



# Uber

Next-Generation Car Service

# Python-EDA Project

f t g

Sakshi Mogal

# Introduction



Uber Technologies, Inc., widely recognized as Uber, is an American multinational corporation specializing in transportation and logistics. Based in San Francisco, California, the company operates in nearly 70 countries and 10,500 cities worldwide. Uber offers services such as ride-hailing, food and courier deliveries, and freight transport. It is the largest ride sharing platform globally, with over 150 million (15 Crores) monthly active users and a workforce of 6 million (60 Lakhs) drivers and couriers. The platform handles an average of 28 million (2.8 Crore) trips every day and has completed 47 billion (4700 Crore) trips since its founding in 2010.

In 2023, Uber achieved a revenue of \$37.2 billion.

# Uber Founding

---

## ✔ Bringing Travis Kalanick on Board

To bring the idea to life, Camp sought the help of Travis Kalanick, a successful entrepreneur who had previously founded and sold two tech companies, including Red Swoosh. Kalanick was initially hesitant but eventually became enthusiastic about the potential of Camp's idea. In early 2009, Camp and Kalanick teamed up to develop the project further.



## Development and Launch

Formation: UberCab was officially founded in March 2009 by Garrett Camp and Travis Kalanick.



## Prototype Development

They worked on building a prototype app that would allow users to book a ride with the tap of a button. The goal was to create a seamless experience for both riders and drivers.



# Uber Founding

---

## ✔ Bringing Travis Kalanick on Board

In 2010, UberCab launched its beta version in San Francisco. The service started with only a few cars, offering a premium black car service. The initial target market was business professionals who valued convenience and were willing to pay a higher fare for a reliable service.

## ✔ First Ride

The first Uber ride took place in San Francisco in June 2010..

## ✔ Positive Reception

The app quickly gained popularity among users who appreciated the convenience of hailing a ride through their smartphones.



In October 2010, the company rebranded itself from UberCab to Uber. This change was partly driven by regulatory pressures, as the company faced legal challenges from taxi regulators who were concerned about the word "cab" in its name.

Expansion to Other Cities: Following the success in San Francisco, Uber expanded to

- ✓ New York City, Chicago, and other major cities in the United States.

International Expansion: By 2012, Uber had launched in Paris, marking its first international market. The company continued its global expansion at a rapid pace, entering mark Europe, Asia, and beyond.

**Ride-Sharing: Uber's core service, allowing users to book rides from drivers using their own vehicles.**

- ✓ **UberX:** Affordable everyday rides.
- ✓ **UberPOOL:** Carpooling with other passengers heading in the same direction.
- ✓ **UberBLACK:** High-end black car service.
- ✓ **UberXL:** Rides for larger groups with spacious vehicles.



# Challenges:



## **ETA Estimation**

Accurately estimating the arrival time of drivers.



## **Dynamic Pricing**

Setting fares based on demand and supply.



## **Driver and Rider Matching**

Efficiently pairing drivers with riders.



## **Fraud Detection**

Identifying and preventing fraudulent activities.



## **Route Optimization**

Finding the most efficient routes

# Questions:

Suppose you work in data analysis at Uber; you need to provide answers to all the questions given below to the company

1) In which category do people book the most Uber rides?

---

2) For which purpose do people book Uber rides the most?

3) At what time do people book cabs the most from Uber?

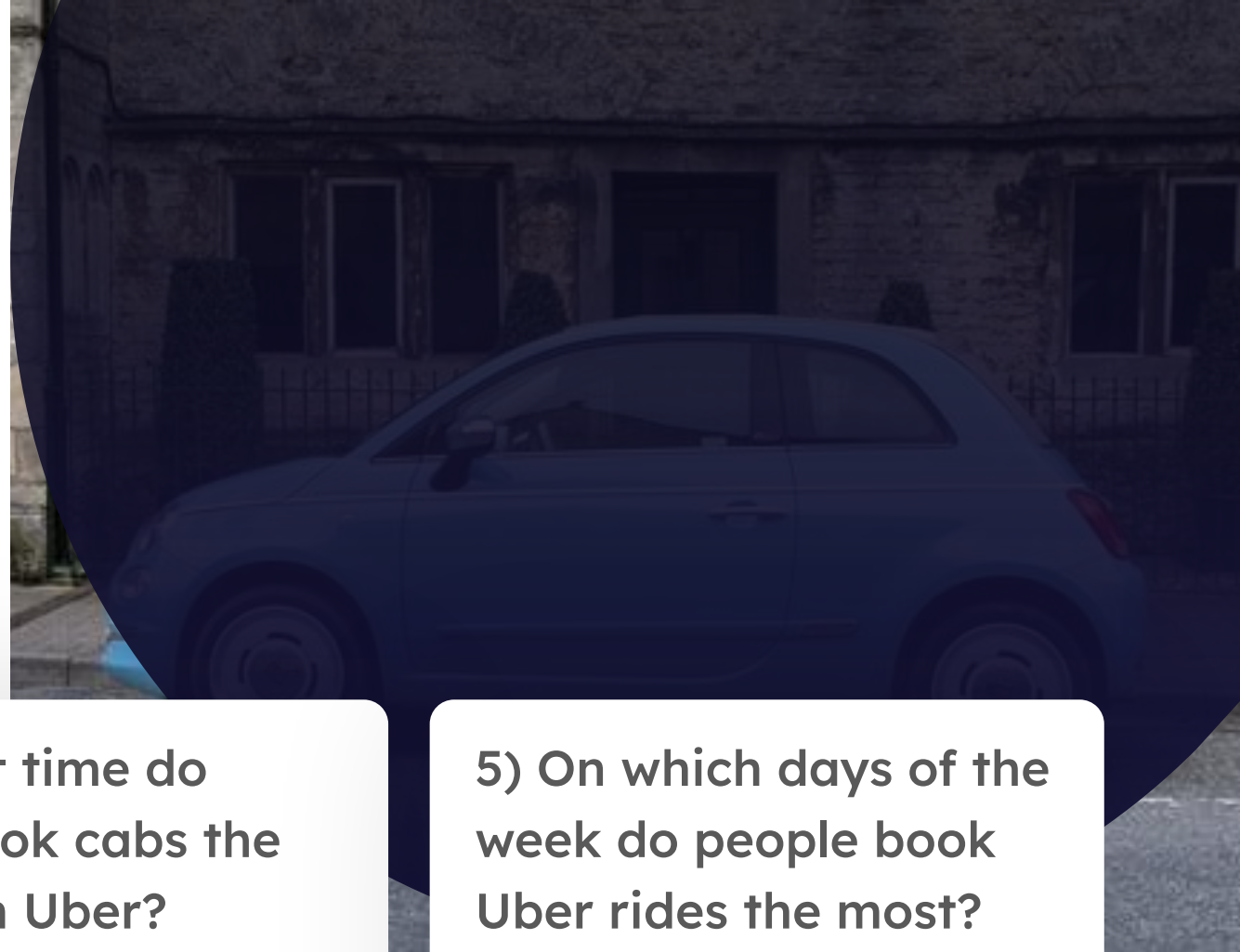
---

4) In which months do people book Uber rides less frequently?

5) On which days of the week do people book Uber rides the most?

---

6) How many miles do people usually book a cab for through Uber?





# Data Preprocessing



**Handling Missing  
Values**



**Change Datatypes  
of Columns**



**Drop All Null  
Values**



**Adding New  
Column**





# Handling missing values



```
dataset['PURPOSE'].fillna("NOT",inplace=True)
```



# Change Datatypes of Columns

```
dataset['START_DATE']=pd.to_datetime(  
    dataset['START_DATE'],errors='coerce')  
dataset['END_DATE']=pd.to_datetime(  
    dataset['END_DATE'],errors='coerce')
```



# Drop All Null Values



```
dataset.dropna(inplace=True)
```

# Adding New Column

(DATE & TIME)



```
from datetime import datetime
dataset['DATE'] = pd.DatetimeIndex(dataset['START_DATE']).date
dataset['TIME'] = pd.DatetimeIndex(dataset['START_DATE']).hour
```

DATE	TIME
2016-01-01	21.0
2016-01-02	1.0
2016-01-02	20.0
2016-01-05	17.0
2016-01-06	14.0



# Adding New Column (DAY\_NIGHT)



```
dataset['DAY_NIGHT'] = pd.cut(x =  
    dataset['TIME'], bins=[0,10,15,19,24]  
    , labels=['Morning', 'Afternoon', 'Evening', 'Night'])
```

DAY_NIGHT
Night
Morning
Night
Evening
Afternoon

# Adding New Column (MONTH)



```
dataset['MONTH'] = pd.DatetimeIndex(dataset['START_DATE']).month
month_label = {1.0: 'Jan', 2.0: 'Feb', 3.0: 'Mar', 4.0: 'April',
               5.0: 'May', 6.0: 'June', 7.0: 'July', 8.0: 'Aug',
               9.0: 'Sep', 10.0: 'Oct', 11.0: 'Nov', 12.0: 'Dec'}
dataset["MONTH"] = dataset.MONTH.map(month_label) |
mon = dataset.MONTH.value_counts(sort=False)
```

## MONTH

Jan

Jan

Jan

Jan

Jan

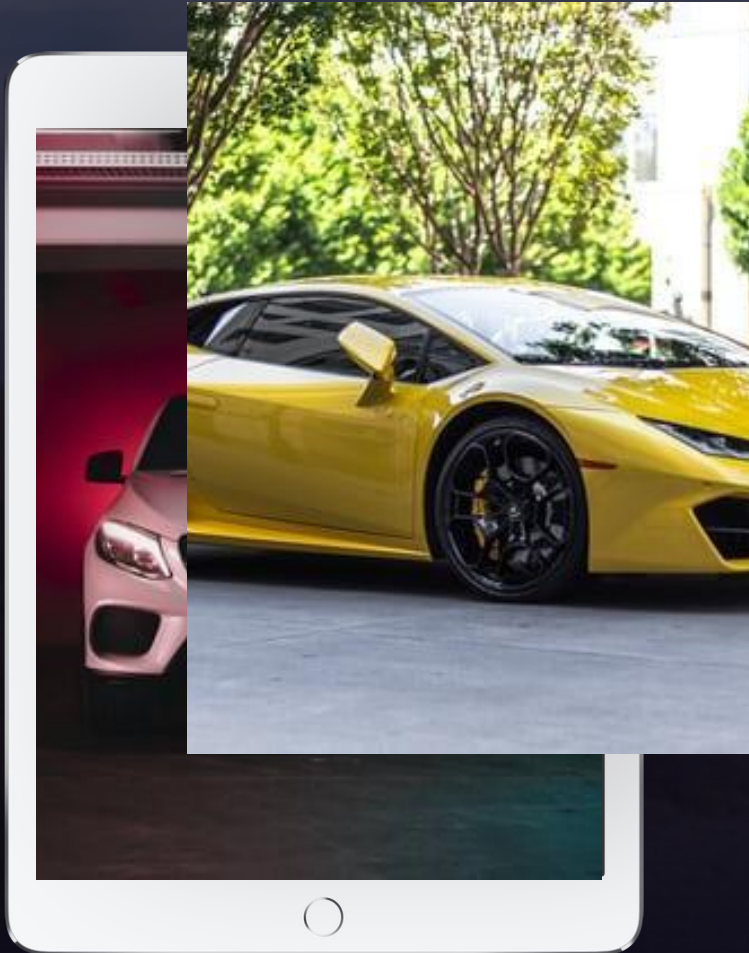


# Adding New Column (DAY)



```
dataset['DAY']=dataset.START_DATE.dt.weekday  
day_label={0:'Mon',1:'Tues',2:'Wed',3:'Thur',4:'Fri',5:'Sat',6:'Sun'}  
dataset['DAY']=dataset['DAY'].map(day_label)
```

DAY
Fri
Sat
Sat
Tues
Wed

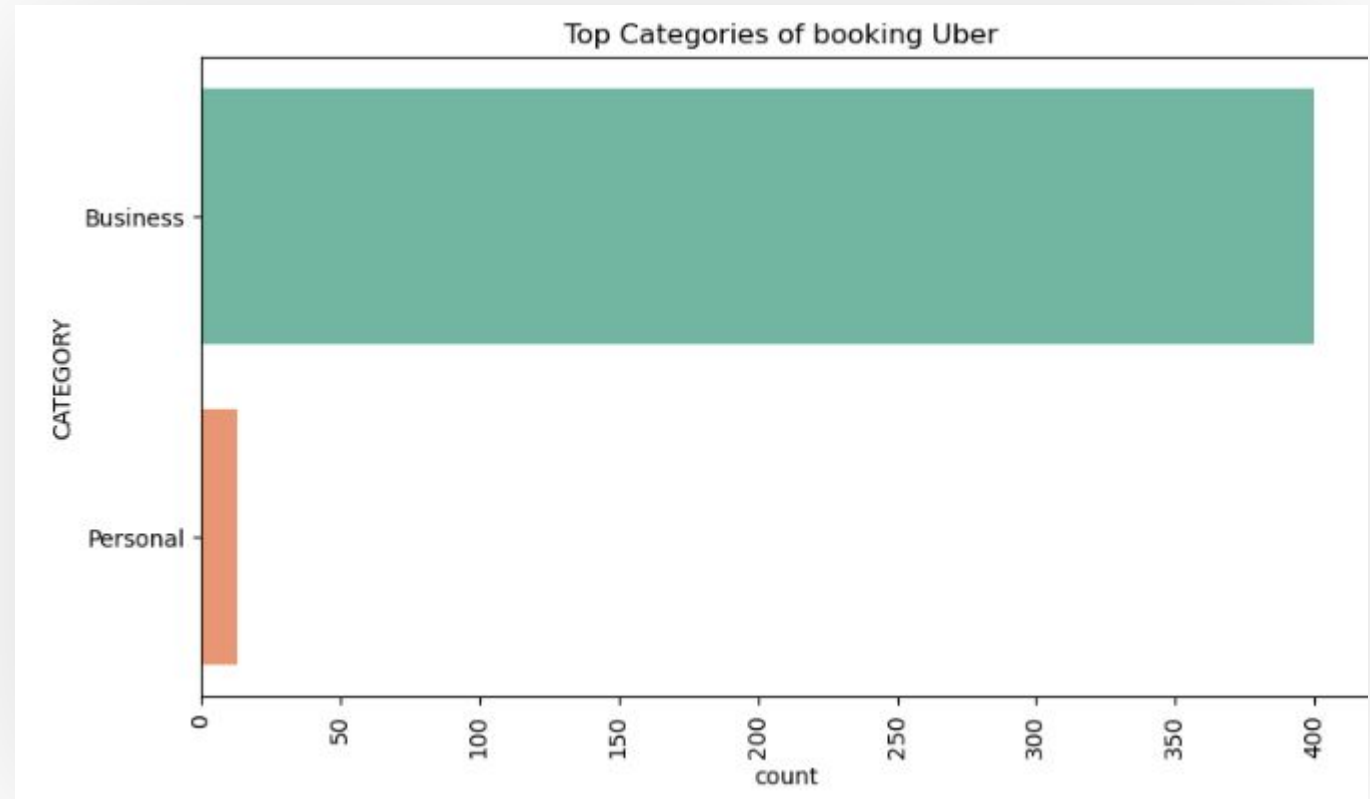




# Data Visualization

1) In which category do people book the most Uber rides?

```
plt.figure(figsize=(20,5))  
plt.subplot(1,2,1)  
sns.countplot(dataset['CATEGORY'], palette='Set2')  
plt.xticks(rotation=90)  
plt.title('Top Categories of booking Uber ')
```

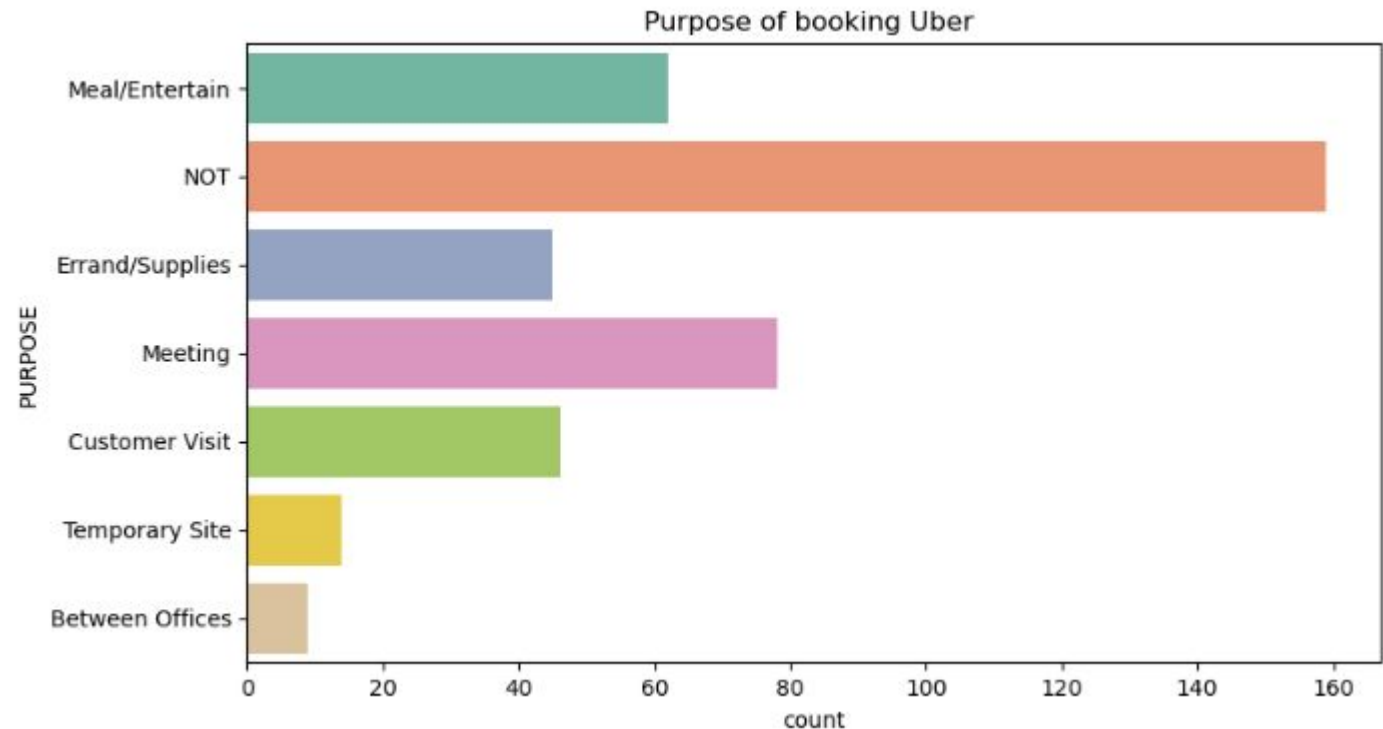


**Business professionals rely on Uber for efficient, reliable, and seamless travel solutions.**

# Data Visualization

2) For which purpose do people book Uber rides the most?

```
plt.subplot(1,2,2)
sns.countplot(dataset['PURPOSE'], palette='Set2')
plt.title('Purpose of booking Uber ')
```

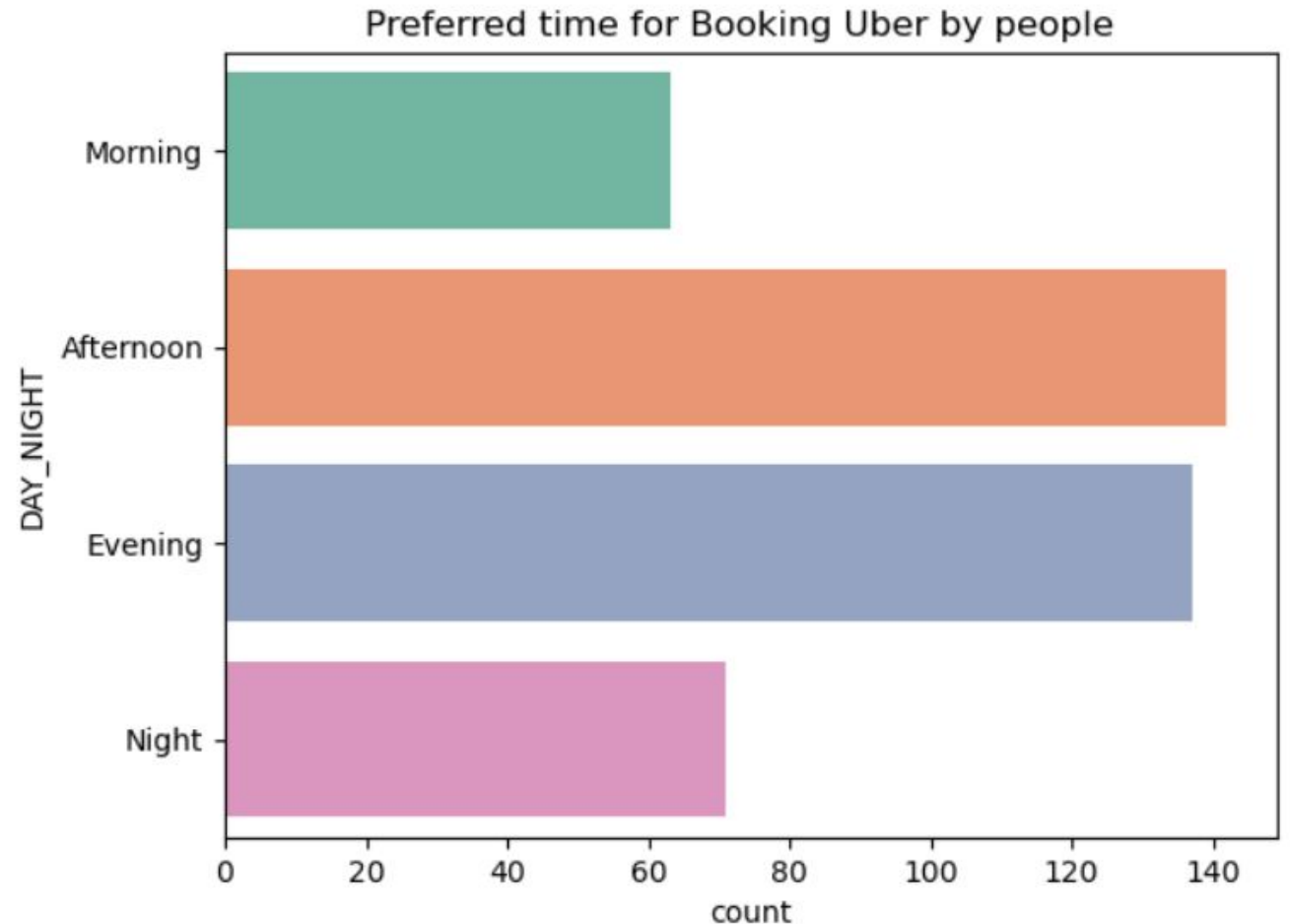


**Uber rides are most often booked for meetings and meal or entertainment purposes.**

# Data Visualization

3) At what time do people book cabs the most from Uber?

```
sns.countplot(dataset['DAY_NIGHT'], palette='Set2')  
plt.title('Preferred time for Booking Uber by people ')
```

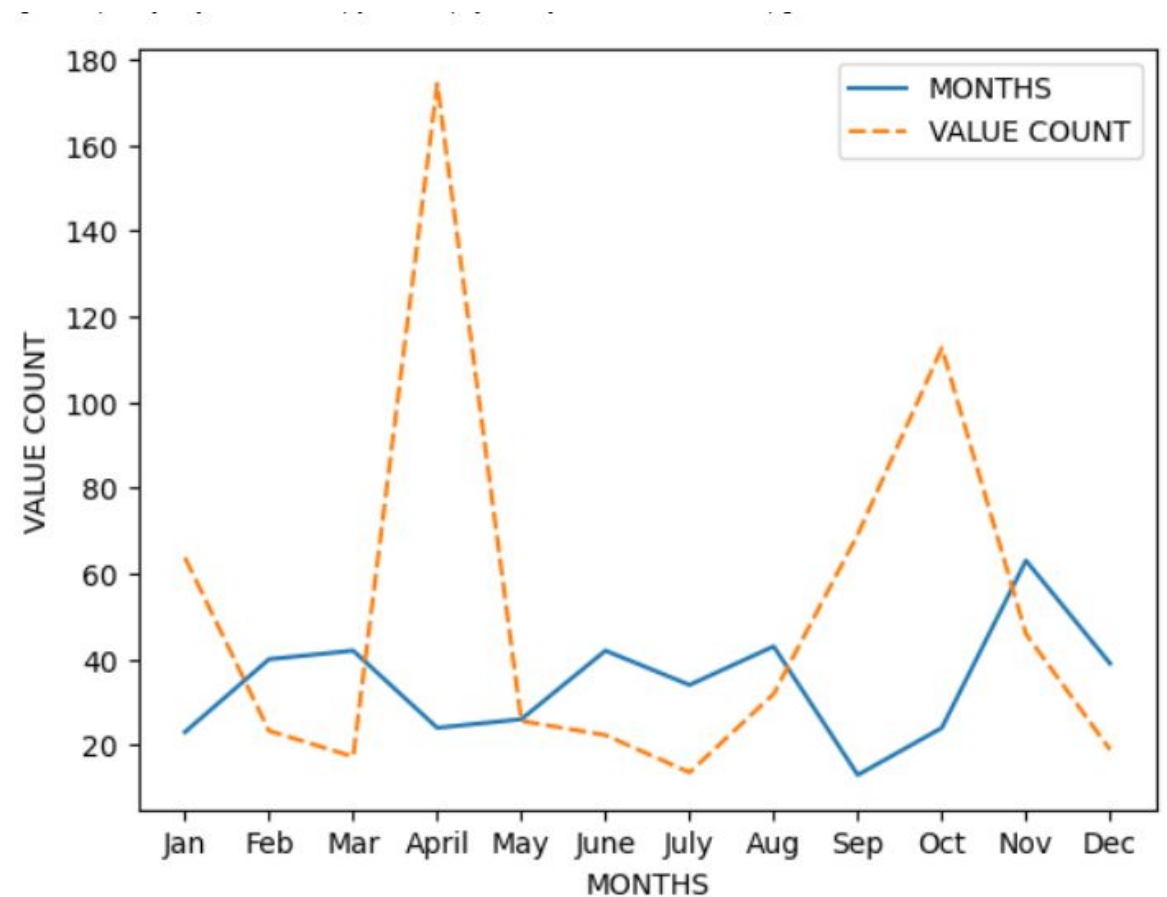


People primarily book Uber rides in the afternoon and evening hours.

# Data Visualization

4) In which months do people book Uber rides less frequently?

```
df = pd.DataFrame({  
    "MONTHS": mon.values,  
    "VALUE COUNT": dataset.groupby('MONTH', sort=False)['MILES'].max()  
})  
p = sns.lineplot(data=df)  
p.set(xlabel="MONTHS", ylabel="VALUE COUNT")
```

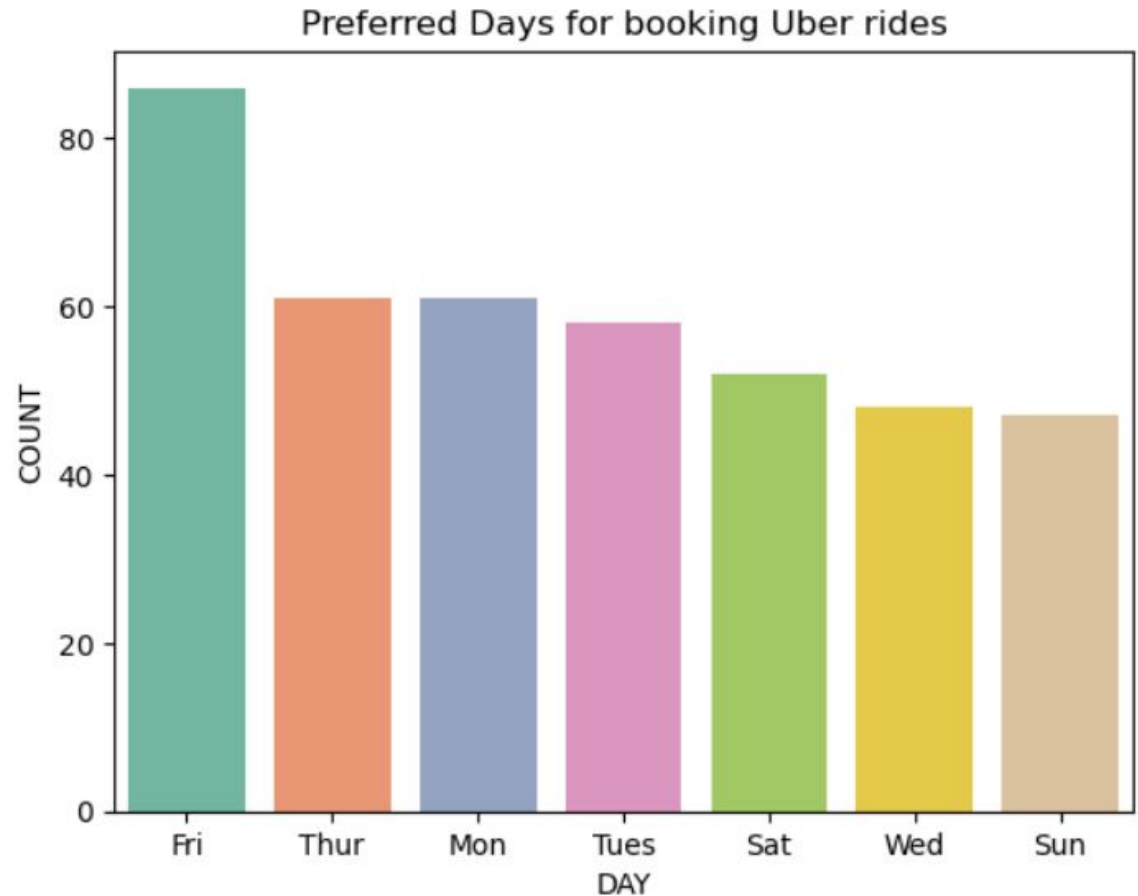


people book Uber rides less frequently in Nov, Dec, Jan .

# Data Visualization

5) On which days of the week do people book Uber rides the most?

```
day_label=dataset.DAY.value_counts()  
sns.barplot(x=day_label.index,  
            y=day_label, palette='Set2')  
plt.xlabel('DAY')  
plt.ylabel('COUNT')  
plt.title('Preferred Days for booking Uber rides')
```



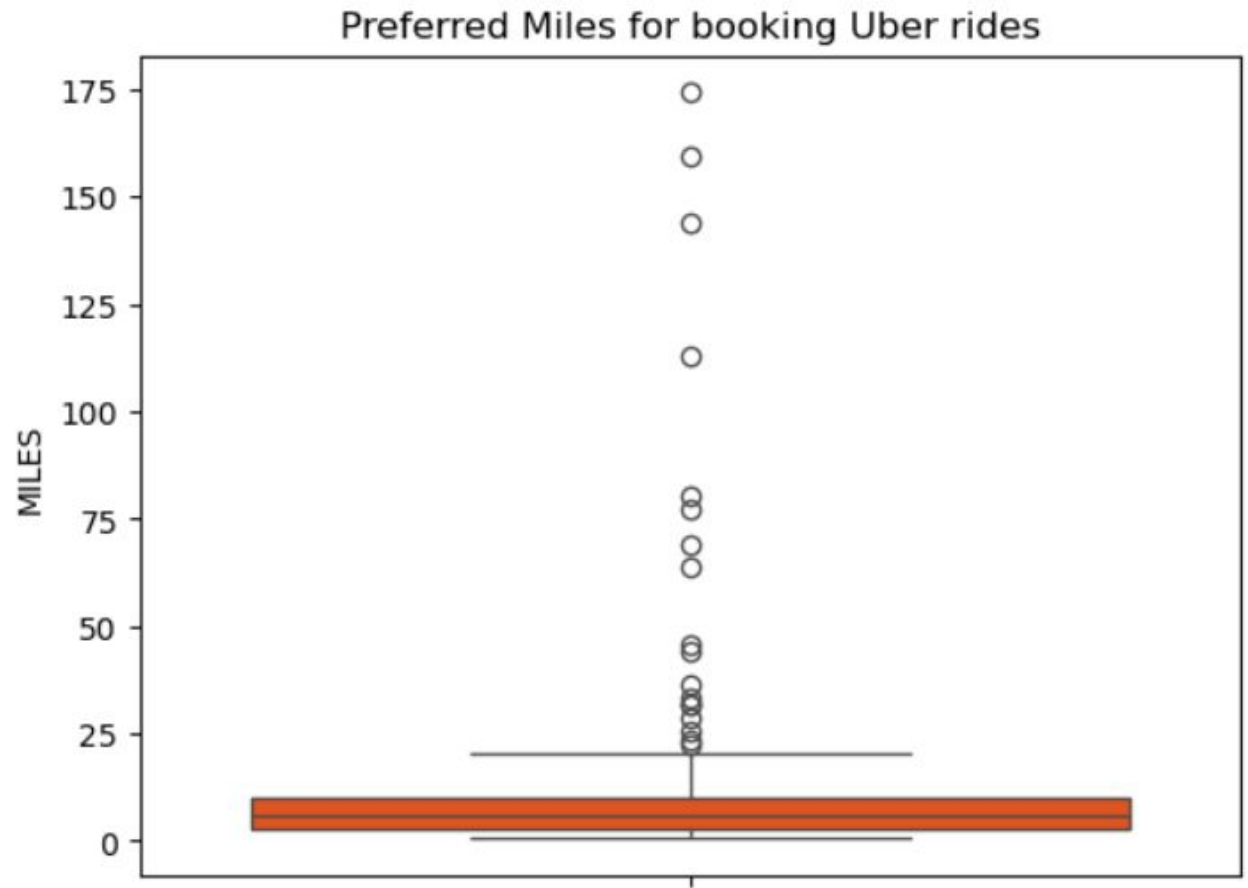
**Friday is the most popular day for booking Uber rides.**



# Data Visualization

6) How many miles do people usually book a cab for through Uber?

```
sns.boxplot(dataset['MILES'],color='orangered')
```

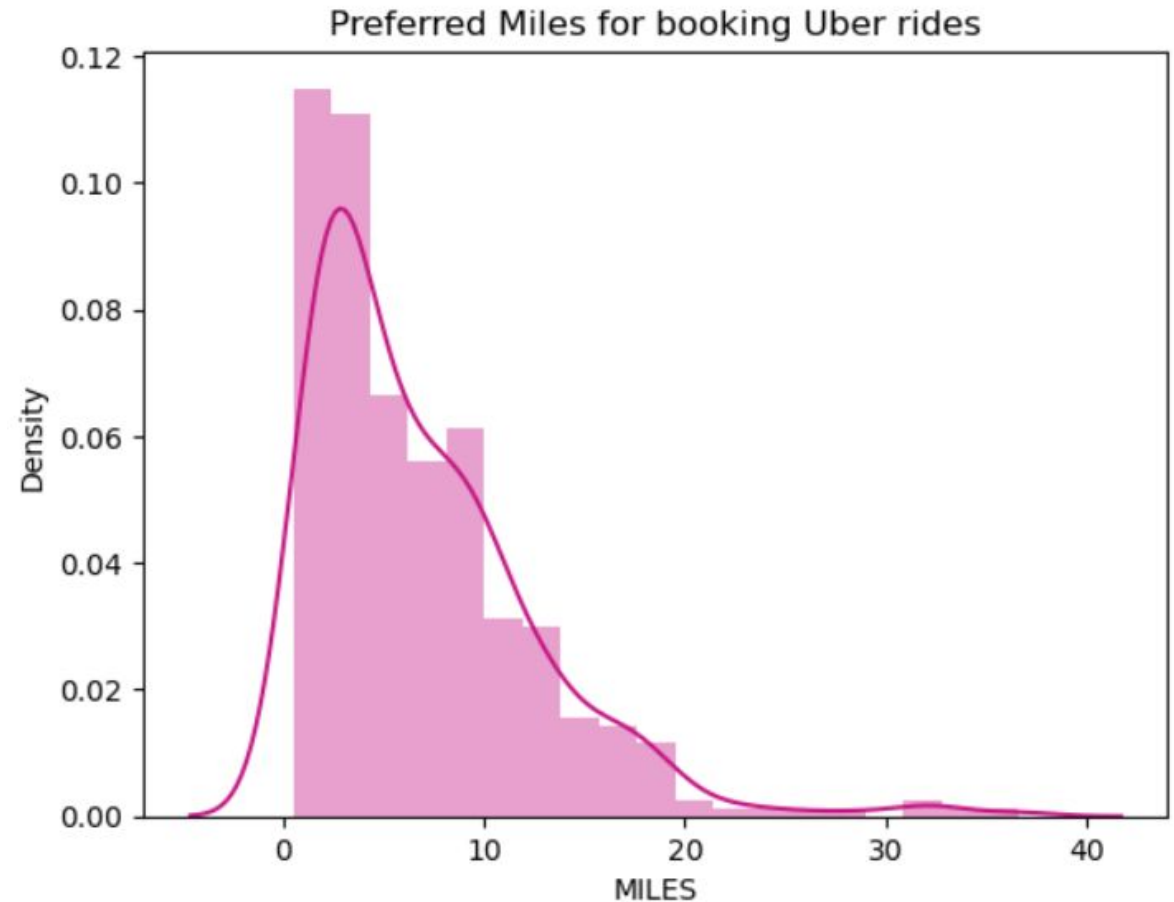


**Most Uber rides are booked for travel distances ranging from 0 to 20 miles.**

# Data Visualization

6) How many miles do people usually book a cab for through Uber?

```
sns.distplot(dataset[dataset['MILES'] < 40]['MILES'],color='mediumvioletred')  
plt.title('Preferred Miles for booking Uber rides')
```



**Most Uber rides are booked for travel distances ranging from 0 to 20 miles.**



# Insights

**Business Use:** Business professionals frequently rely on Uber for efficient, reliable, and seamless travel solutions.

**Common Purposes:** The majority of Uber rides are booked for meetings or meal and entertainment purposes.

**Peak Booking Hours:** Afternoon and evening are the most popular times for booking Uber rides.

**Seasonal Trends:** Uber rides are booked less frequently during November, December, and January.

**Popular Day:** Friday stands out as the most popular day for booking Uber rides.

**Travel Distance:** Most Uber rides cover travel distances ranging between 0 to 20 miles.

# THANK YOU



[GITHUB](#)



[LINKEDIN](#)