Exploratory Data Analysis

G2M Case Study

9/21/2023

Jinghao Shen

Agenda

Data Glacier
Your Deep Learning Partner
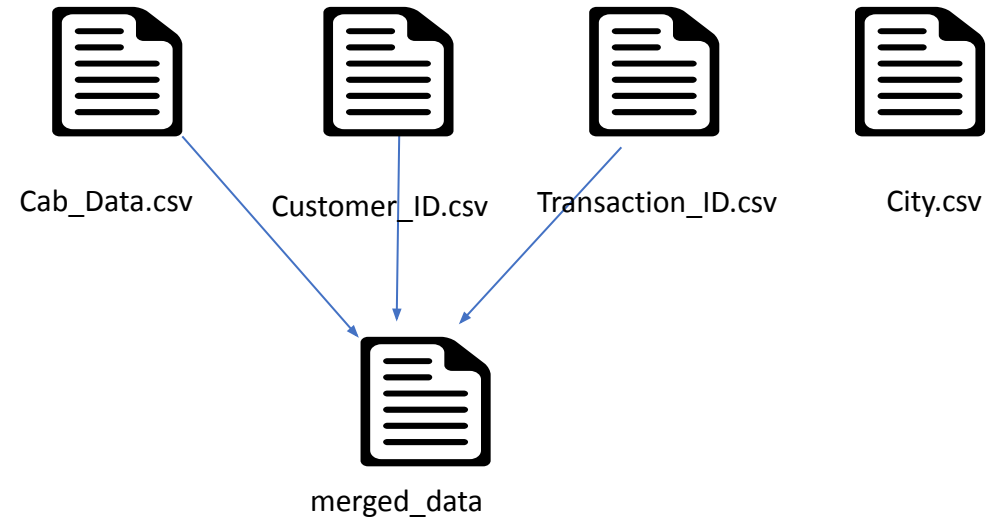
Problem Statement

- XYZ is a private 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 and as per their Go-to-Market(G2M) strategy they want to understand the market before taking final decision.

- Object: XYZ is interested in using your actionable insights to help them identify the right company to make their investment.
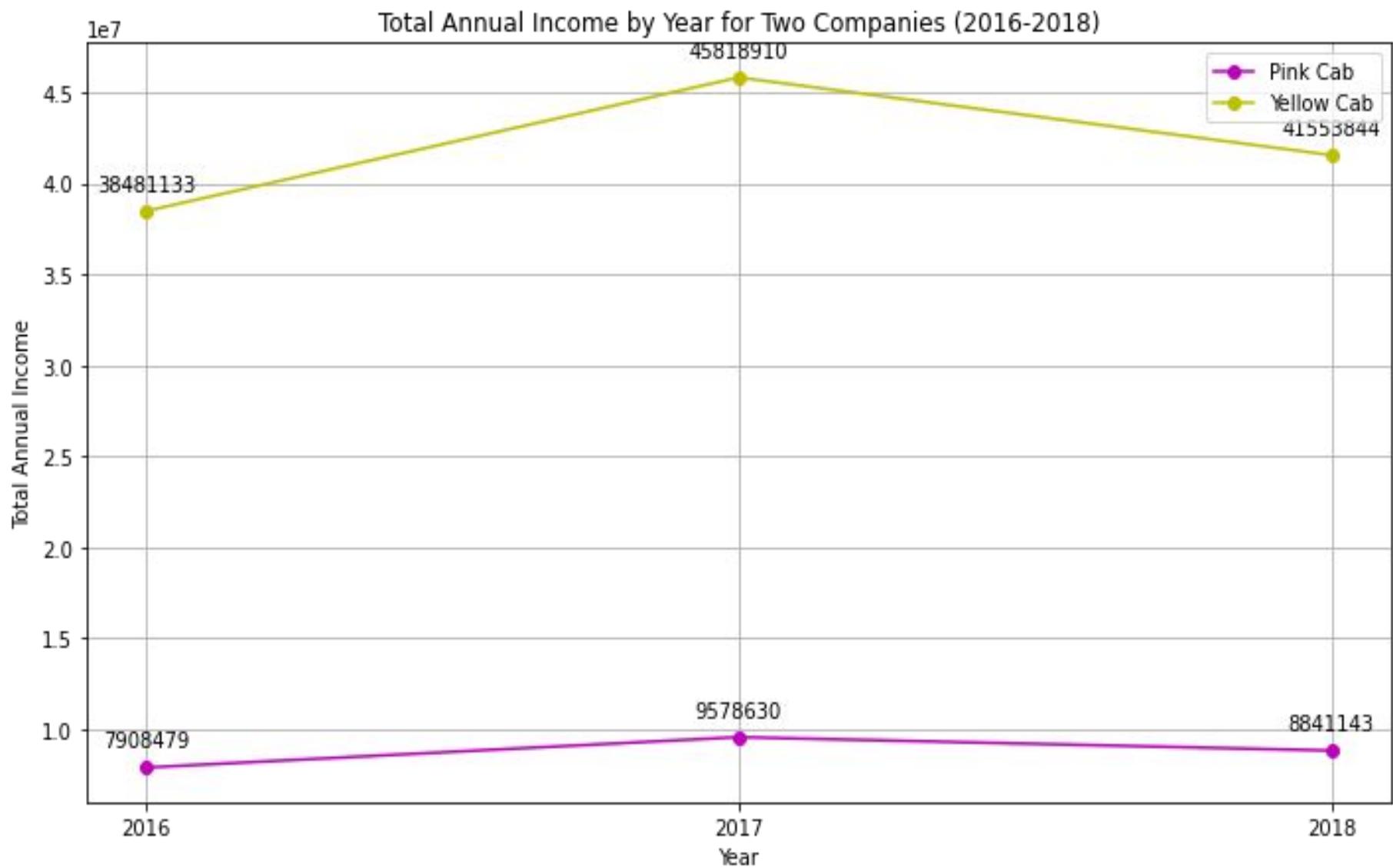
Data Exploration

- Features: 12
- Time period of data: 31/01/2016 to 31/12/2018
- Total number of observations: 359392

Assumptions:

1. The difference between 'Price charged' and 'Cost of Trip' is the 'Profit' for each transaction.
2. Users can user of Yellow Cab and Pink Cab at the same time.
3. Users with 'Income (USD/Month)' > 15000 as 'High income' class, 'Income (USD/Month)' < 3500 as 'Low income' class, the rest as 'Middle Income' class.

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

merged_data

Annually Profit Analysis

Total Annual Income by Year for Two Companies (2016-2018)

Companies total annual income in three years.

From 2016 to 2017:

Yellow Cab increased 19.09%.

Pink Cab increased 21.18%.
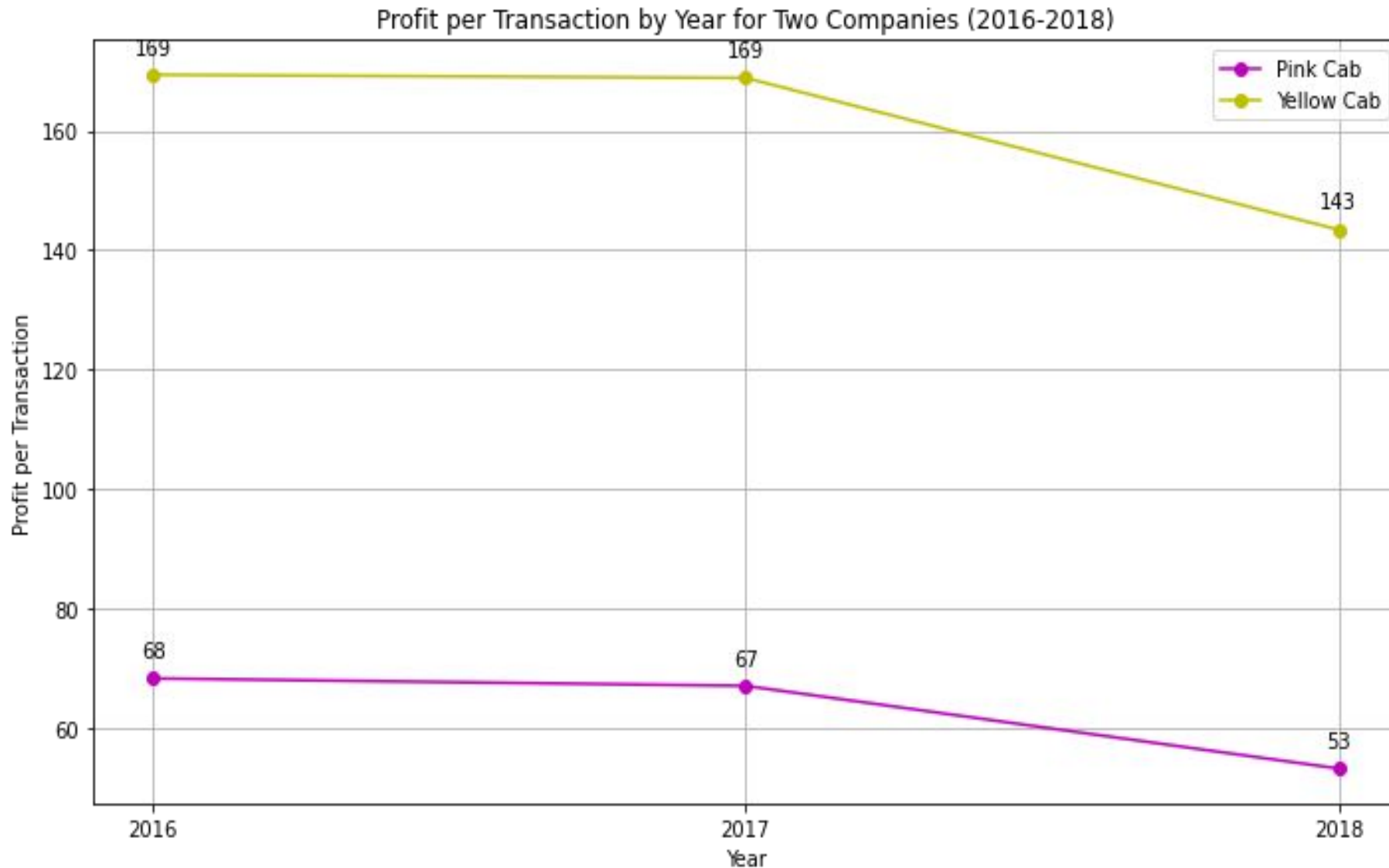
From 2017 to 2018:

Yellow Cab decreased 9.3%

Pink Cab decreased 7.69%

Overall increase:

Yellow Cab: 7.98%

Pink Cab: 11.79%

Annually Profit Analysis

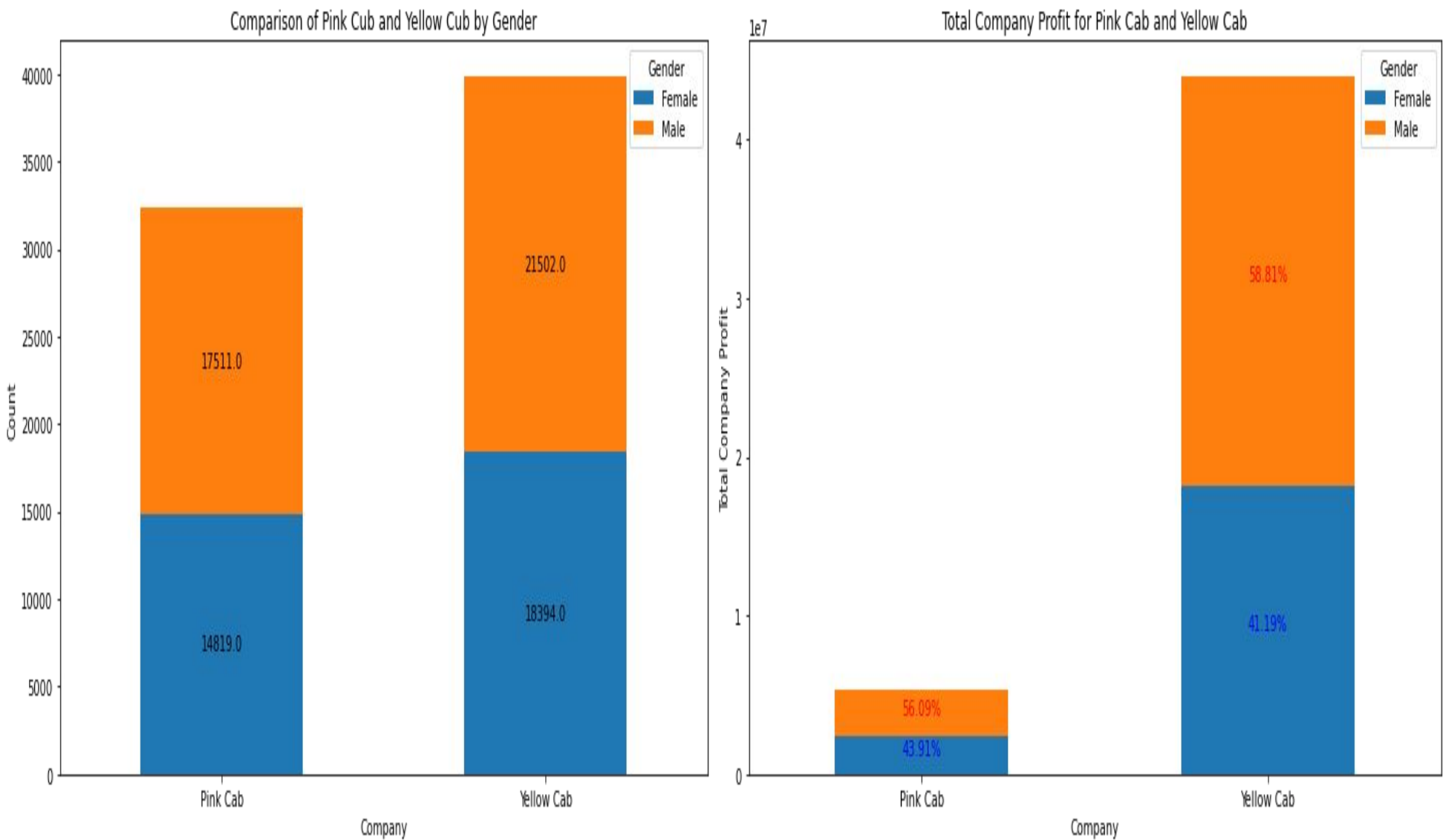Profit per Transaction by Year for Two Companies (2016-2018)

Average profit per transaction for each year.

Overall:

Yellow Cab decreased 15.38%

Pink Cab decreased 22.05%

# Profit and Customer base Analysis Gender wise

There is a similar gender distribution for the two companies.

In Pink Cab company, Male users contribute 12.18% more than Female users to the total company income.

In Yellow Cab company, Male users contribute 17.62% more than Female users to the total company income.

There are 7566 more users using Yellow Cab

Income Class wise Profit and Customer base Analysis

Number of Transactions by Income Class
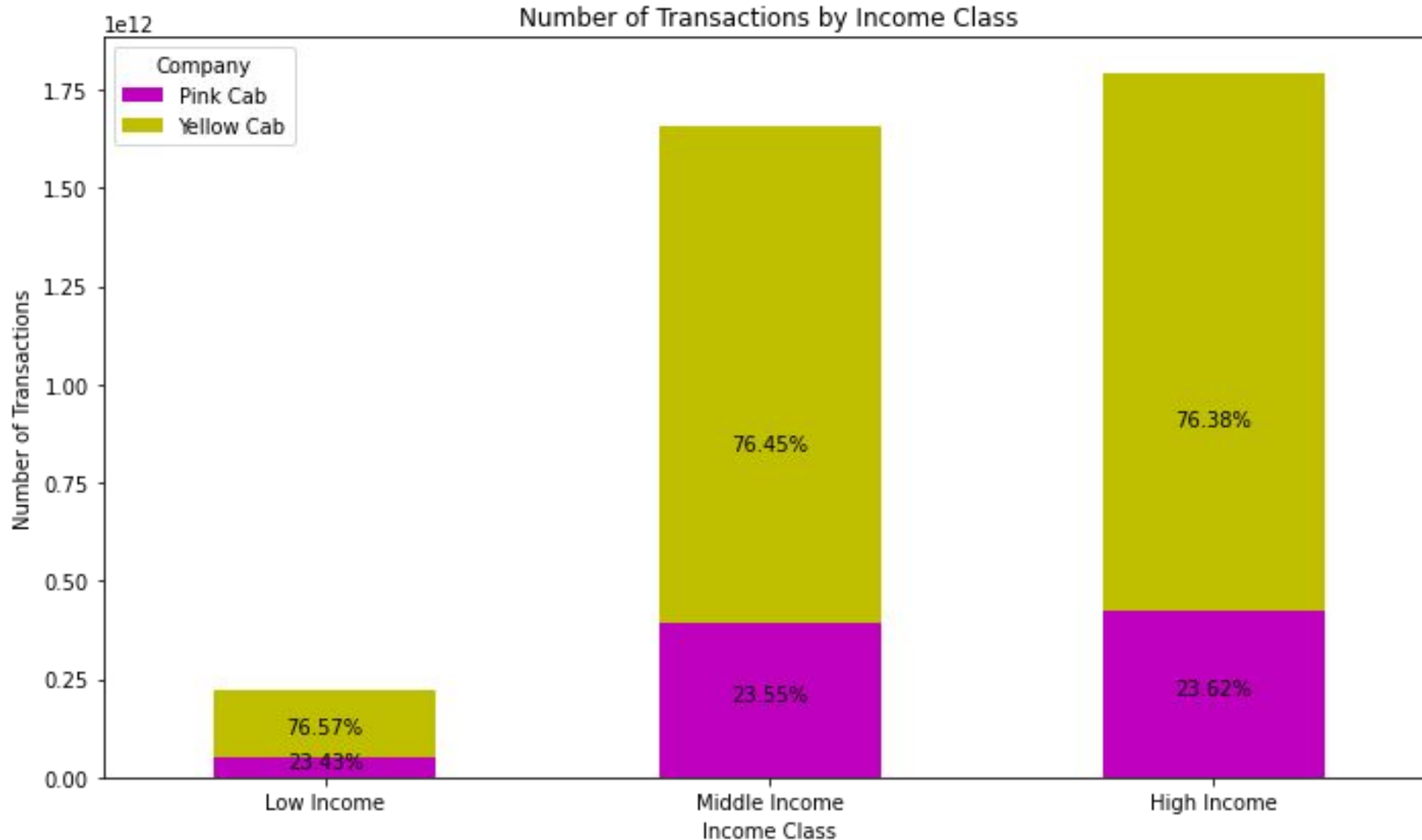
High and middle income classes take cab much more than low income class users.
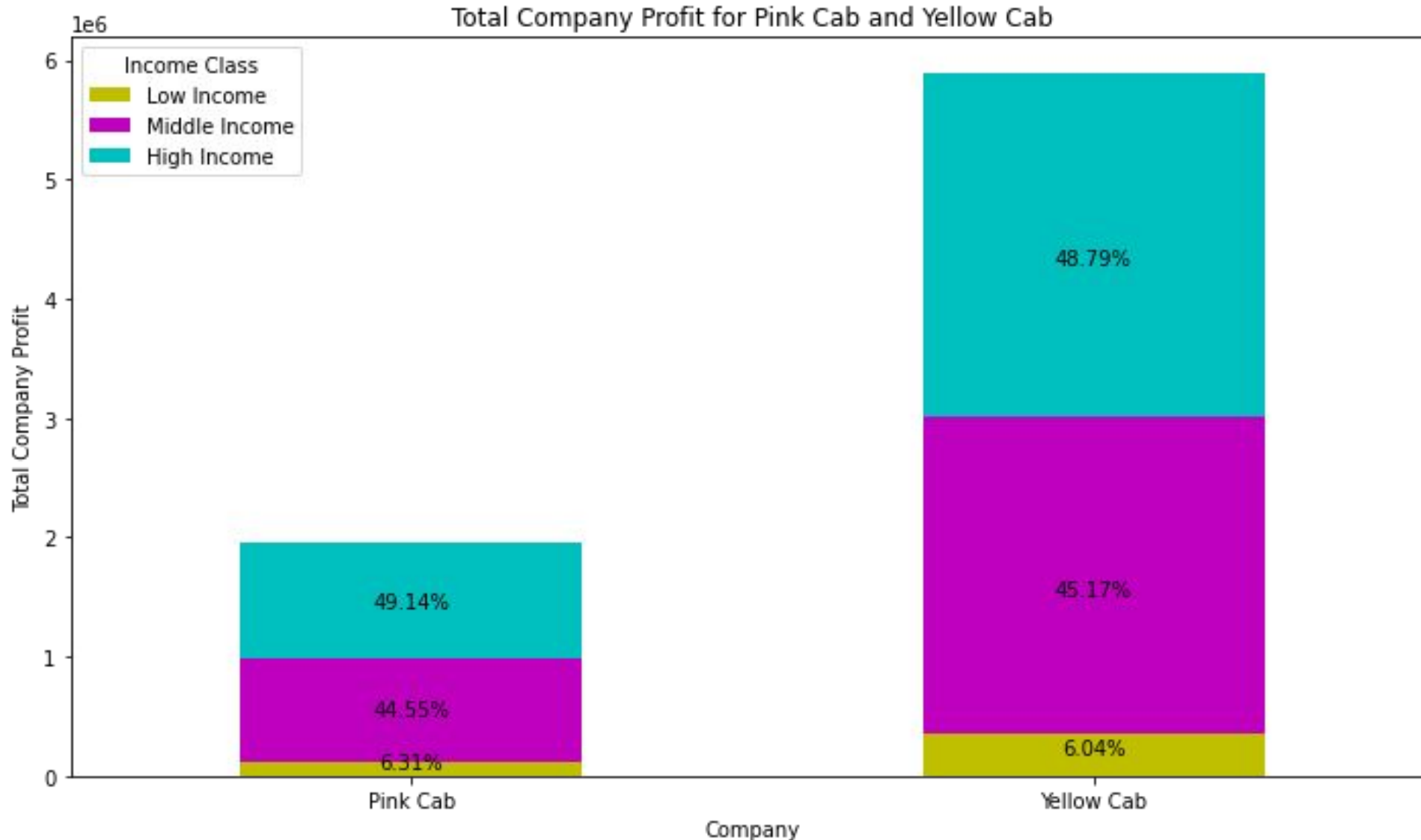
They contribute to most of the transactions.

Assuming:

High Income: Income (USD/Month) > 15000

Middle Income: Income (USD/Month) <= 15000 && >= 3500

Low Income: Income (USD/Month) < 3500

Income Class wise Profit and Customer base Analysis
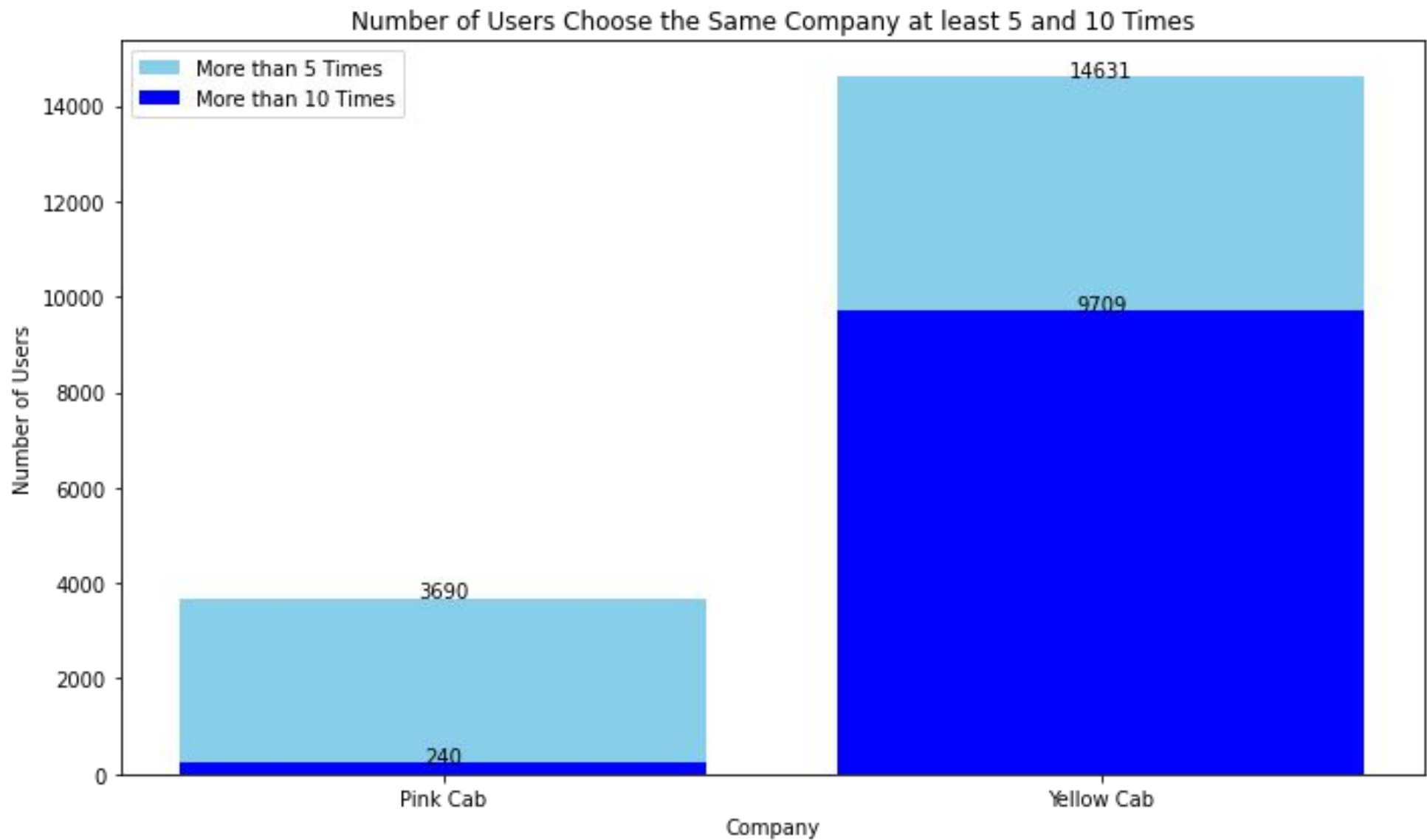

Total Company Profit for Pink Cab and Yellow Cab

High income class user contribute most in the company profit for both companies.

For Yellow Cab, 93.96% of the profit are come from Middle and Upper classes.

For Pink Cab, 93.69% of the profit are come from Middle and Upper classes.

Customer Retention

Number of Users Choose the Same Company at least 5 and 10 Times

We count the number of individual user who choose the same company for more than 5 times and 10 times.

There are 296.5% more users choose Yellow Cab more than 5 times.

There are 3945.41% more users choose Yellow Can more than 10 times.
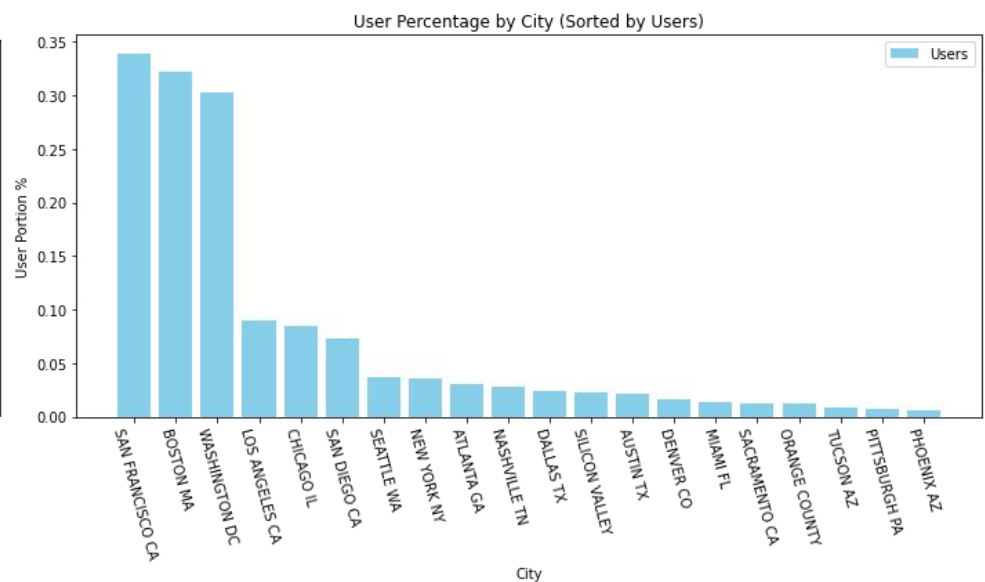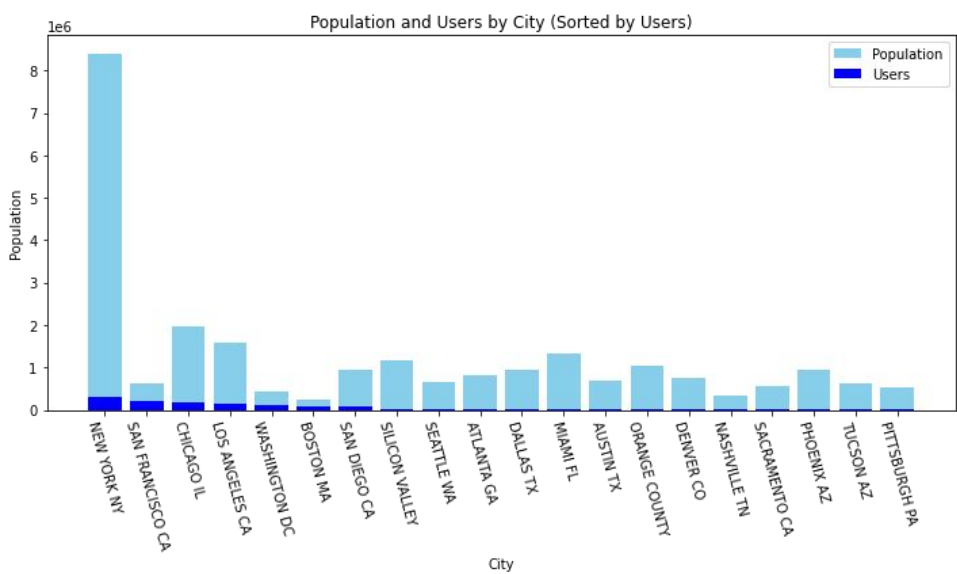
Profit and Travel Distance study by Company

Relationship between Profit and KM Travelled
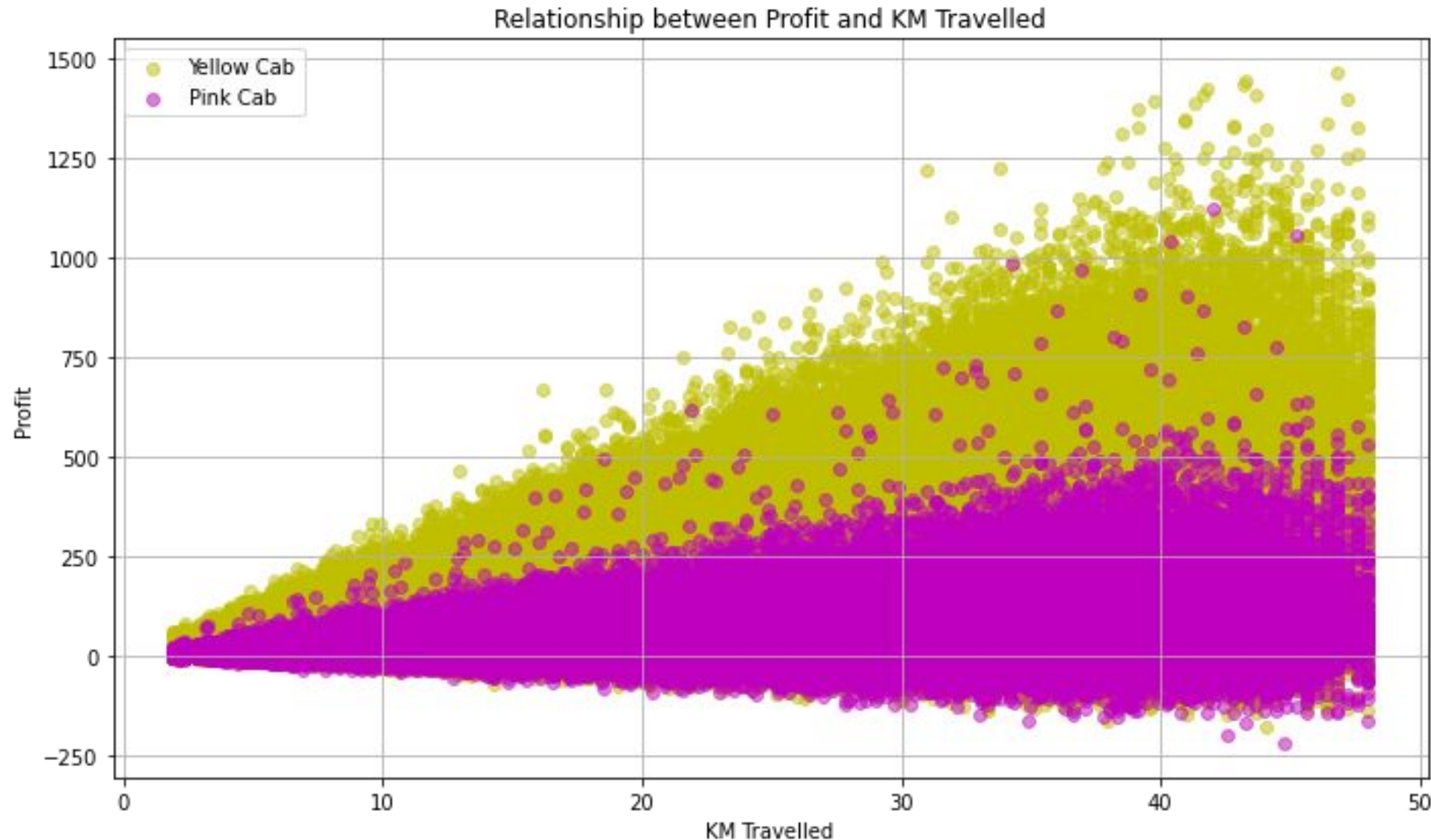
There is a positive relation between KM Travelled and Profit.

Yellow Cab profit more than Pink Cab under the same travel distance.

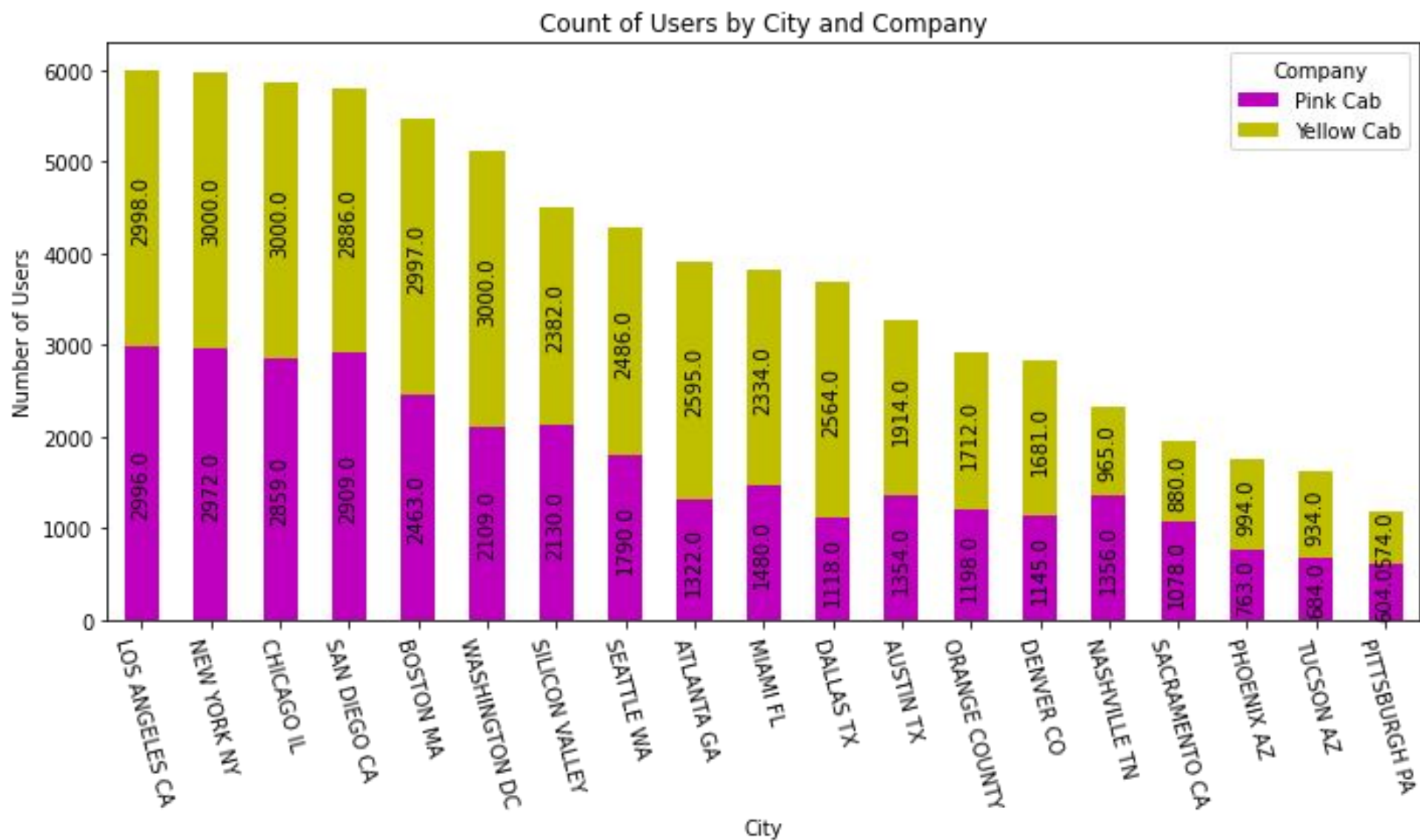* Assuming Transaction with profit below 0 means the user used a discount.

City Wise Cab Users

User number in each city.

User portion in each city.

New York has the most users.

San Francisco, Boston, and Washington DC are outstanding in user portions.

City Wise Cab Users Covered By Company
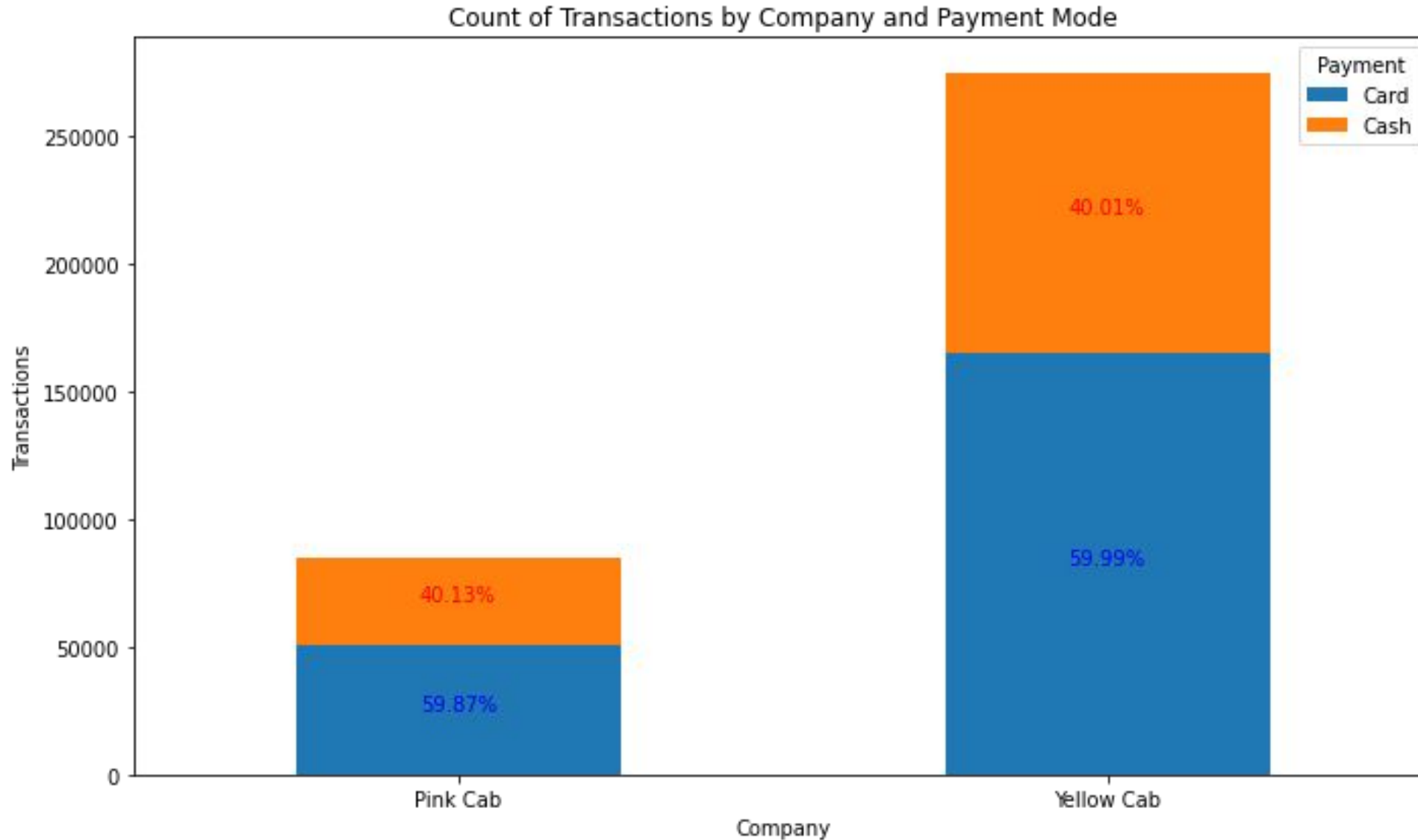

Count of Users by City and Company

In top four city, Pink Cab and Yellow Cab have similar number of users.

Out of 19 cities, in 15 cities, Yellow Cab has more users than Pink Cab.

* There is no transaction data for San Francisco.

Payment Analysis

Count of Transactions by Company and Payment Mode

In both company, more users use card than cash for transactions.

Summary and Recommendations

- Customer retention: It is found that Yellow Cab has better customer retention rate than Pink Cab, in both cases of more than 5 and 10 drives.

- City wise customer reach: Out of 19 cities, Yellow Cab has more users in 15 cities than Pink Cab.

- Annual income: Both companies have an increase in 2017 and a decrease in 2018. However, in profit per transaction, Pink Cab decreased more than Yellow Cab in three years.

- Distance profit: Yellow Cab profit more than Pink Cab under same travel distance in average.

- Income class reach: Both companies mainly reach middle and high income classes. However, Yellow Cab gets more transactions from low income class as well.

- Payment pattern: About 60% the users pay with card for both companies.

Recommendation:

We will recommend Yellow Cab for investment.

Thank You