



Data Glacier

Your Deep Learning Partner

G2M Study for Cab Investment

A Detailed Study of the US Cab Market

14.4.2023

Agenda

Problem Statement

Executive Summary

Approach

EDA

Detailed Analysis

Recommendations

Problem Statement

- XYZ is a private firm in the US. Due to remarkable growth in the Cab Industry in the last few years and multiple key players in the market, it is planning for an investment in the Cab industry.
- The Objective of this study is to provide insights into the Cab market that would help XYZ to find the right company to invest in.

Executive Summary

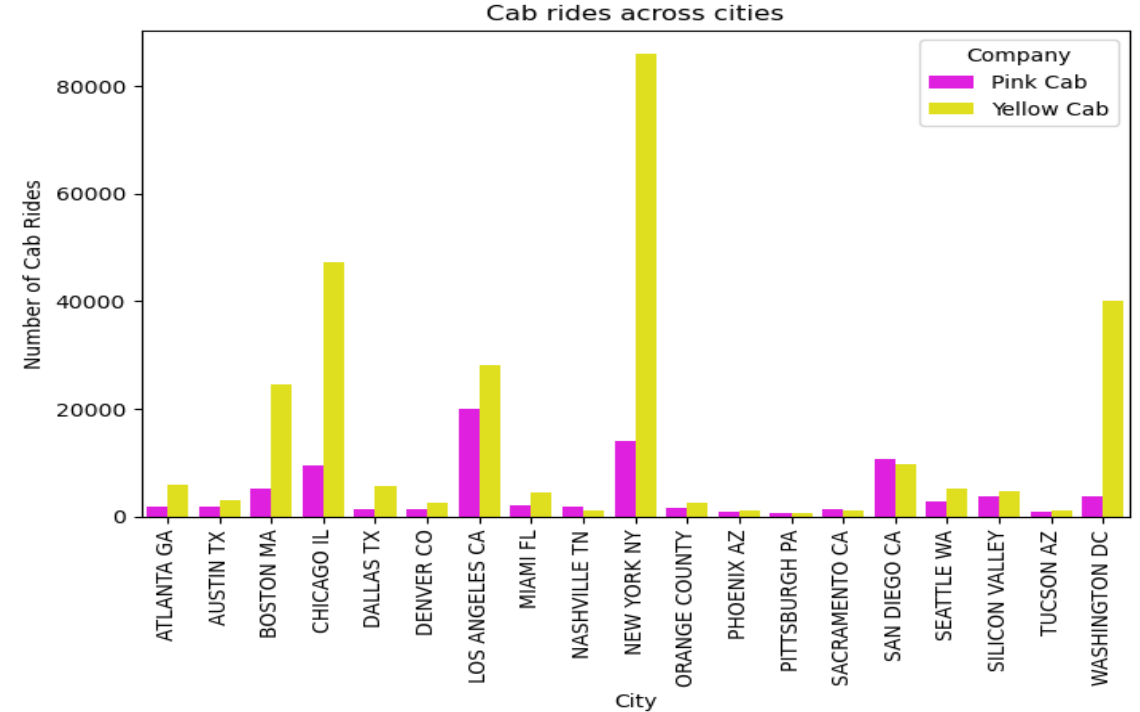
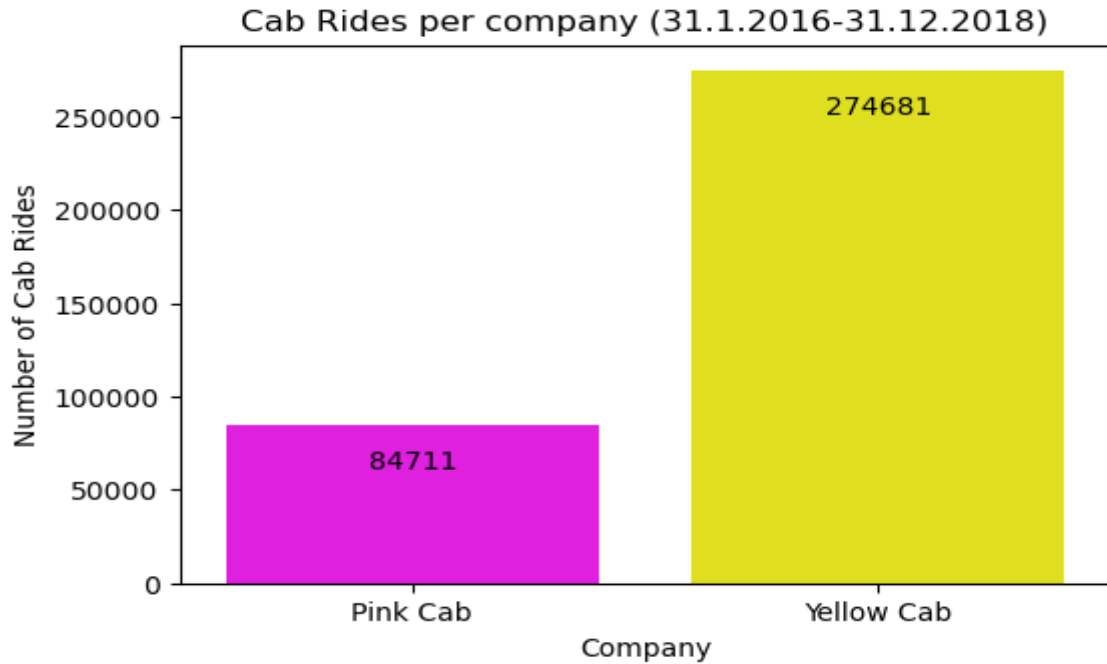
- This study was undertaken to derive insights into the US cab market based on the data provided. The data includes information about the cab rides undertaken by the two companies present between the period 31.1.2016 – 31.12.2018.
- The first part of the study aims to understand the provided data by creating various visuals that help in identifying patterns that would help during the analysis.
- The second part delves deeper into the data trying to answer why XYZ should invest in a certain cab company.

Approach

- There were 4 data files provided. Please refer to the Data Intake Report for a detailed description of the datasets.
 1. Cab_data.csv
 2. City_data.csv
 3. Customer_data.csv
 4. Transaction_data.csv
- The data provided was cleaned and processed to extract more information (years and profit).
- An initial summary of all data was documented and can be found in the DI report.
- All steps involved in conducting the analysis have been noted in the notebook along with the code.

Exploratory Data Analysis

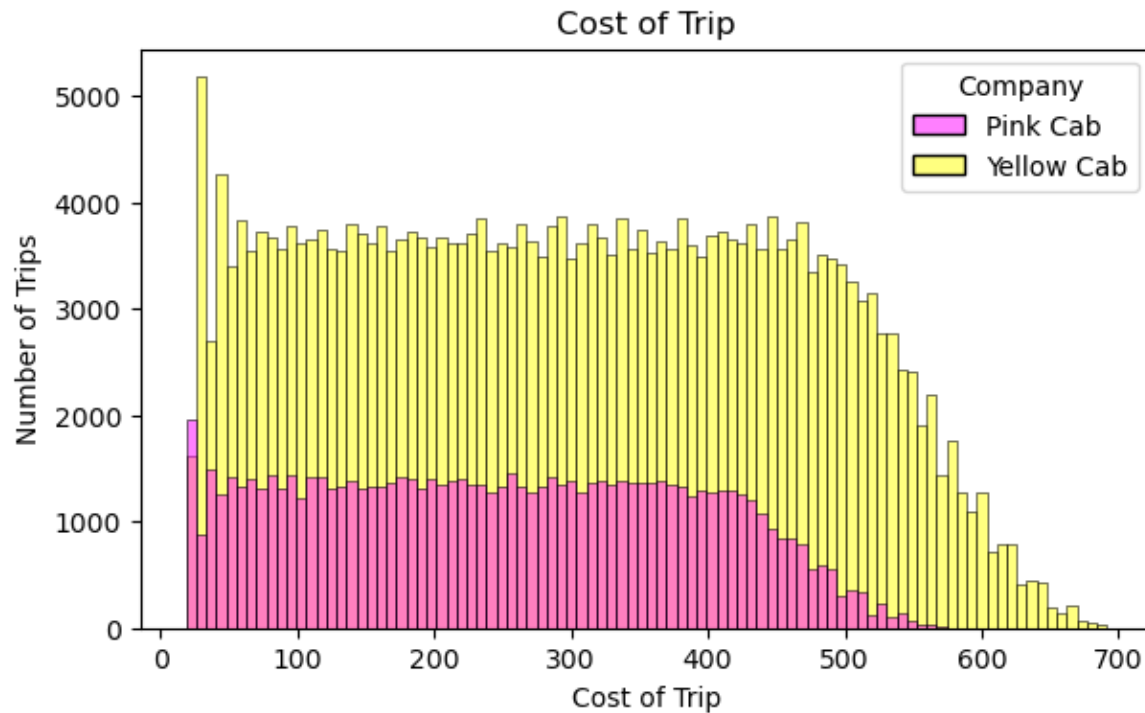
Understanding the datasets: cab data



- From both figures we can understand that the Yellow cab company has a much larger reach than the Pink cab company. The Yellow Cab company dominates in most cities.

Exploratory Data Analysis

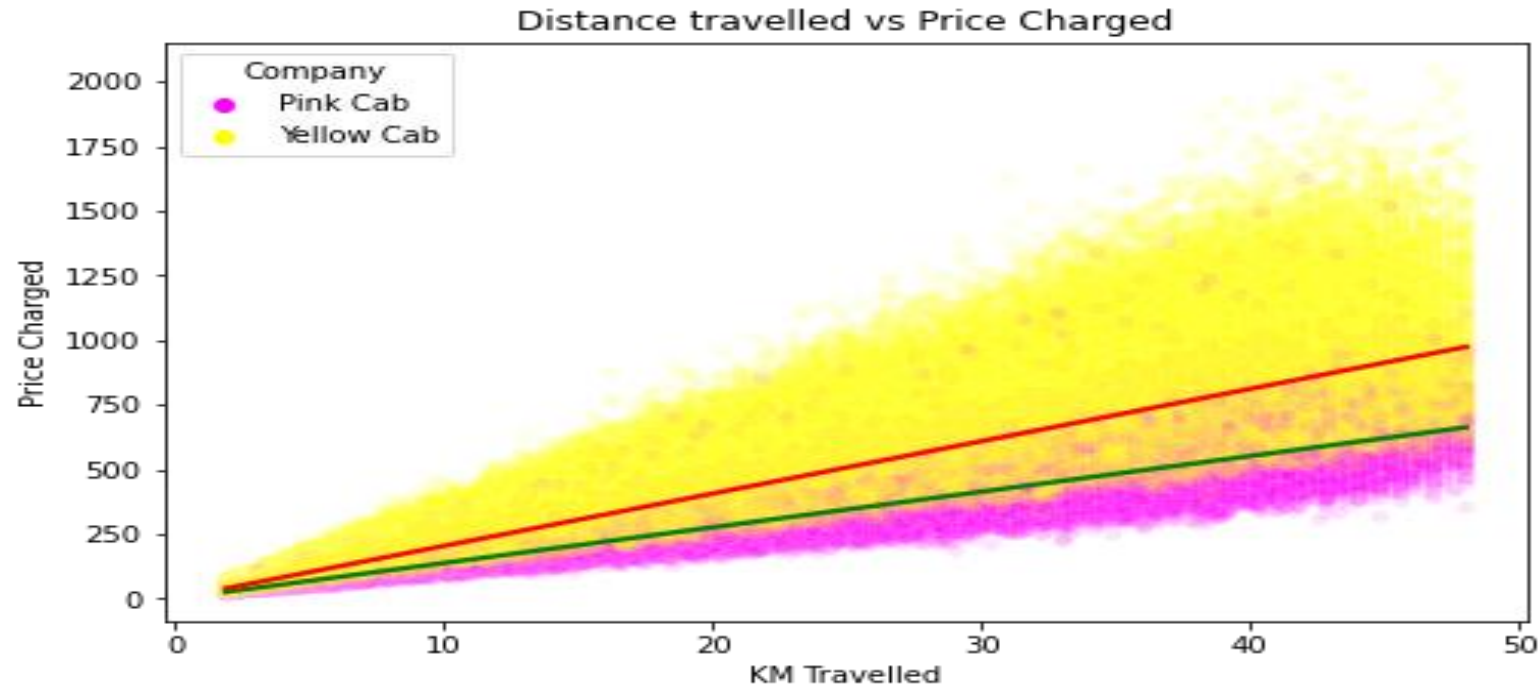
Understanding the datasets: cab data



- The cost of a trip and the prices charged by the Yellow Cab company seem to be higher. We will analyse this in a later section.

Exploratory Data Analysis

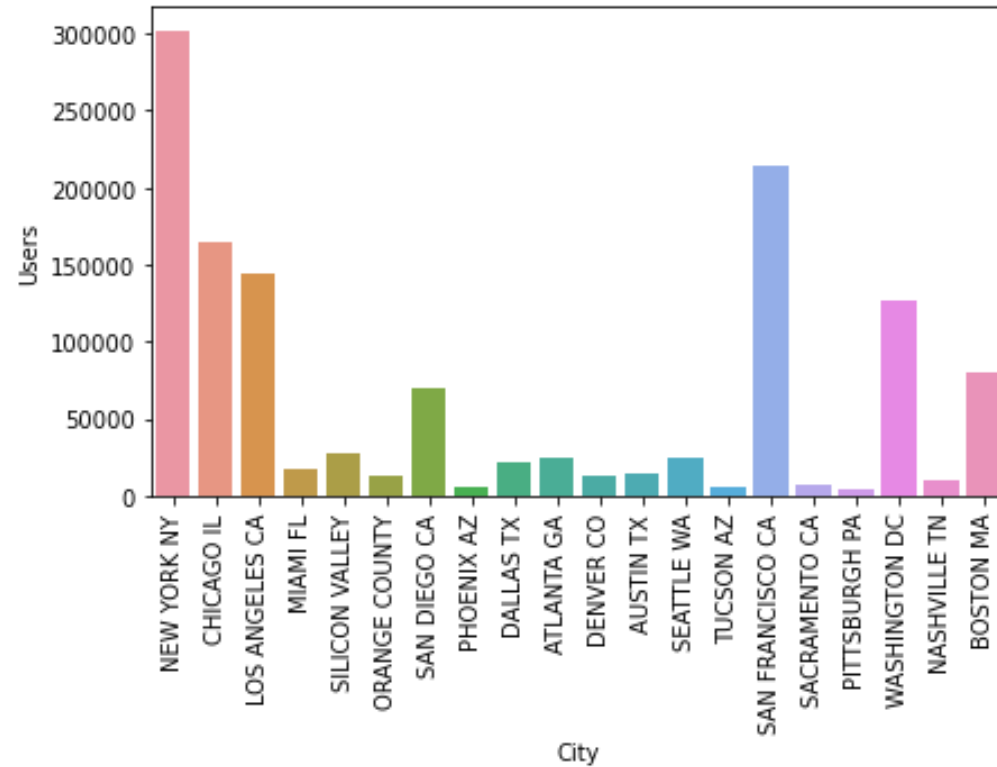
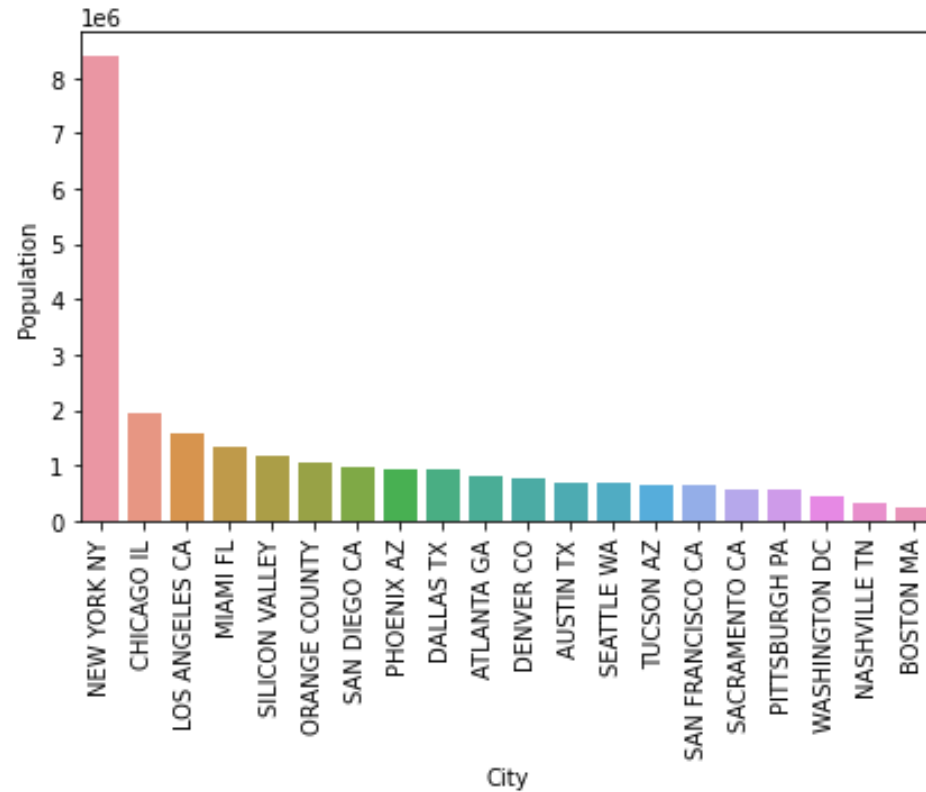
Understanding the datasets: cab data



- As seen in the previous slide, we can see from the regression lines (Red for the Yellow cab and green for the Pink Cab) that the Yellow Cab tends to charge more per KM travelled. This may be due to the fact that they offer a more pleasant user experience.

Exploratory Data Analysis

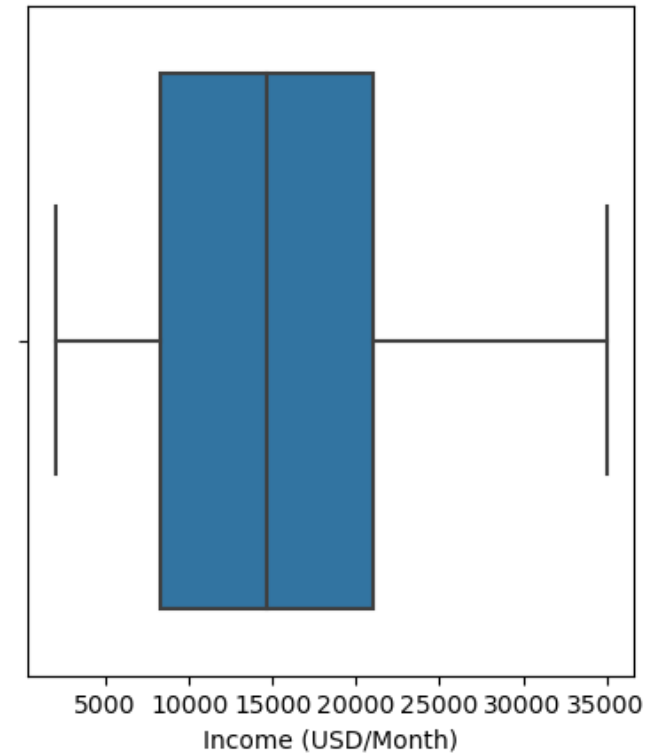
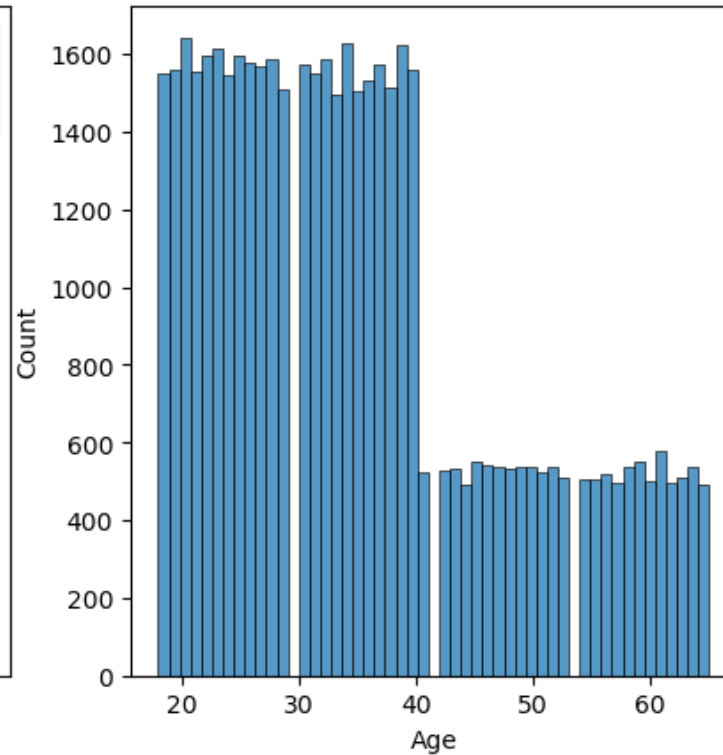
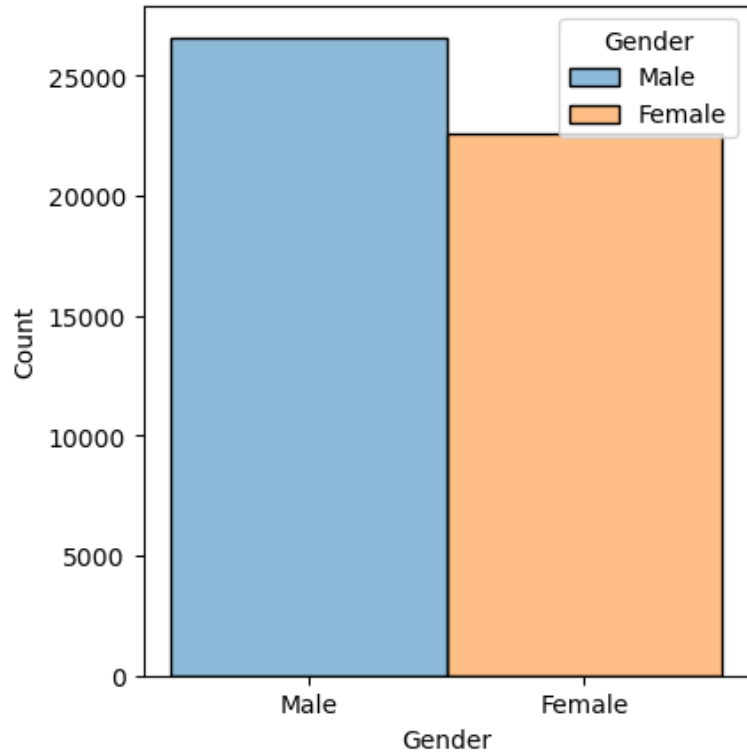
Understanding the datasets: City data



- From the above graphs we can understand that only a small portion of the population are cab users.

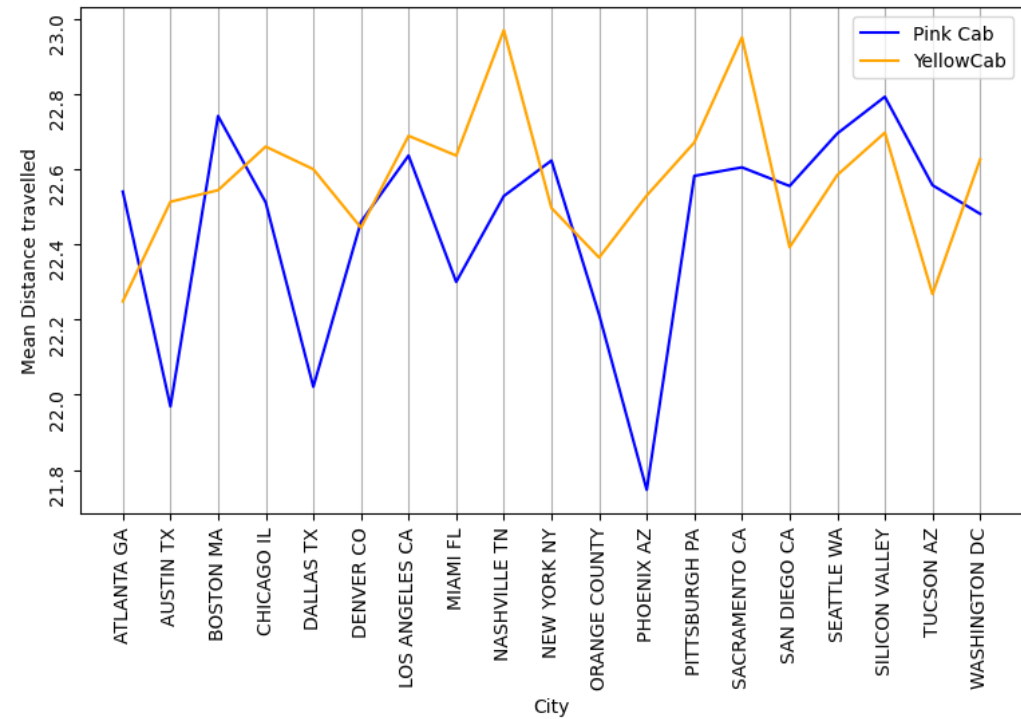
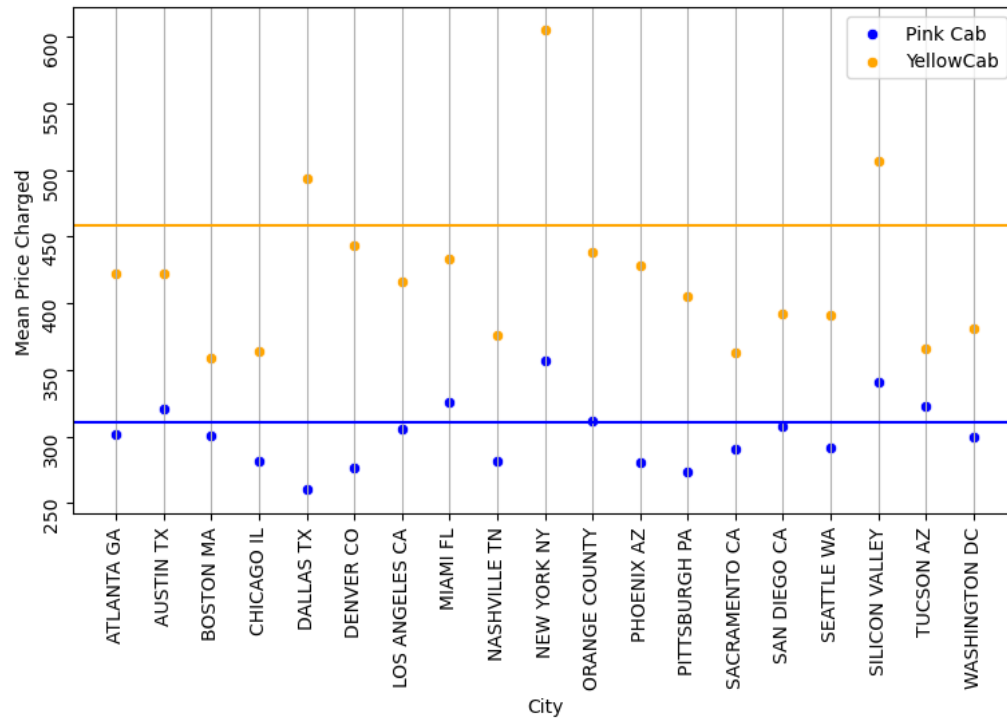
Exploratory Data Analysis

Understanding the datasets: Customer data



- There are almost as many female customers as male customers.
- Most customers are below the age of 40
- The average income of a customer is 15000 USD per month

Price Charged vs Distance Travelled



- Across all cities, the mean price charged by the Yellow Cab company is higher than that of the Pink Cab company while the mean distance travelled is almost the same around 21 and 23 km. The horizontal line in the first graph gives the overall mean price charged.

Profit Analysis

Analysing the profits made by the companies

	Pink Cab	Yellow Cab
Total Profit	5.307328e+06	4.402037e+07
Average Profit/Trip	6.265217e+01	1.602600e+02
Average Profit/KM	2.777146e+00	7.100727e+00

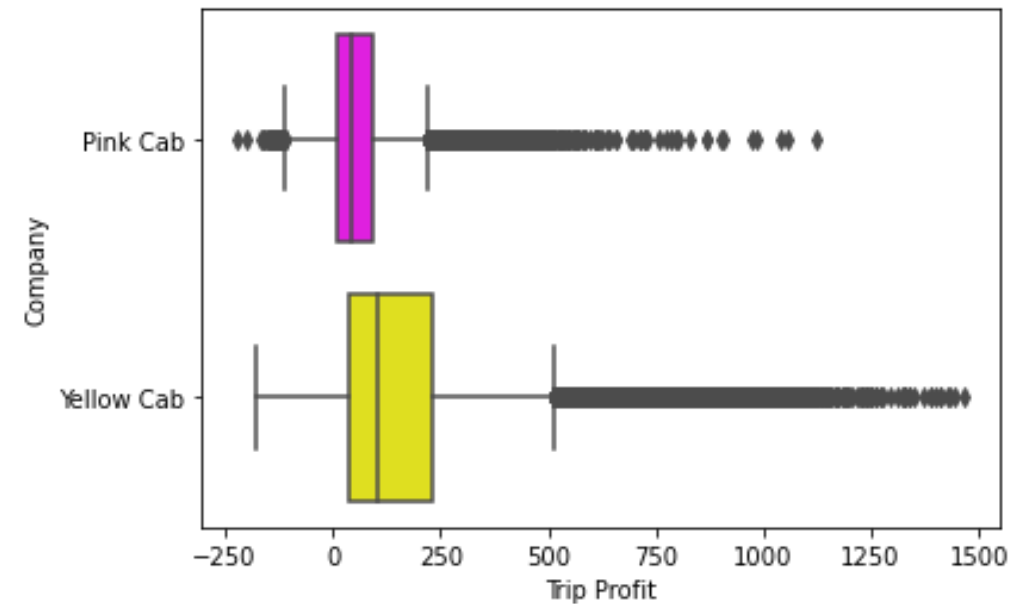
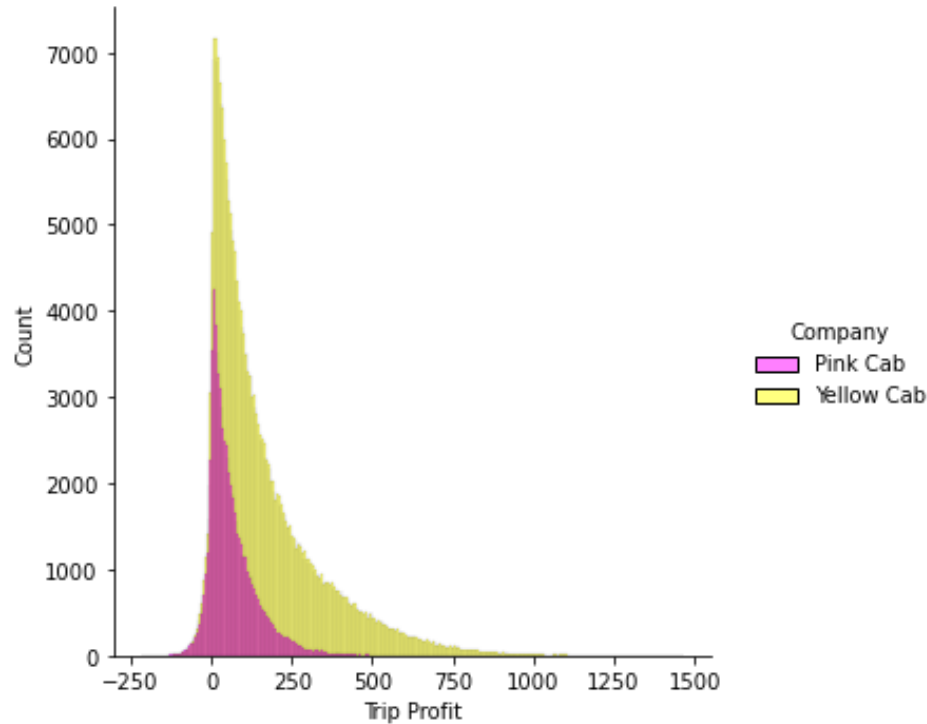
- Profit is calculated as:
Price Charged – Cost of Trip

	Pink Cab	Yellow Cab
Total Profit	5.3 Million USD	44.02 Million USD
Average Profit/Trip	62.6 USD	160.2 USD
Average Profit/KM	2.77 USD	7.1 USD

- The Data in the Table shows that the Yellow cab company manages to make higher profits. This is could be due to the fact that they charge higher prices as seen earlier.

Detailed Analysis

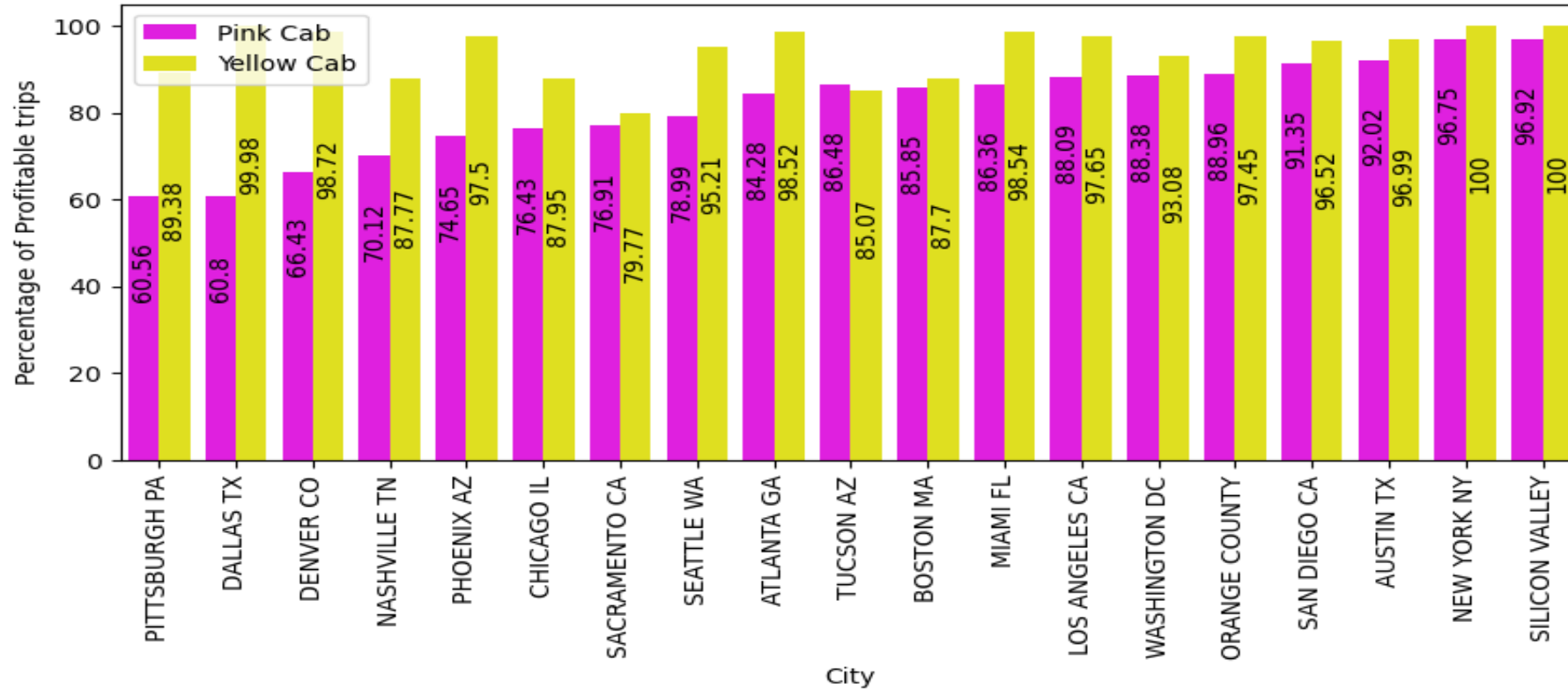
Analysing the profits made by the companies



- The Above graphs show the distribution of the profits. It can be seen that there are a lot of outliers. This is because of our earlier assumption.
- It can be noted that there are many rides that turned out non profitable.

Detailed Analysis

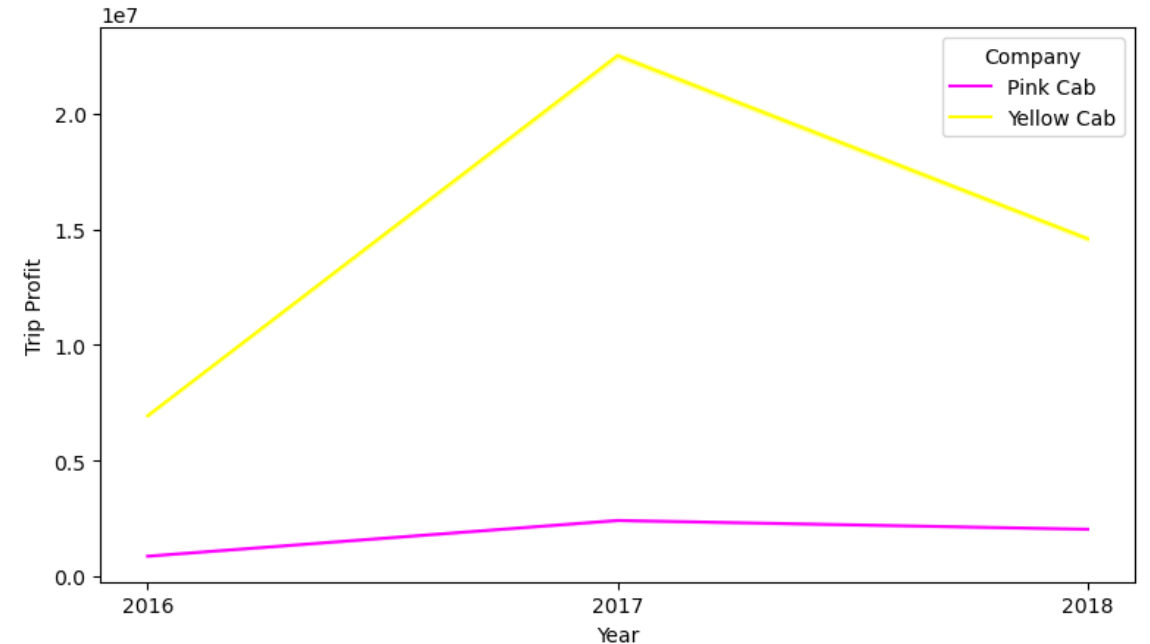
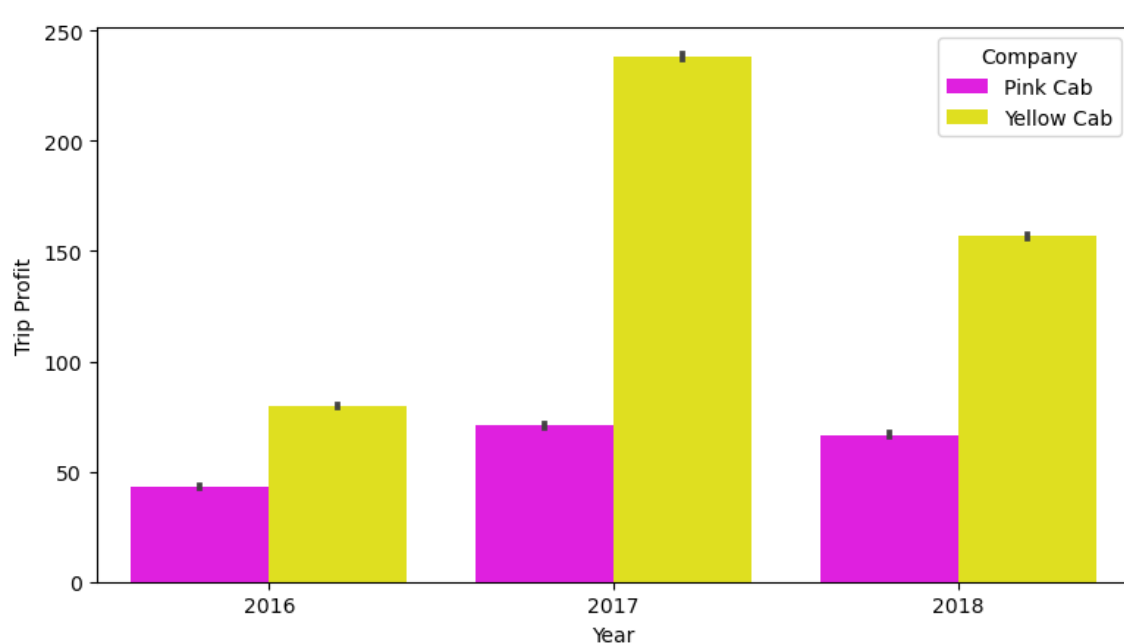
The percentage of rides that returned a profit across every city.



- We see that the Yellow cab company has a higher percentage of profitable trips than the Pink cab company.

Yearly Profit

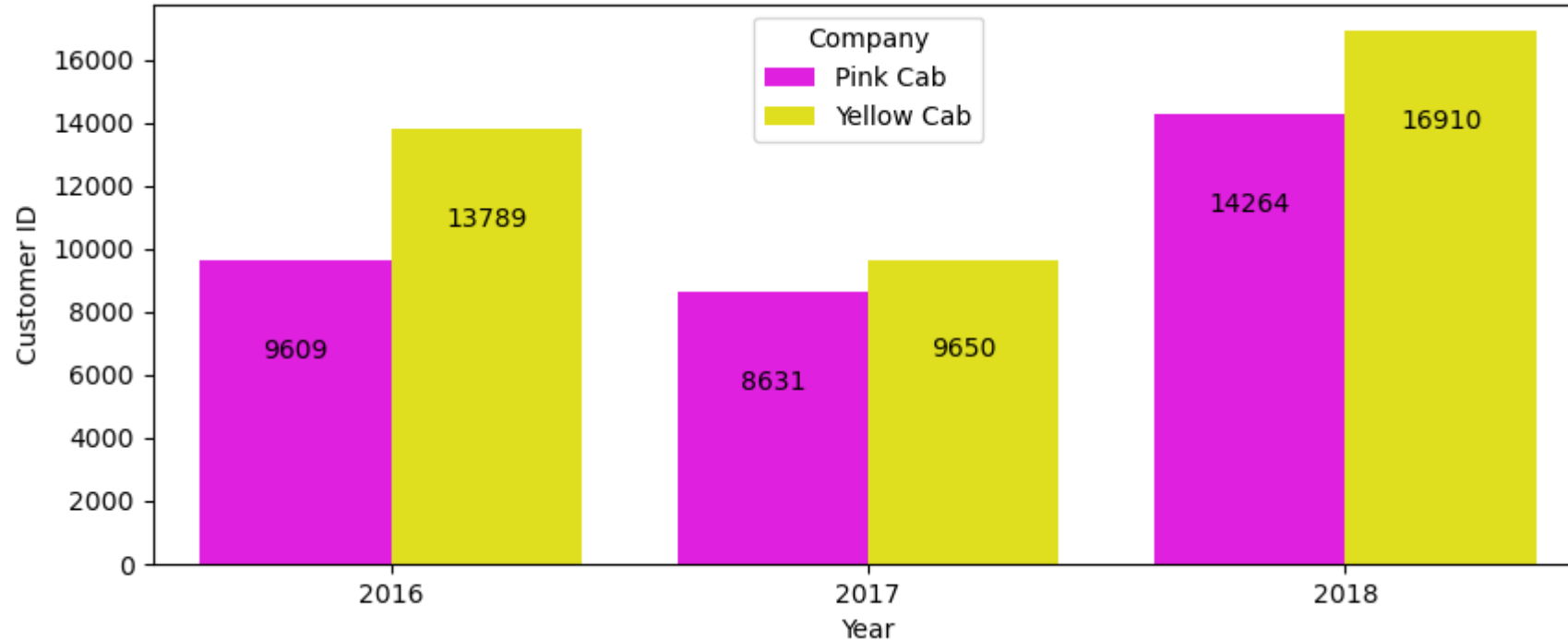
Analysing the profits made by the companies across the three years.



- The Yellow Cab company has made higher profits across all three years. We cannot say exactly why there is a decrease in the profit in 2018 without further analysis using more data.

Customer Count Per Year

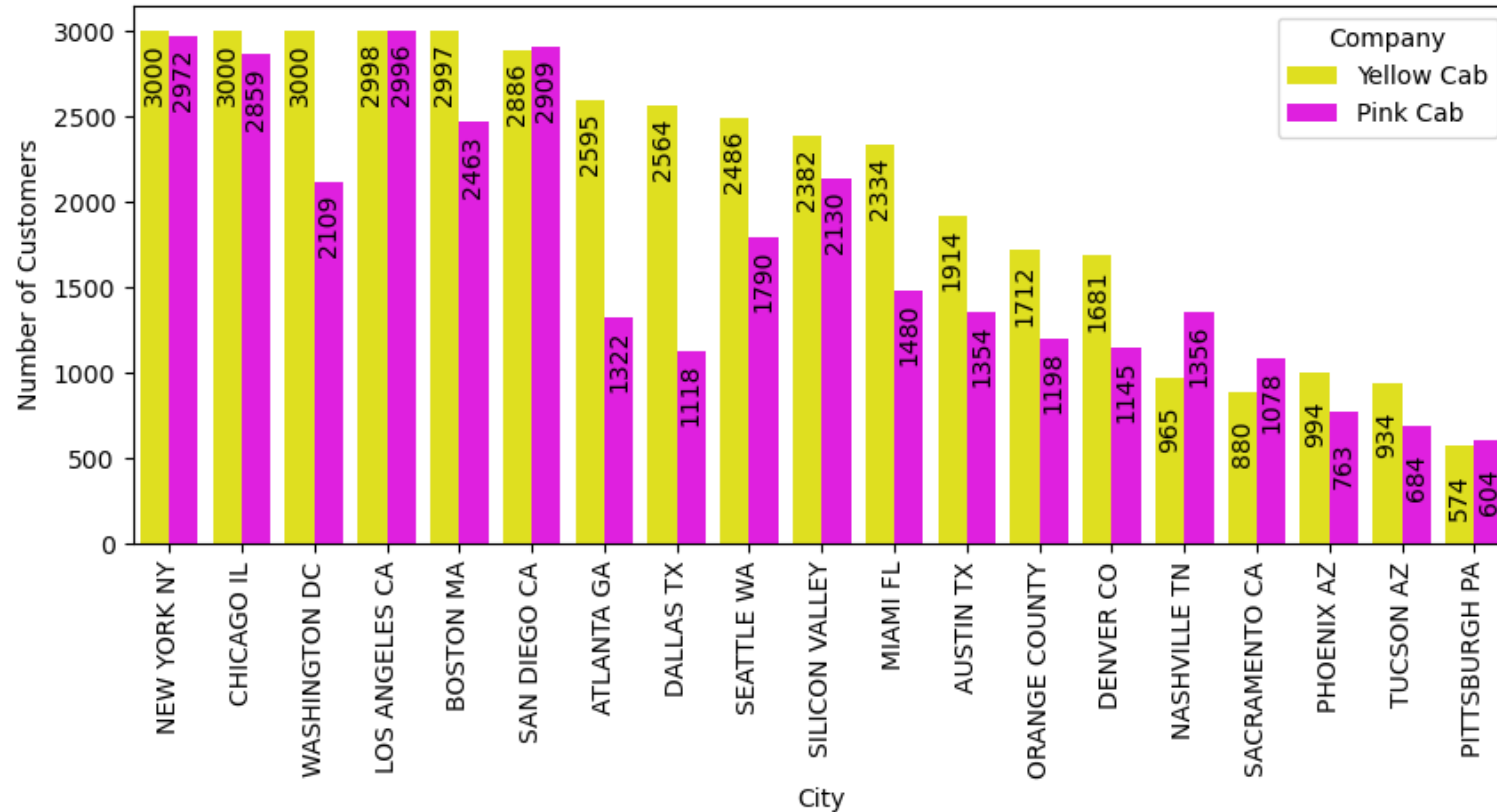
The number of customers handled by both companies each year.



- There has been an increase in the number of customers for both companies in 2018. This however has not led to an increase in profit.

City-wise customer share between the companies

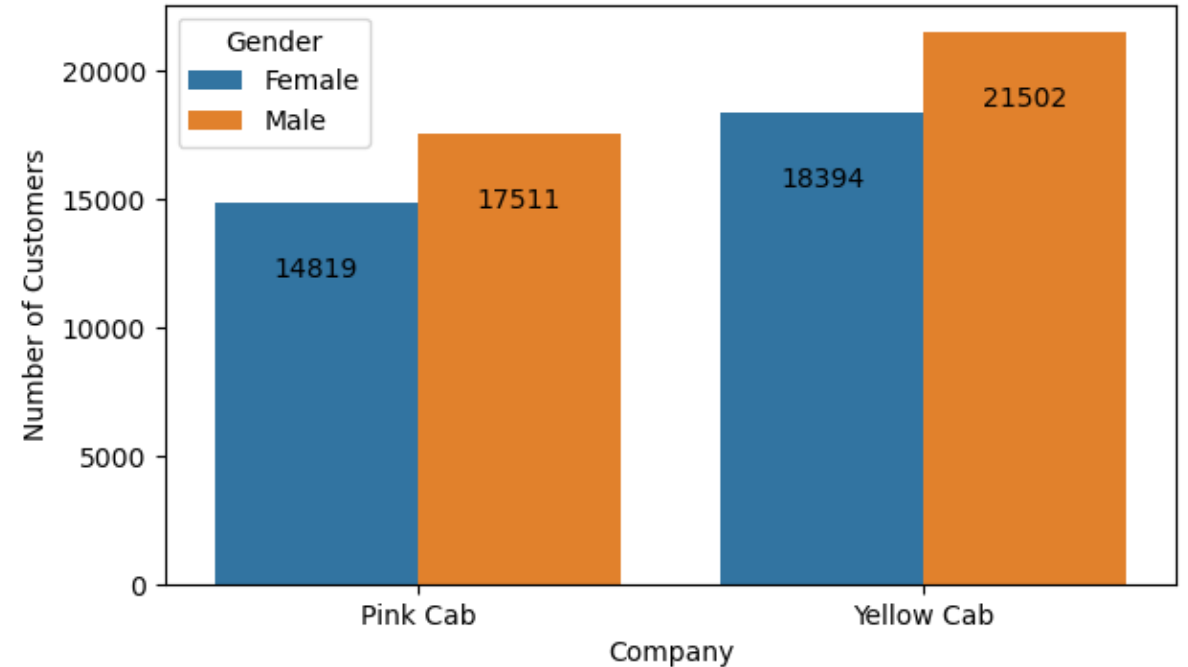
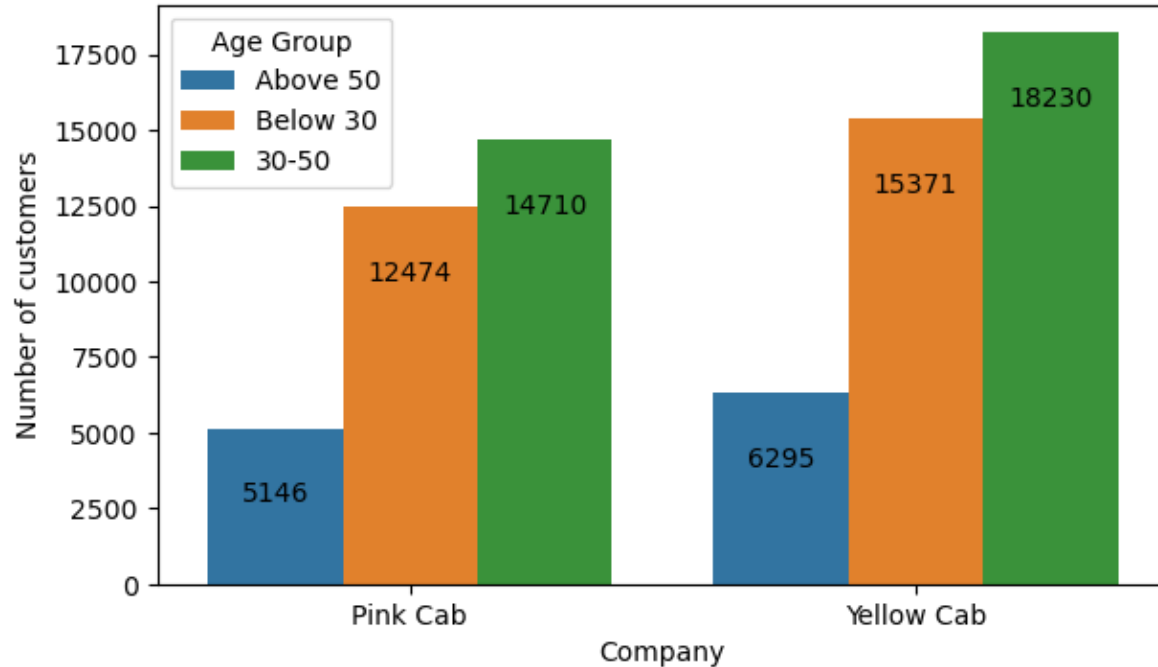
Analysing the ride share between the two companies in every city.



- The Pink Cab company has had a higher share of customers in the 4 cities of San Diego, Nashville, Sacramento and Pittsburgh

Customer share between the companies

Analysing ride share based on age group and gender.



- There is no indication that age or gender plays a role in a customer choosing a cab company. Both cab companies have a similar user pattern based on age and gender.

Recommendations

Based on the analysis done, the following recommendations are being put forward

- The Yellow Cab company seems to be more popular among customers.
- It has a higher customer count than the Pink Cab company.
- The Yellow Cab company performs better as it has a higher percentage of profitable trips across all the cities.
- The Yellow Cab company also has had higher profits overall and per trip/KM than its competitor across all three years and is therefore most like to remain more profitable than the Pink Cab company in the near future.

Based on these observations, we recommend that XYZ invest in the Yellow Cab Company.

Thank You