



UBER TECHNOLOGIES INC



UBER DATA ANALYSIS PROJECT

INSIGHTS FROM DATA EXORATION



INTRODUCTION

- Uber is a leading ride-sharing platform revolutionizing transportation.
- **Objective:** Analyze user booking behavior using Uber data.
- **Tools used:** Jupyter Notebook and Python libraries:
 - Pandas: For data cleaning and manipulation.
 - Numpy: For mathematical operations.
 - Matplotlib & Seaborn: For data visualization.

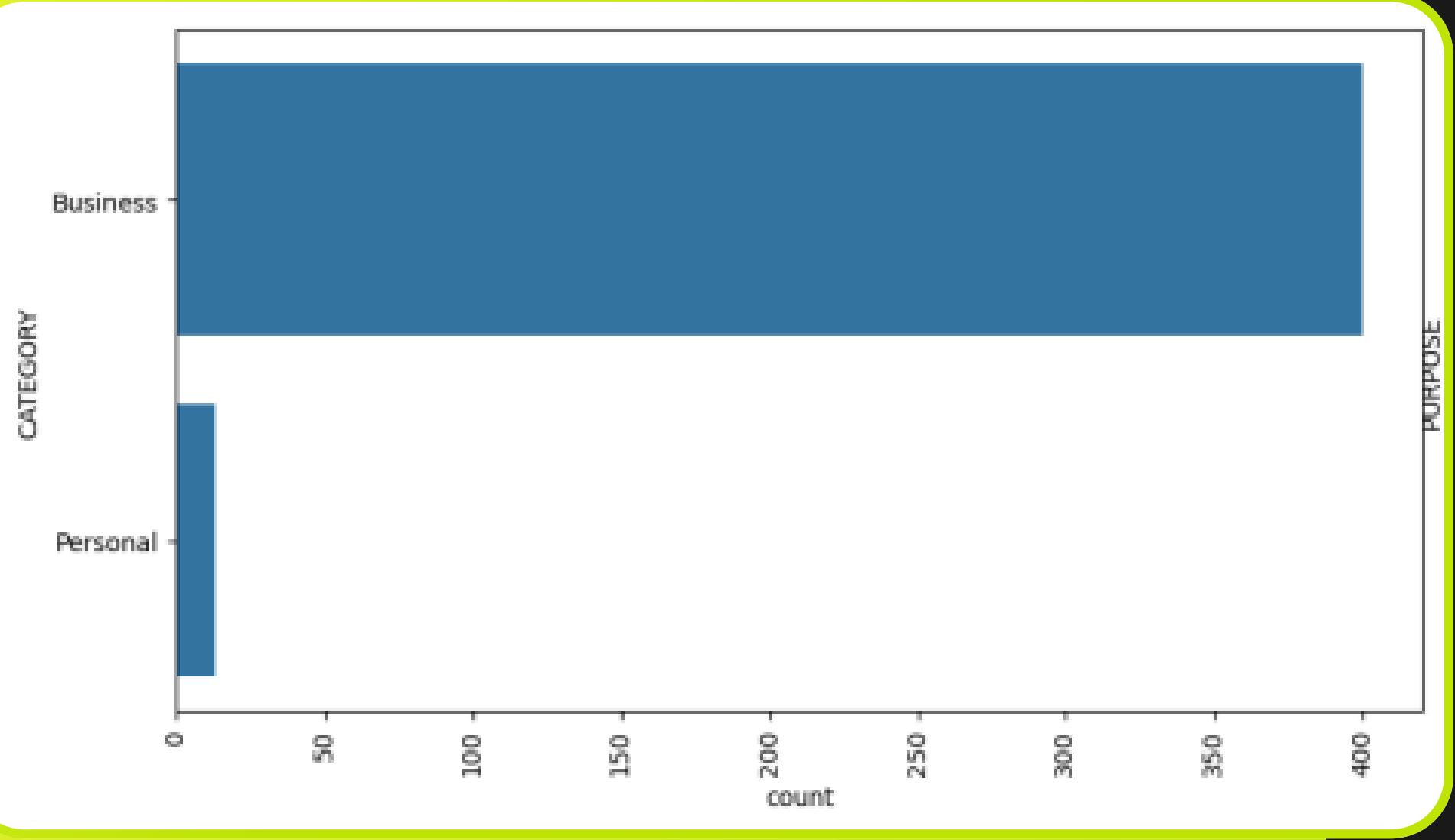
x x x
x x x
x x x
x x x

x x x
x x x
x x x
x x x

RESEARCH QUESTIONS

- In which category do people book the most Uber rides?
- At what time do people book cabs the most?
- On which days of the week do people book Uber rides the most?
- For which purpose do people book Uber rides the most?
- In which months do people book Uber rides less frequently?
- How many miles do people usually book a cab for?

× × ×
× × ×
× × ×
× × ×



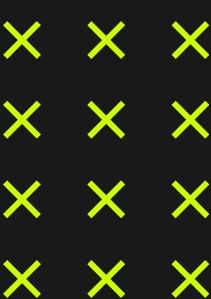
RIDE CATEGORY ANALYSIS

- Most Uber rides are booked under the Business category.
- Visualization: Graph showing ride distribution across categories.

```
plt.figure(figsize=(20,5))

plt.subplot(1,2,1)

sns.countplot(dataset['CATEGORY'])
plt.xticks(rotation =90)
```



PURPOSE OF BOOKING

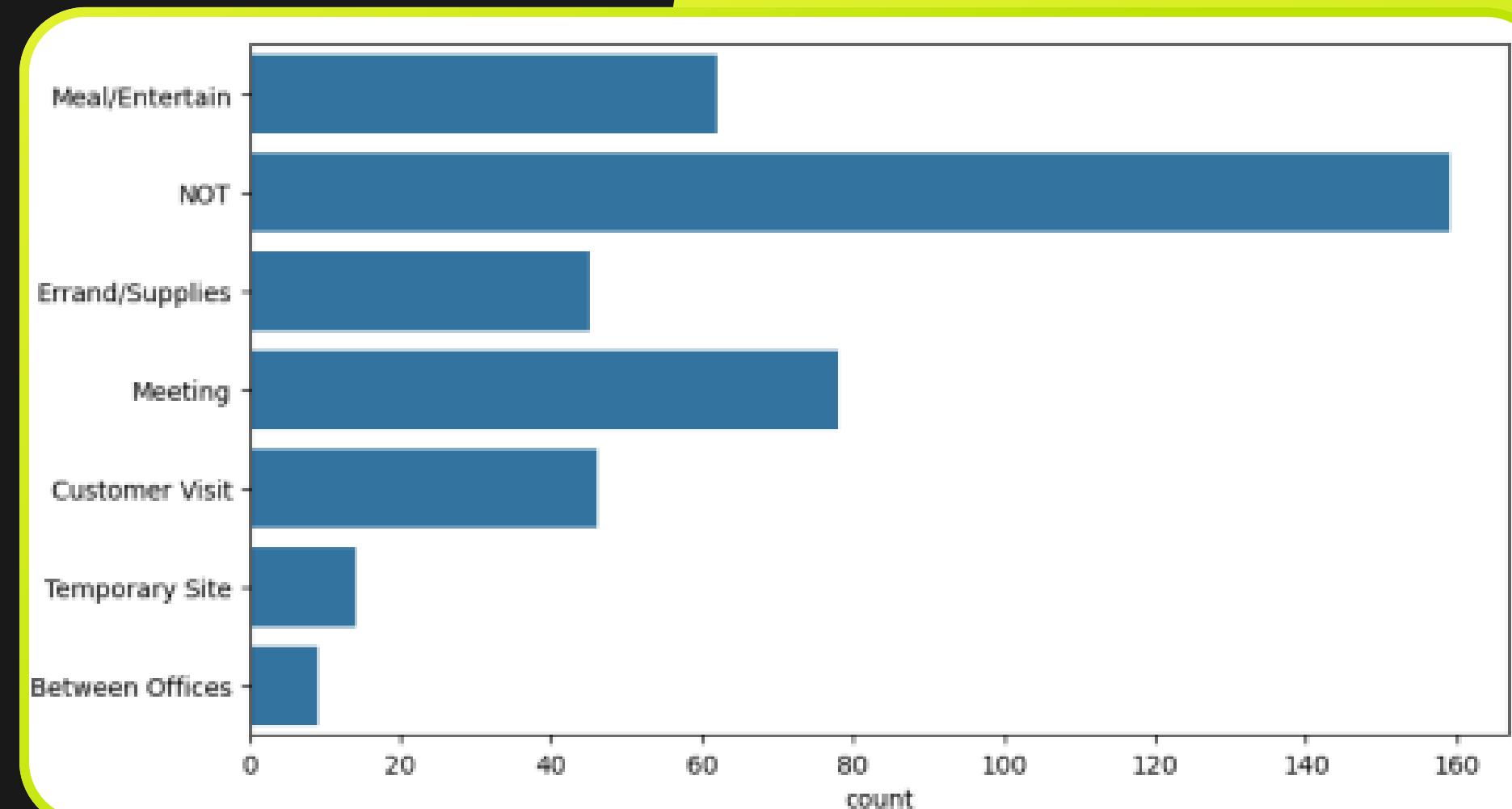
- The most common reason for booking an Uber ride is Meeting.
- Visualization: Bar graph representation.

```
plt.figure(figsize=(20,5))

plt.subplot(1,2,1)

sns.countplot(dataset['CATEGORY'])
plt.xticks(rotation =90)

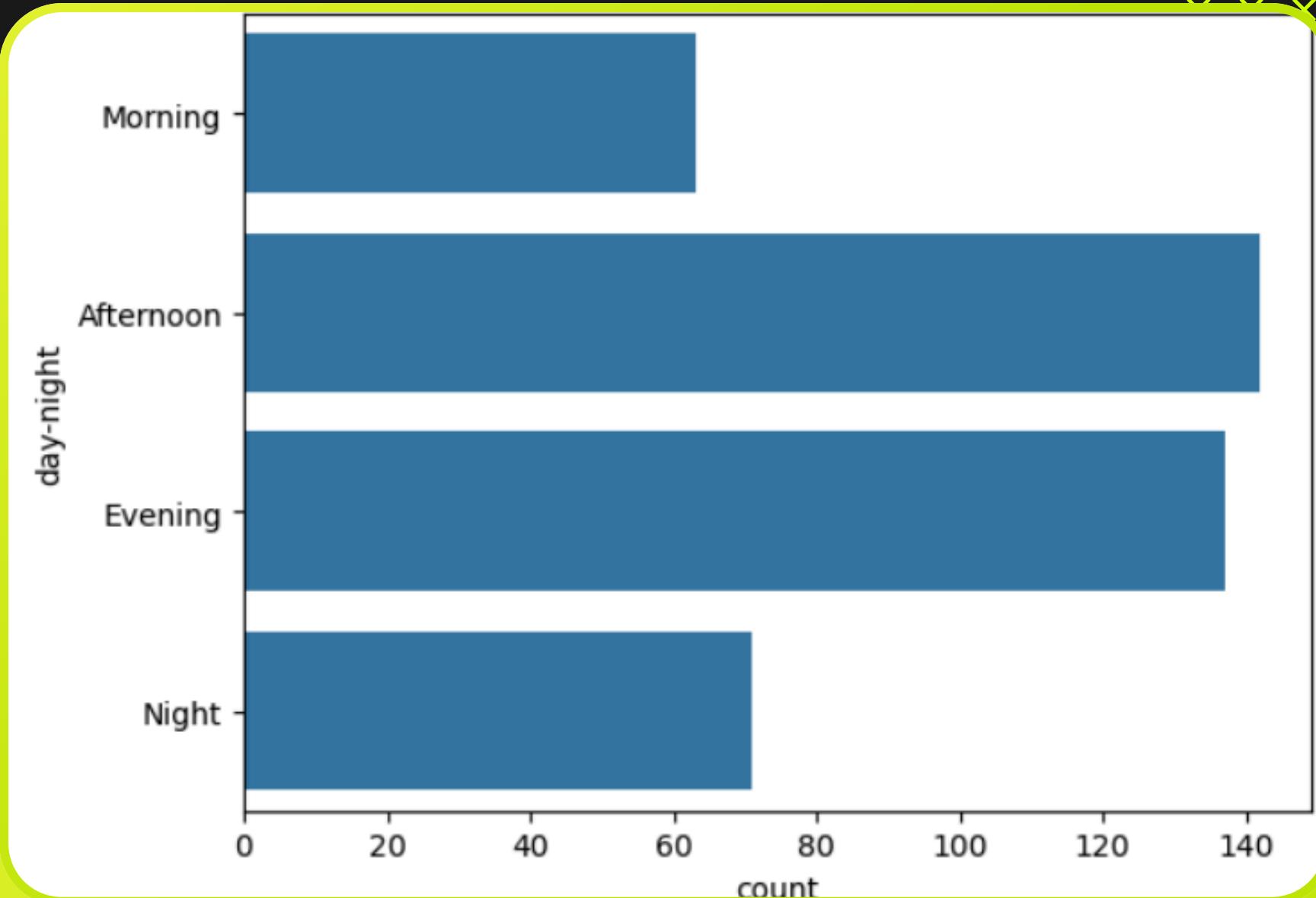
plt.subplot(1,2,2)
sns.countplot(dataset['PURPOSE'])
```

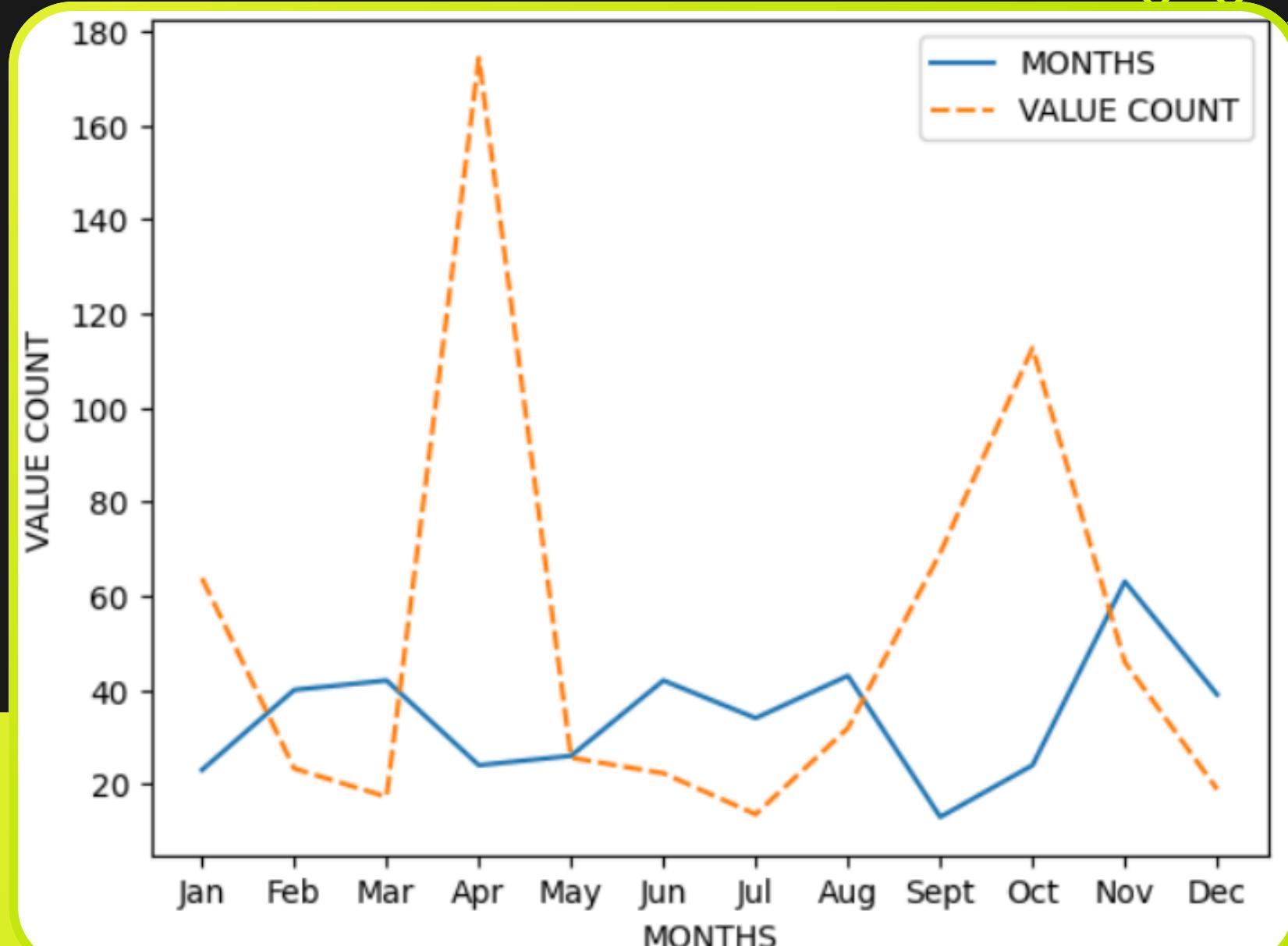


TIME & FREQUENCY ANALYSIS

- Afternoon, followed by evening, night, and morning.

```
dataset['day-night'] = pd.cut(x=dataset['time'],  
bins =[0,10,15,19,24],labels = ['Morning', 'Afternoon', 'Evening', 'Night'])  
sns.countplot(dataset['day-night'])
```





TIME & FREQUENCY ANALYSIS

- January, November, and December.
- Visualization: Line graph showing trends.

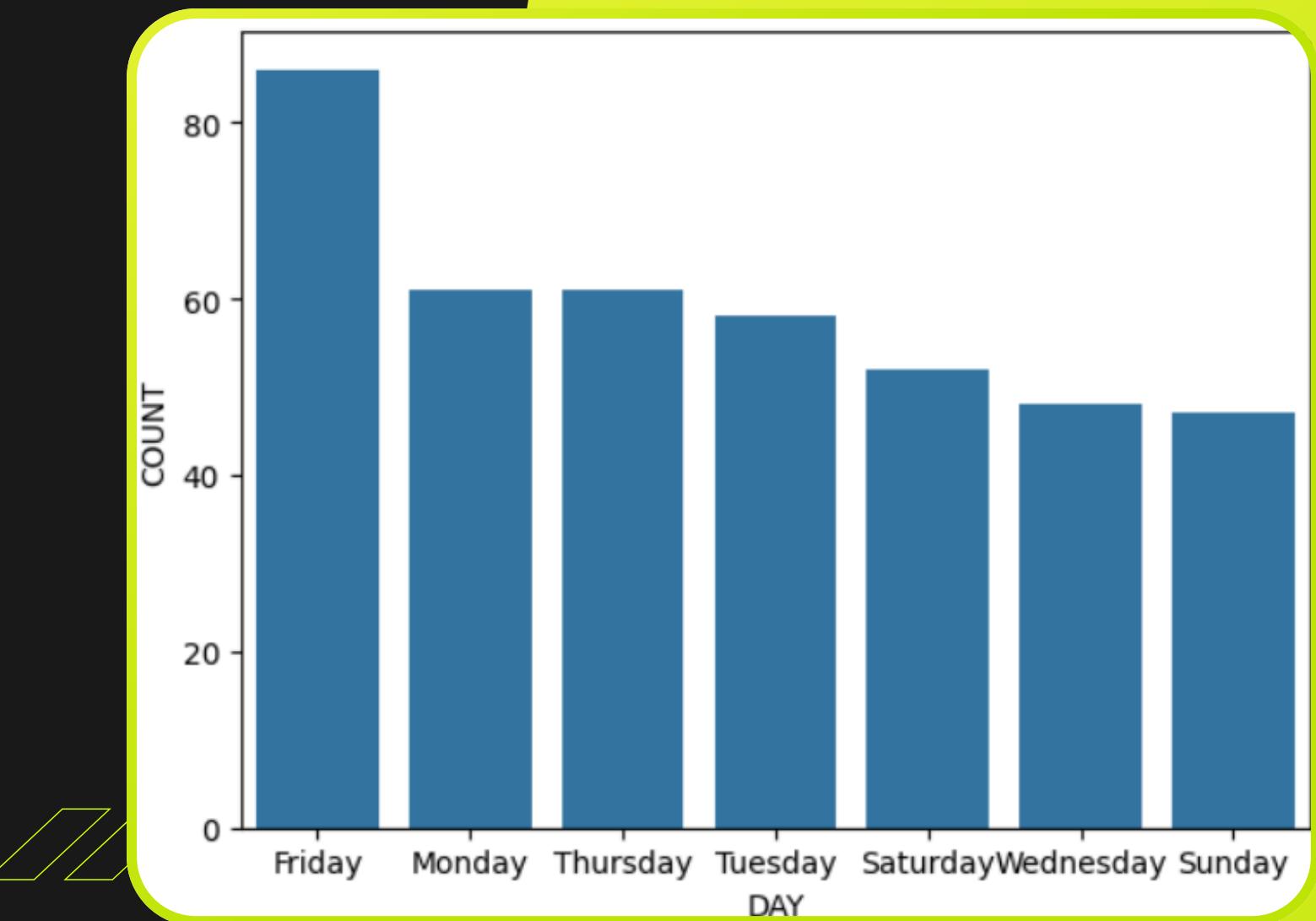
```
df = pd.DataFrame({  
    "MONTHS": mon.values, # Total count of each month  
    "VALUE COUNT": dataset.groupby('MONTH', sort=False)[ 'MILES' ].max()  
})  
  
p = sns.lineplot(data=df) #making lineplot  
p.set(xlabel="MONTHS", ylabel="VALUE COUNT") #For set Axis Labels
```



TIME & FREQUENCY ANALYSIS

- Friday people book Uber rides the most
- Visualization: Bar graph showing trends.

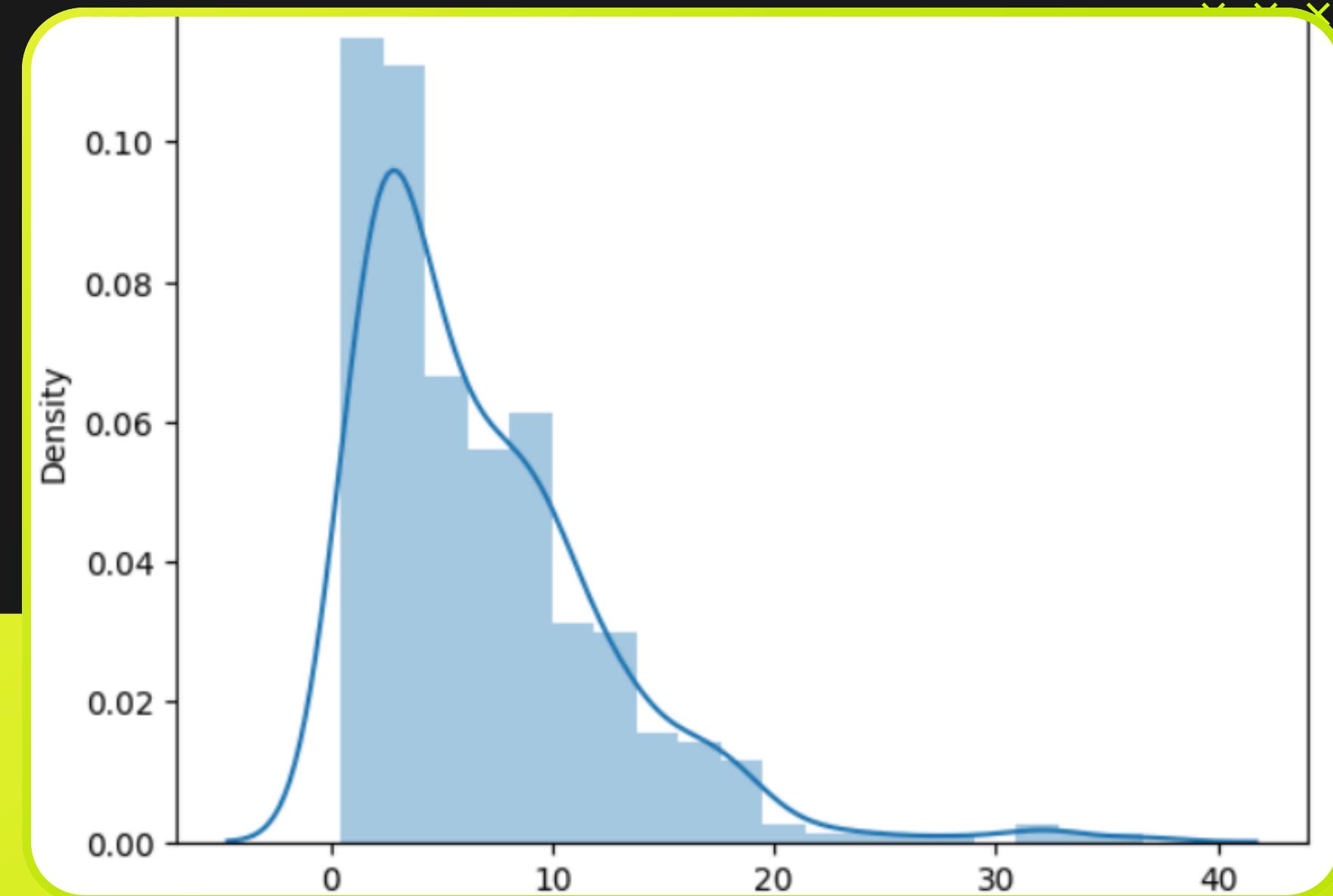
```
sns.barplot(x=day_label.index, y= day_label)
plt.xlabel('DAY')
plt.ylabel('COUNT')
```



DISTANCE ANALYSIS

- Most Uber rides are booked for 0 to 20 miles.
- Visualization: Distplot representation.

```
sns.distplot(dataset[dataset['MILES']<40]['MILES'])
```



CONCLUSION & INSIGHTS

- **Key Findings:**
 - Business rides are the most common.
 - Meetings are the main purpose.
 - Afternoon is the peak booking time.
 - January, November, and December have the lowest demand.
 - Friday sees the most bookings.
 - Most trips range from 0 to 20 miles.
- **Future Scope:** Further analysis on price trends, city-wise demand, and seasonal patterns.