G2m Case Study

20-Aug-2024

Background –G2M(cab industry) case study

- XYZ is a private equity firm in US. Due to remarkable growth in the Cab Industry in last few years and multiple key players in the market, it is planning for an investment in Cab industry.

- Objective : Provide actionable insights to help XYZ firm in identifying the right company for making investment.

The analysis has been divided into four parts:

- Data Exploration

- Hypothesis Testing

- Summary of what we have learned

Data Exploration

- 14 Features
- Timeframe of the data: 2016-01-31 to 2018-12-31
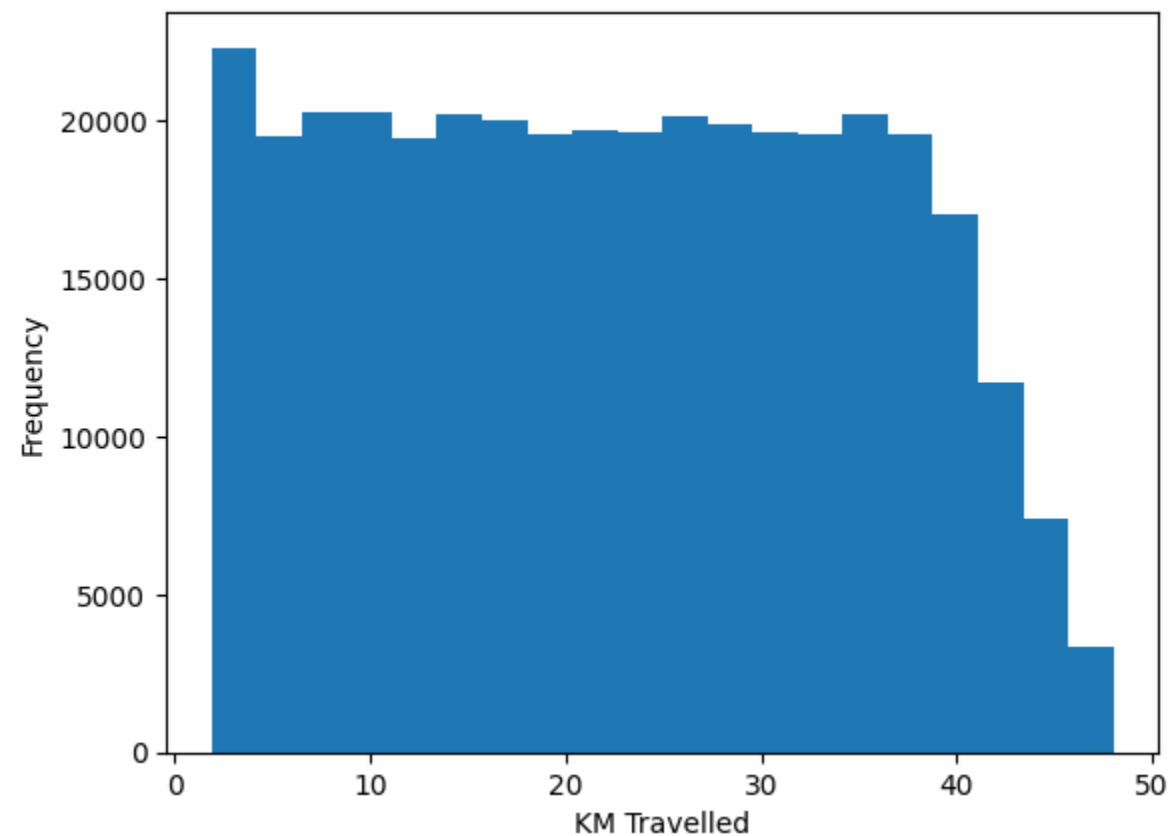- Total data points : 359,392

Assumptions:

- Outliers are present in Price_Charged feature so these are replaced with upper third quartile values.
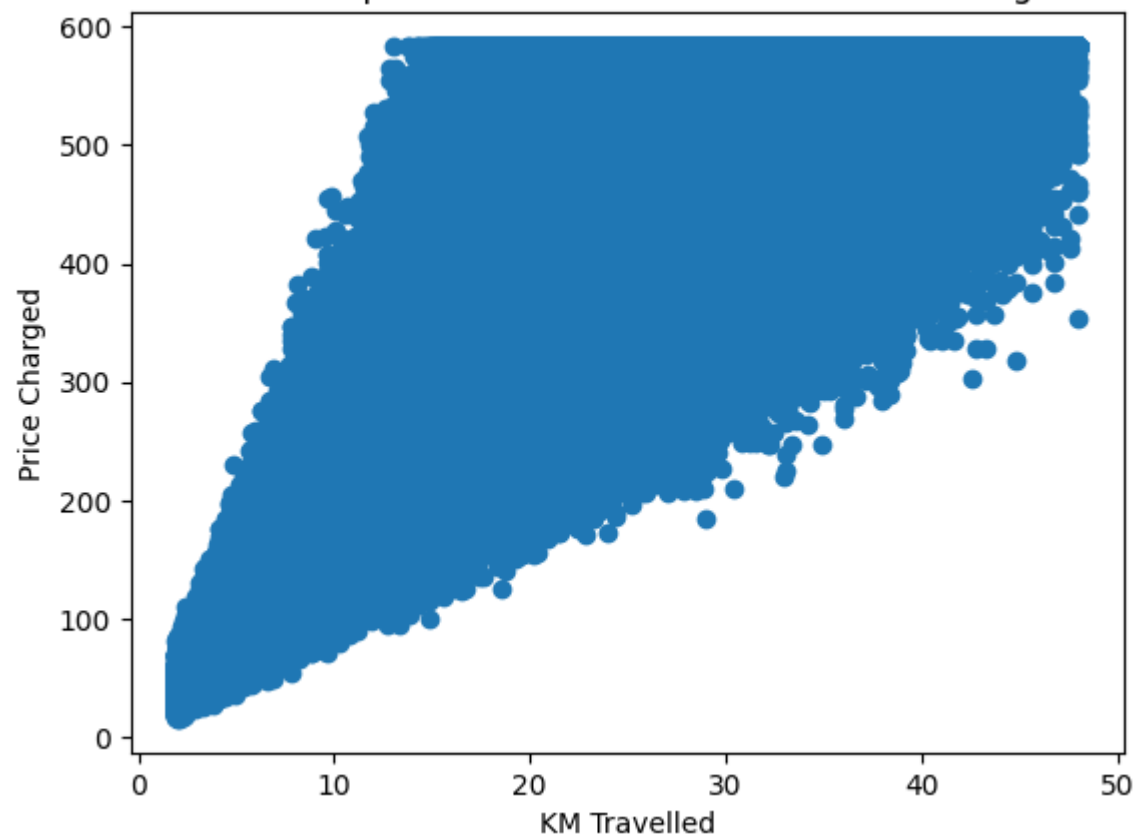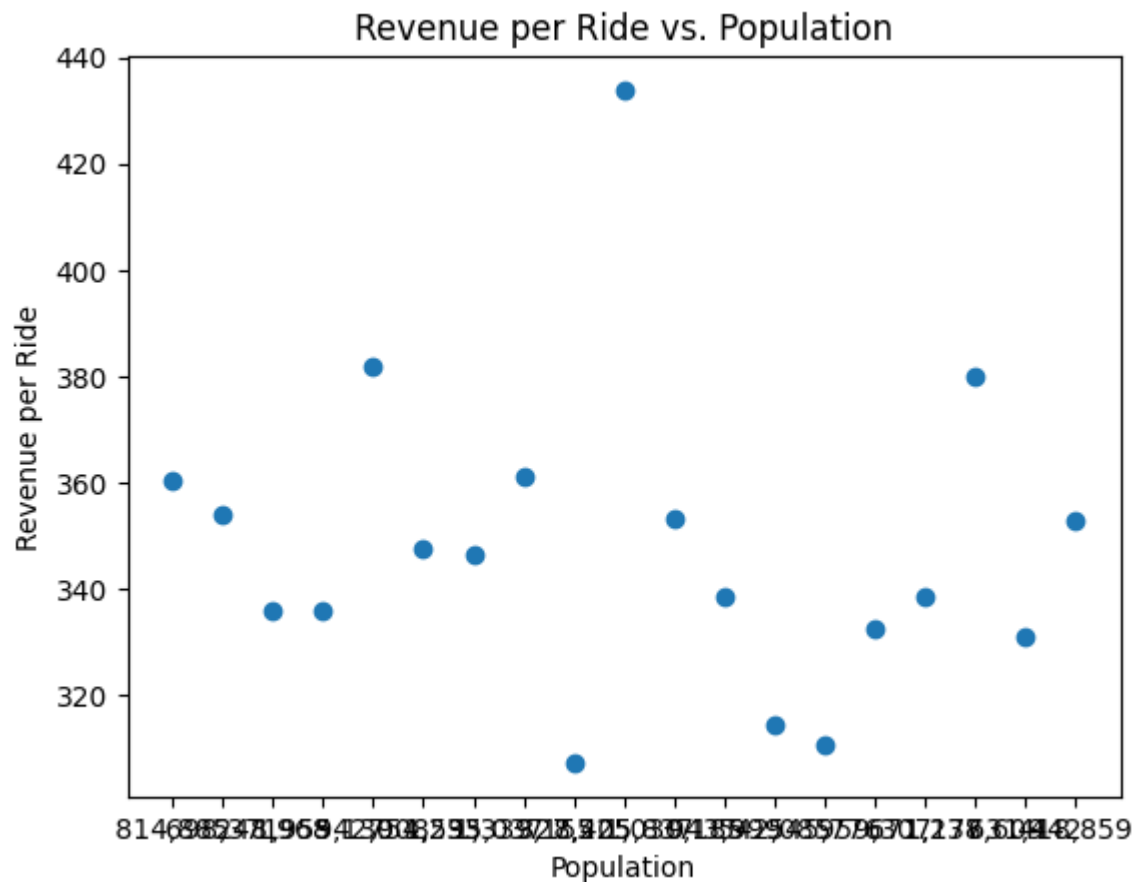
- No missing values were found.

Cab_Data.csv   Customer_ID.csv   Transaction_ID.csv   City.csv

Final cab data
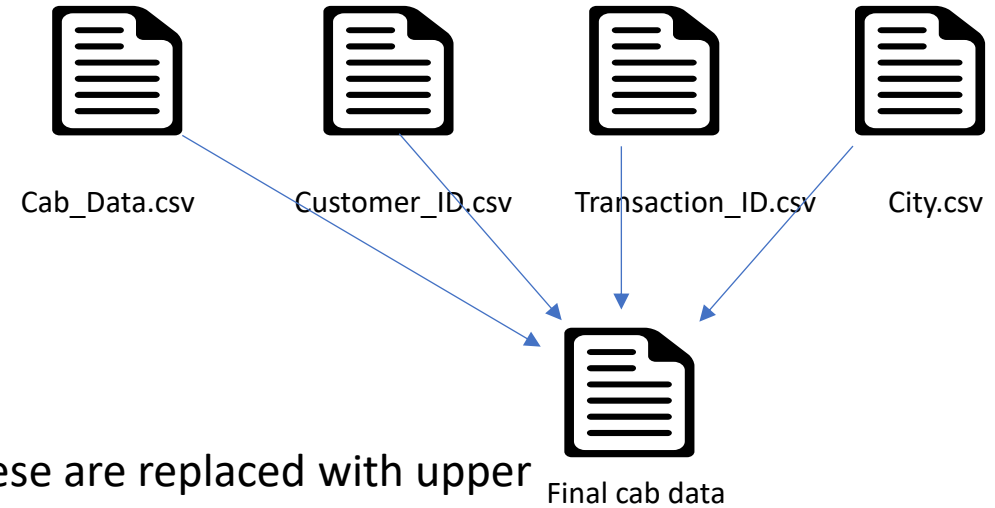
Travel Distance

Distribution of KM Travelled

Relationship between KM Travelled and Price Charged

Population and Price

Revenue per Ride vs. Population

Price Charged
(0, 100]   36356
(100, 200]   50113
(200, 300]   50390
(300, 400]   49585
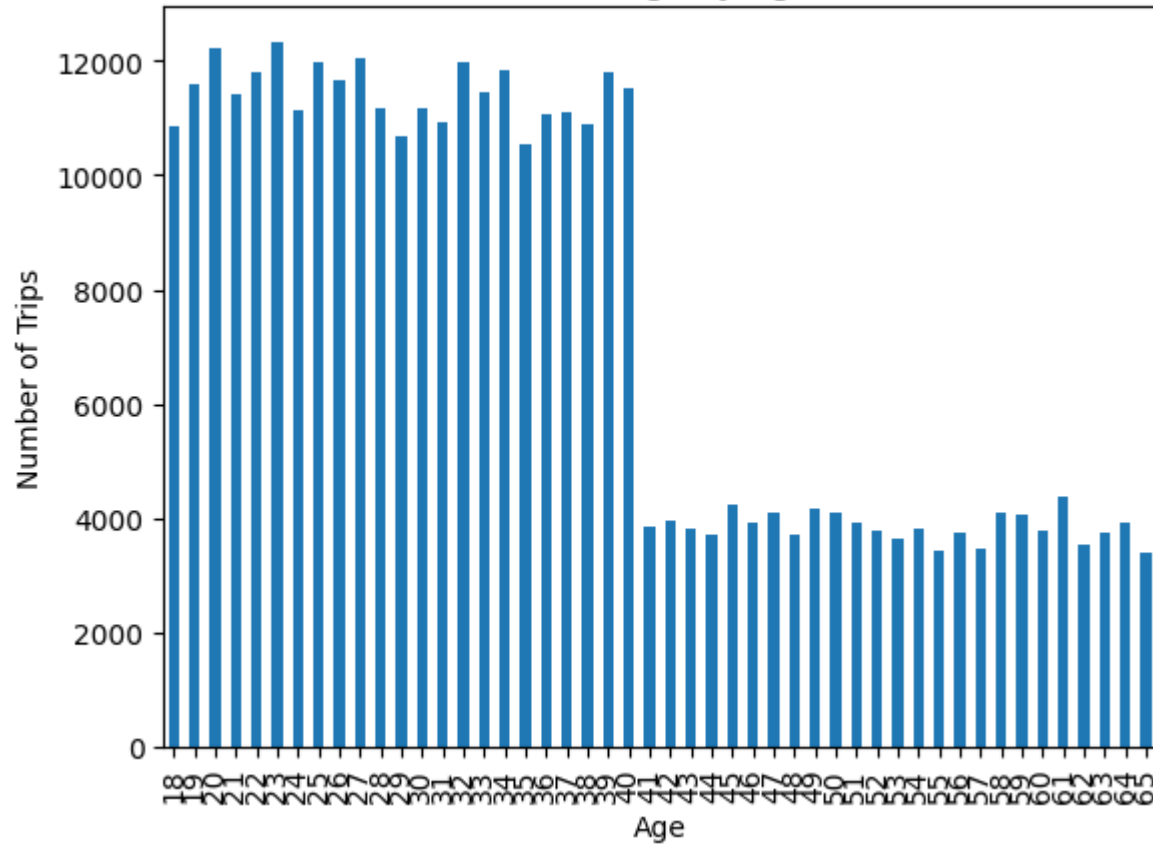(400, 500]   48012
(500, 600]   124936
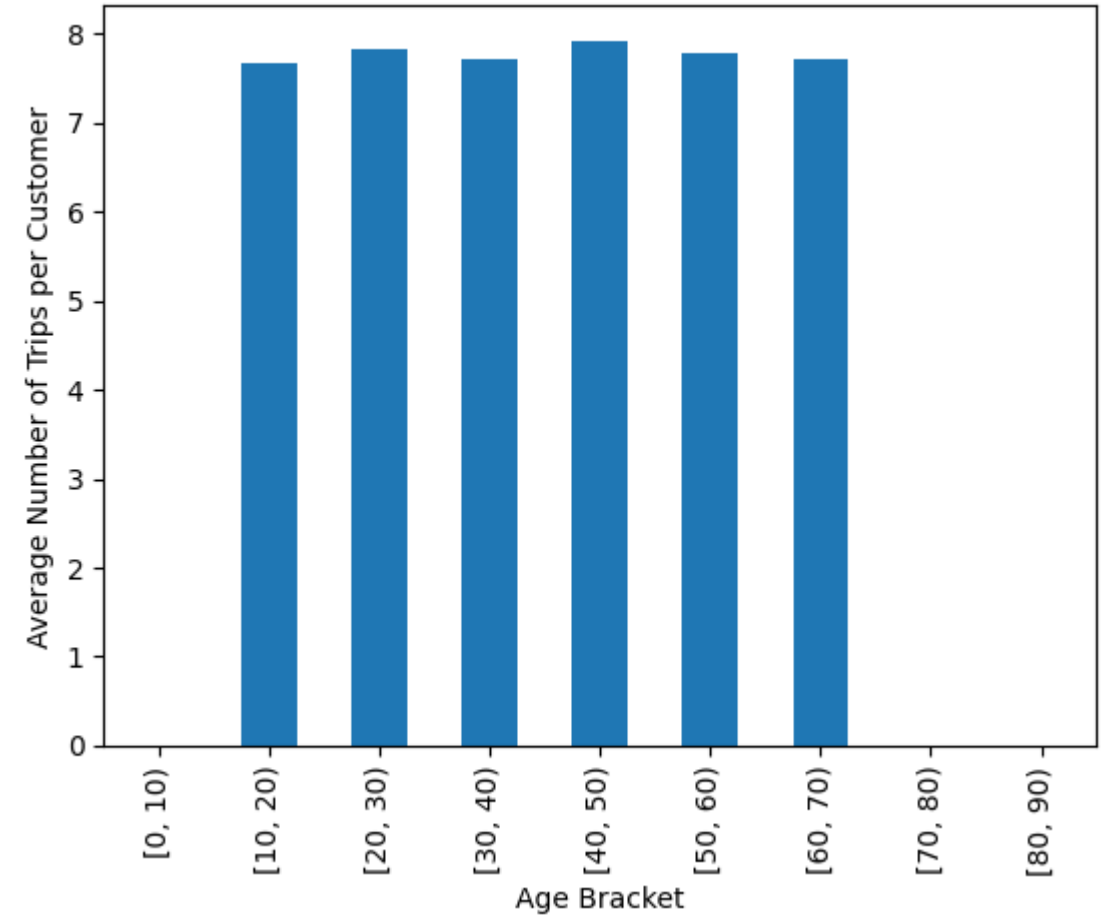
Payment_Mode
Card   215504
Cash   143888
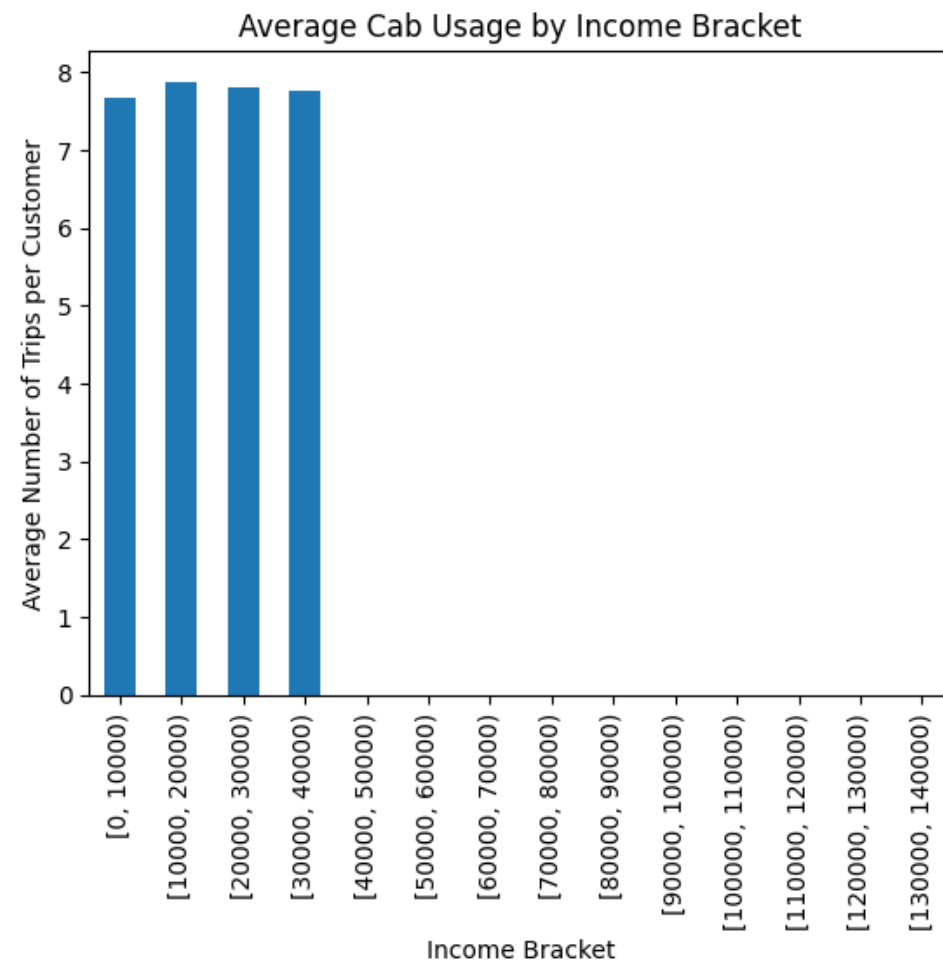Name: Transaction ID, dtype: int64

Age

Income

Average Cab Usage by Income Bracket

Time
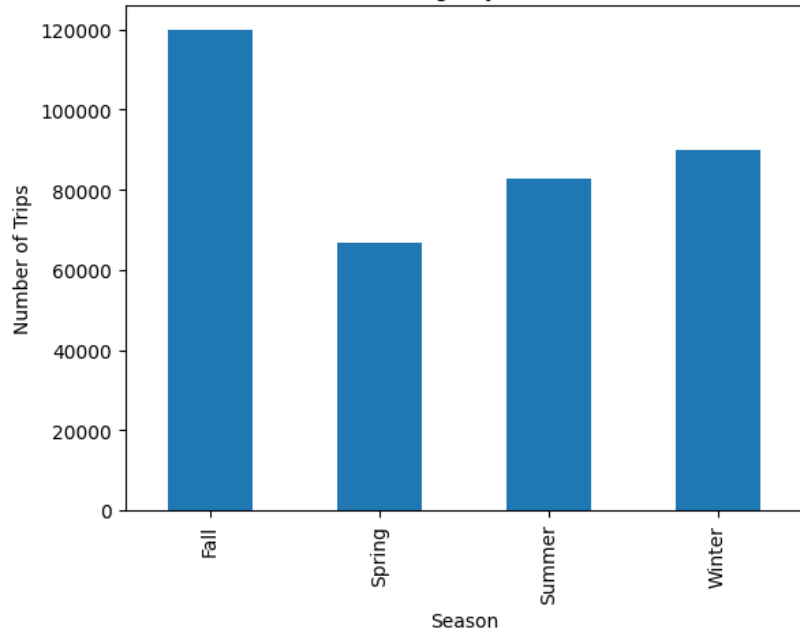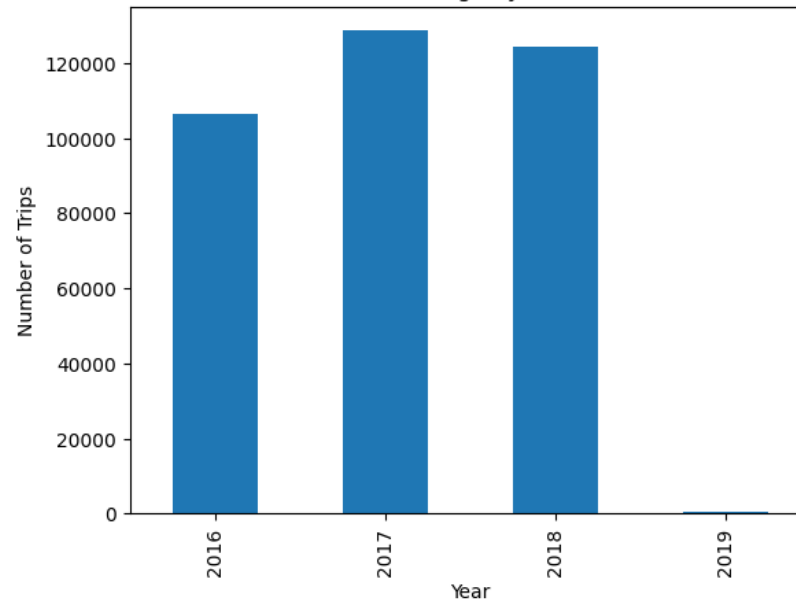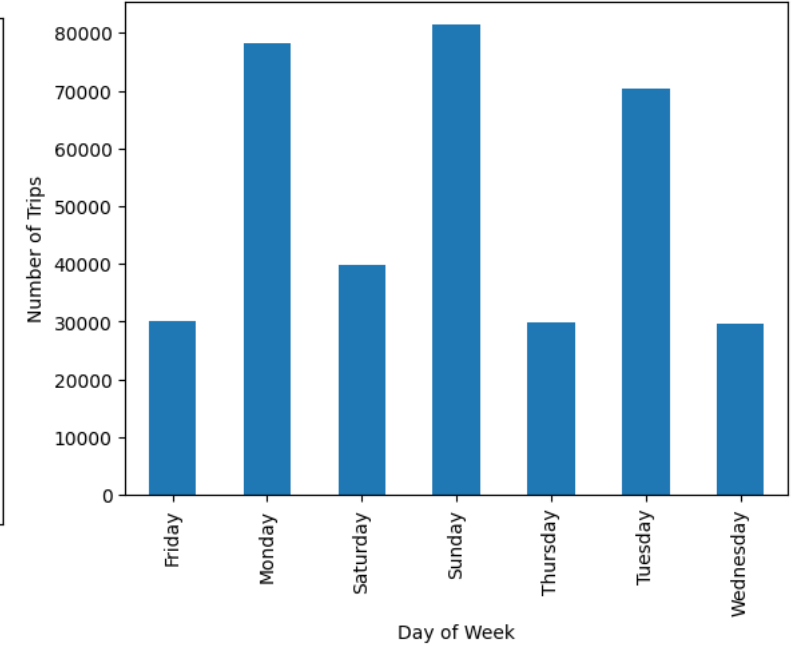
Understandings

We have evaluated both the dataset on usage statistics on following points:

- **Travel Distance:** We see a roughly equal distribution in travel distance and a linear relationship between cost and travel distance.

- **Population:** We determine no relationship between the population of a city and cab usage.

- **Age:** A roughly even age distribution, with much more users being young. However, all users have roughly equal cab usage.

- **Income:** No relationship between income and cab usage.

- **Time:** Cab usage peaks in fall. Cab usage generally increasing by year. Sunday-Tuesday show the highest cab usage.

Thank You