Tip Amount vs Total Bill (Categorized by Day)')
plt.legend()

# 10. How does the total bill vary with the size of the group?
plt.figure(figsize=(8, 6))
plt.scatter(df['size'], df['total_bill'])
plt.xlabel('Group Size')
plt.ylabel('Total Bill')
plt.title('Total Bill vs Group Size')

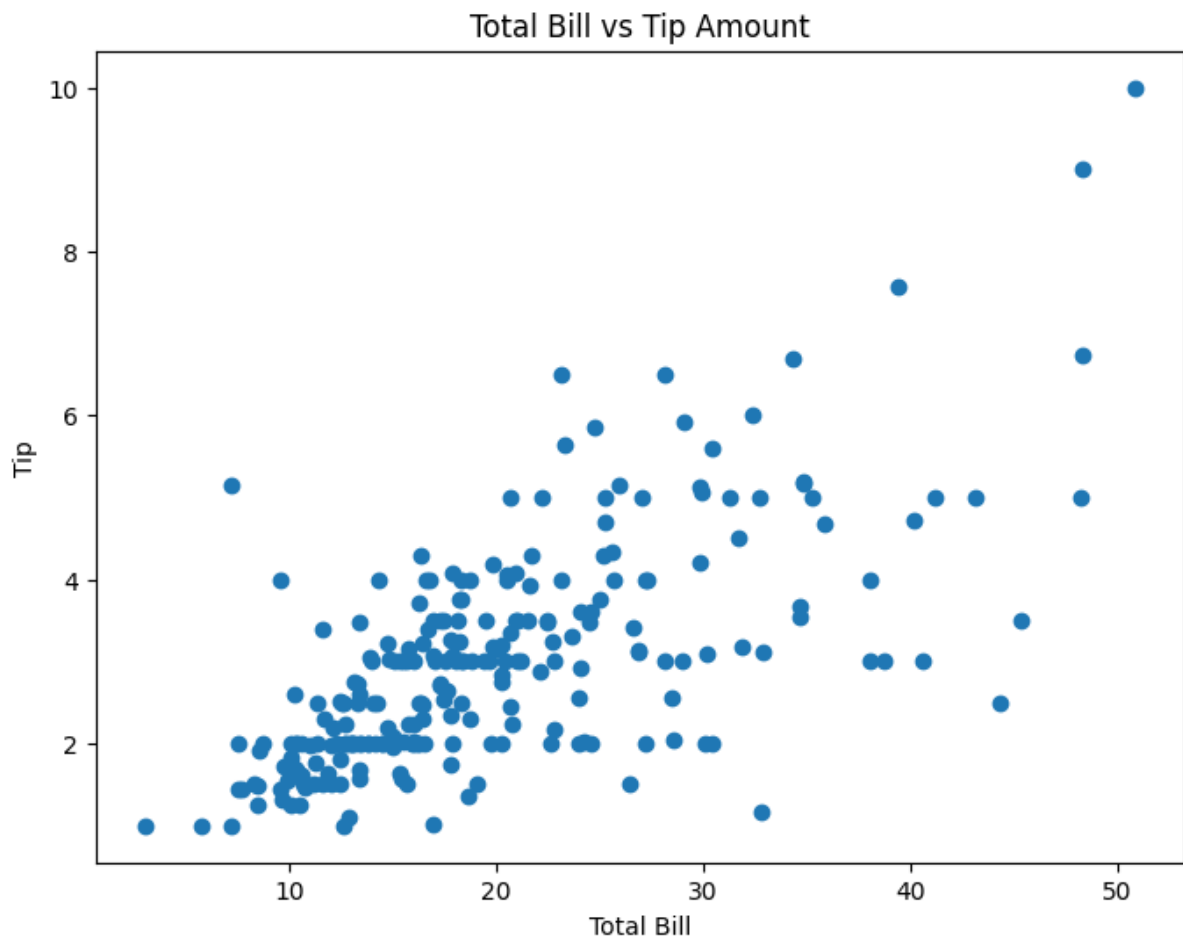
# 11. What is the percentage distribution of meals by time of day?
plt.figure(figsize=(8, 6))

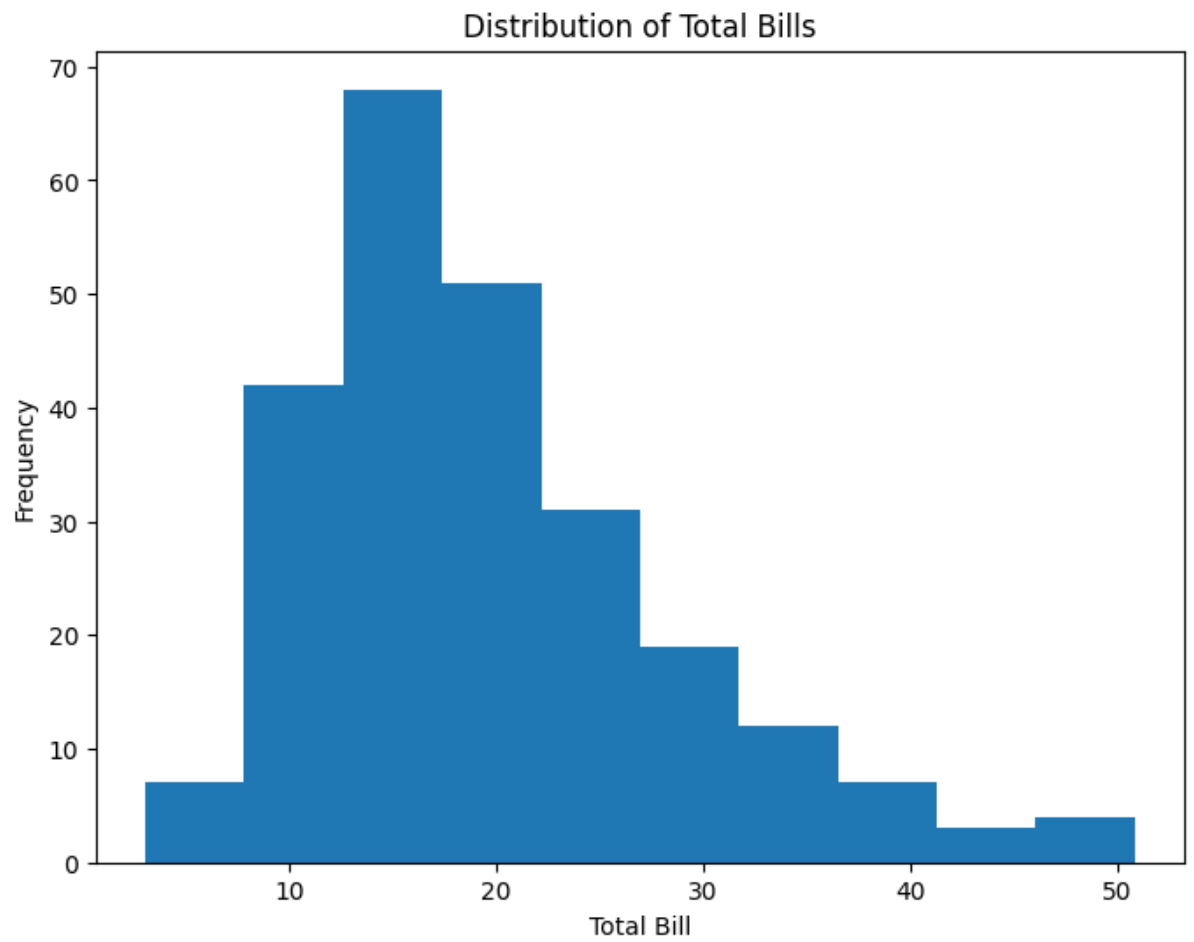
```

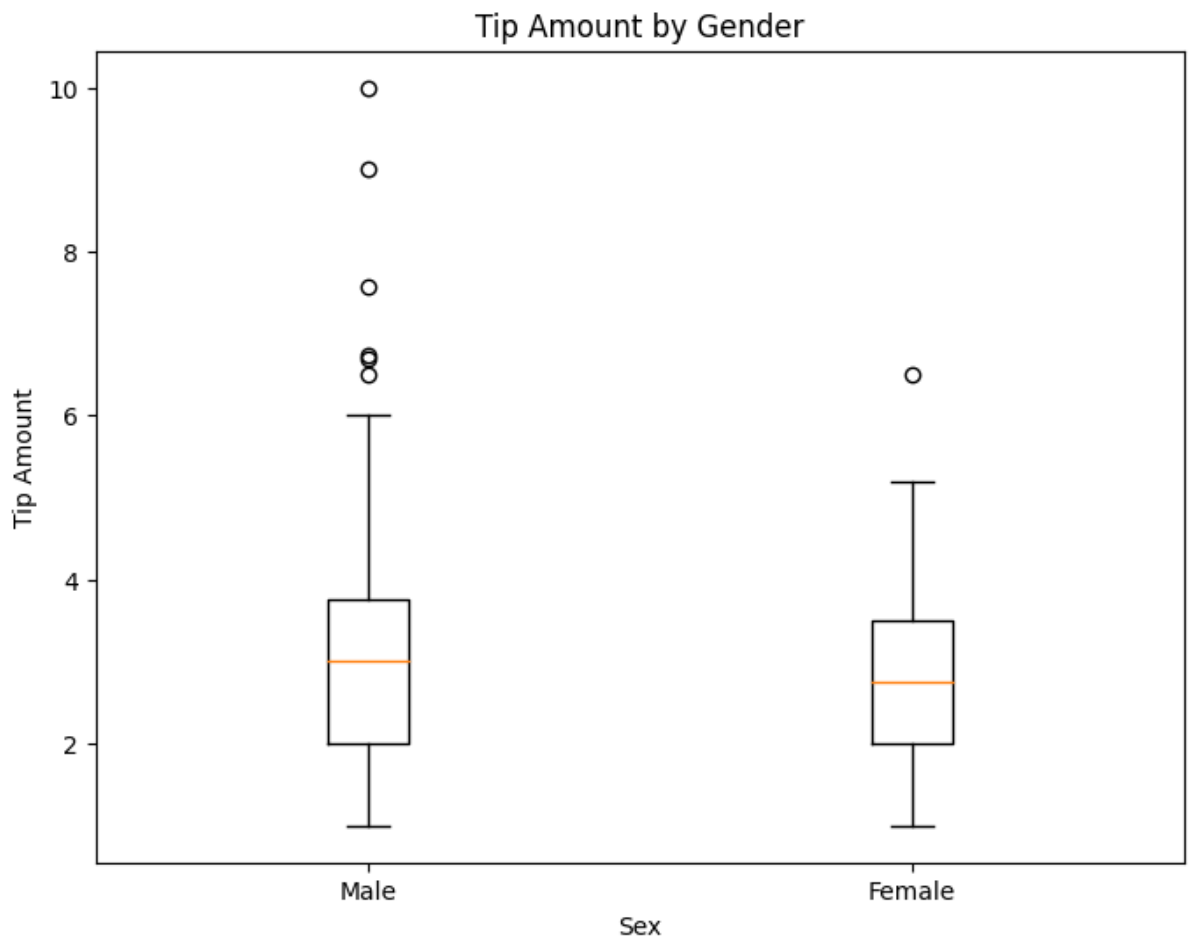
```
time_counts = df['time'].value_counts()
plt.pie(time_counts, labels=time_counts.index, autopct='%1.1f%%')
plt.title('Meal Distribution by Time of Day')

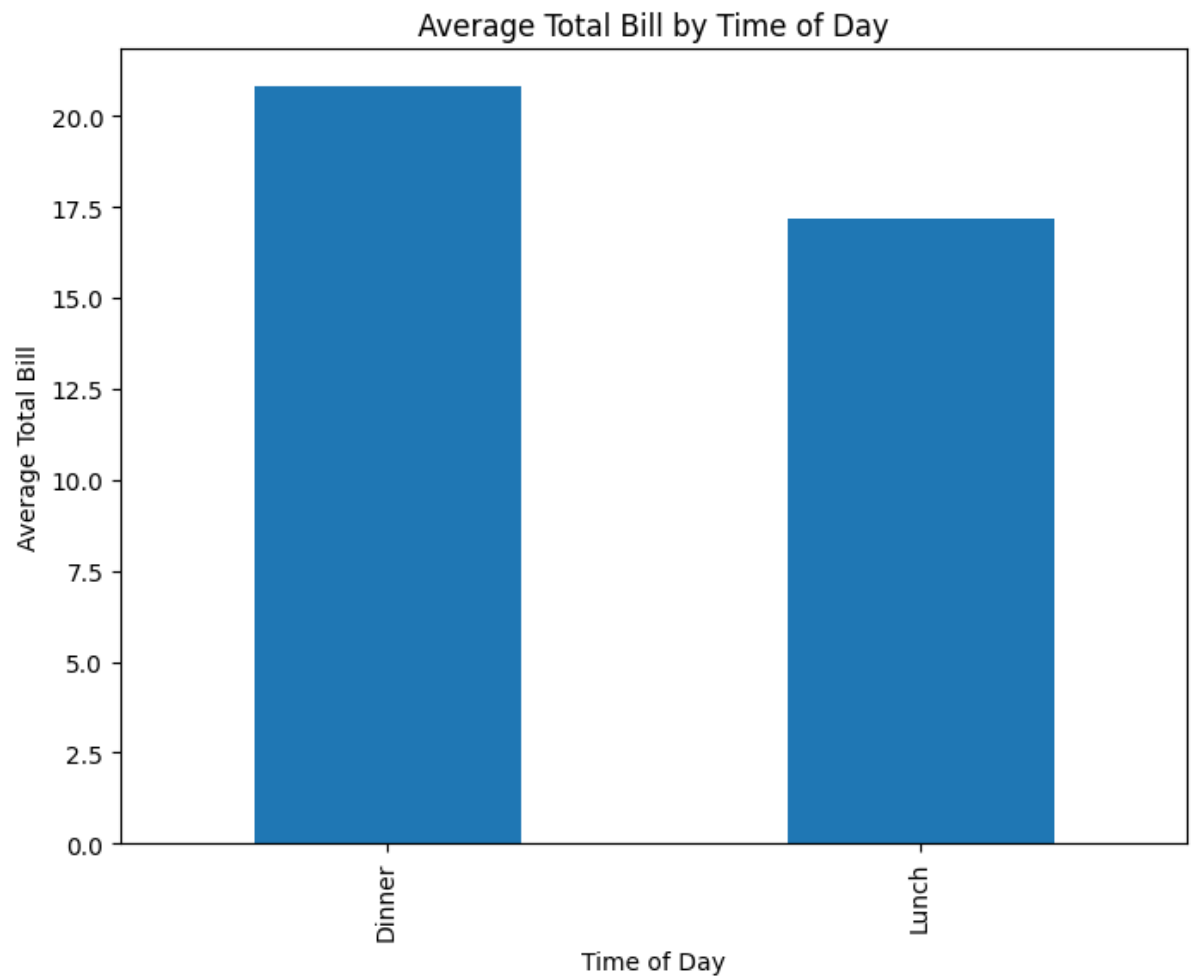
# Display all the plots
plt.show()
```

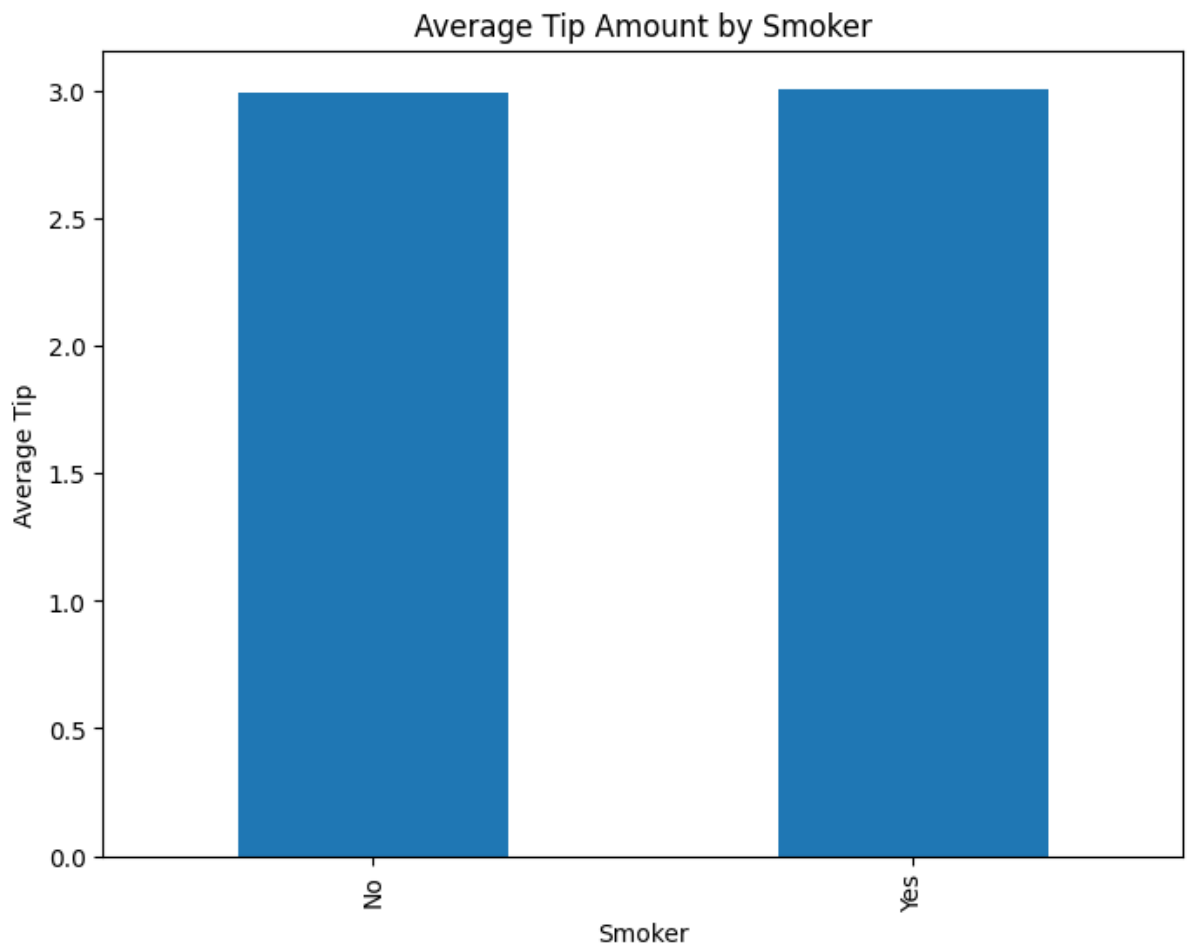
**Result:**

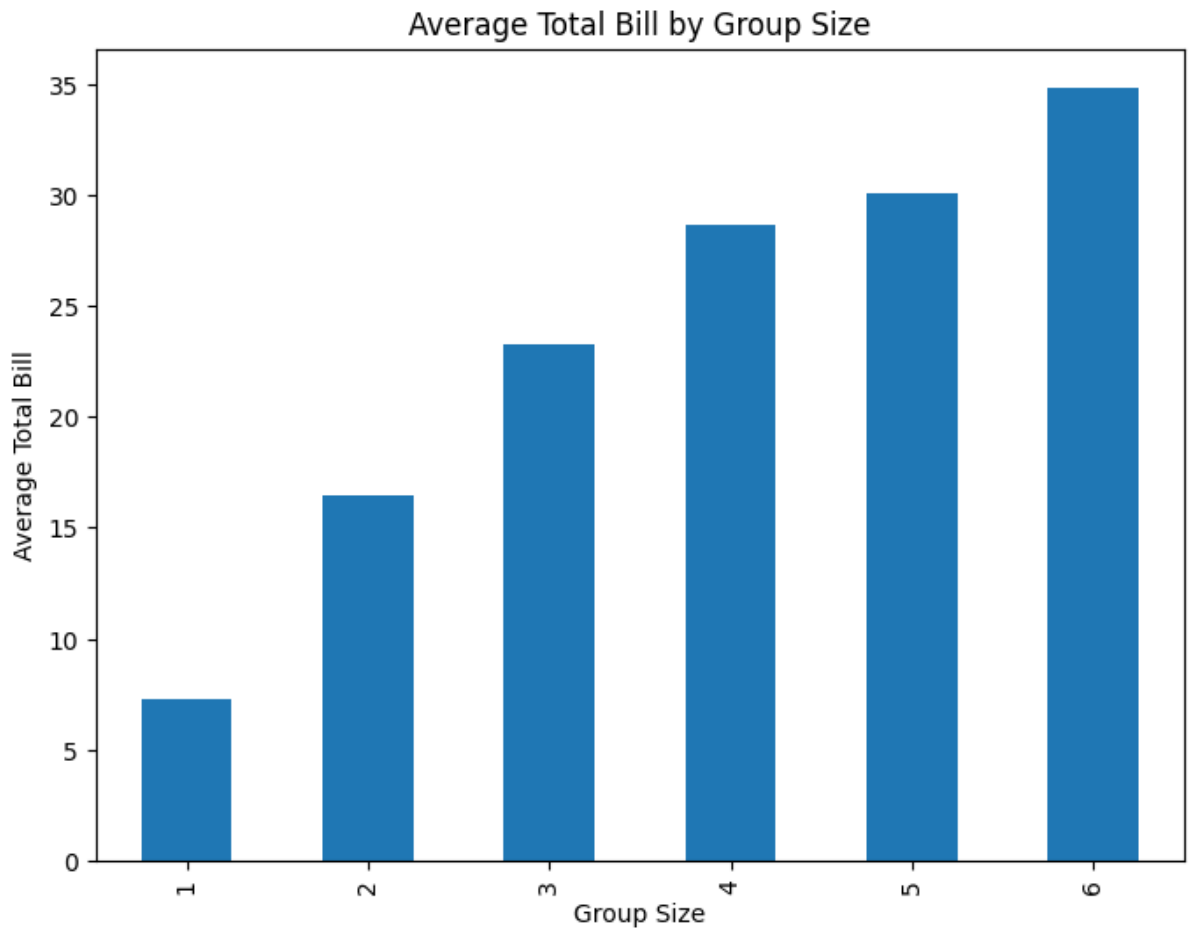




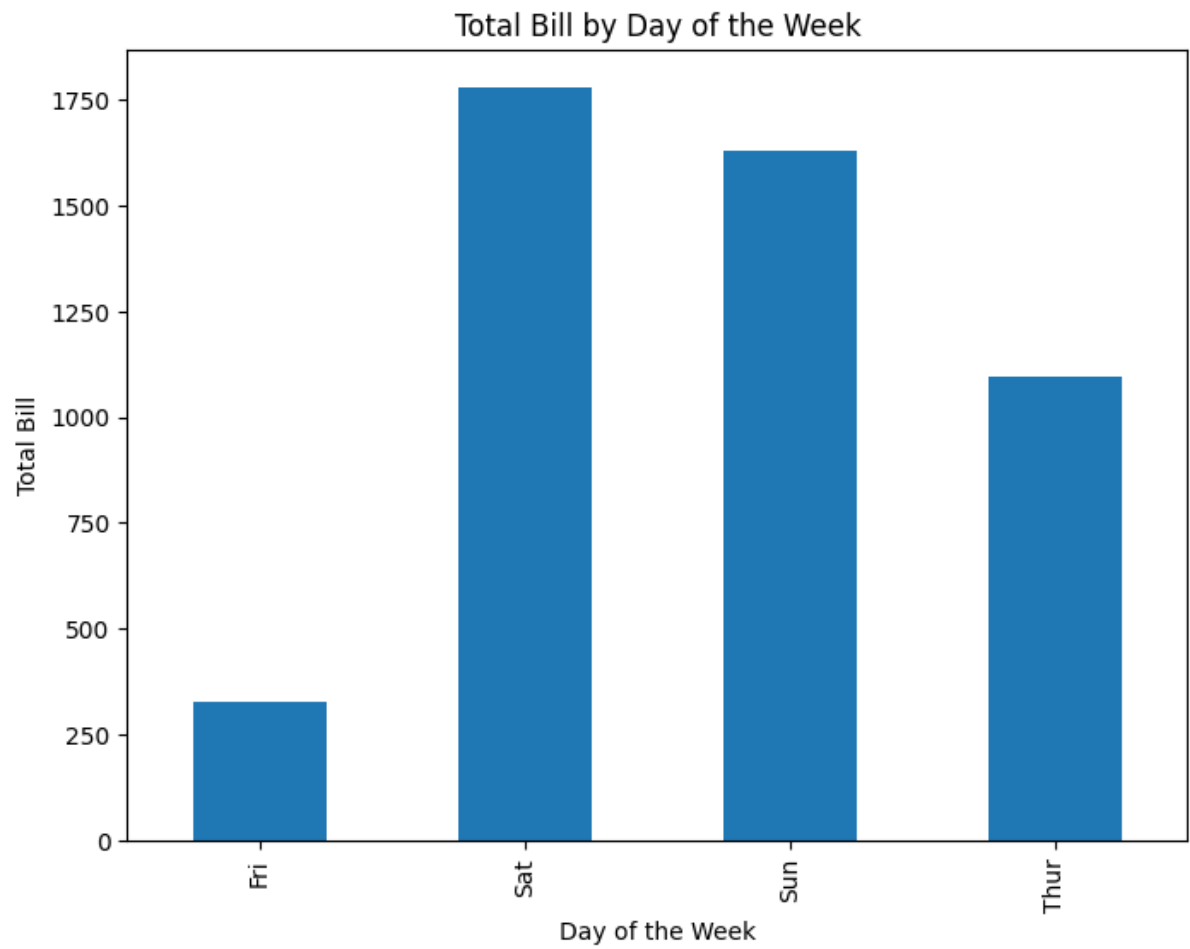


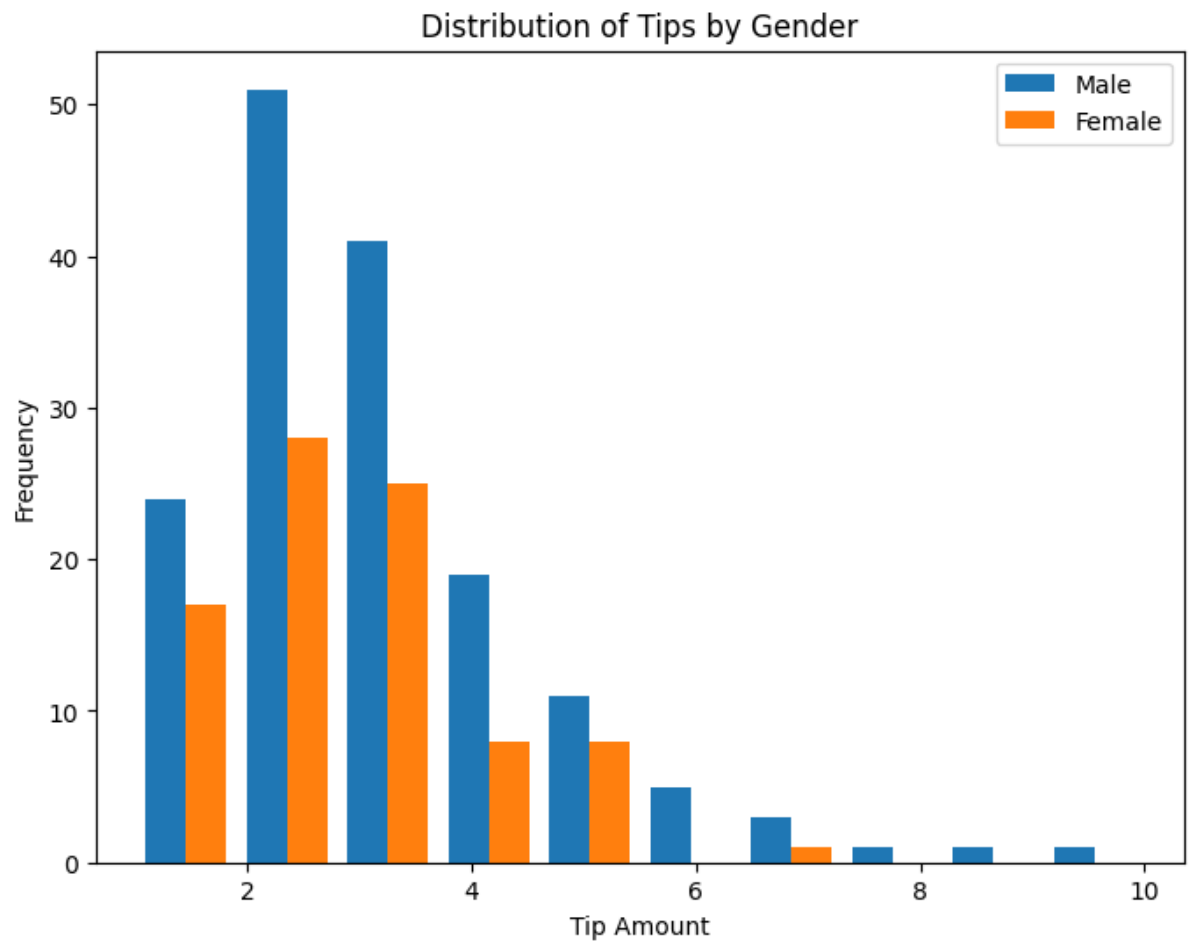


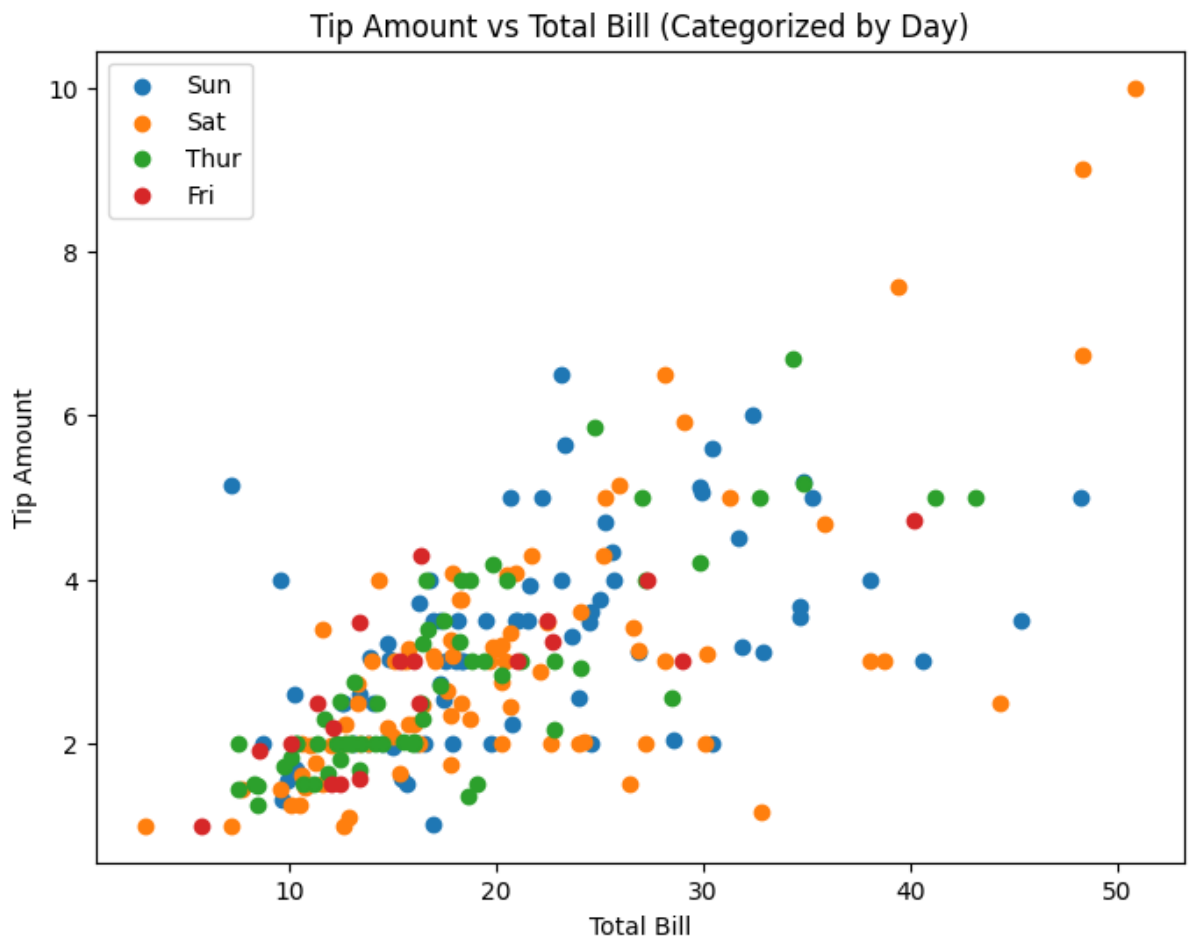


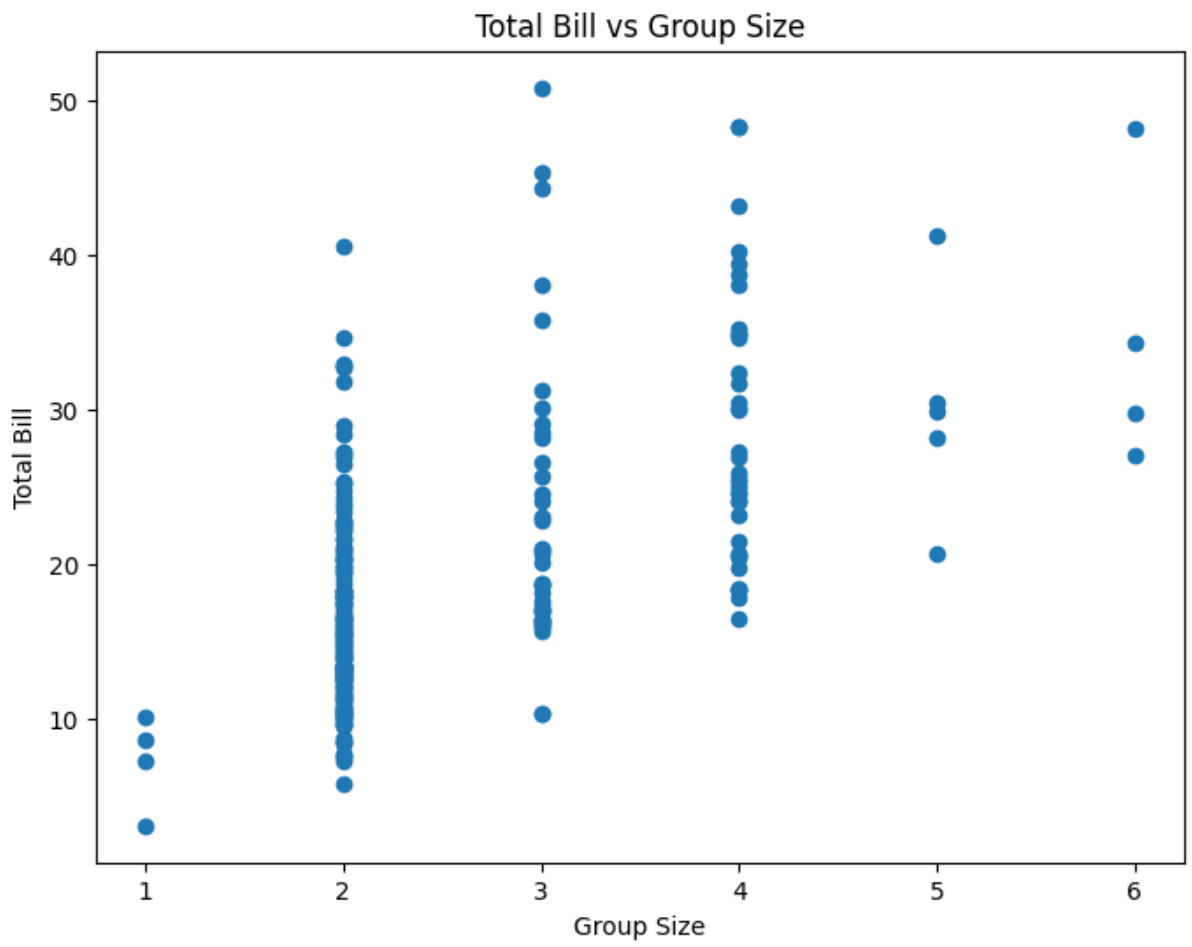












Meal Distribution by Time of Day

