FY B.tech 2022-23

Name: Saurabh singh Roll: 583
PRN: 202201090160
Div: E4
Sub: Assignment 5

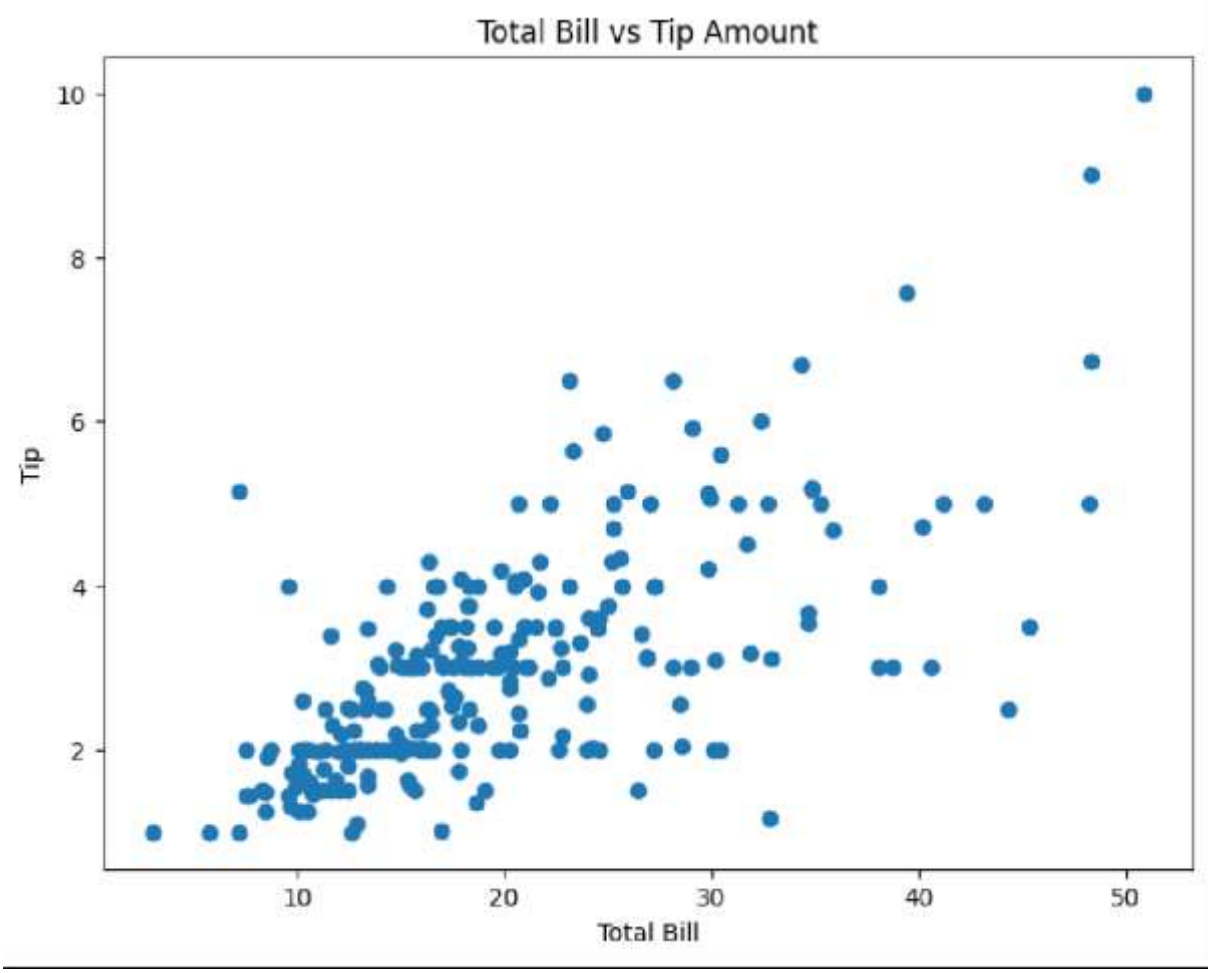
Code:

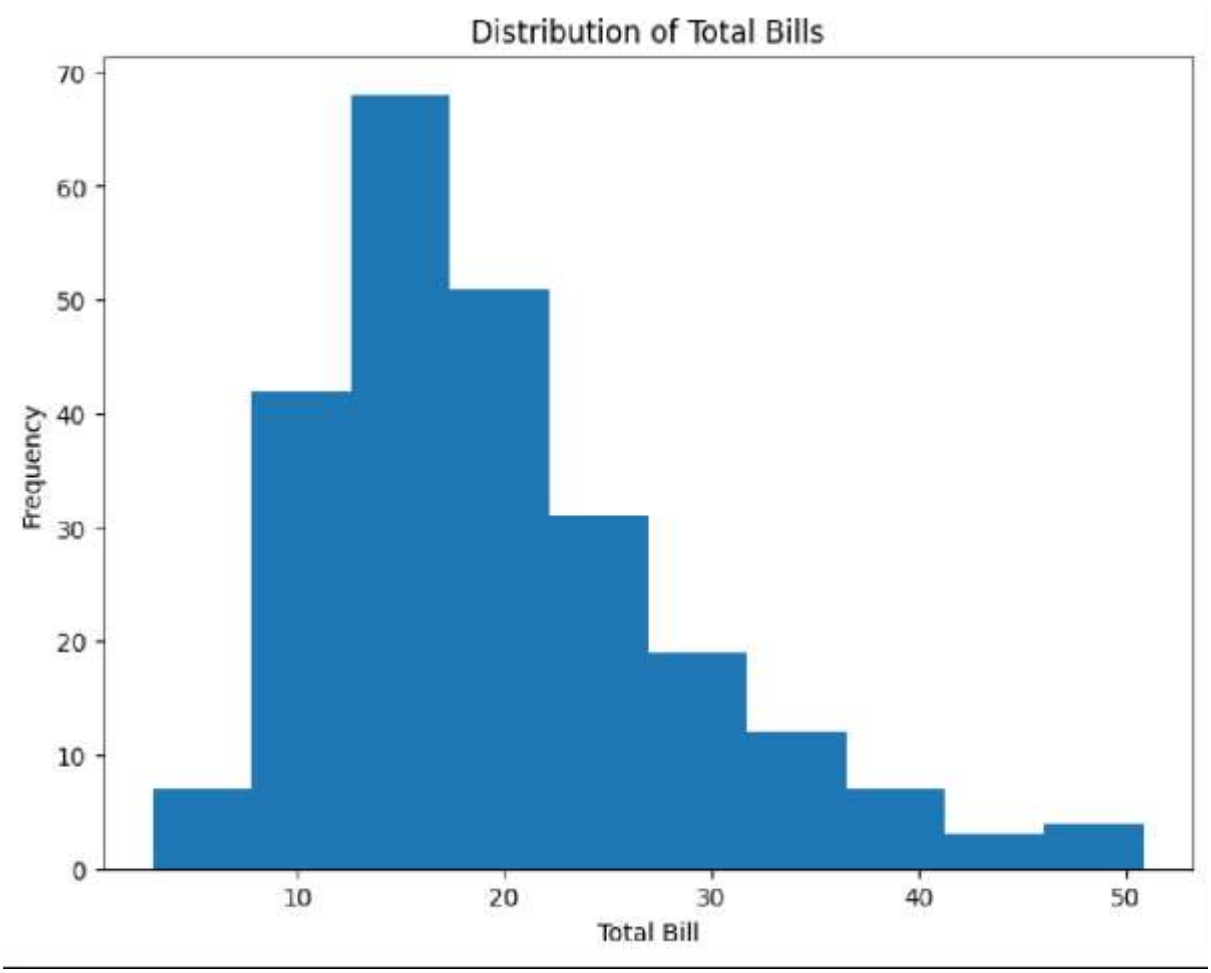
```
import pandas as pd
import matplotlib.pyplot as plt

# 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))
```

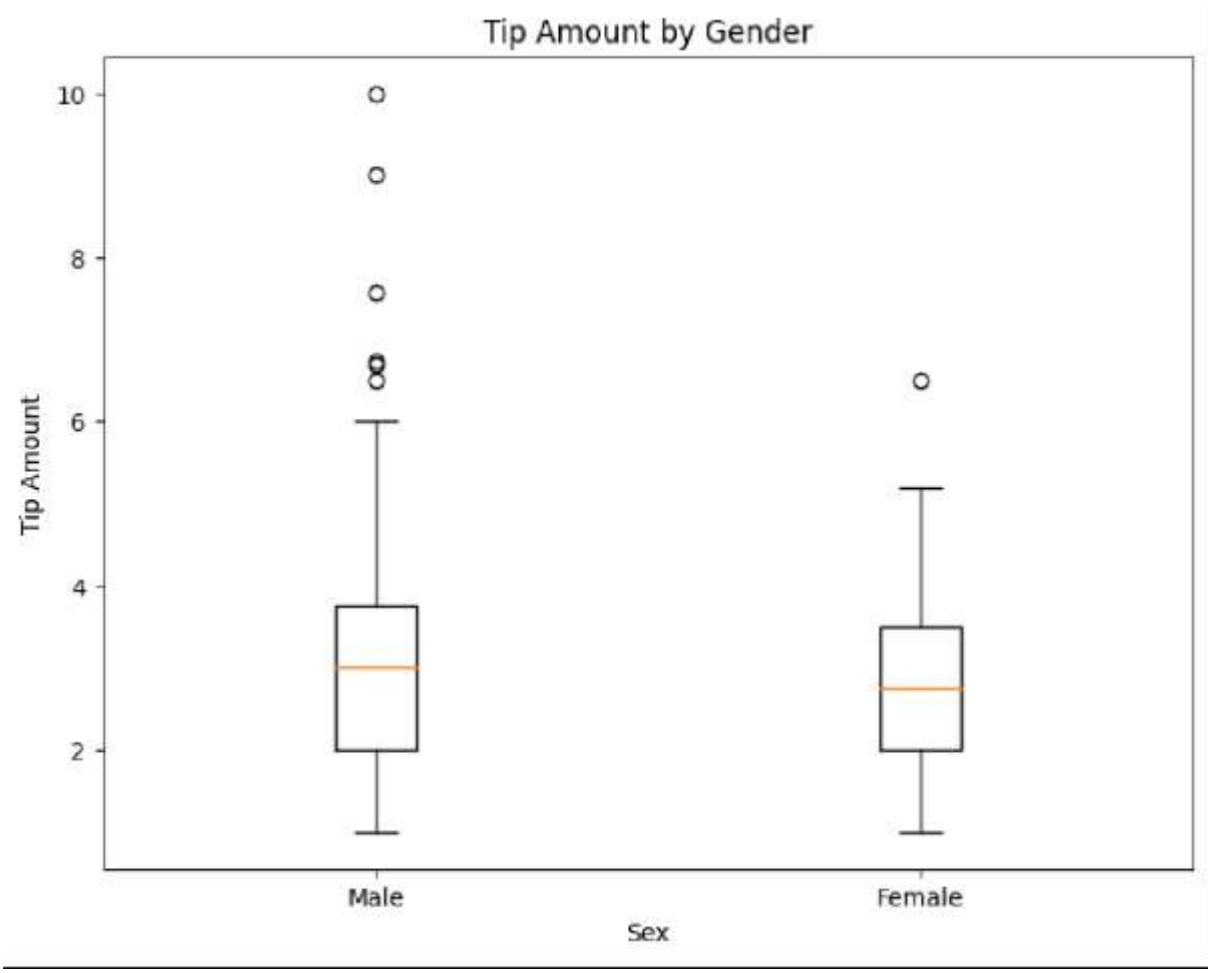
```
df.groupby('smoker')['tip'].mean().plot(kind='bar')
plt.xlabel('Smoker')
plt.ylabel('Average Tip')
plt.title('Average Tip Amount by Smoker')

# 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))
```

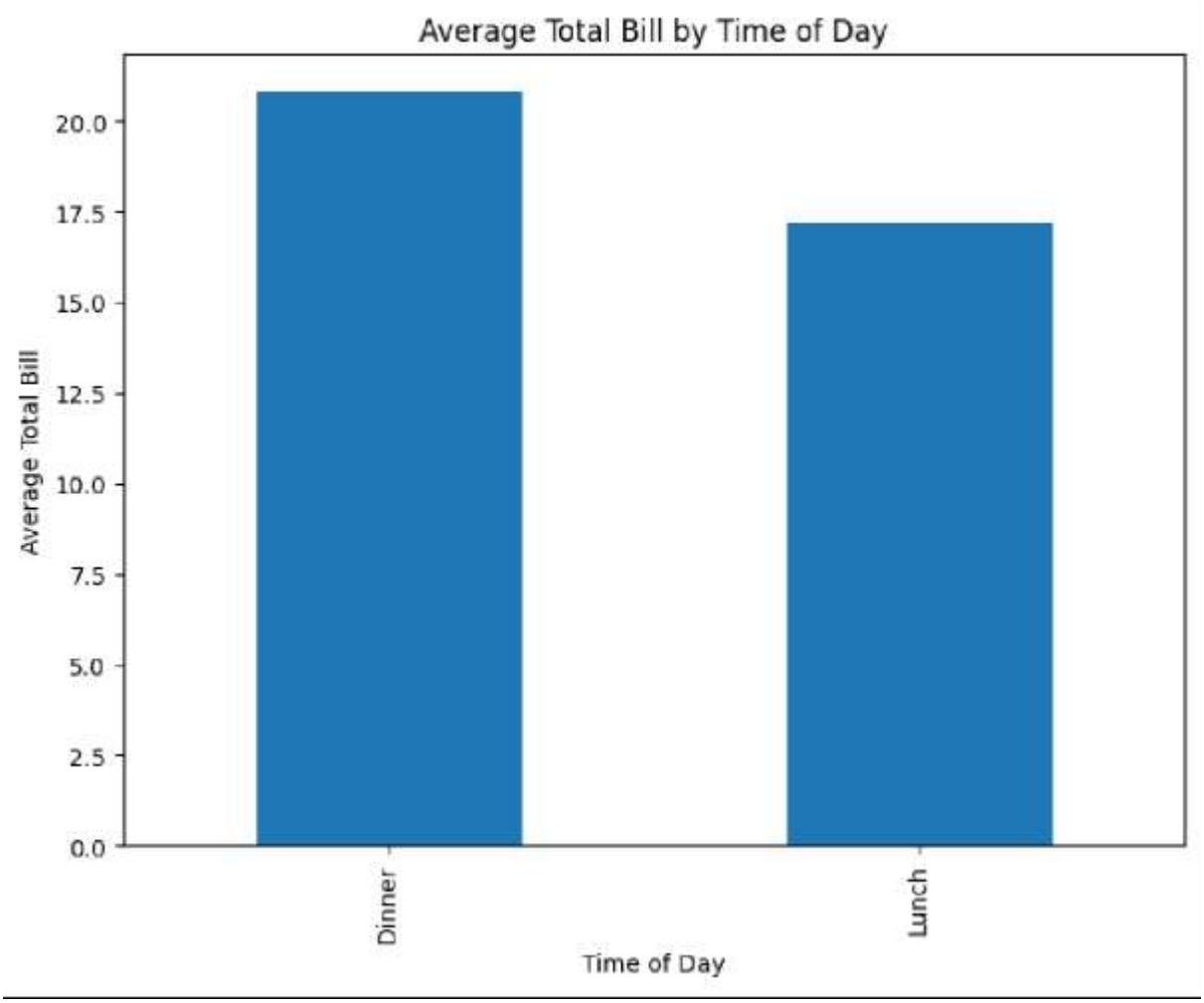
```
time_counts = df['time'].value_counts()
plt.pie(time_counts, labels=time_counts.index, autopct='%1.1f%%')
plt.title('Meal Distribution by Time of Day')

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

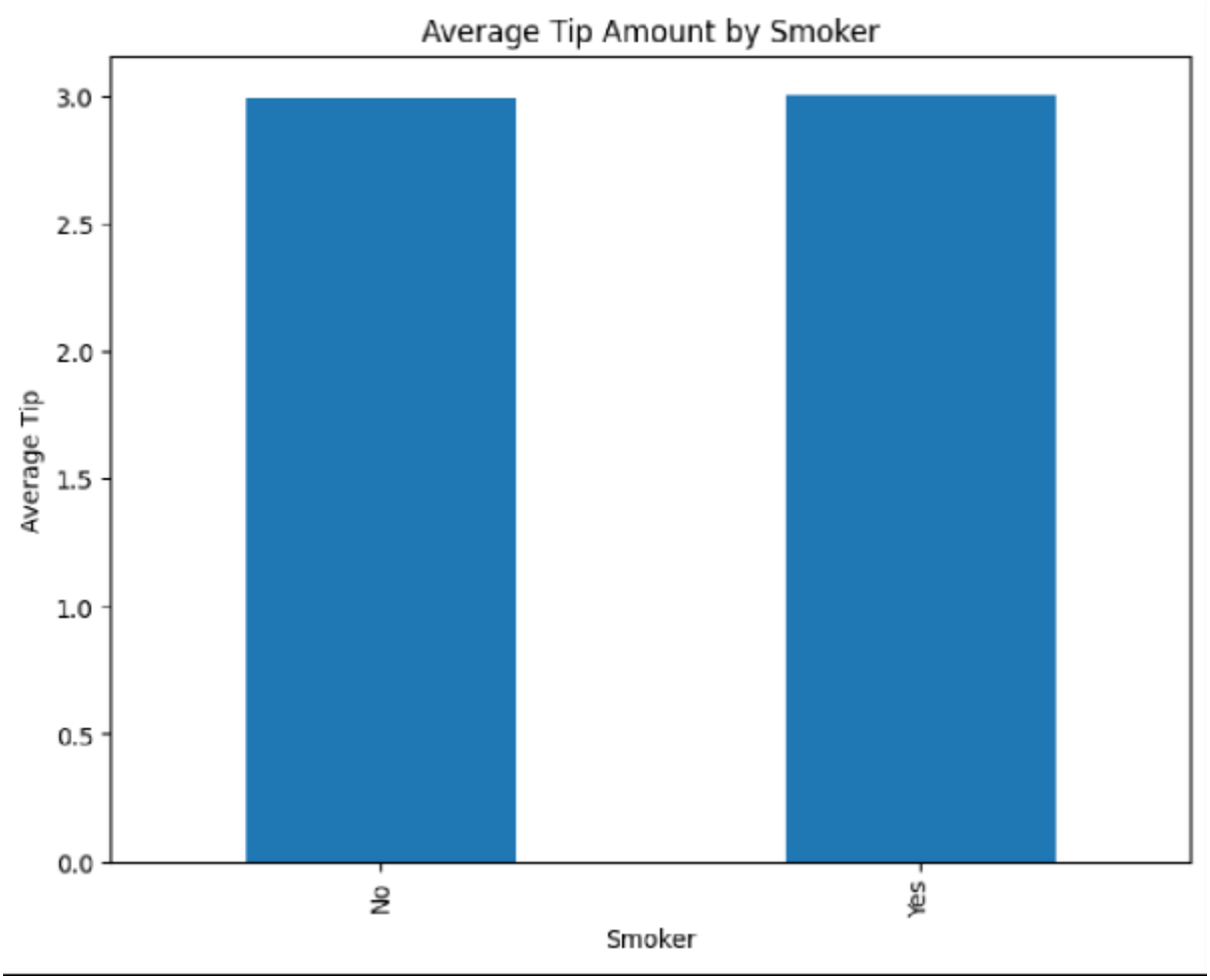
Result:

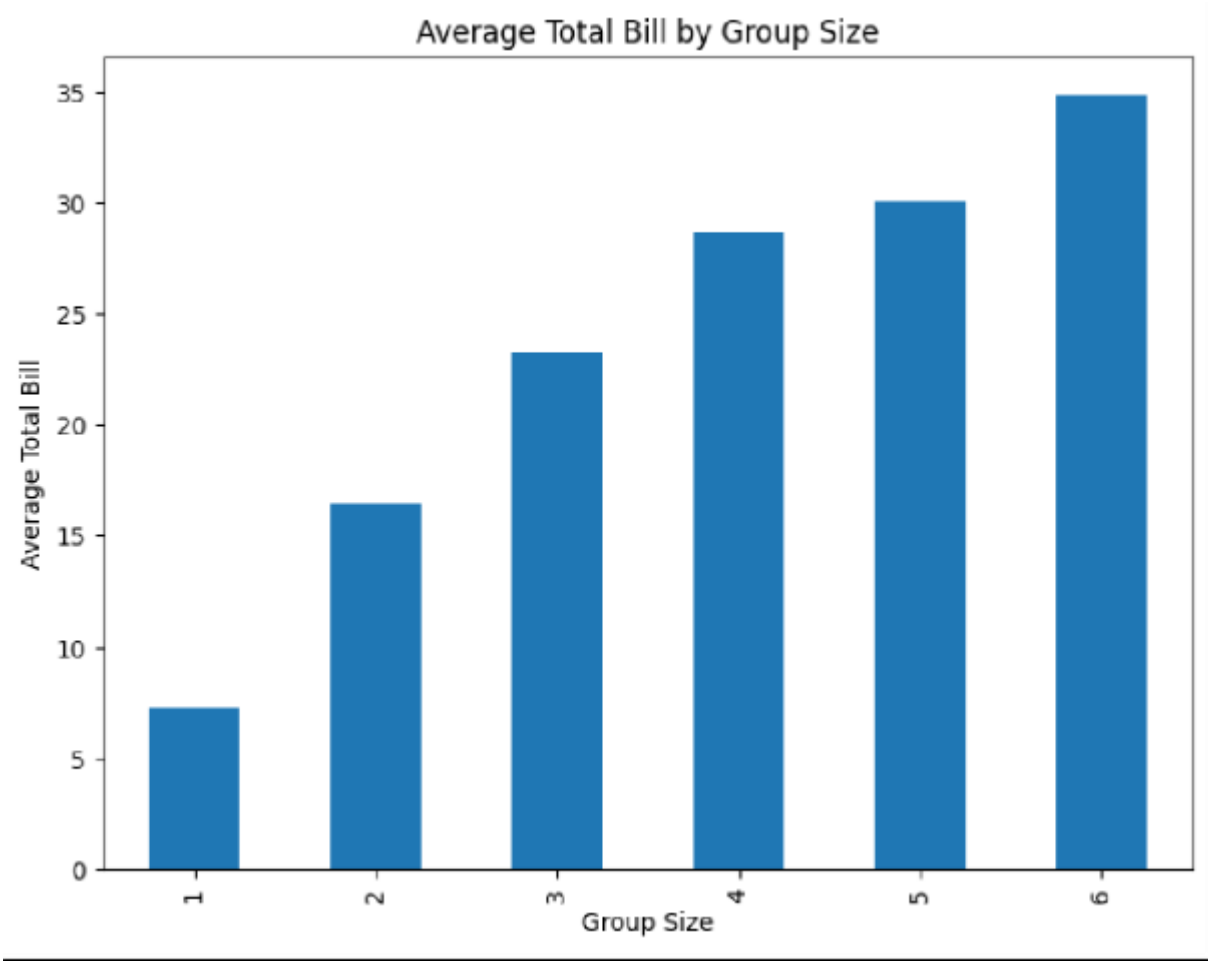


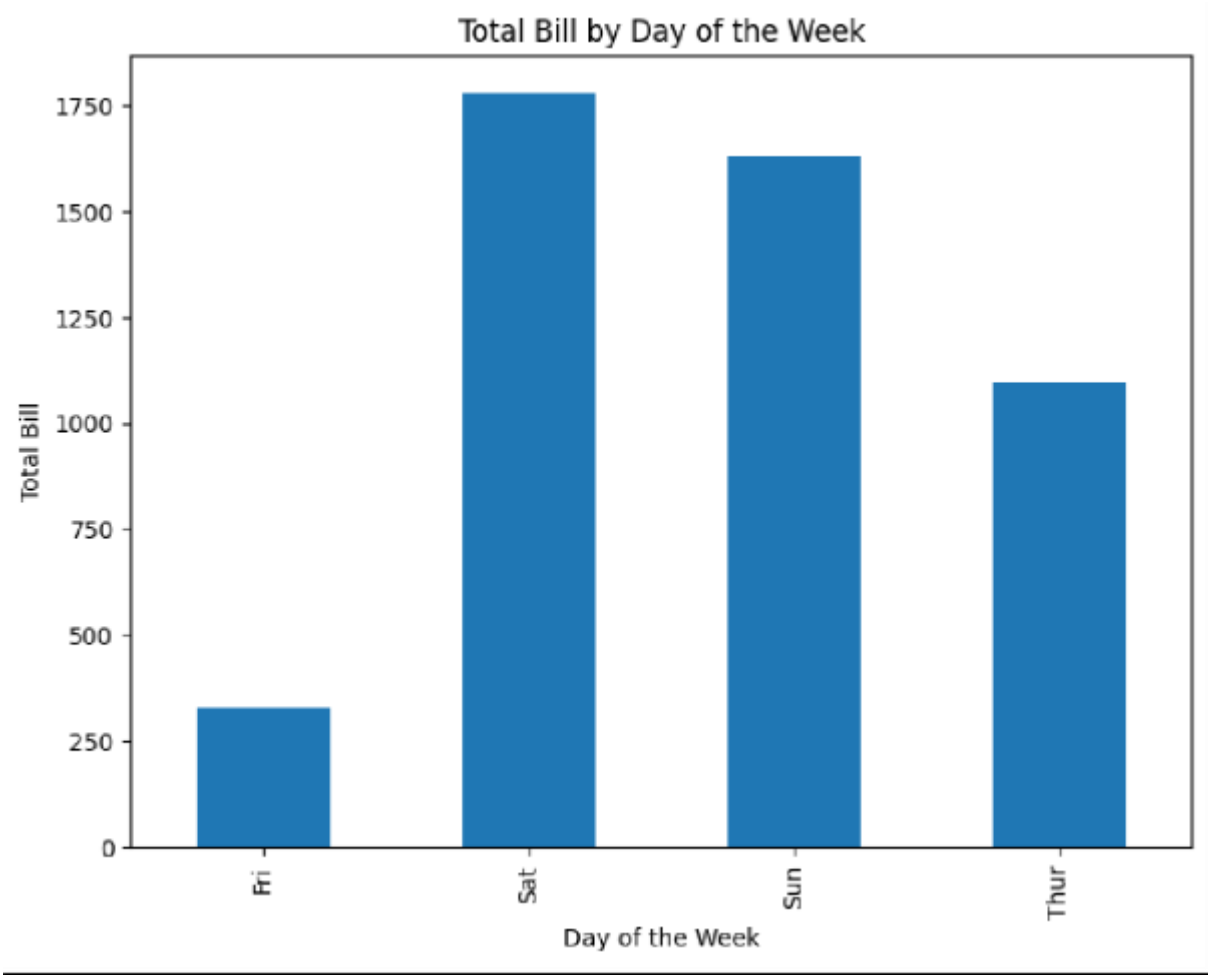

Distribution of Total Bills
Frequency
Total Bill
0
10
20
30
40
50
60
70
10
20
30
40
50

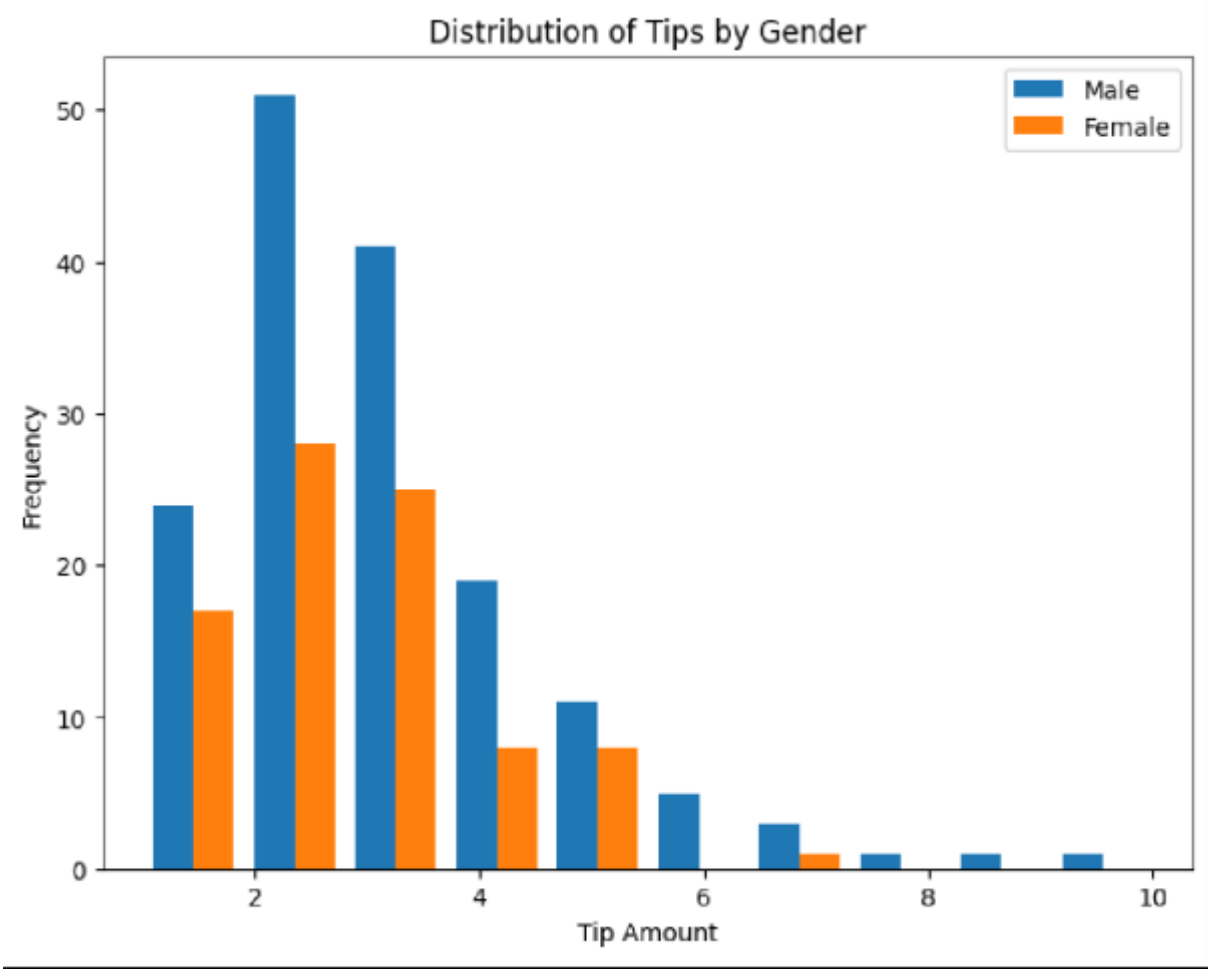

Tip Amount by Gender
Tip Amount
10
8
6
4
2
Male
Female
Sex

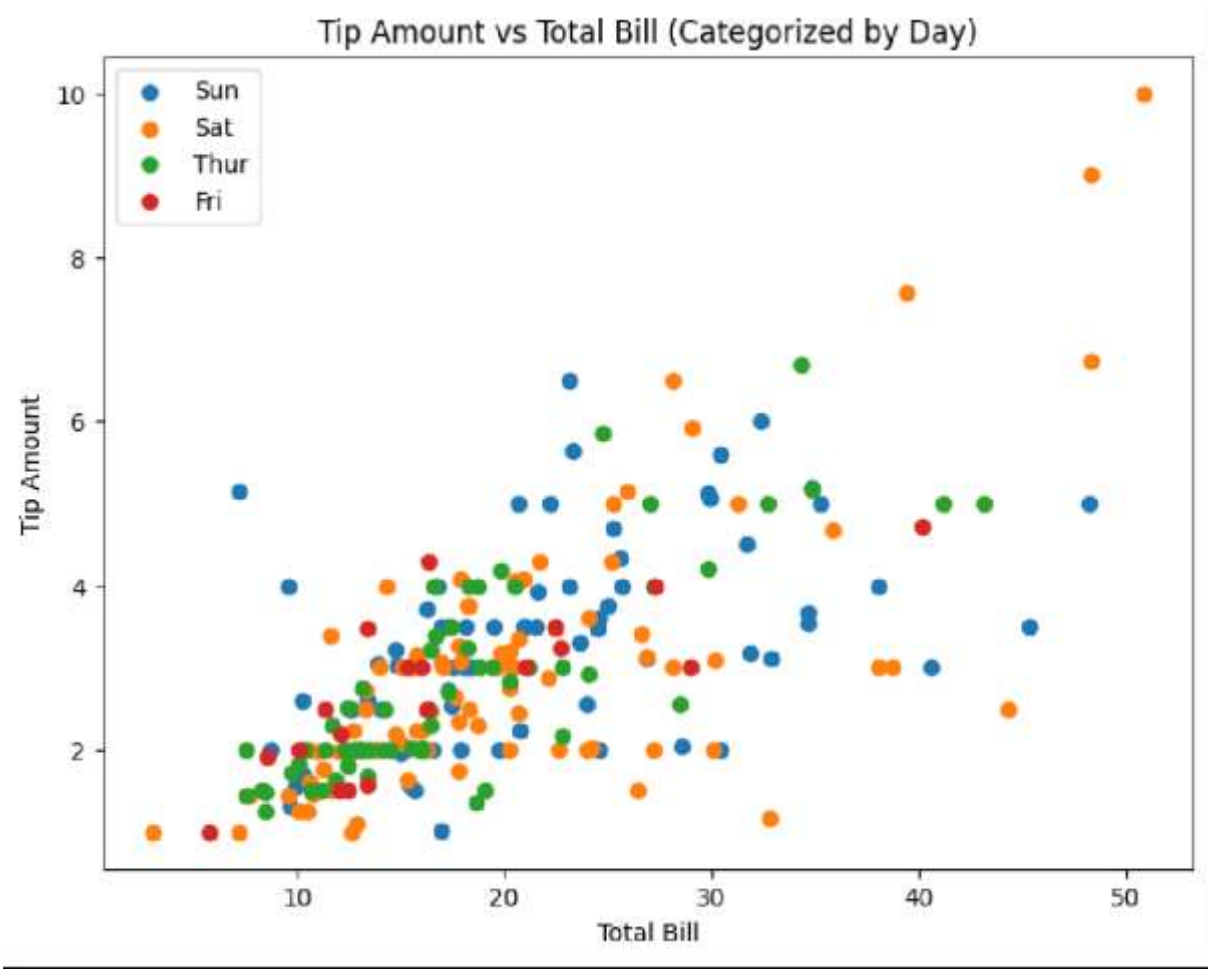

Average Total Bill by Time of Day
Average Total Bill
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
Dinner
Lunch
Time of Day

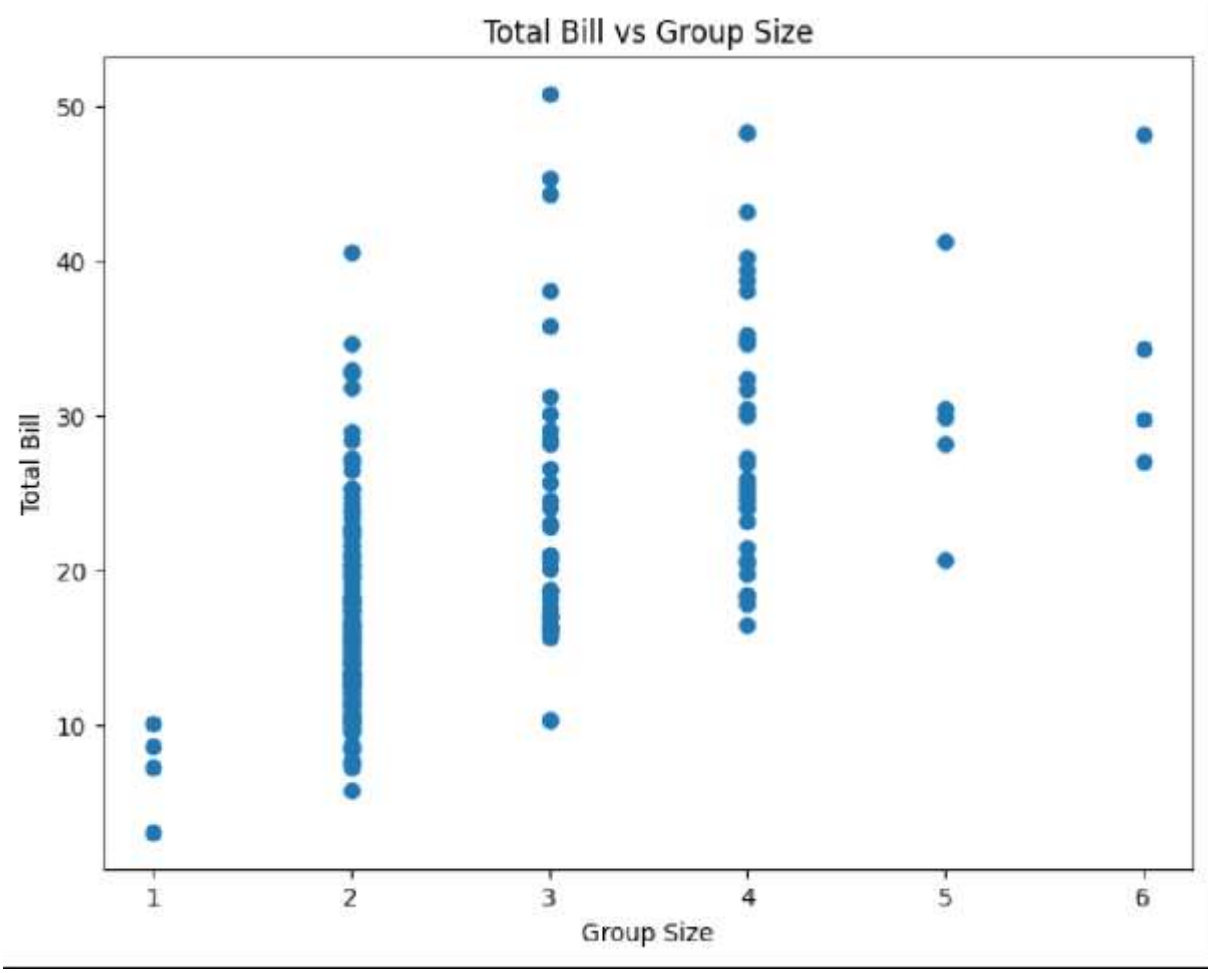

Average Tip Amount by Smoker
Average Tip
3.0
2.5
2.0
1.5
1.0
0.5
0.0
No
Yes
Smoker

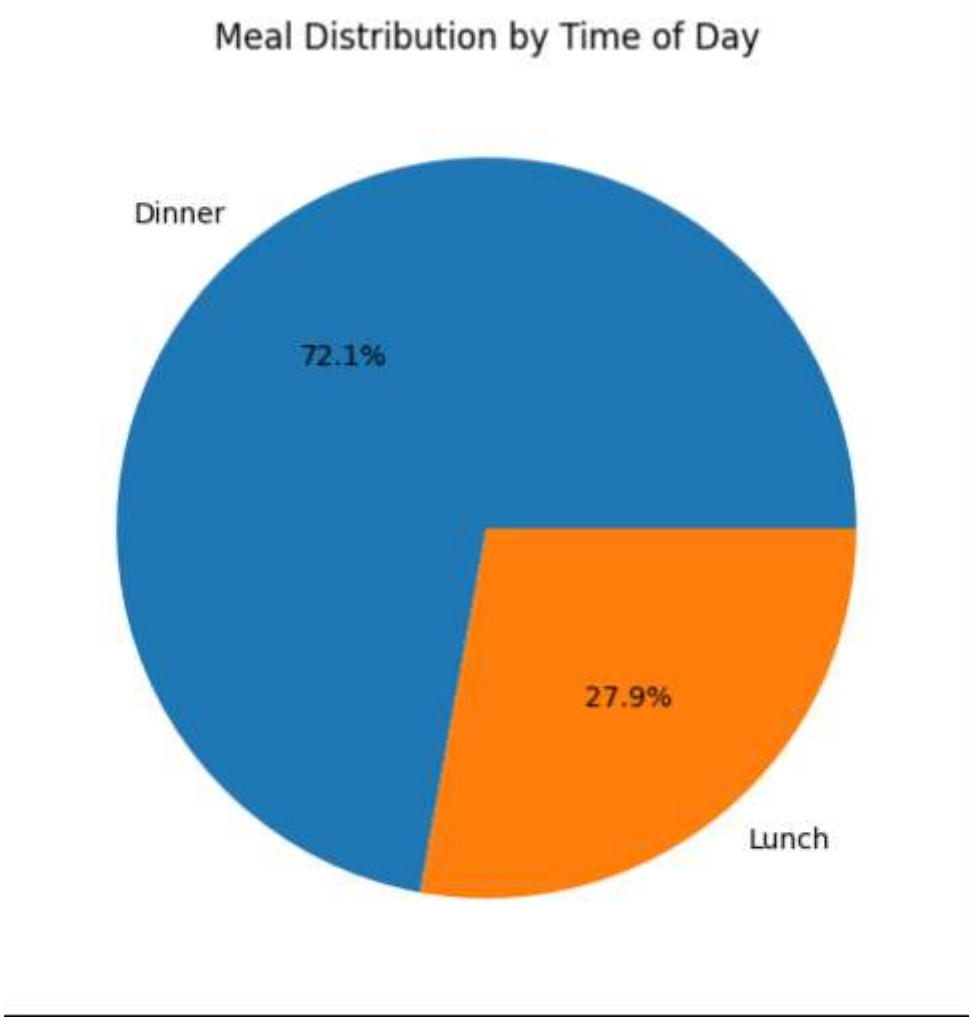

Average Total Bill by Group Size
Average Total Bill
0
5
10
15
20
25
30
35
1
2
3
4
5
6
Group Size


Total Bill by Day of the Week
Total Bill
1750
1500
1250
1000
750
500
250
0
Fri
Sat
Sun
Thur
Day of the Week


Distribution of Tips by Gender
Male
Female
Frequency
Tip Amount
0
10
20
30
40
50
2
4
6
8
10


Tip Amount vs Total Bill (Categorized by Day)
Sun
Sat
Thur
Fri
Tip Amount
Total Bill
10
8
6
4
2
10
20
30
40
50


Total Bill vs Group Size
Total Bill
50
40
30
20
10
1
2
3
4
5
6
Group Size


Meal Distribution by Time of Day
Dinner
72.1%
27.9%
Lunch