Exploratory Data Analysis
G2M Insight for Cab Investment Firm

20-06-2024

NAZRI

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Data Glacier
Your Deep Learning Partner

Executive Summary

- XYZ, a private firm in US, is planning for an investment in the cab industry due to its remarkable growth in recent years.

- It is seeking actionable insights to inform their investment decisions. This project aims to provide actionable insights to help XYZ identify the right company for their investment.

Problem Statement

- XYZ aims to invest in the cab industry, but they need to understand the market better before making a final decision as per their Go-to-Market (G2M) strategy. Specifically, they want to determine which cab company is a better investment opportunity. To address this, we conducted a comprehensive data analysis.

Approach

- Here we merge the datasets 'cab_data' and 'transaction_data' with the 'Transaction ID' column and the 'customer_data' dataset will be merged to it with the 'Customer ID' column. Then we merge the 'city_data' dataset to it with the 'city' column and hence form 'final_cab_data' dataset.

- We then perform Exploratory data Analysis in the final_cab_data dataset.

Exploratory Data Analysis

- Data analysis covers the period from 31/01/2016 to 31/12/2018, using multiple datasets, including Cab_Data, Customer_ID, Transaction_ID, and City that contains information on 2 cab companies- Yellow Cab and Pink Cab

➢ Cab_Data.csv – this file includes details of transaction for 2 cab companies
➢ Customer_ID.csv – this is a mapping table that contains a unique identifier which links the customer's demographic details
➢ Transaction_ID.csv – this is a mapping table that contains transaction to customer mapping and payment mode
➢ City.csv – this file contains list of US cities, their population and number of cab users

- Final_Cab_Data dataset is formed by merging these datasets
- Total Data points: 359392 rows and 14 columns

Relationship Between Variables
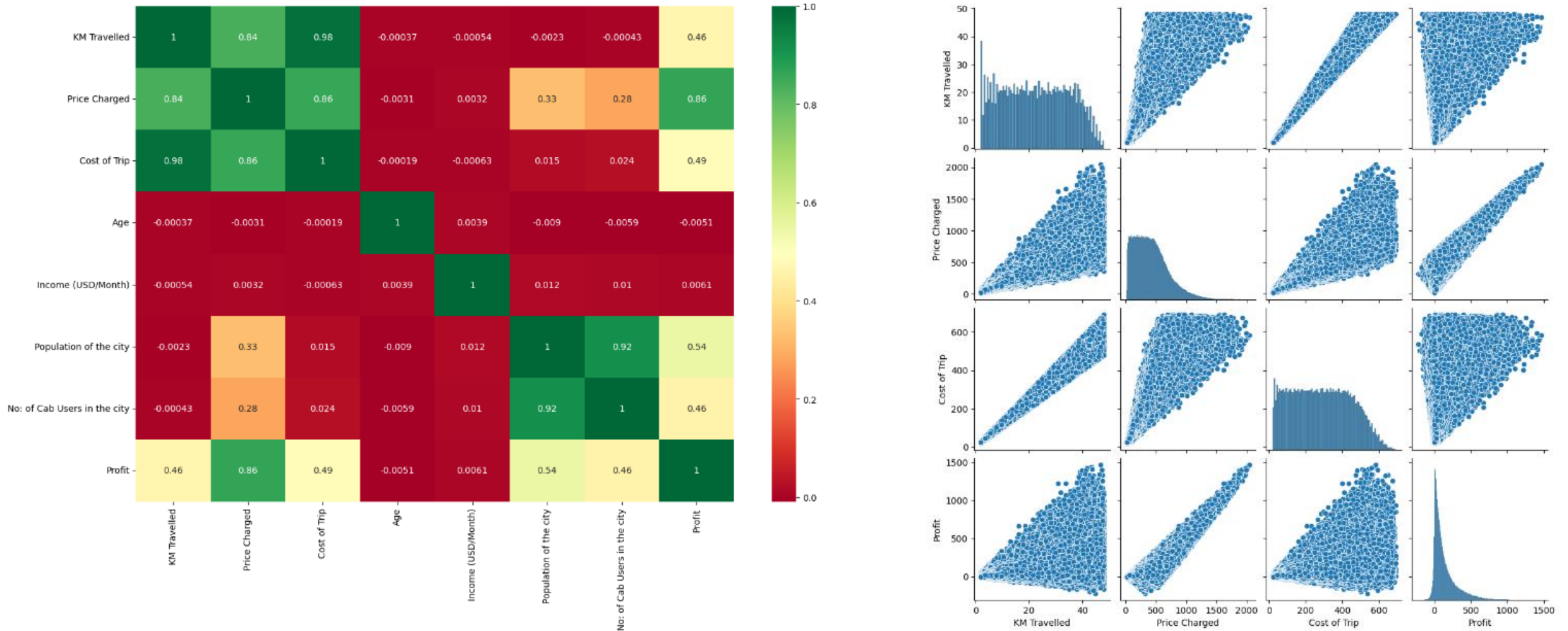
- From the pair plot and heat map, we can get to know that km travelled, price charged, cost of trip and profit has strong correlation between each other
- Also No: of cab Users in the City and Population of the city has direct relationship/variation.

Number of Rides for each Cab Company
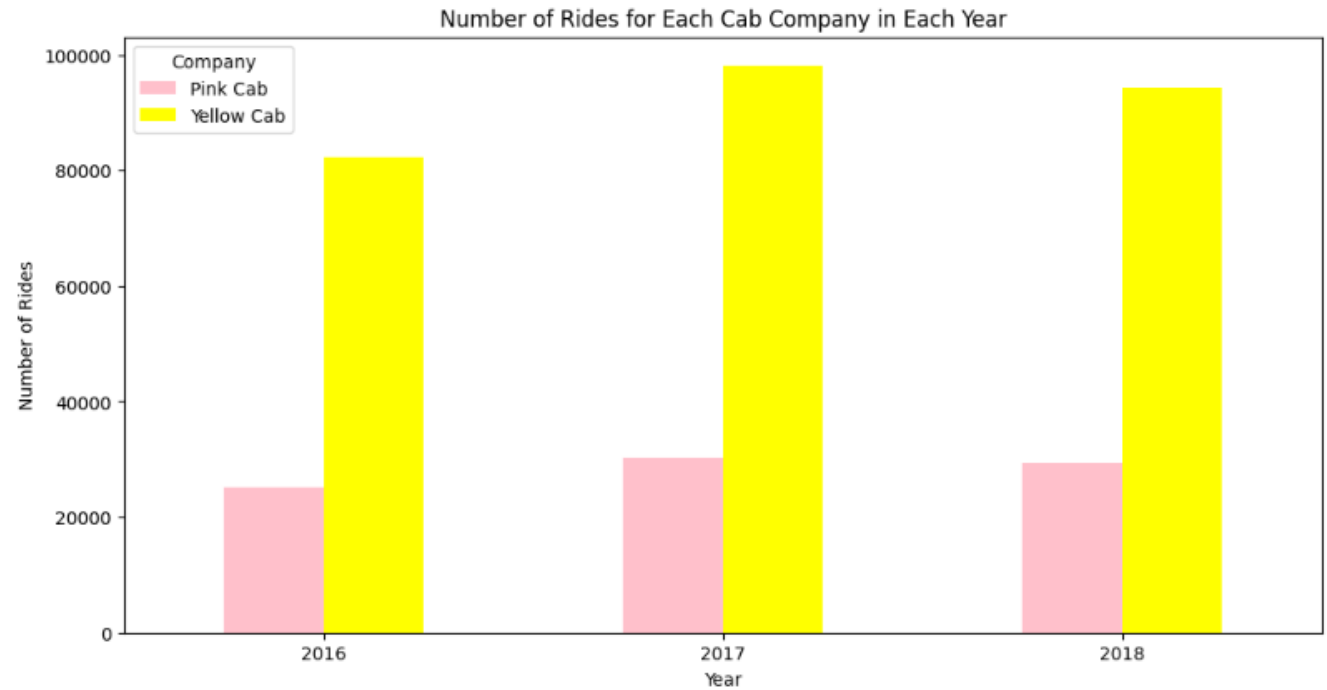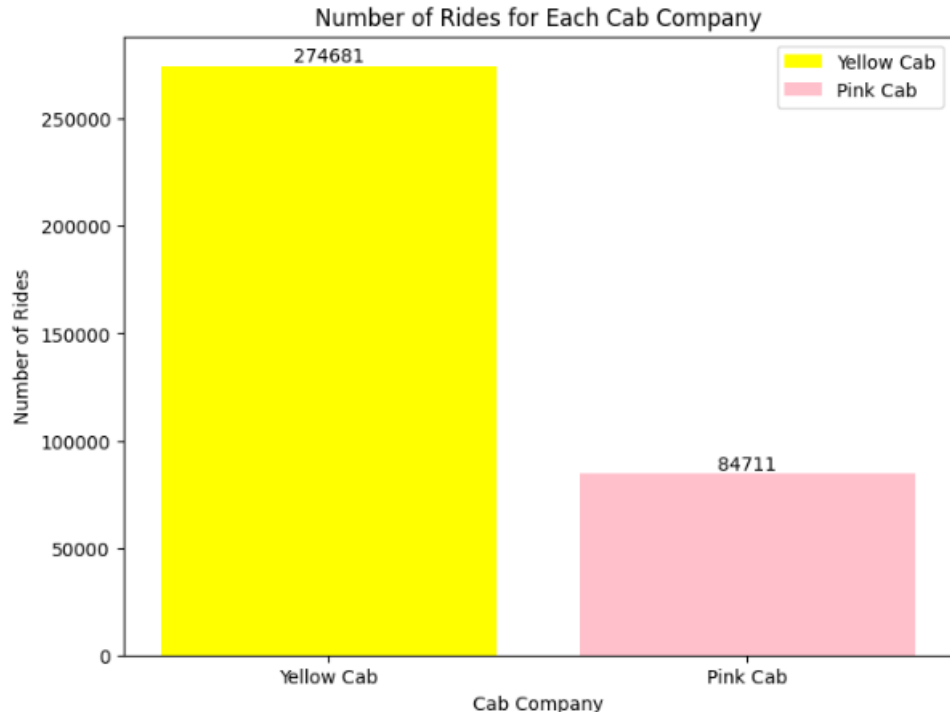
Insights:

- 'Yellow Cab' has maximum number of rides compared to pink cab. That means 'Yellow Cab' has maximum cab users and most of the cab users prefer to choose services offered by 'Yellow Cab'.

- Yellow cab has the highest number of rides in the year 2017.

Number of Rides for each Cab Company City Wise

Number of Rides for Each Cab Company in Each City
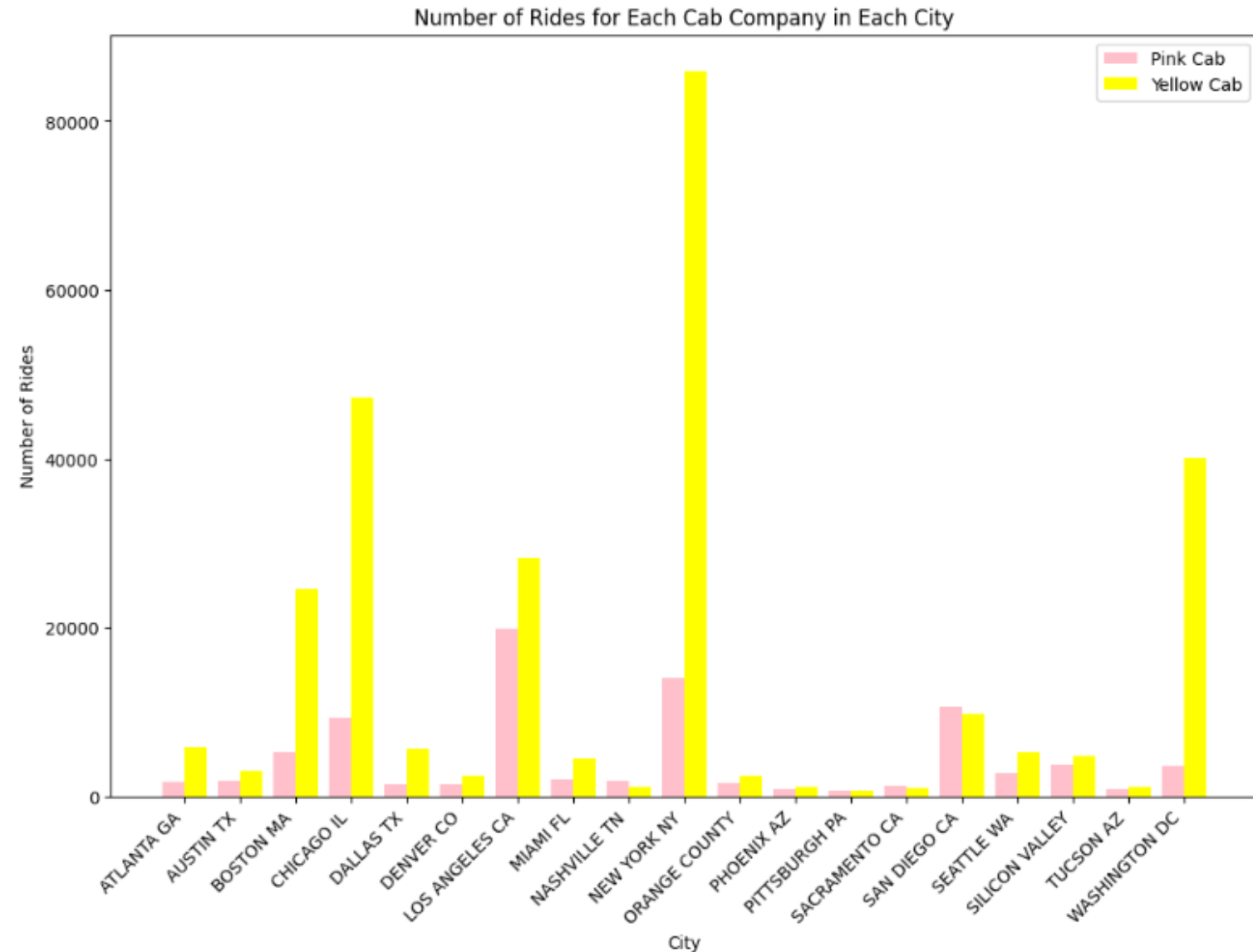
Insights:

•Yellow Cab Has highest number of rides in the cities NewYork NY, Chicago IL, Washington DC respectively

Outliers:

Pink cab has got more number of rides when compared to yellow cabs in cities like:

•Nashville TN(Pink cab:1841, Yellow cab:1169)

•Pittsburgh PA(Pink cab:682, Yellow cab: 631)

•Sacramento CA(Pink cab:1134, Yellow cab:1033)

•San Deigo CA(Pink cab:10672, Yellow cab:9816)

These exceptions we can ignore when compared to the yellow cab rides in other major cities. But we need to focus on these cities to increase the rides and customers for yellow cab. We need to give promotional offers to yellow cab users in these cities

Number of Customers for Each Cab Company

Trend of Customers for Each Cab Company - Year Wise
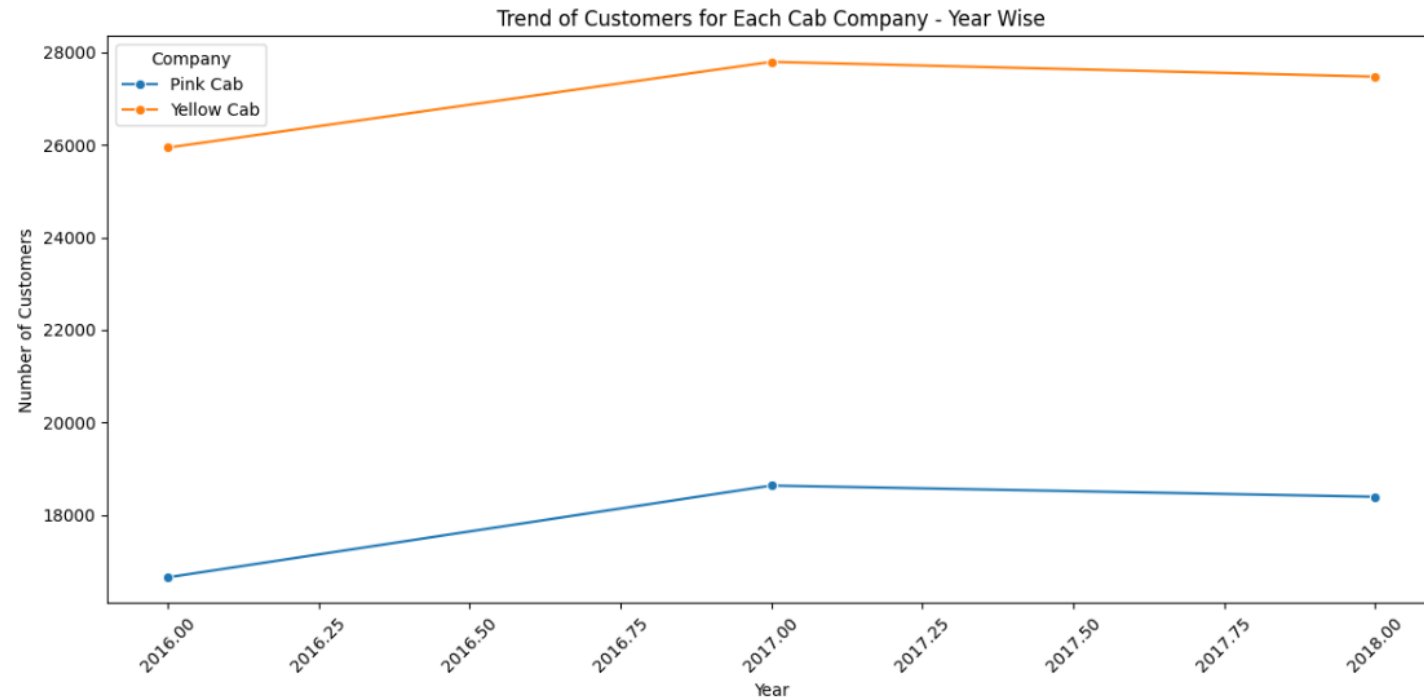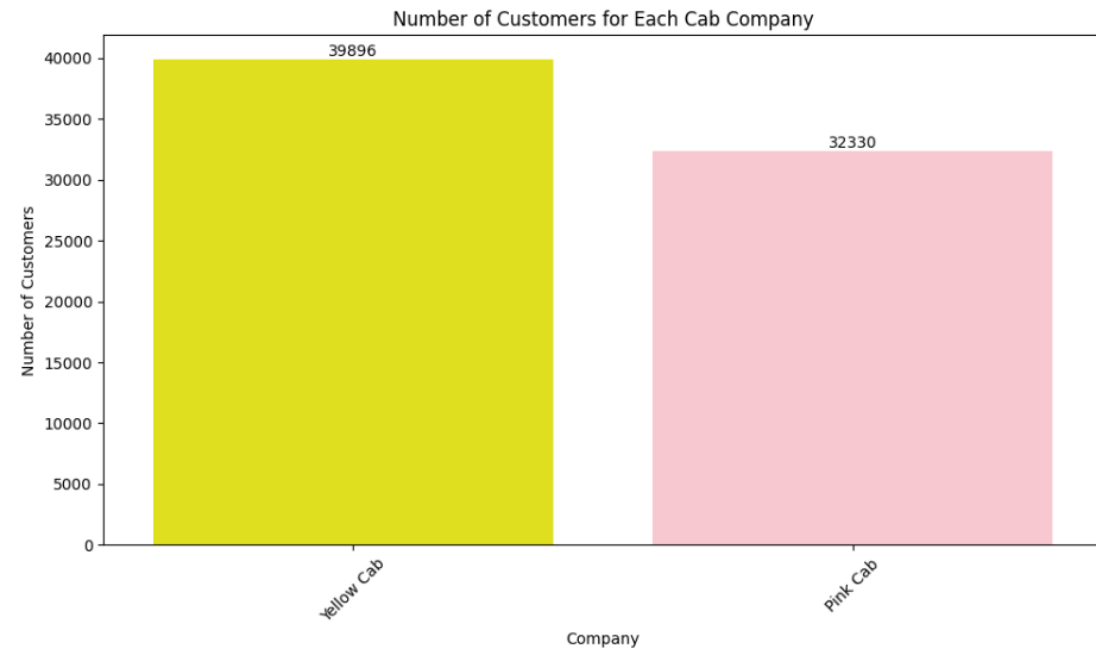
Insights:

•Yellow Cab has the highest number of customers than pink cab.
• Still there is high competition among both cab companies in some cities. Yellow cab can increase more customers in those cities by bringing promotional offers and discounts to new customers.
•Both the cab companies have more number of customers in the year 2017 and after that there is a slight decrease in them. This can be due to the increase in price charged for the cost of trip. This should be considered immediately otherwise there is chances of losing more customers.

Number of Customers for each Cab Company – City Wise

Number of Customers for Each Cab Company - City Wise
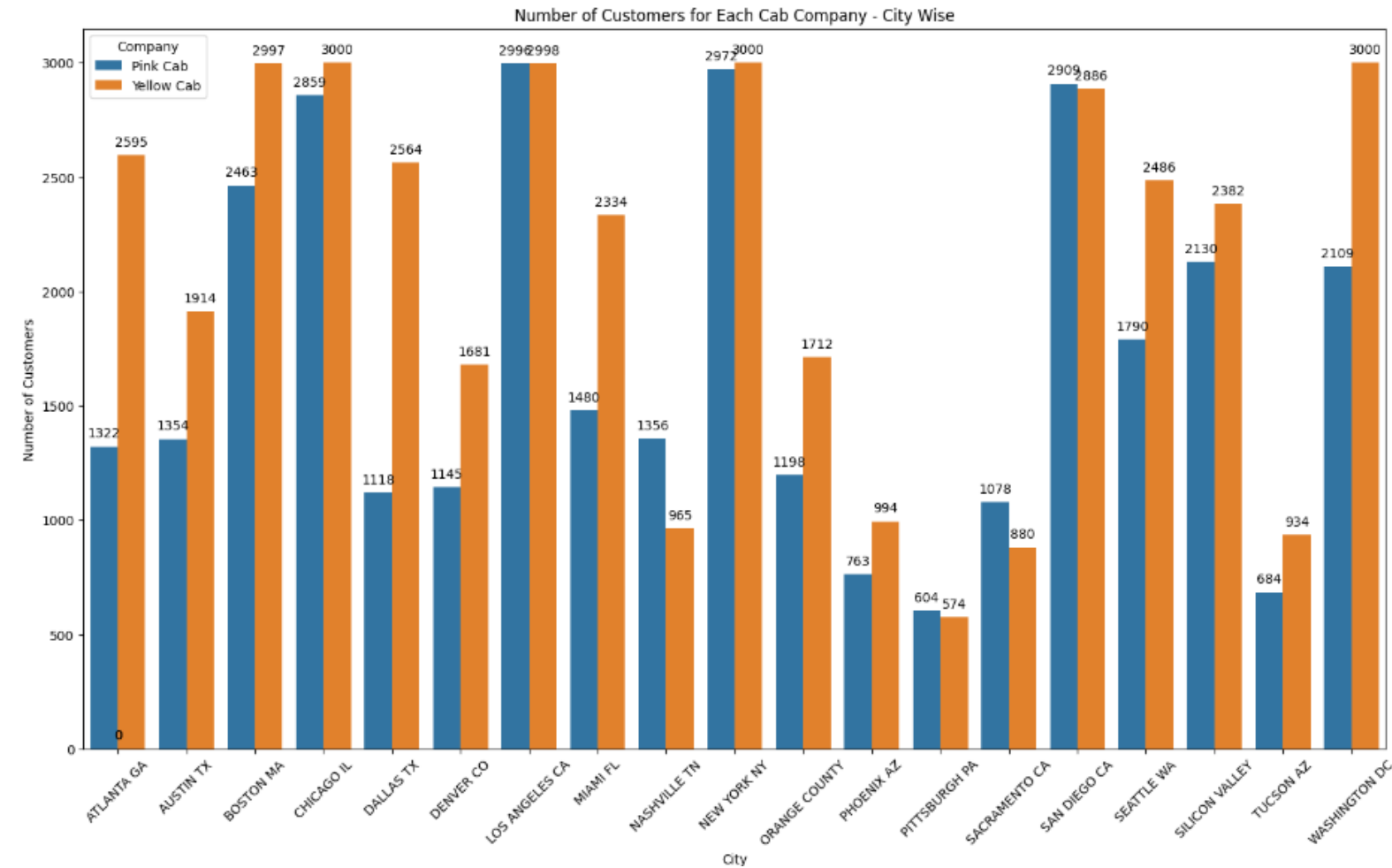
Insights:

- Yellow Cab users are high in most of the cities.
- In the cities like Los Angeles CO, New York NY, Pitturburgh PA, SAN DIEGO CA, there is a high competition among Yellow Cab and Pink Cab. We need to offer dicounts and promotional offers to customers, so that we can increase the rides of Yellow cabs.
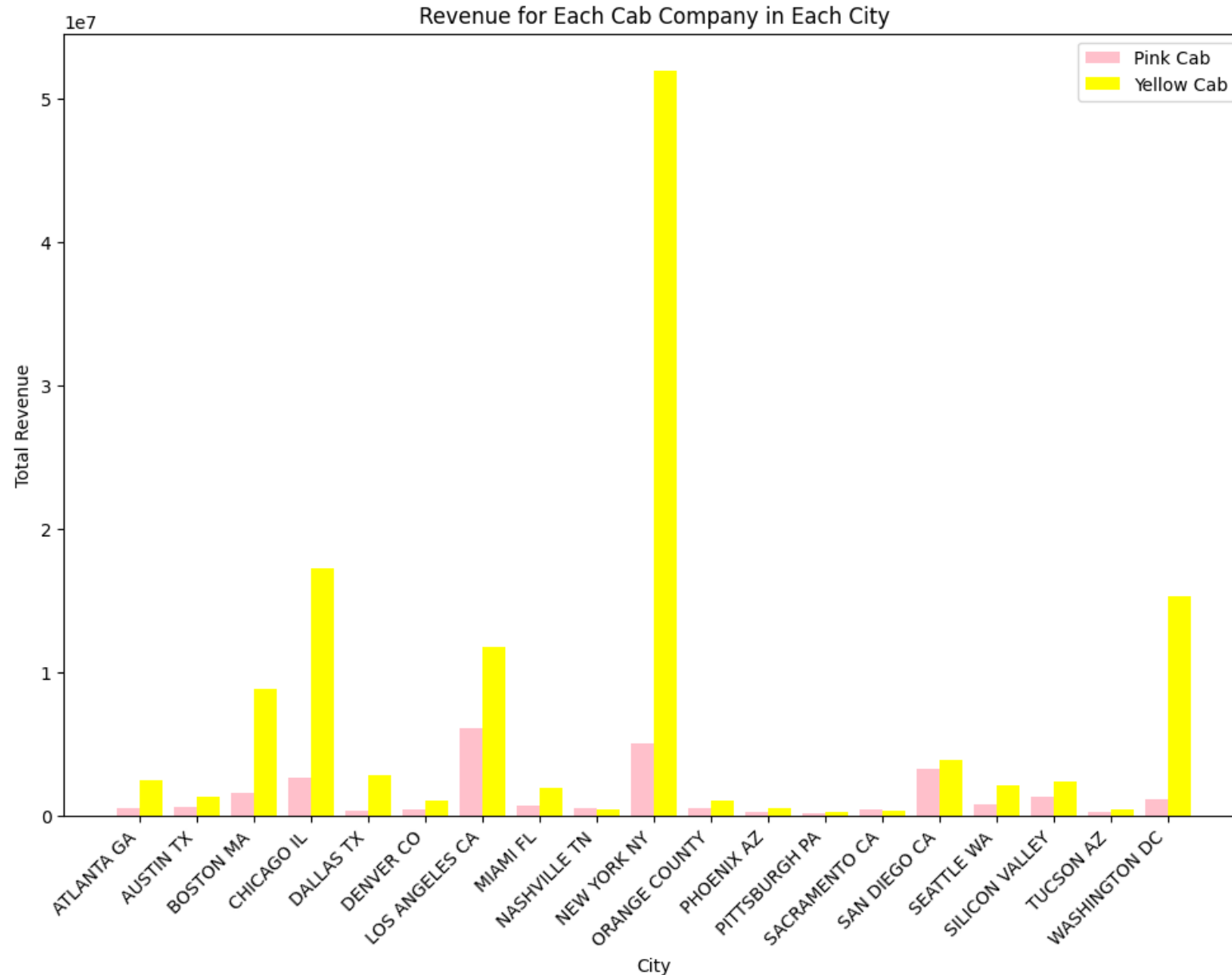
Total Revenue for each Cab Company
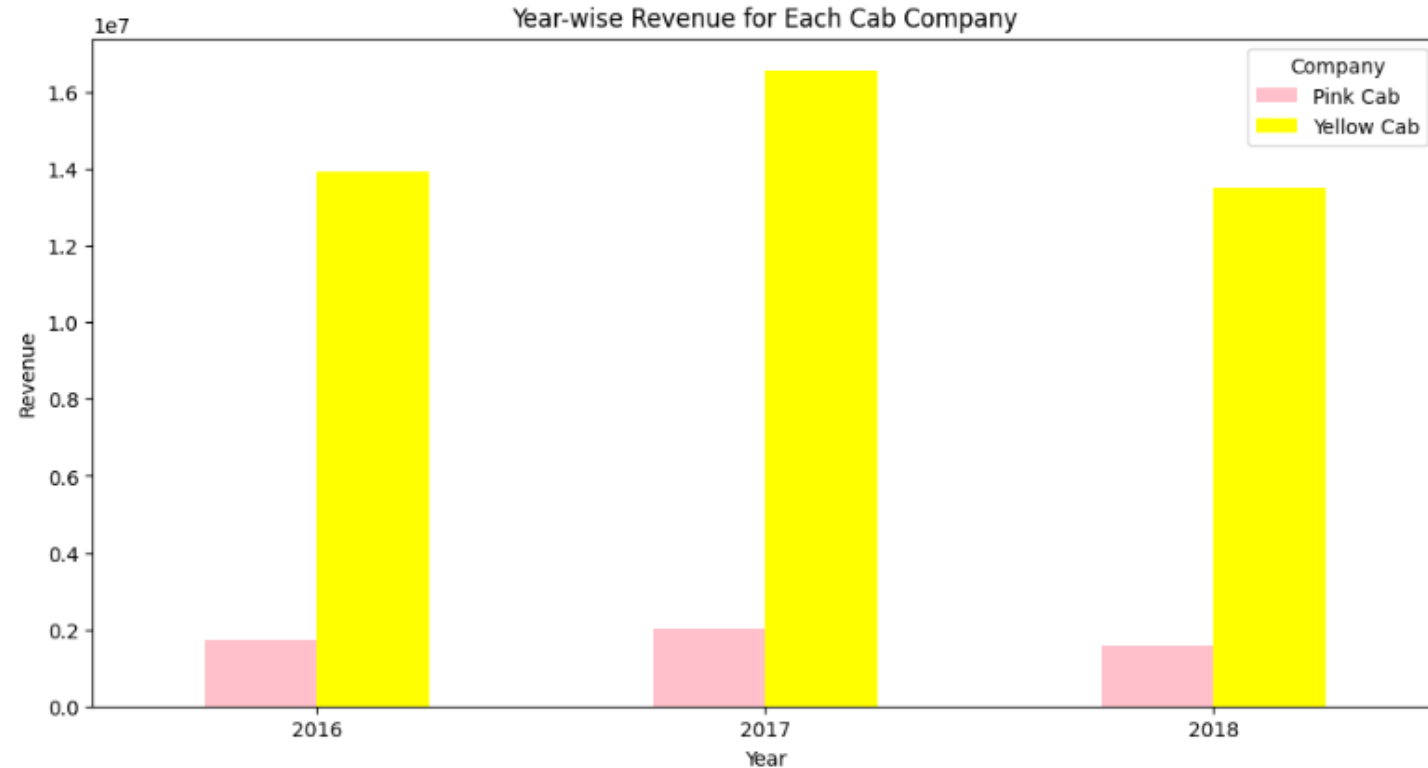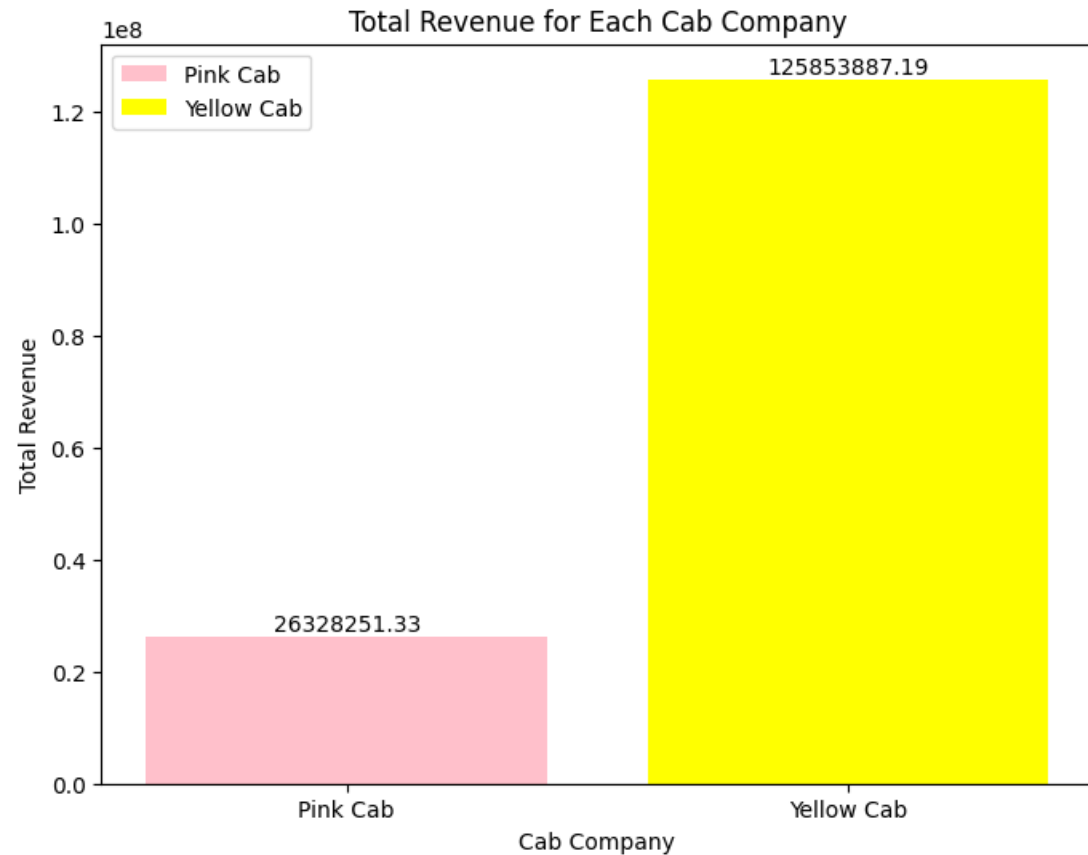
Insights:
- Yellow Cab has earned more revenue when compared to Pink Cab in each Year.
- Yellow Cab has earned more revenue in the year 2017.

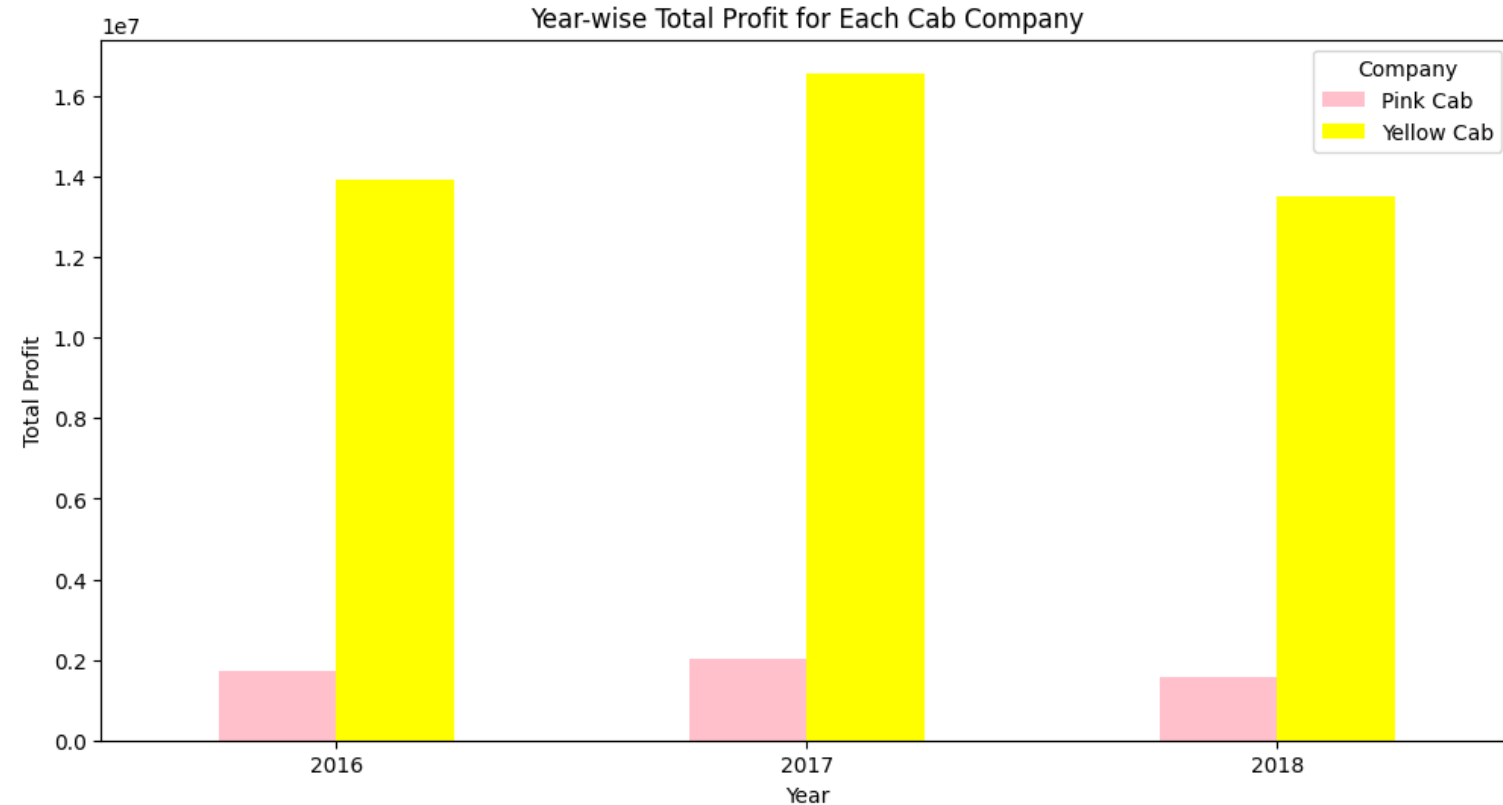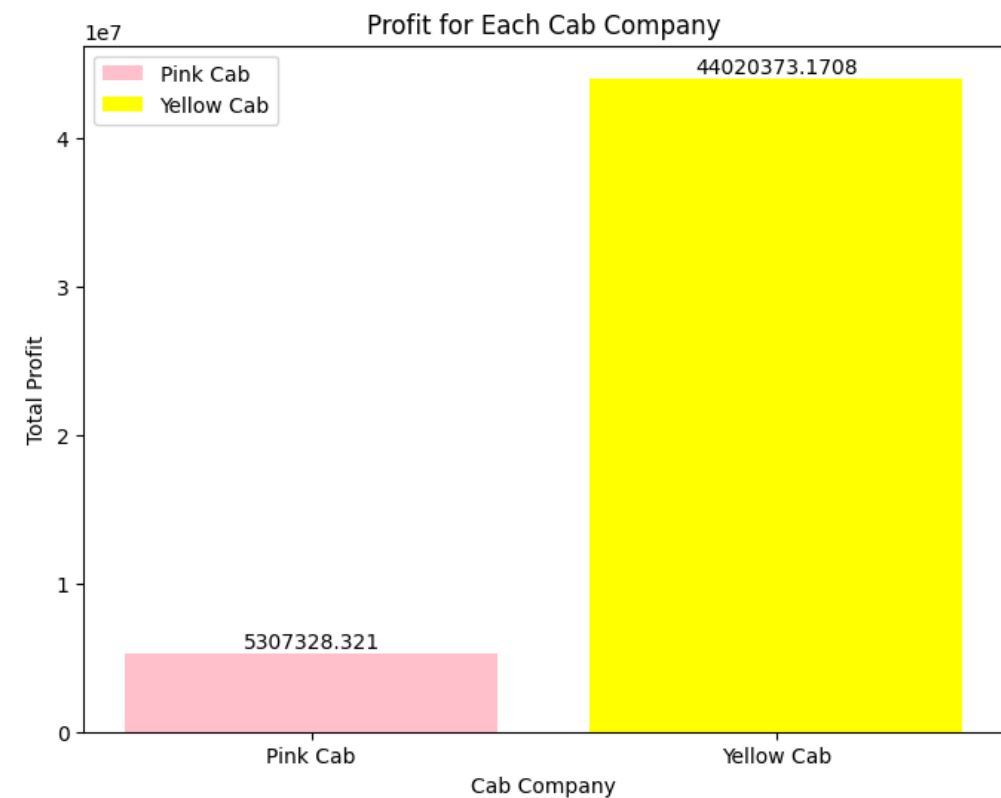# Total Revenue for each Cab Company City Wise

Revenue for Each Cab Company in Each City

Outliers:

Total Revenue generated by Pink cab is more when compared to yellow cab in cities like
•Nashville TN
•Pittsburgh PA
•Sacramento CA

These exceptions we can ignore when compared to the total revenue generated by yellow cab in other major cities.
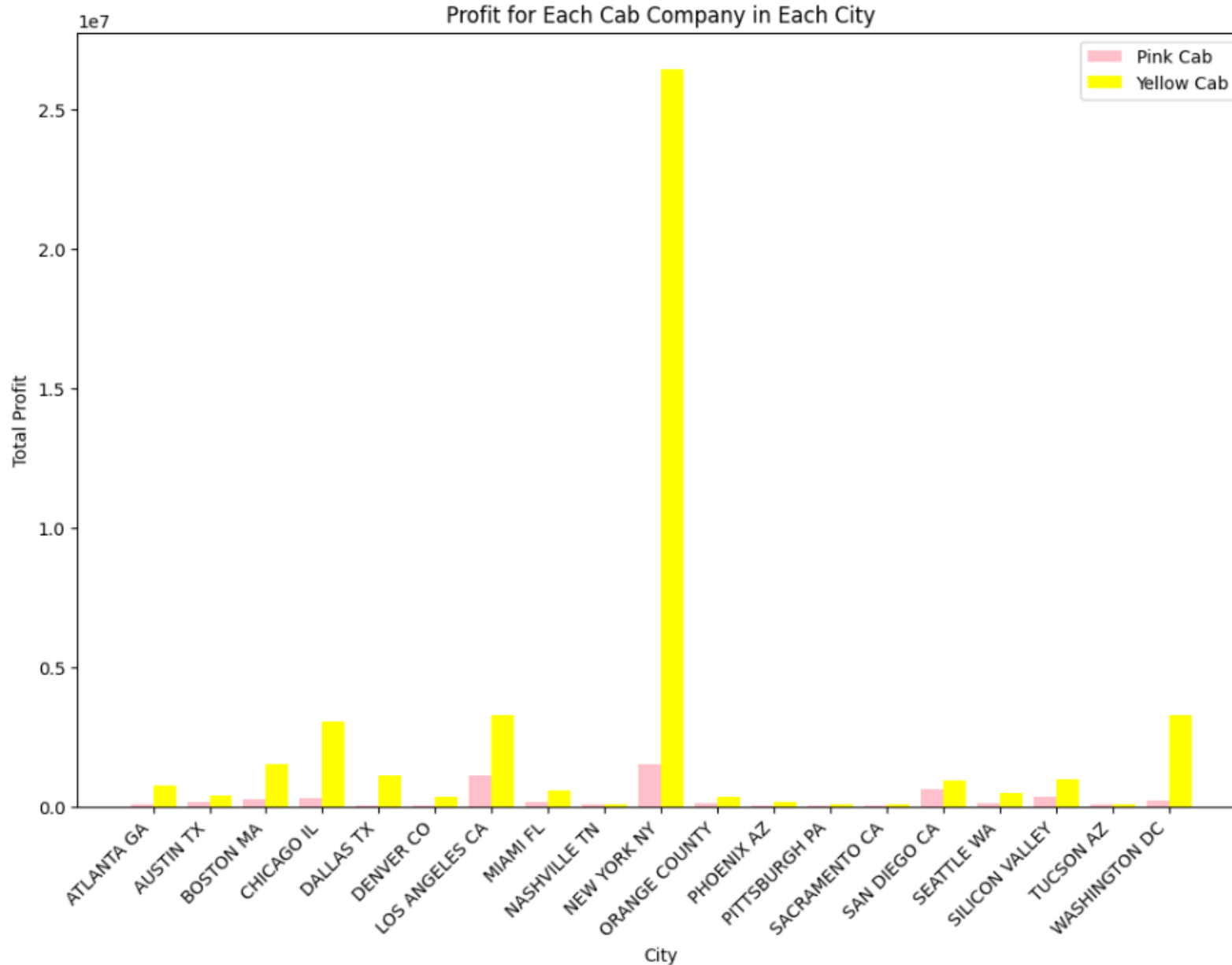
Total Profit for each Cab Company

Insights:
- Yellow Cab has earned more profit when compared to Pink Cab in each Year.
- Yellow Cab has earned more profit in the year 2017.

Total Profit for each Cab Company City Wise

Profit for Each Cab Company in Each City

Insights

- Yellow cab has earned more profit in the city New York NY. Yellow cab has got highest number of rides and Kms travelled in this city.
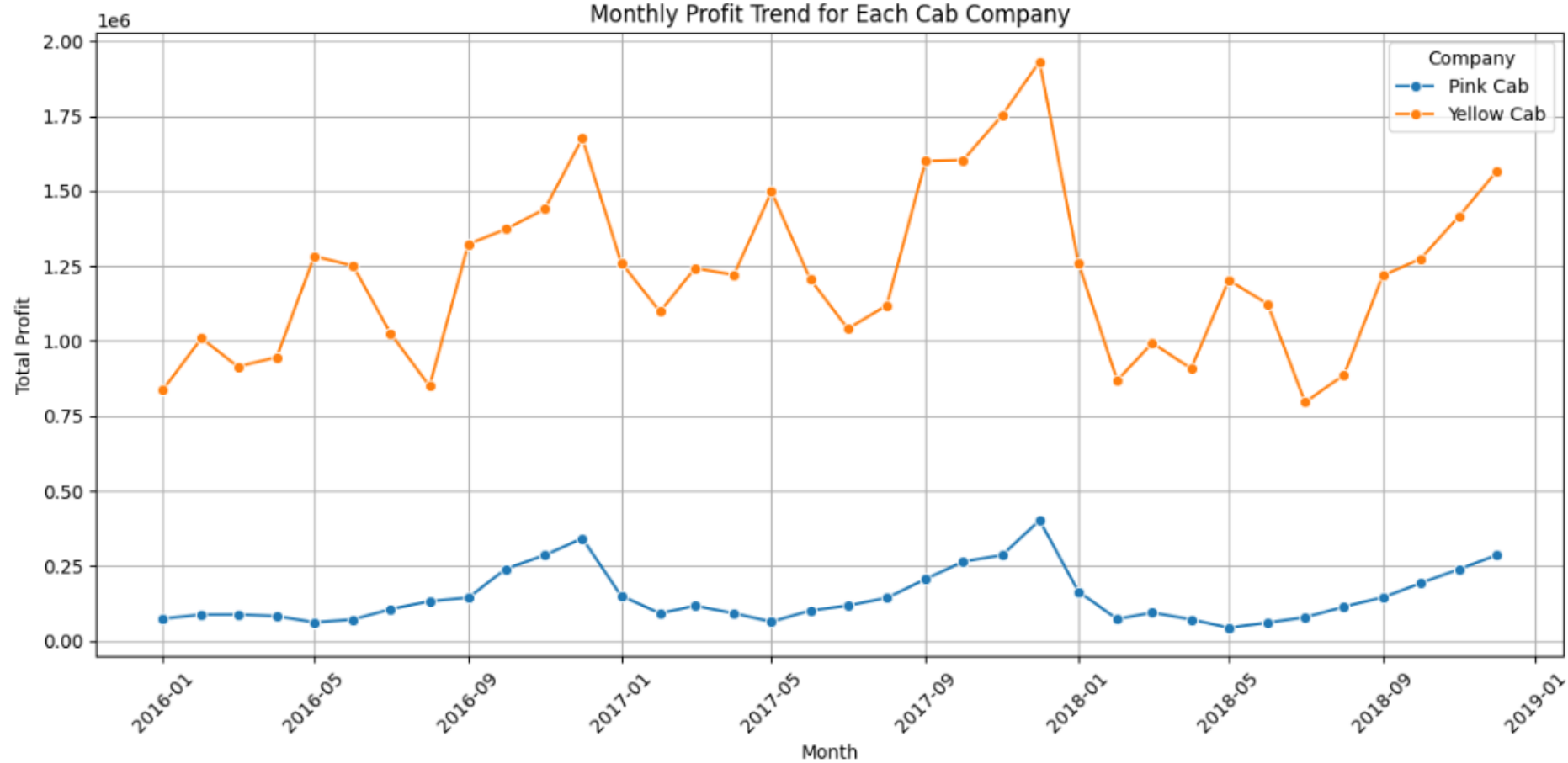
Outliers:

Total Profit earned by Pink cab is more when compared to yellow cab in cities like
•Nashville TN
•Tucson AZ

These exceptions we can ignore when compared to the total profit earned by yellow cab in other major cities.

Trend Line on Monthly profit for each Cab Company

Monthly Profit Trend for Each Cab Company

Insights:

•Yellow cab has earned more profit than Pink Cab.
•The profit level of Yellow cab has reached the maximum by the end of the year 2017 and it got suddenly reduced in the beginning of 2018.
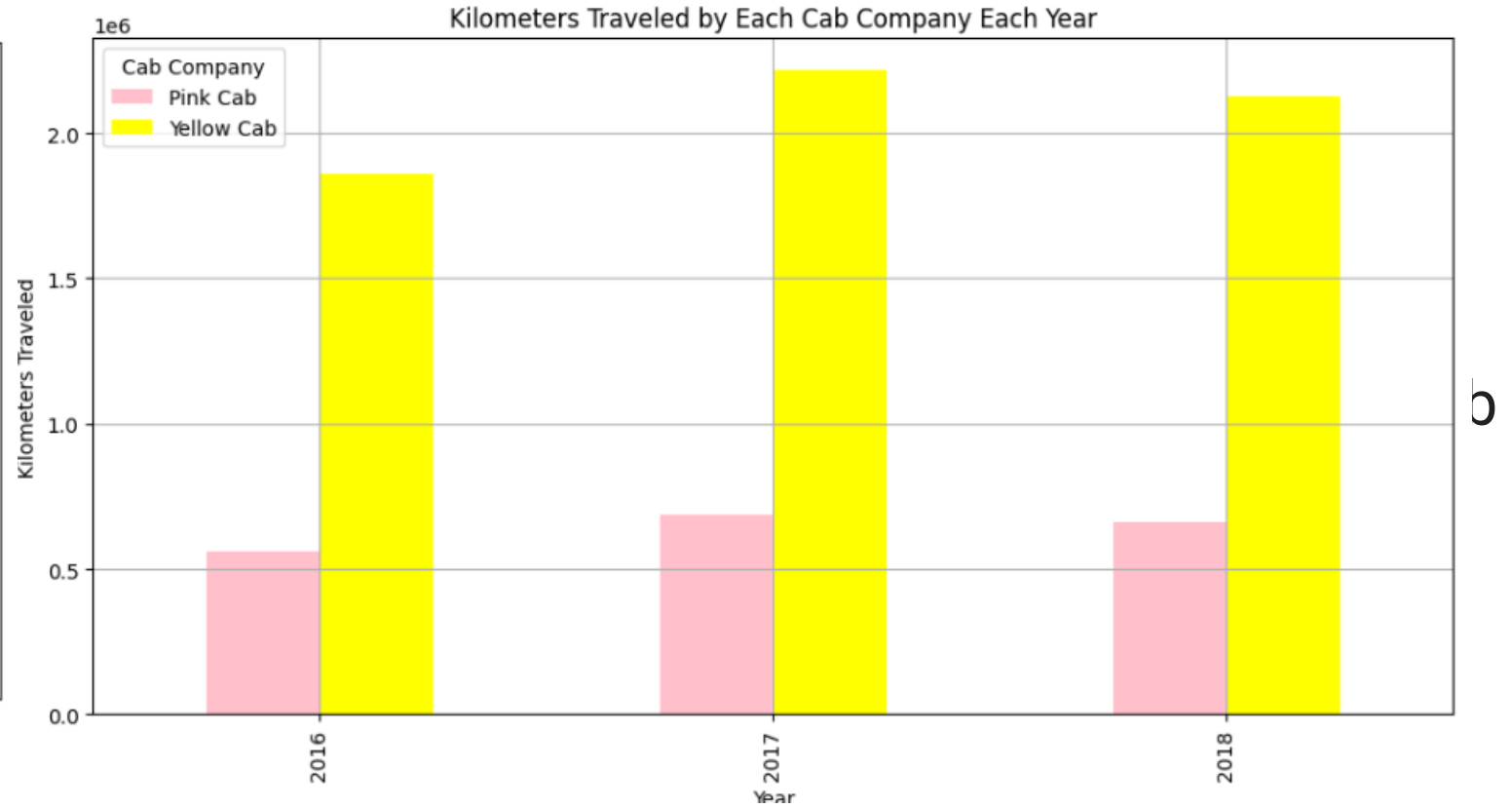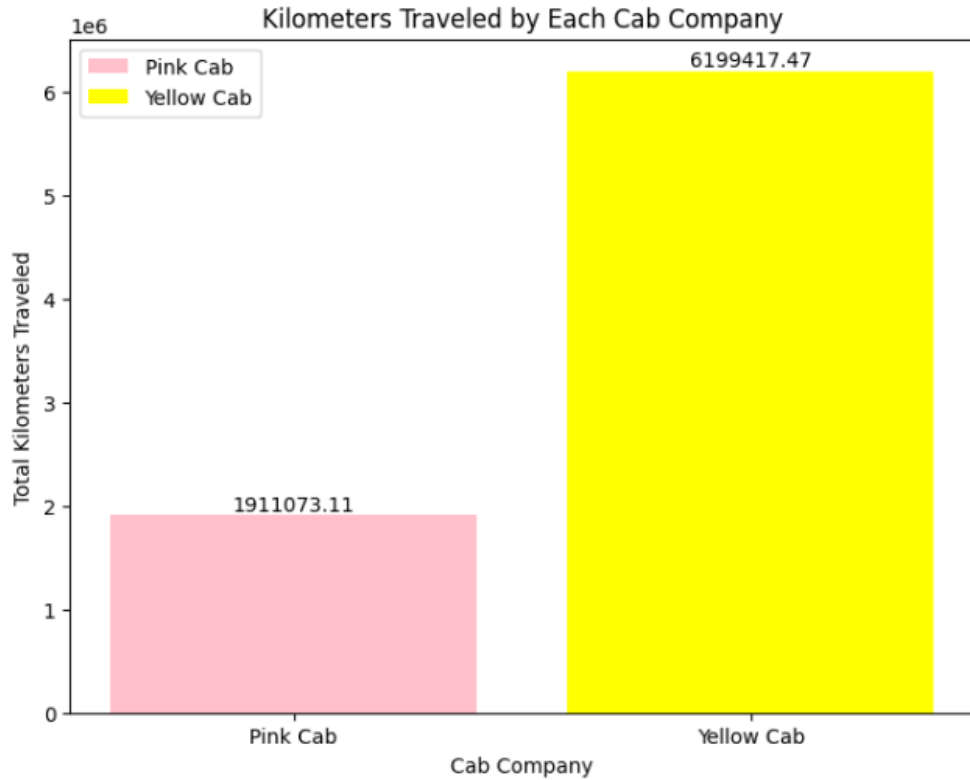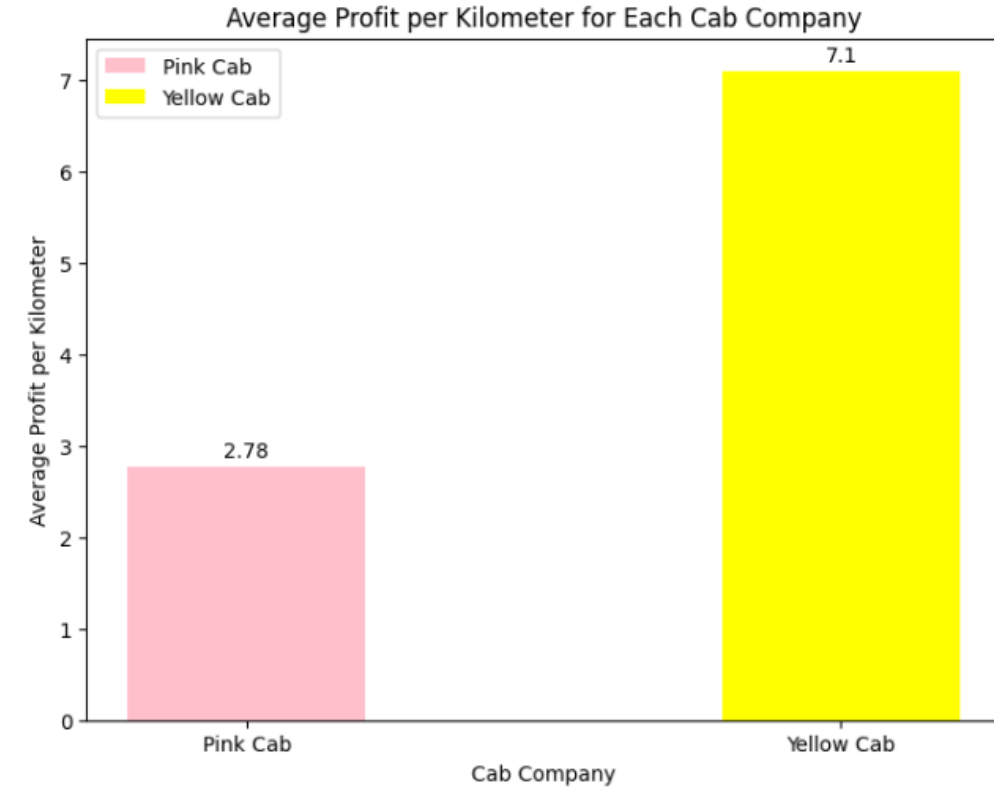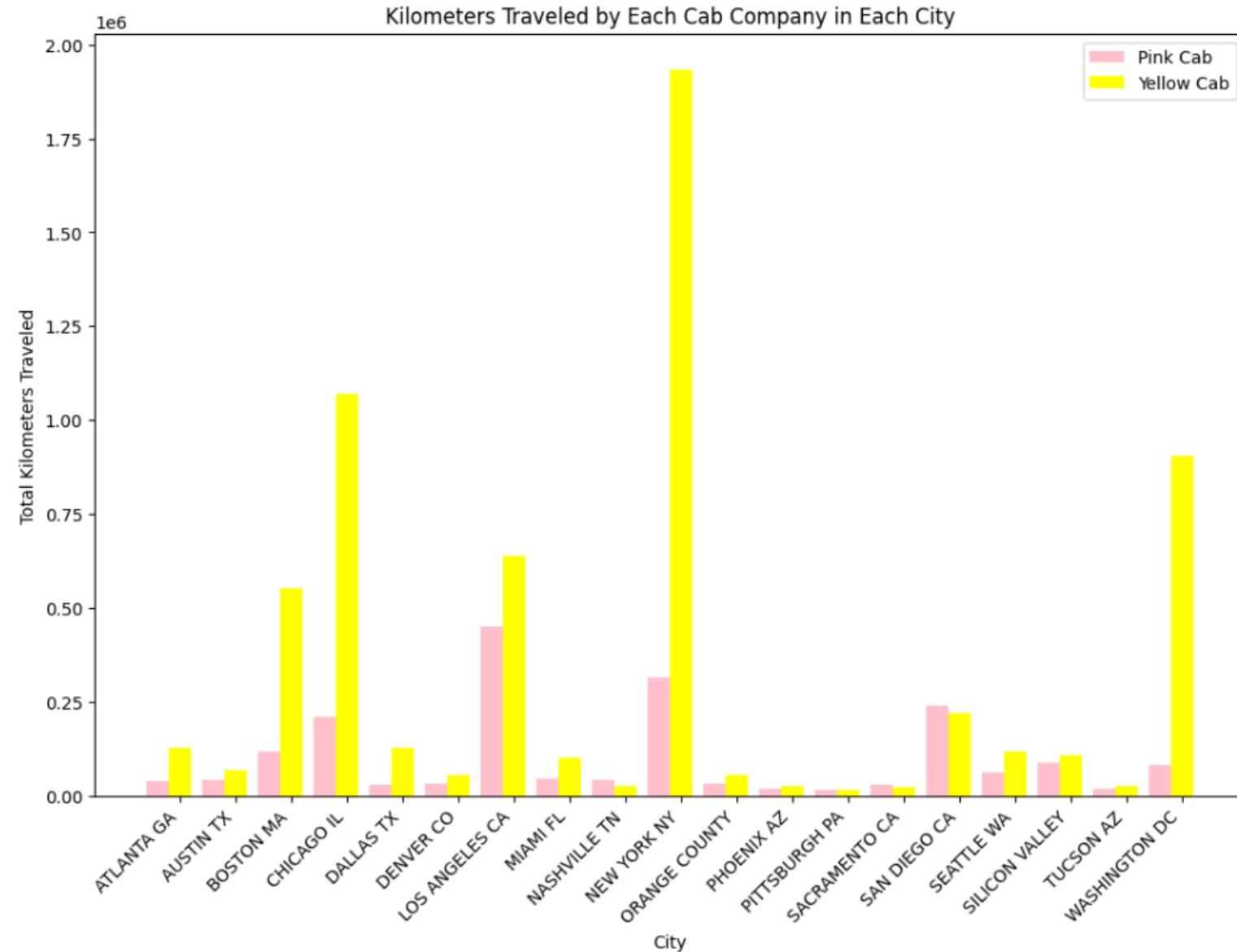
Insights:

- Kilometres travelled by 'Yellow Cabs' are more when compared to the 'Pink cab'.

Average Profit per Kilometre

Insights:

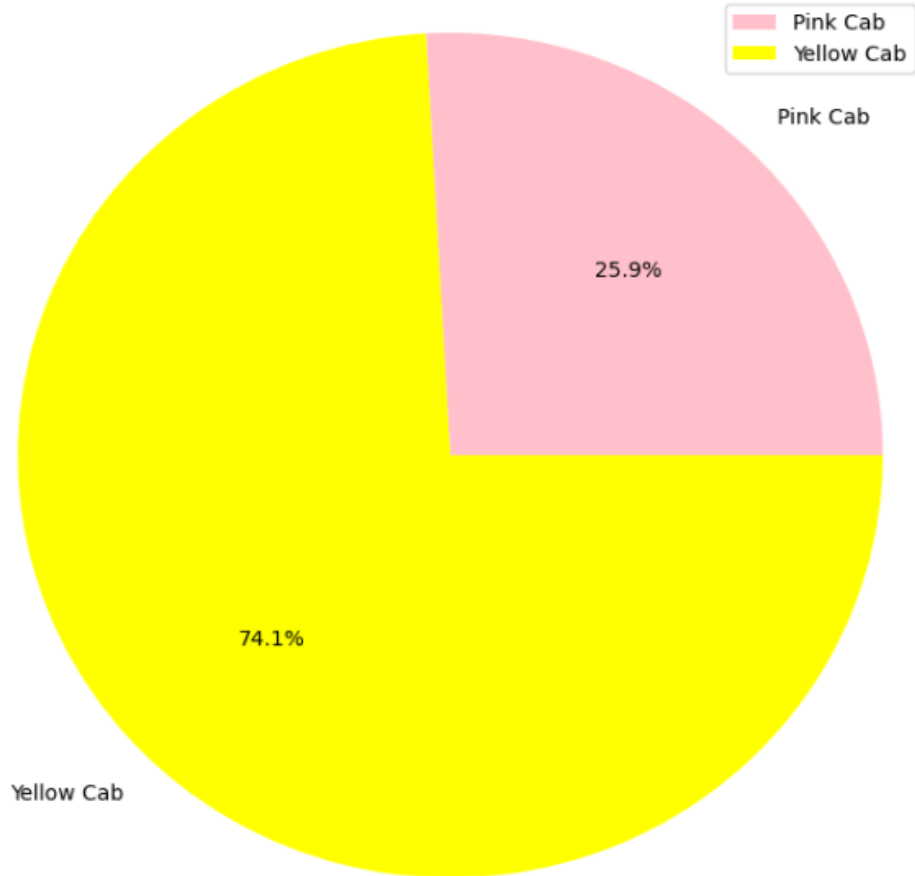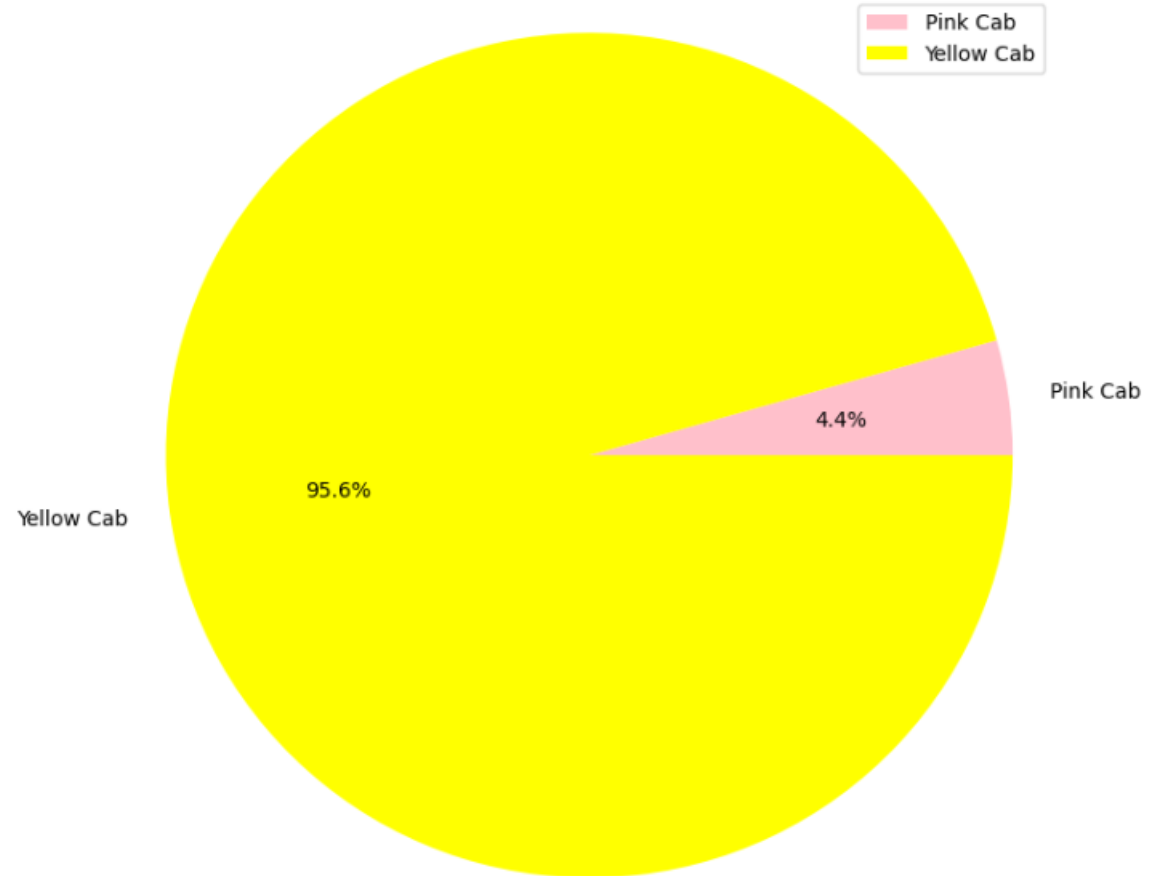- Average Profit per Kilometre is greater for Yellow Cab Company

Customer Retention for Yellow Cab and Pink Cab (At Least 5 Rides)

Customer Retention for Yellow Cab and Pink Cab (At Least 10 Rides)

Insights:

- Customer Retention is more for choosing services offered by Yellow Cab when they have undergone a journey with it.

EDA Summary & Recommendation

- Yellow Cab' has more number of rides compared to 'Pink Cab'. That means most of the cab users prefer to choose services offered by 'Yellow Cab' and it has maximum cab users.

- Total Revenue generated by 'Yellow Cabs' are more when compared to the 'Pink cab'.

- Total Profit earned by 'Yellow Cabs' are more when compared to the 'Pink cab'.

- Kms Travelled by 'Yellow Cabs' are more when compared to the 'Pink cab'.

- Profit per Km of 'Yellow Cabs' are more when compared to the 'Pink cab'.

- Customers prefer to choose 'Yellow Cabs' during their next journey once they have undergone a ride with it. Customer Retention is more.

Recommendation

I recommend that XYZ company should invest in **'Yellow Cab'** company to gain more profit

In the cities with the exceptions, we need to focus on services offered to customers to get more rides for the Yellow cabs.

Thank You