G2M Case Study

Data Science Virtual Internship

10-Oct-2021

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 Understanding

- Forecasting profit and number of rides for each cab type

- Finding the most profitable Cab company

- Recommendations for investment

Data Exploration

- 20 Features(including 6 derived 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 but due to unavailability of trip duration details ,we are not treating this as outlier.

- Profit of rides are calculated keeping other factors constant and only Price_Charged and Cost_of_Trip features used to calculate profit.

- Users feature of city dataset is treated as number of cab users in the city. we have assumed that this can be other cab users as well(including Yellow and Pink cab)
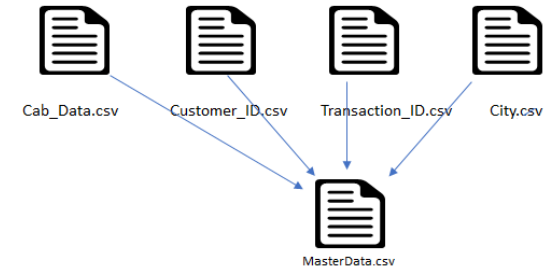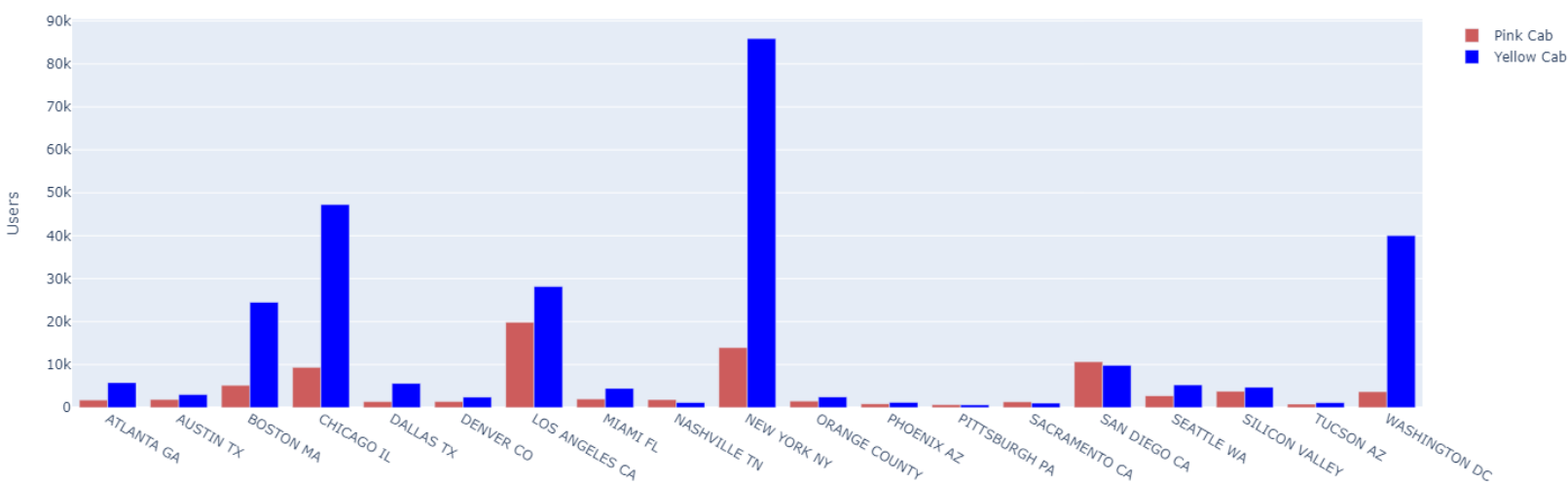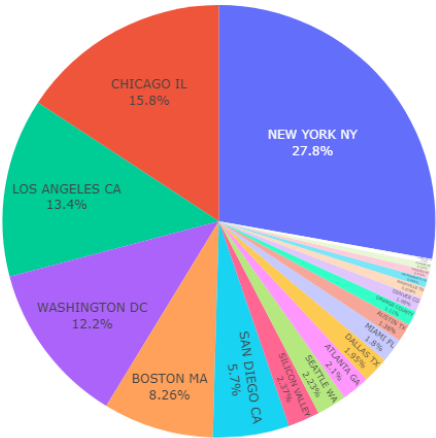
Total User Profile Analysis

Pink & Yellow Cab Firm Total Users Overview
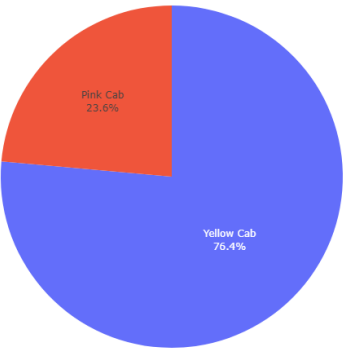
Pink & Yellow Cab Firm Users Distribution Over City


Total Users Overview by Cities

● As seen from this Pie Chart; The total number of users of Yellow Cab is approximately 3 times that of Pink Cab.

● As seen from this Bar Chart; For the Yellow Cab Company, the highest number of users on a city basis are in New York, Washington and Chicago, while for the Pink Cab Company, the most are in Los Angeles, New York and San Diego.
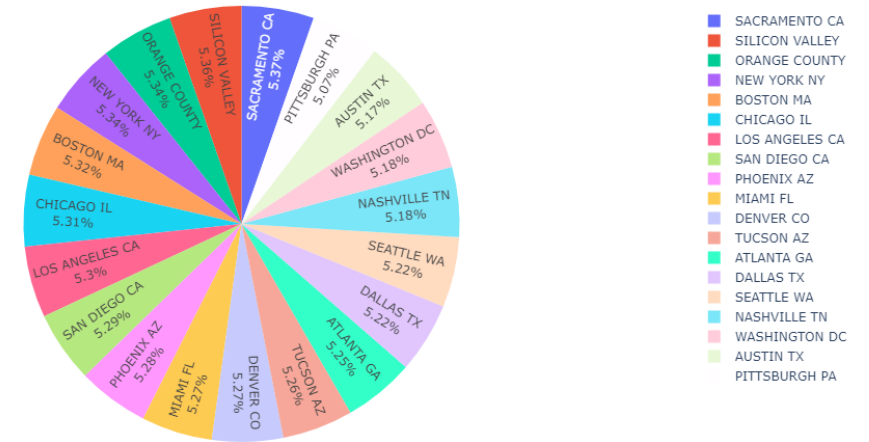
● As seen from this Pie Chart; On the basis of cities, the highest number of total users are in New York, Chicago, Los Angeles, Washington and Boston.

Users Average Income Analysis
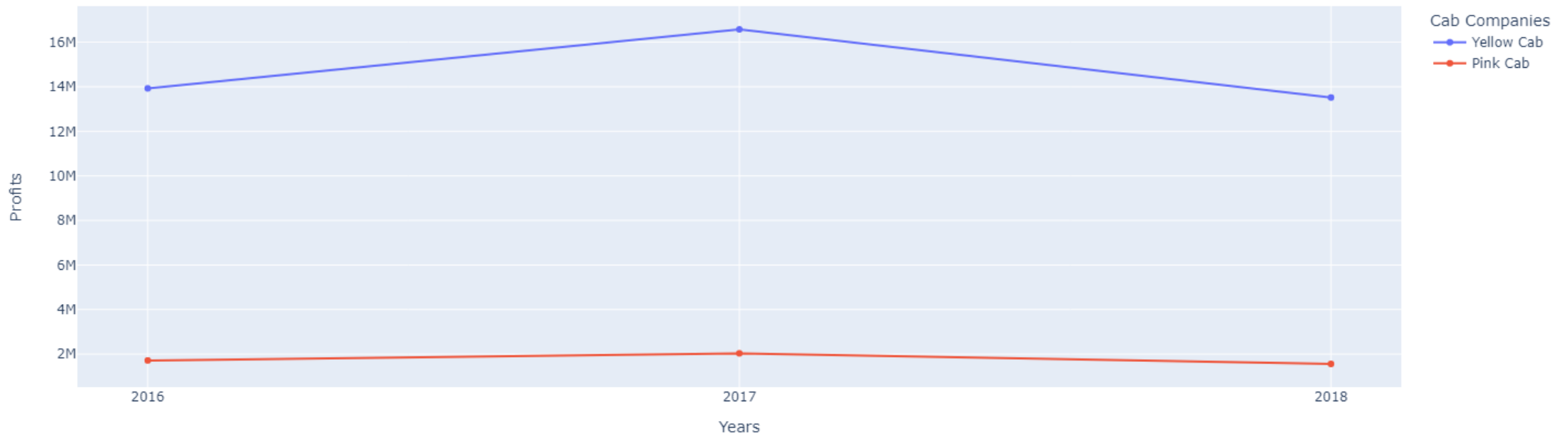

Average Income by Cab Firm
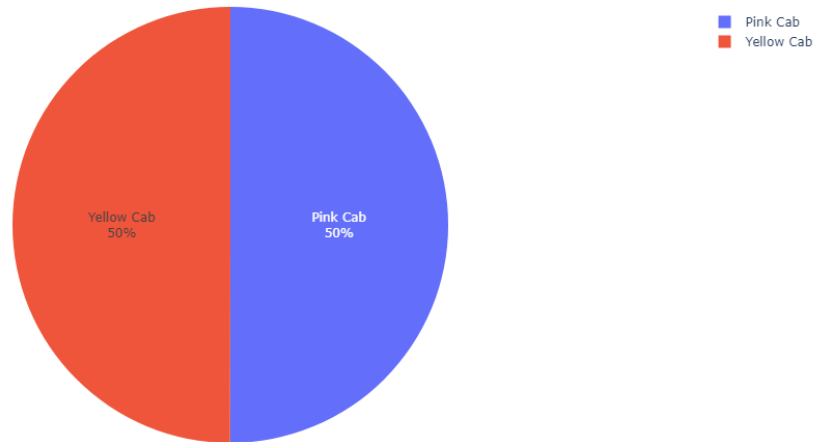
Average Income by Cities

● As seen from this Pie Chart; The average income of all users by city is approximately equal.
● As seen from this Pie Chart; The average income of all users by companies is approximately equal.

Yearly Profit Analysis

Total Profit per year by Cab Firm
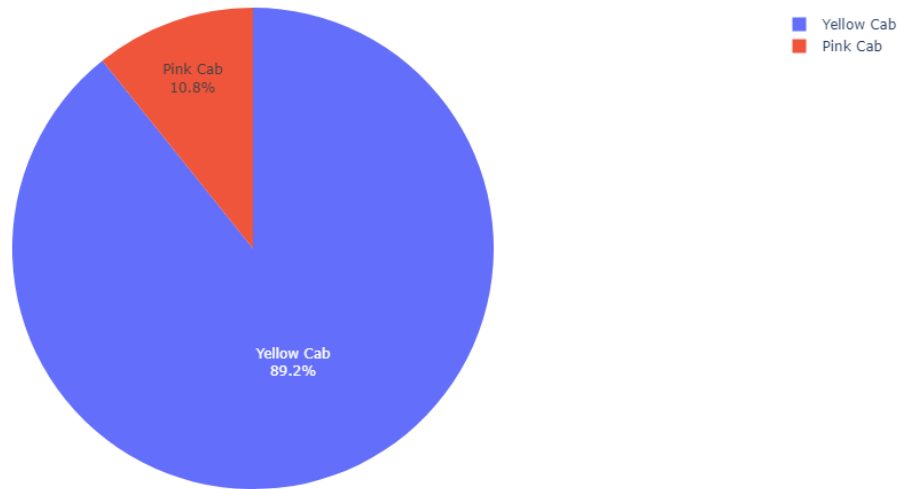

- As seen from this Pie Chart;
  In 2016 , The **total market profit share** of Yellow Cab is **approximately 8.15 times** that of Pink Cab.
  In 2017 , The **total market profit share** of Yellow Cab is **approximately 8.16 times** that of Pink Cab.
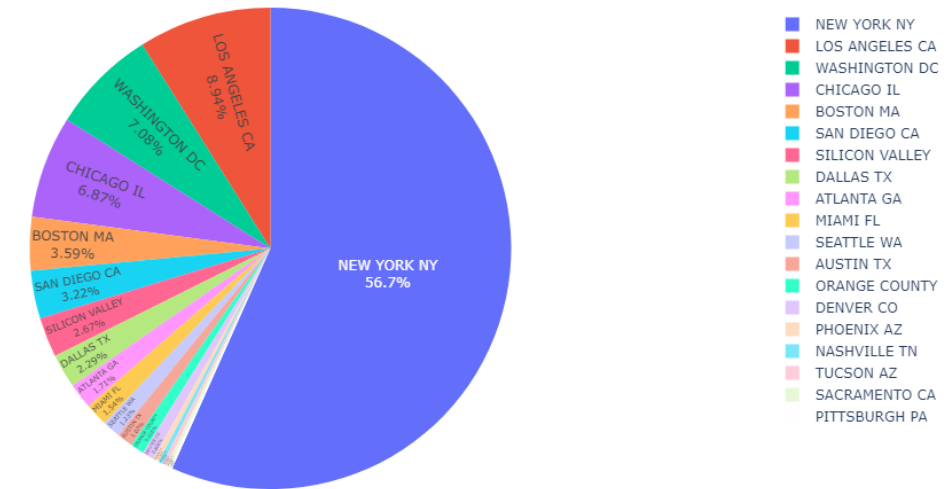 In 2018 , The **total market profit share** of Yellow Cab is **approximately 8.66 times** that of Pink Cab

Total Market Profit Share Analysis

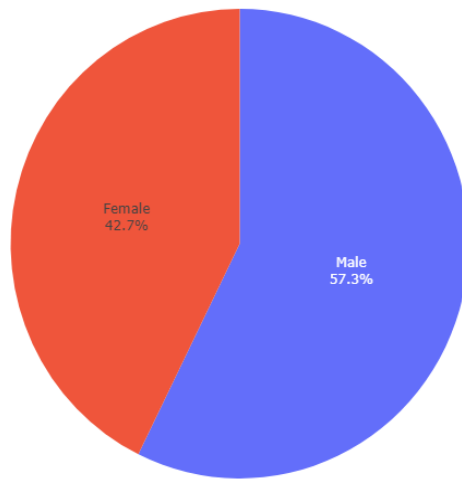Total Market Profit Share by Cab Firms

Total Market Profit Share by Cities

● As seen from this Pie Chart; The total market profit share of Yellow Cab is approximately 9 times that of Pink Cab.

● As seen from this Pie Chart; More than half of the total market profit share on the basis of cities belongs to New York.

Total User Gender Analysis

Total Users Overview by Gender

Pink & Yellow Cab Firm Users Distribution Over Gender
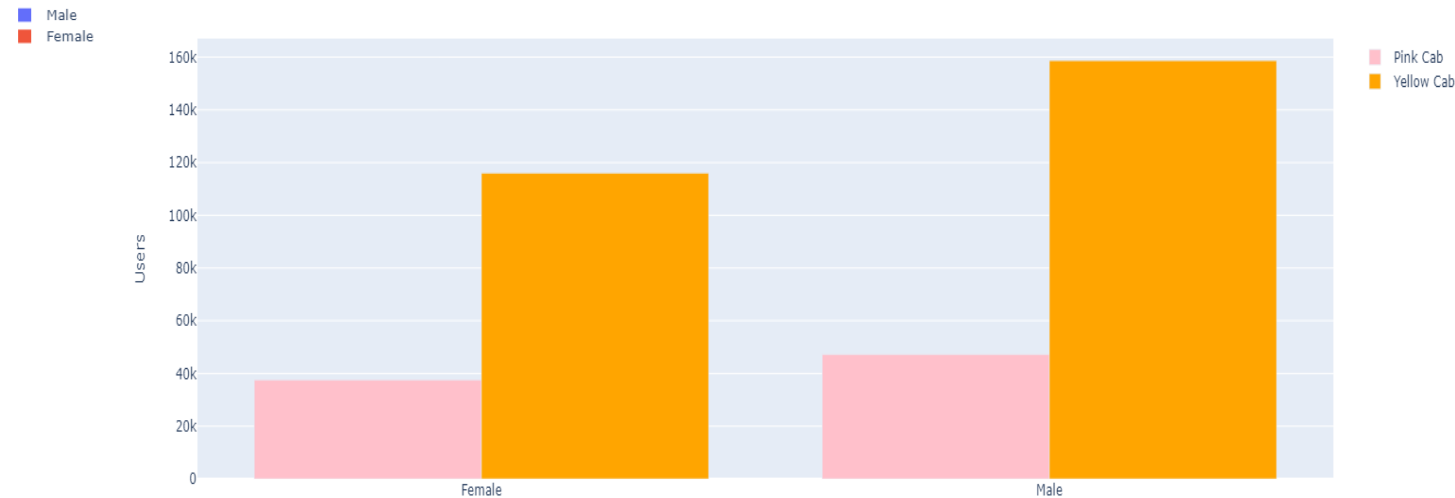
● As seen from this Pie Chart; In the distribution of users by gender, there is an approximate 3 to 2 ratio for men and women.

● As seen from this Bar Chart; When the distribution of users by gender is analyzed on a company basis, while the male-female ratio is 57.6% - 42.4% in Yellow Cab Company, the male-female ratio is 55.9% - 44.1% in Pink Cab Company.

Total User Age Group Analysis

Total Users Overview by Age Groups

■ 40 > Age >= 25 (MIDDLE)
■ 65 >= Age >= 40 (OLD)
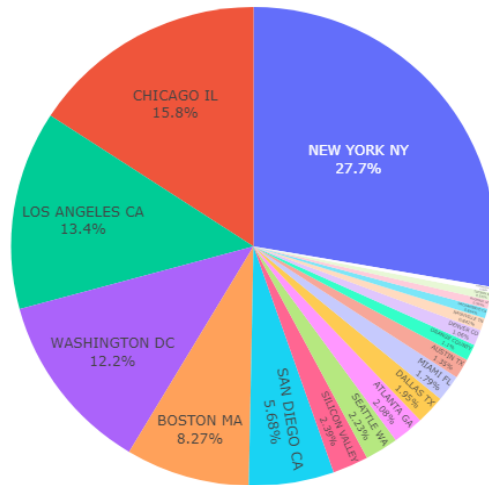■ 25 > Age >= 18 (YOUNG)
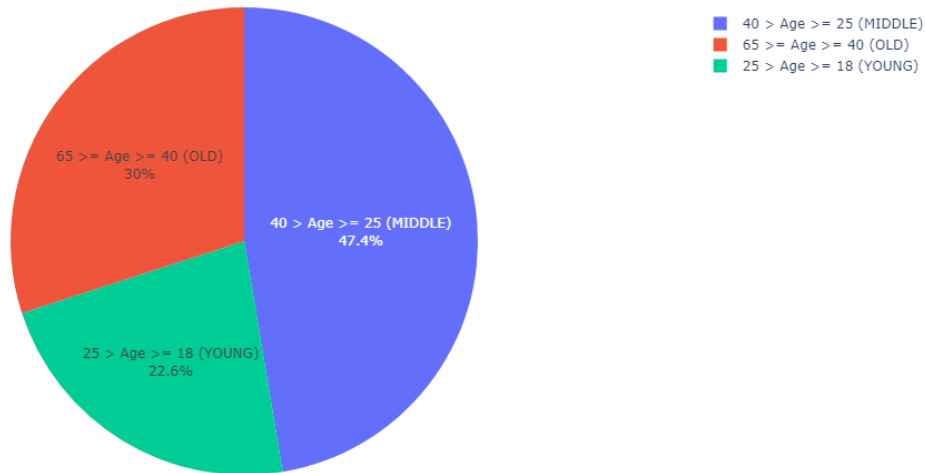
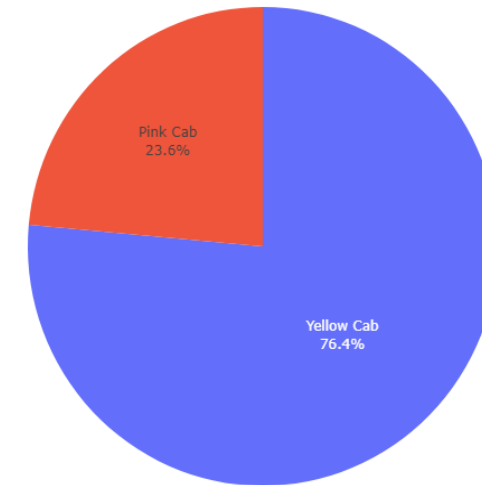Pink & Yellow Cab Firm Users Distributions by Age Groups
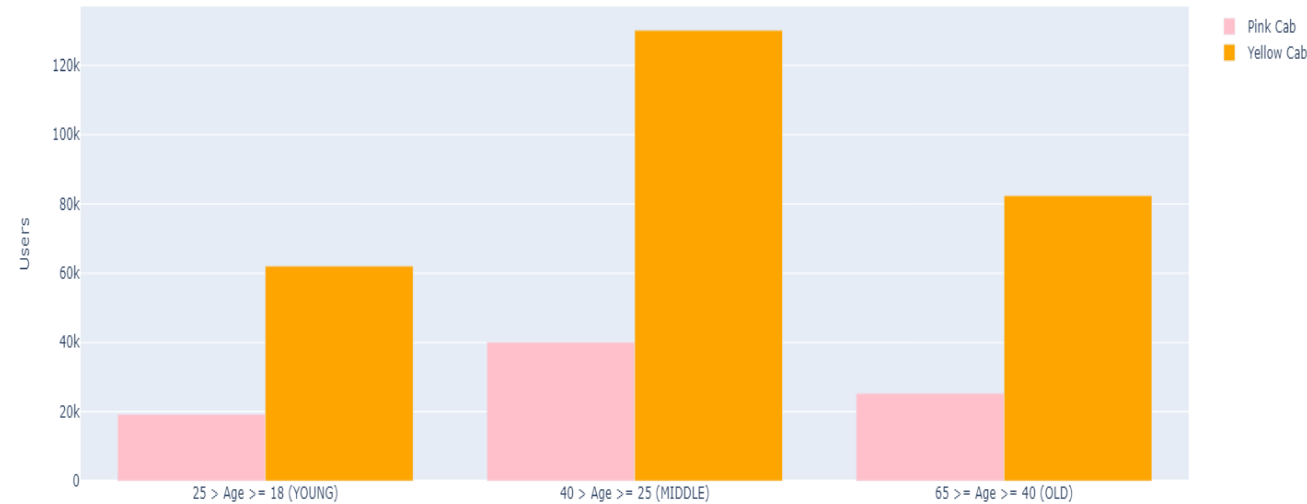
● As seen from this Pie Chart; Looking at the age distribution of all users, it is seen that approximately half of them are between the ages of 18-25.

● As seen from this Bar Chart; Looking at the age distribution of all users in the basis of companies, it is seen that both have the same percentage distribution for every age group.

Total Travelled in KM Analysis

Total KM Travelled by Cities

NEW YORK NY
27.7%

CHICAGO IL
15.8%

LOS ANGELES CA
13.4%

WASHINGTON DC
12.2%

BOSTON MA
8.27%

SAN DIEGO CA
5.68%

SILICON VALLEY
2.3%

SEATTLE WA
2.2%

ATLANTA GA
1.06%

DALLAS TX
1.99%

MIAMI FL
1.79%

- NEW YORK NY
- CHICAGO IL
- LOS ANGELES CA
- WASHINGTON DC
- BOSTON MA
- SAN DIEGO CA
- SILICON VALLEY
- SEATTLE WA
- ATLANTA GA
- DALLAS TX
- MIAMI FL
- AUSTIN TX
- ORANGE COUNTY
- DENVER CO
- NASHVILLE TN
- SACRAMENTO CA
- PHOENIX AZ
- TUCSON AZ
- PITTSBURGH PA

Total KM Travelled by Cab Firm

Pink Cab
23.6%

Yellow Cab
76.4%

- Yellow Cab
- Pink Cab

● As seen from this Pie Chart; On the basis of cities, the most travelled in KM are New York, Chicago, Los Angeles, Washington and Boston.

● As seen from this Pie Chart; The total travelled in KM for Yellow Cab is approximately 3 times that of Pink Cab.

Average Profit per Travelled in KM Analysis

Average Profit per KM Travelled by Cab Firm
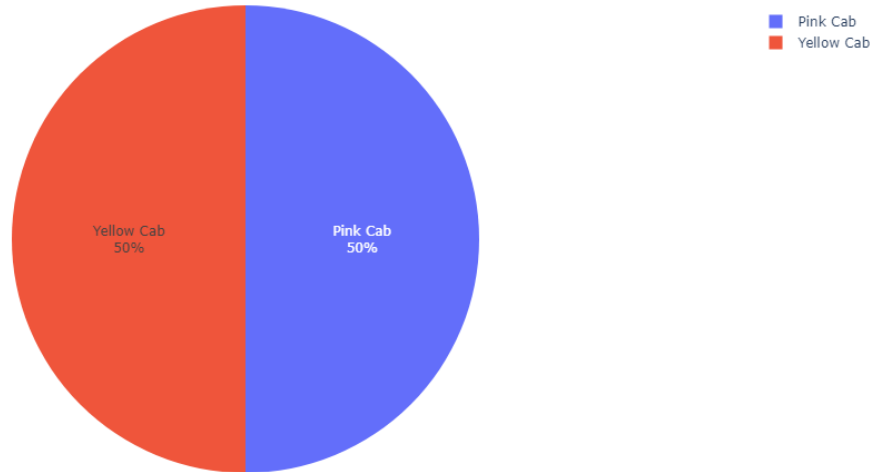
Average Profit per KM Travelled by Cities
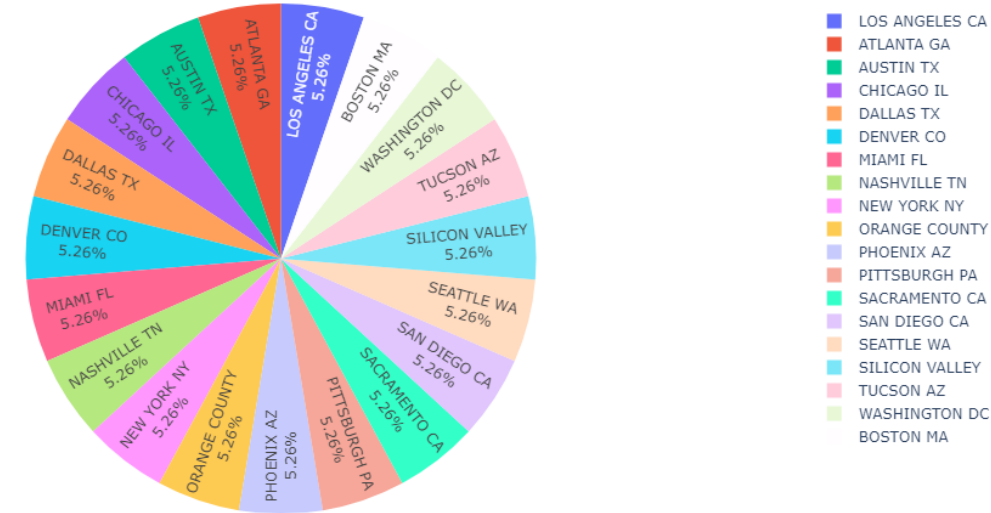
- As seen from this Pie Chart; The average profit per travelled in KM by companies is approximately equal.

- As seen from this Pie Chart; The average profit per travelled in KM by cities is approximately equal.

Users Payment Preference Analysis

Total Users Overview by Payment Method
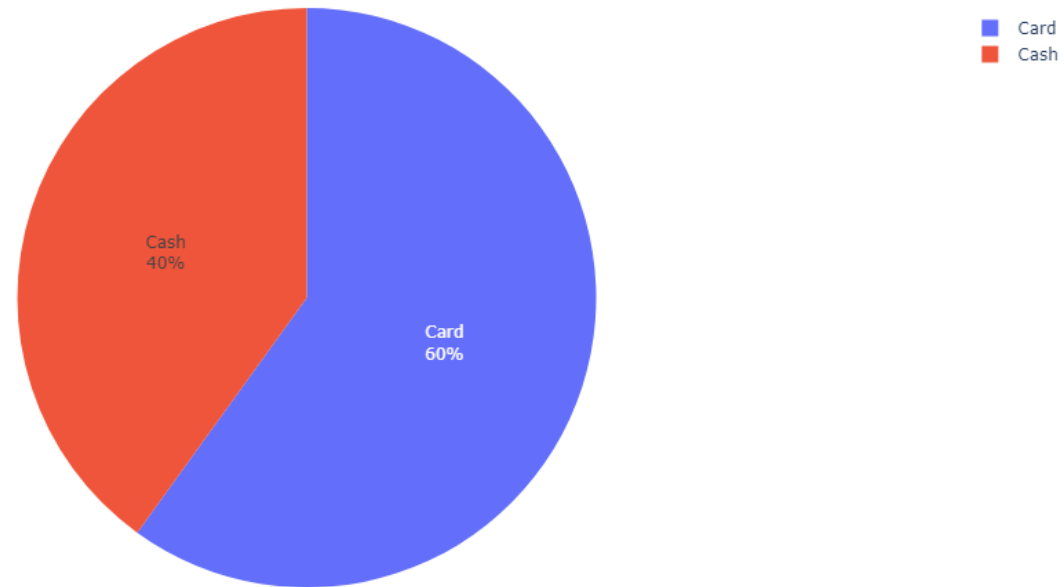
- Card
- Cash

● As seen from this Pie Chart; Considering the payment preferences of all users, the credit card- cash payment ratio is 3 to 2.

Thank You

Data Glacier

Your Deep Learning Partner