



Data Glacier

Your Deep Learning Partner

G2M insight for Cab Investment firm

Company Name : XYZ

Location: USA

Name: Refka Mejri

Date: 21/10/2021

Plan –G2M(cab industry) case study

Objective :

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

The analysis has been divided into 5 parts:

- Data Understanding and cleaning.
- EDA.
- Analyzing the more profitable Cab Company.
- Hypothesis Analysis results.
- Recommendations for investment.

Data Information

D

There are 4 datasets:

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.

Data Exploration

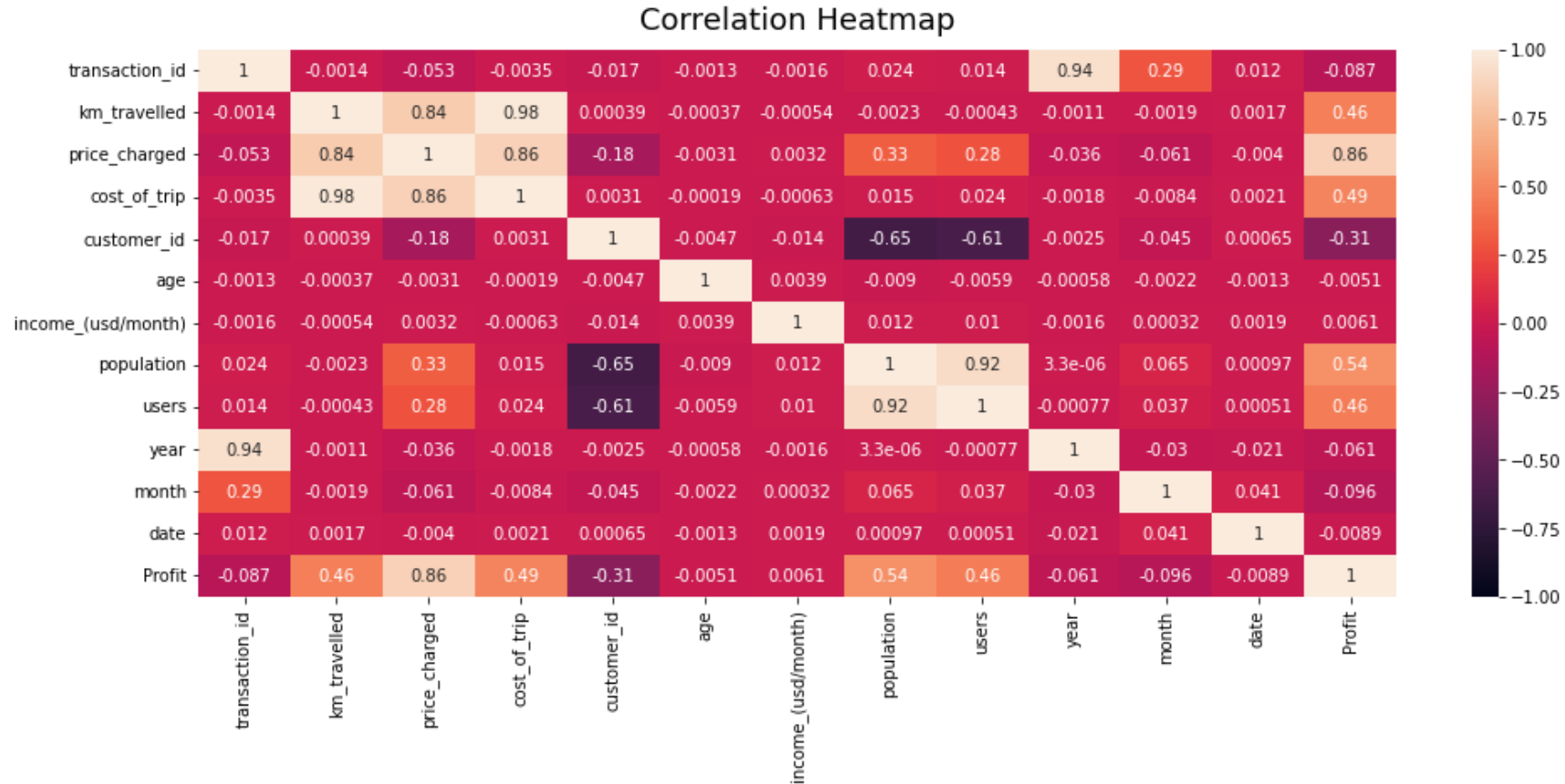
Assumptions:

Main_Data.Csv : We obtain this dataset by merging the 4 datasets

- 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) .

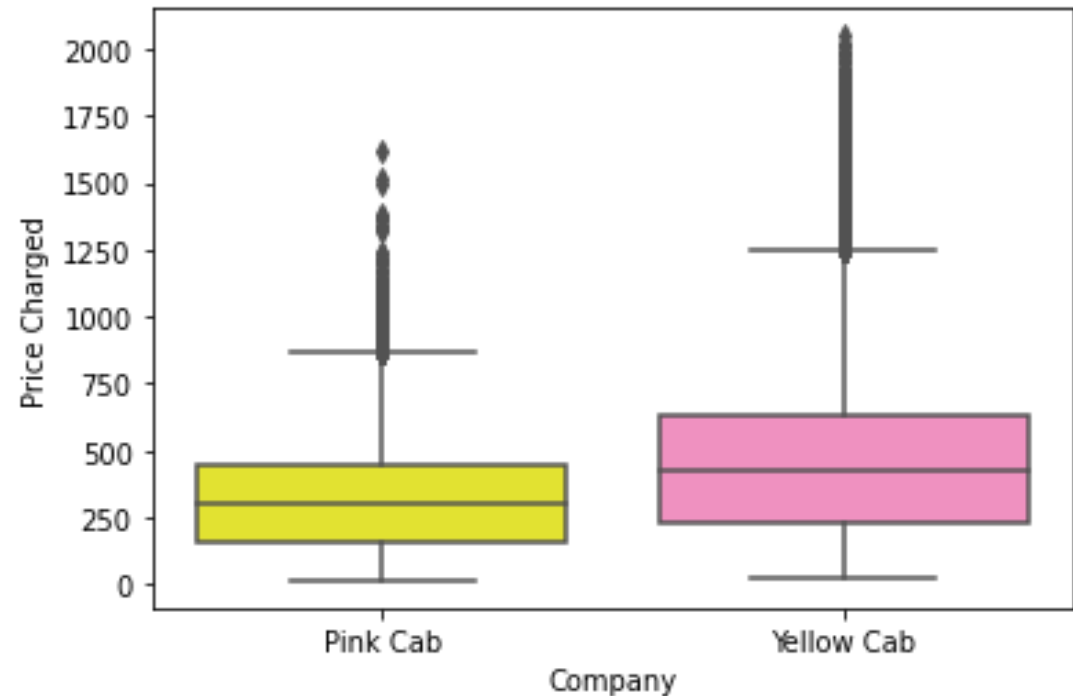
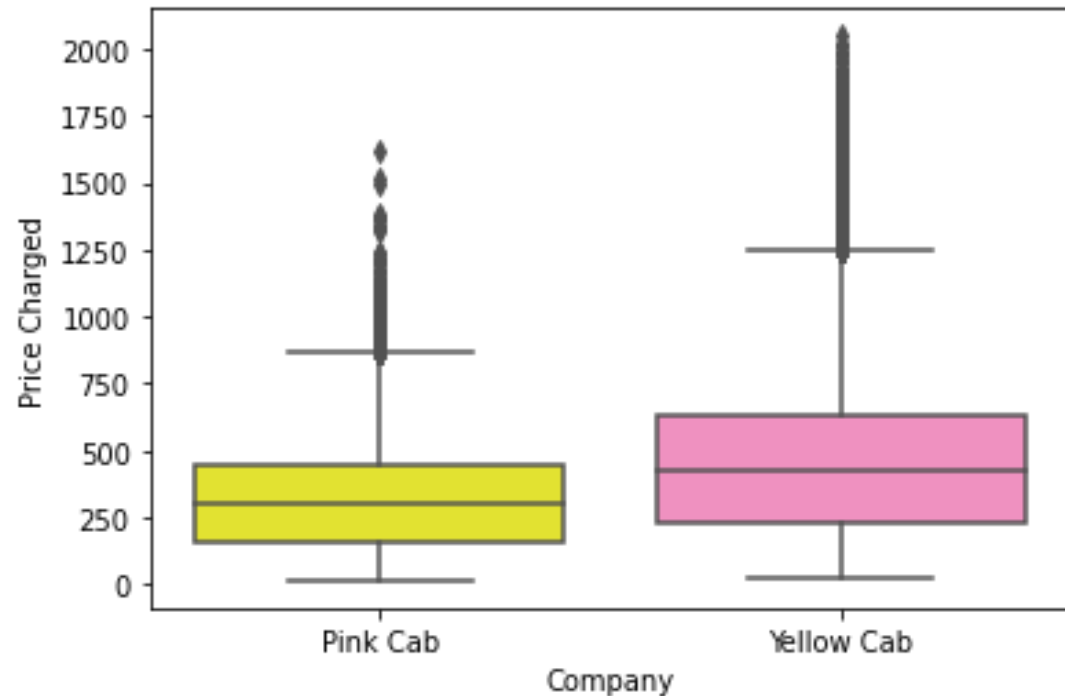
Time frame of the data: 2016-01-31 to 2018-12-31

Feature Correlation



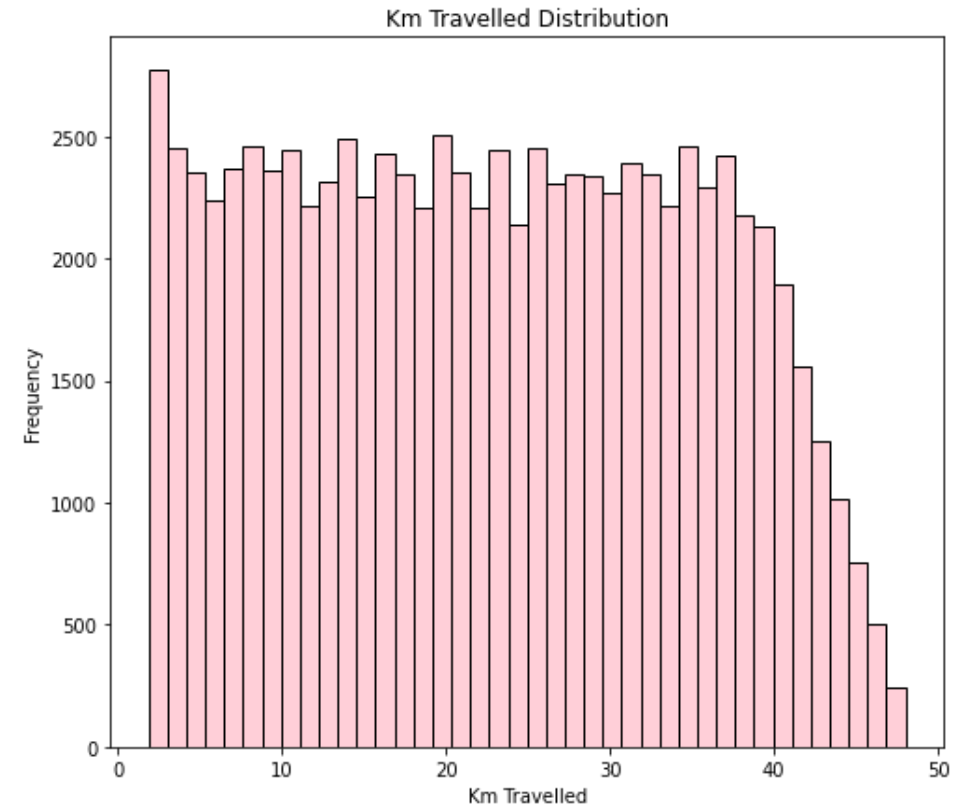
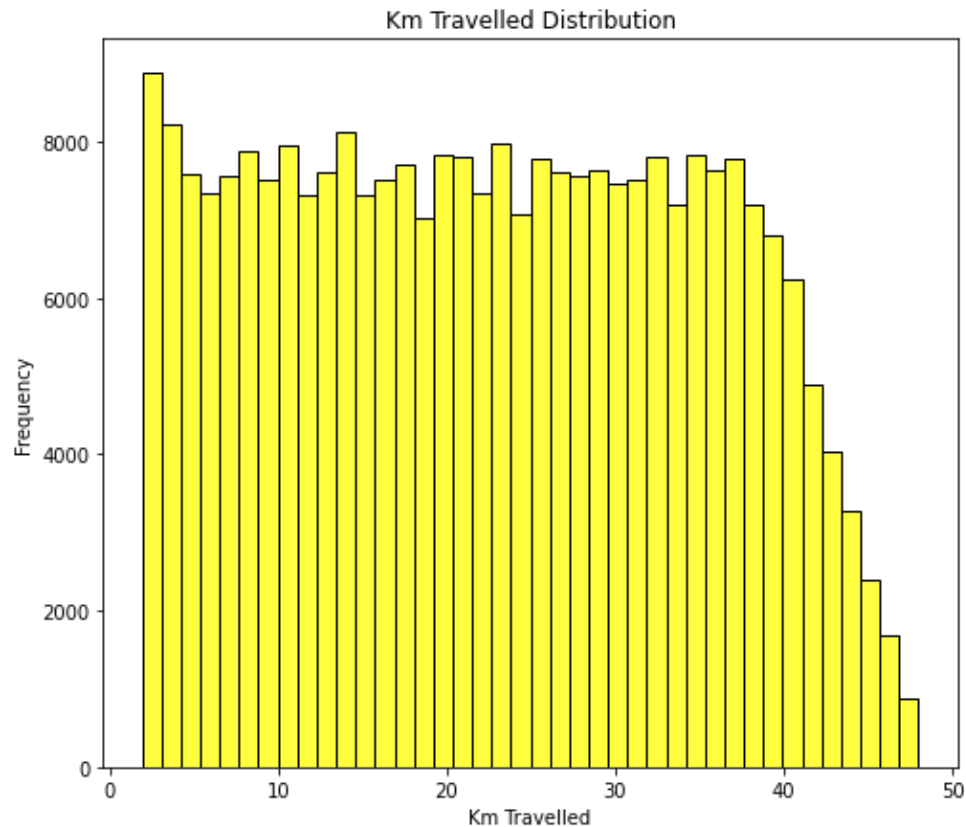
There is a strong positive correlation between Profit & Price Charged around 0.86, Cost of Trips and KM Travelled: 0.98, Price Charged and KM travelled: 0.84.

Outlier for significant feature



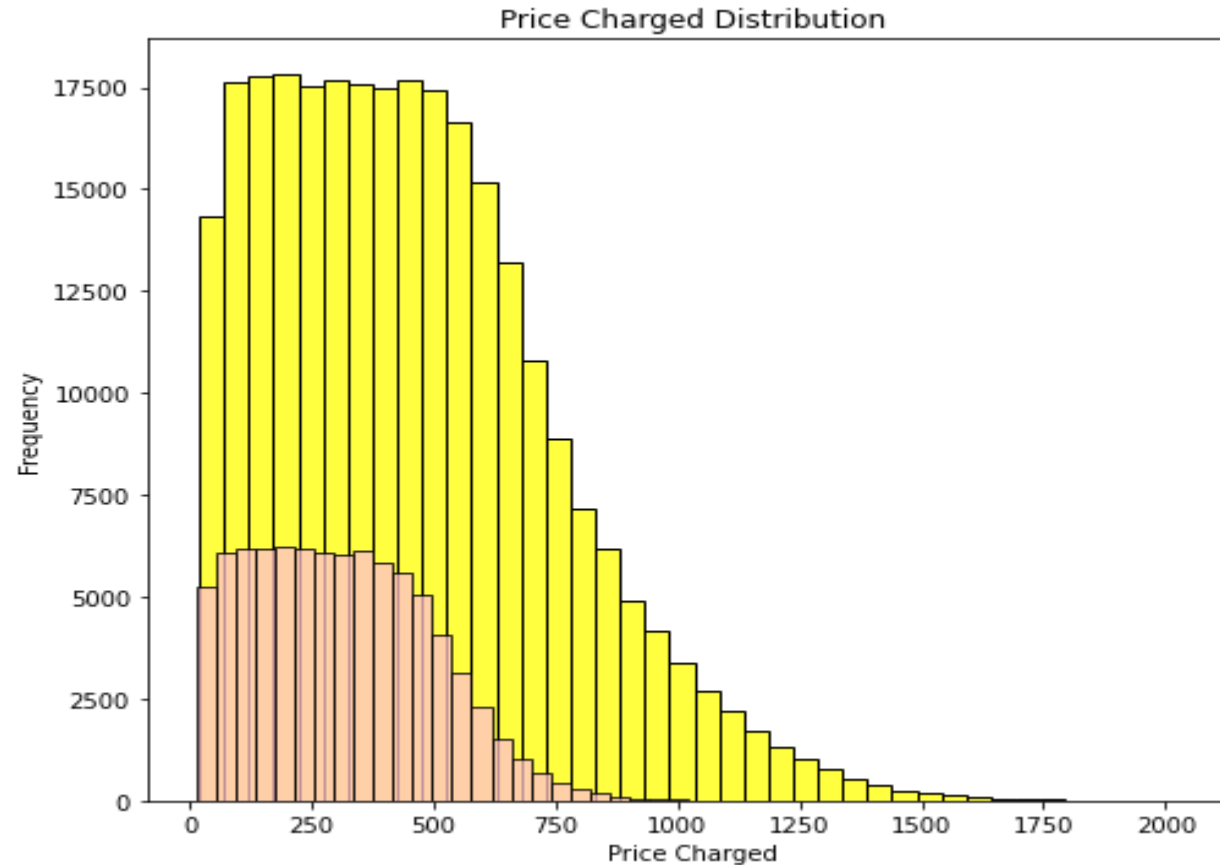
Outliers are present in Price_Charged feature but due to unavailability of trip duration details ,I am not treating this as outlier.

Distribution of KM Travelled for Yellow and Pink Cab



I see that for both pink and yellow Cab most of the ride are in the range of approximately 2 to 48 KM

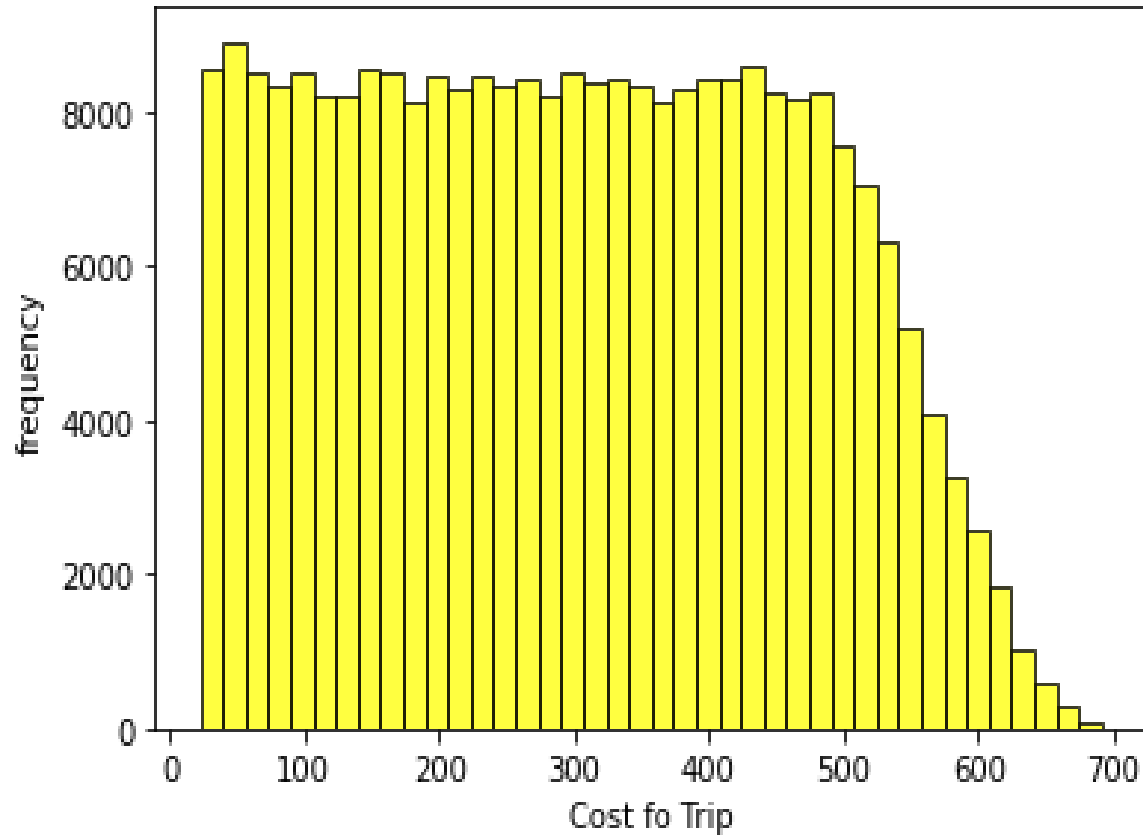
Distribution of Price Charged By Company



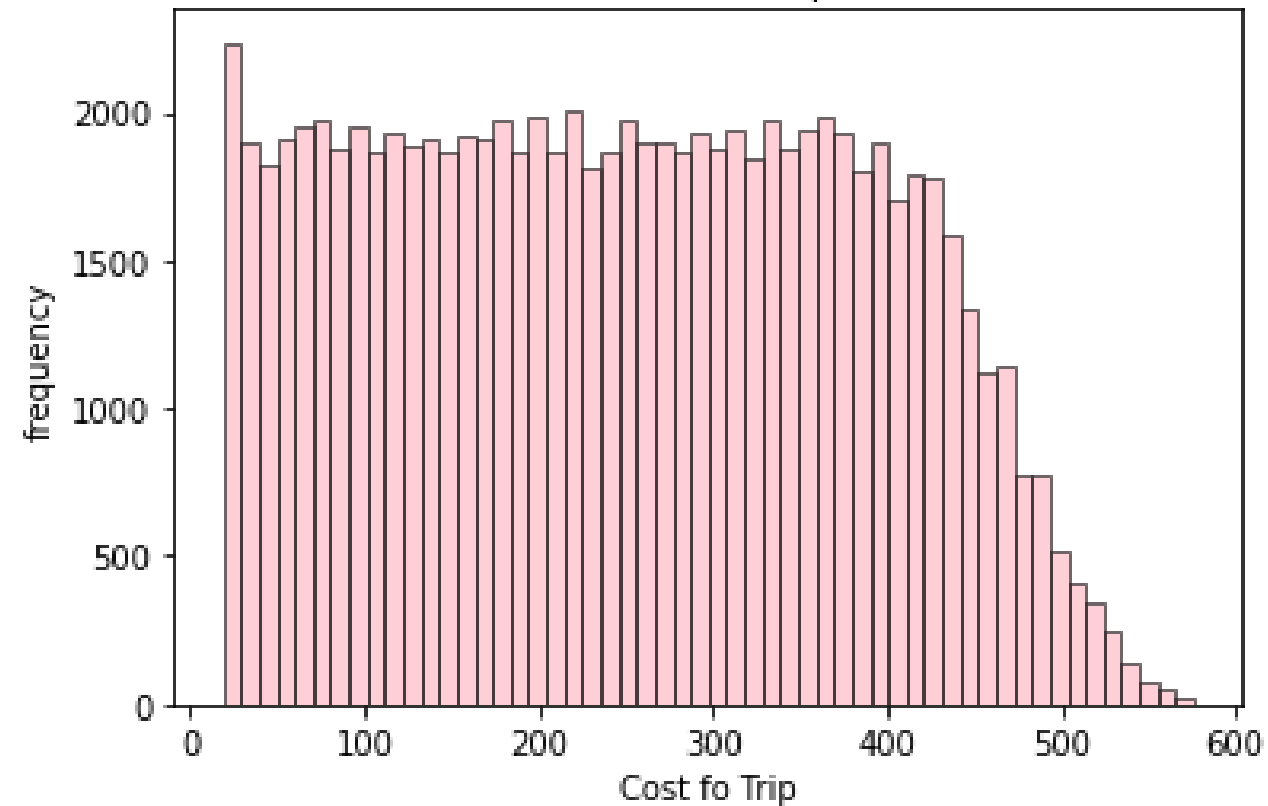
The price charged range for the Yellow Company is higher than the pink Cap Company

Cost of Trip repartition for both Pink and Yellow Company

Distribution of cost for Trip for Yellow Cab

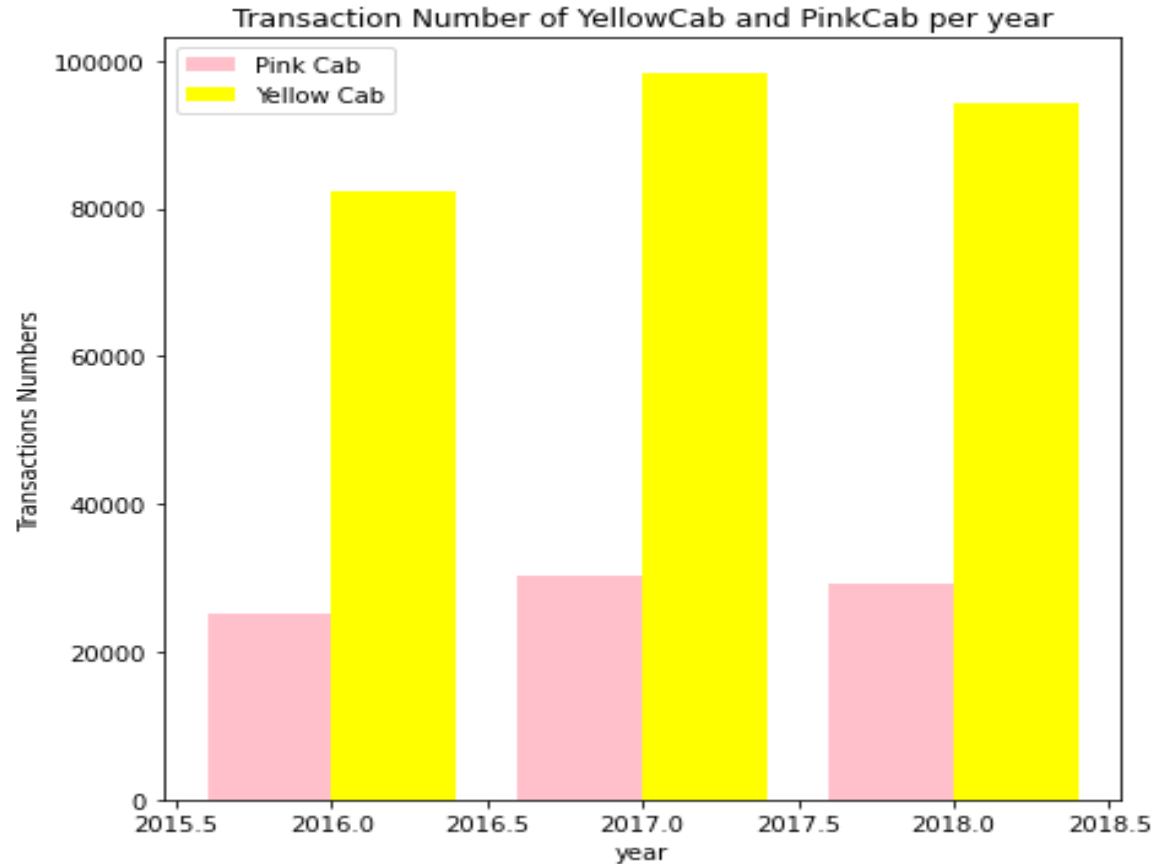


Distribution of cost fo Trip for Pink Cab



The Cost of Trip range for the Yellow Company is higher than the pink Cap Company

Transaction repartition per number of each company per Year



- We see from this graph that the number of transaction of Yellow Cab company is higher than the Pink Cab company numbers.

- We conclude that Yellow cab company is more active and the users prefer Yellow cab company.

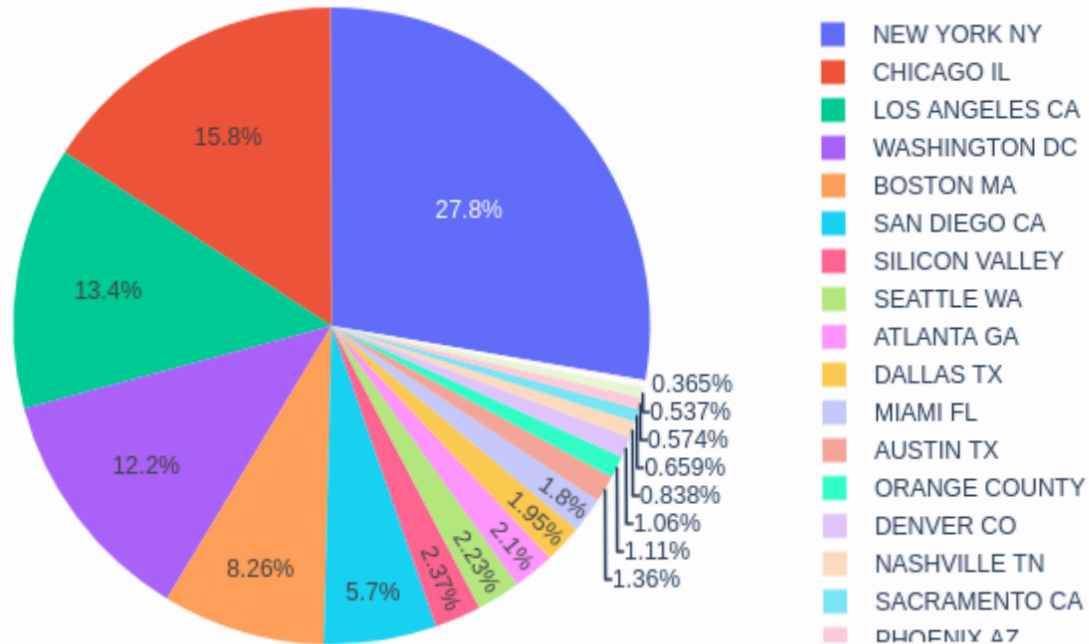
Transactions repartition per number of each company per month



The number of transaction of Yellow Cab company is higher than the Pink Cab company.

Customer repartition per city

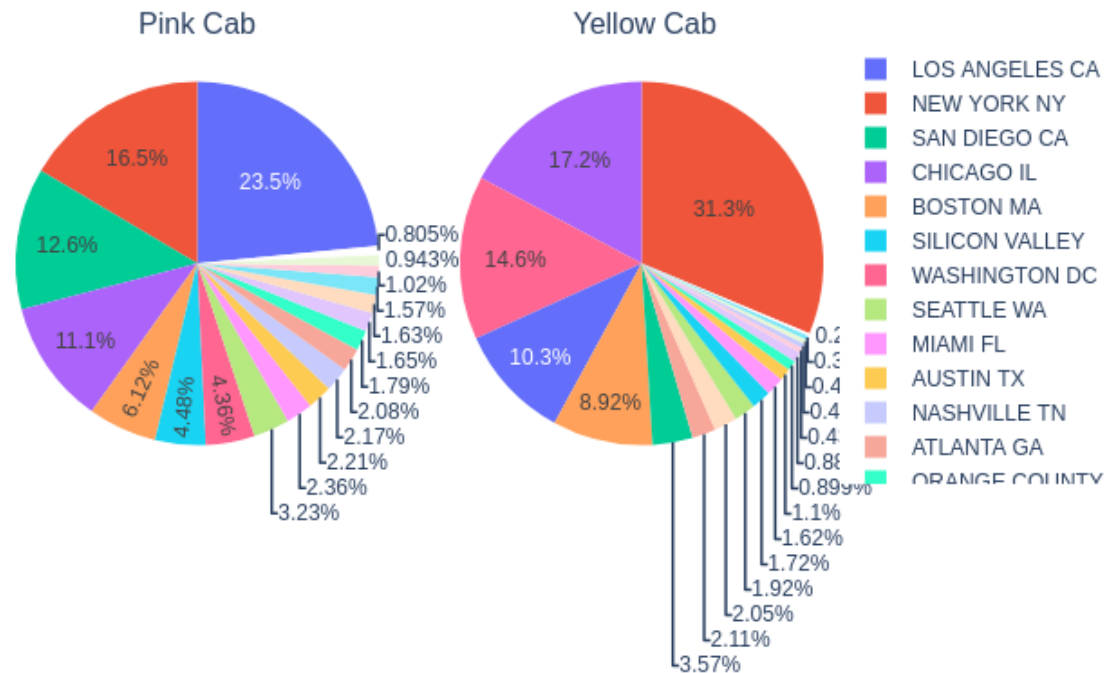
Percentage of Users per cities



- New York City has the highest cab users with 27,8 %.
- Chicago and Los Angeles has also a higher number of users with 15,8 % and 13,4%.

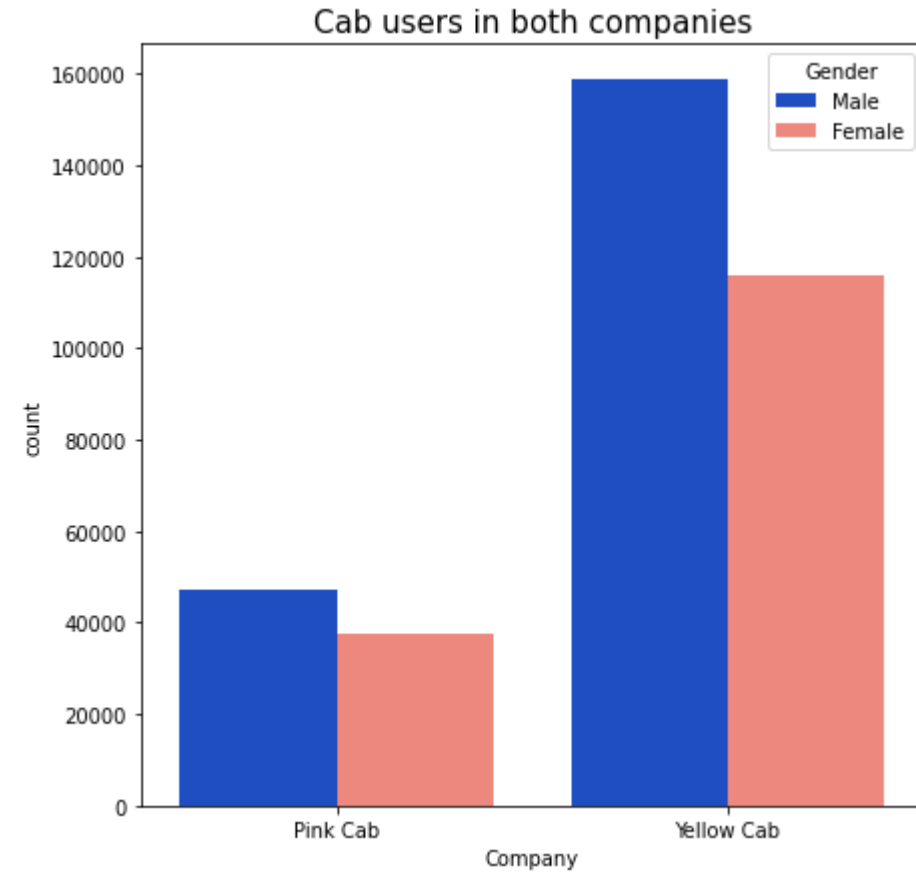
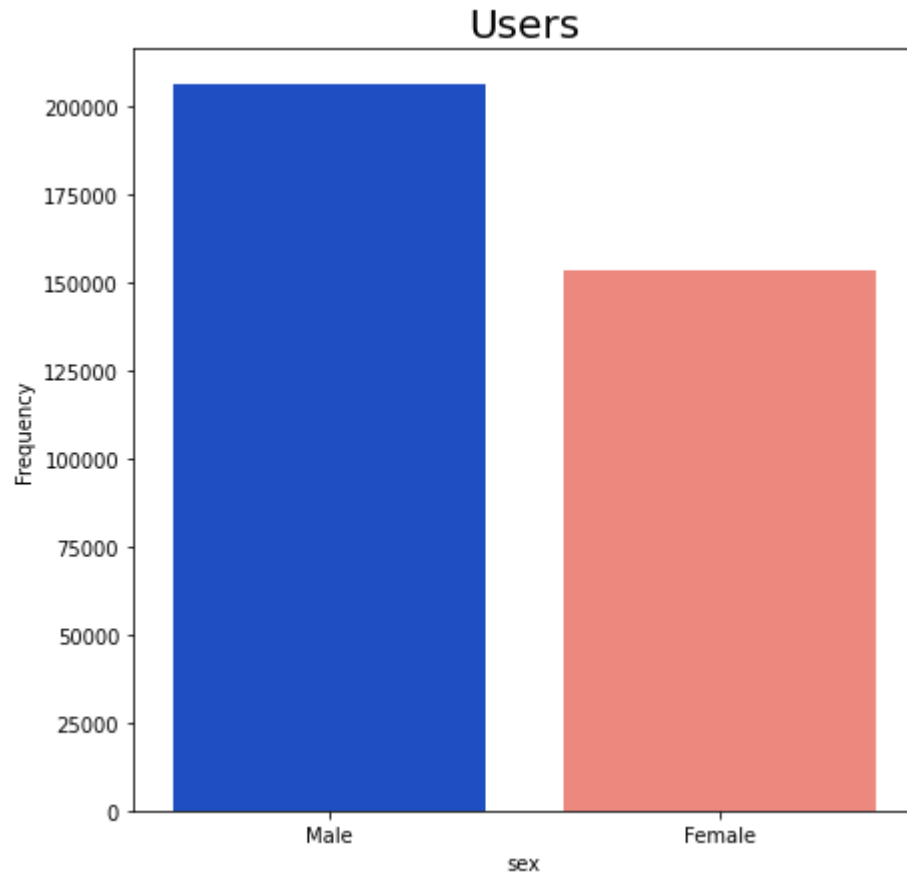
Customer repartition per city for both Yellow and Pink Cab

Percentage of Customers per City for both Yellow and Pink Cab



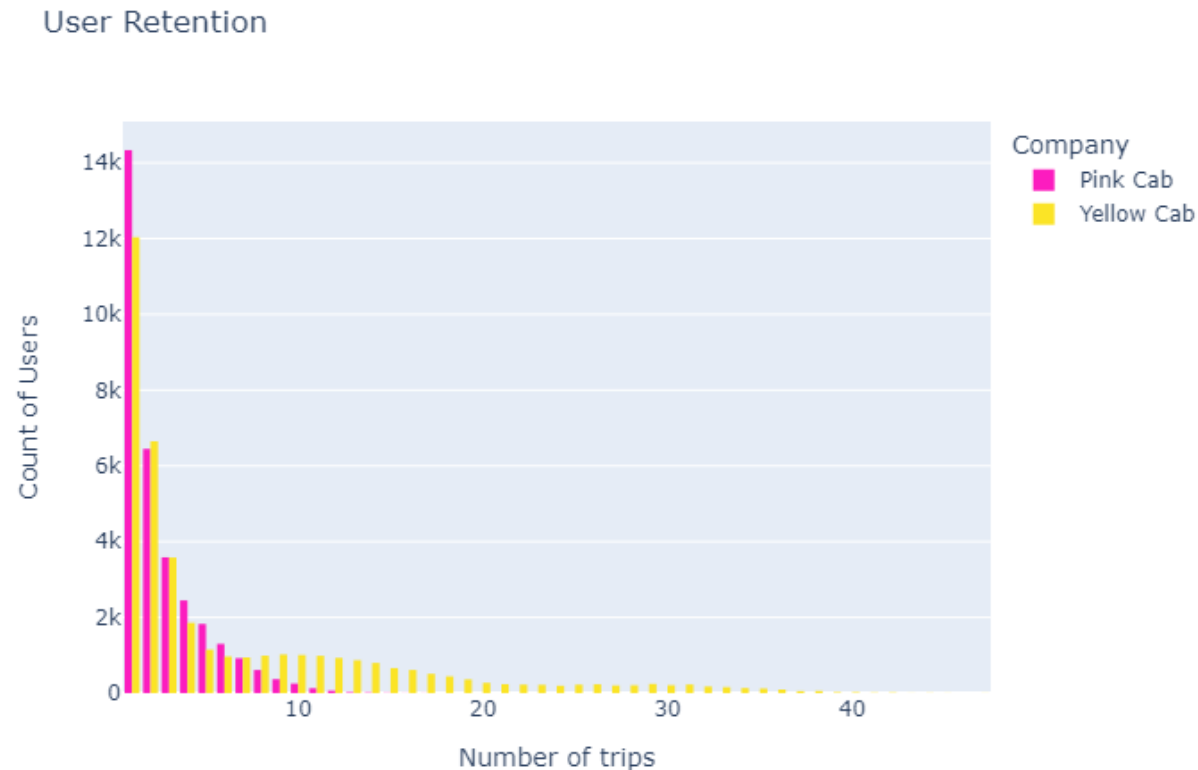
- The highest numebrs of customers is in New York for both company, with 31,% for yellow cab and 23,5 % for Pink Cab.
- The number of customer riding with Pink Cab in Los Angelos and San Diego is higher than Yellow cab.

Customer Frequency per Gender



The highest number of femal customers use the Yellow Cab

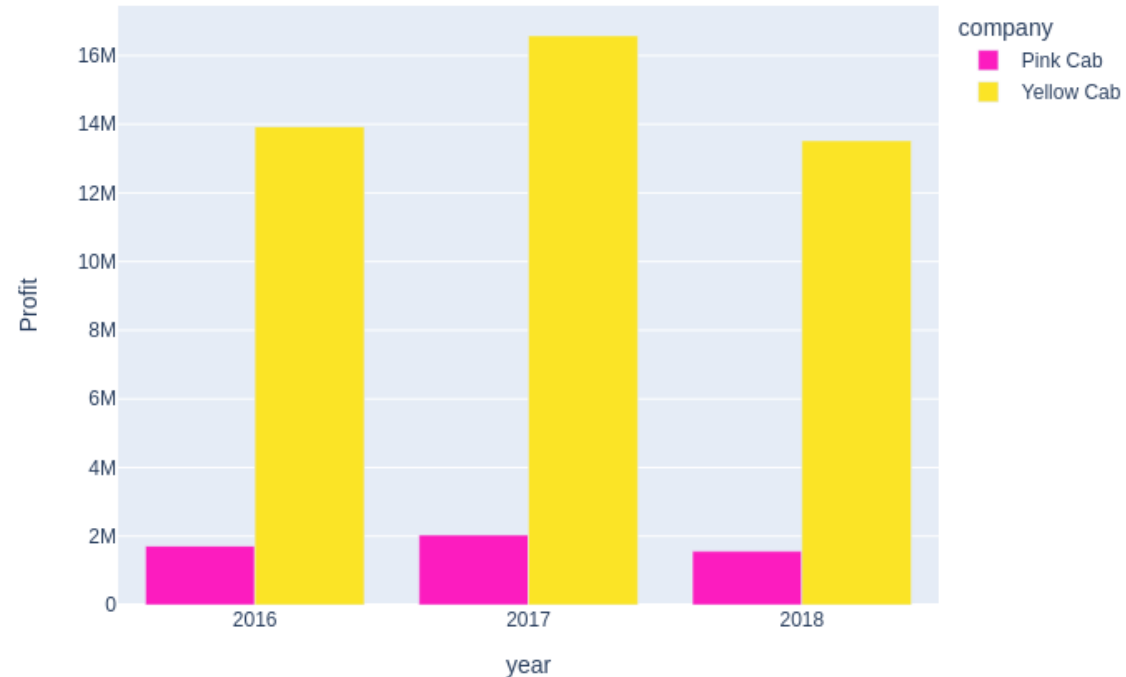
Customer Repartition per number of trip



For the first time, Pink Cab seems the best choice for customers with more than 14k users and 12k for Yellow cab. From the 7 trips we see that people tend to ride with yellow cab.

Profit repartition for both company

Profit between 2016 and 2018 for Pink and Yellow Cab Company

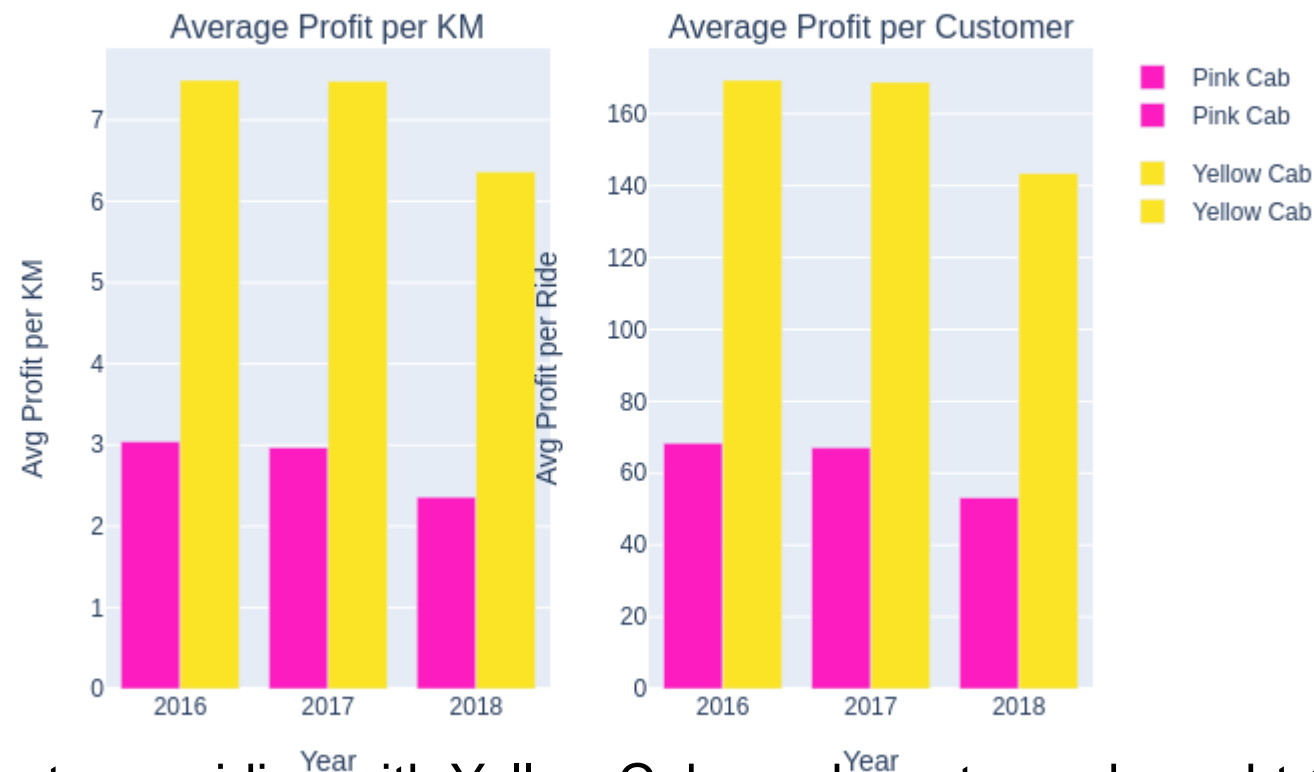


In general, Yellow Cab has higher profit from 2016 to 2018 compared to Pink Cab.

From 2016 to 2017, both Yellow and Pink Cab had been experienced the increase in Profit. For Yellow Cab, 2017 was the most successful year with roughly 16.57 million dollars compared to nearly 2 million dollars from Pink Cab.

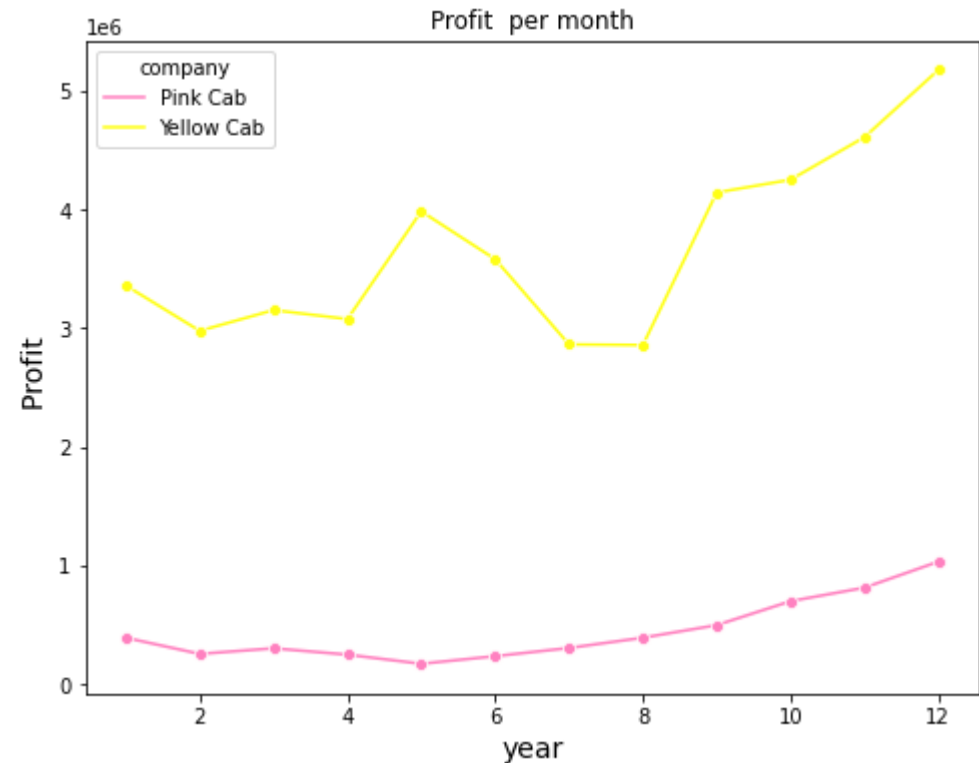
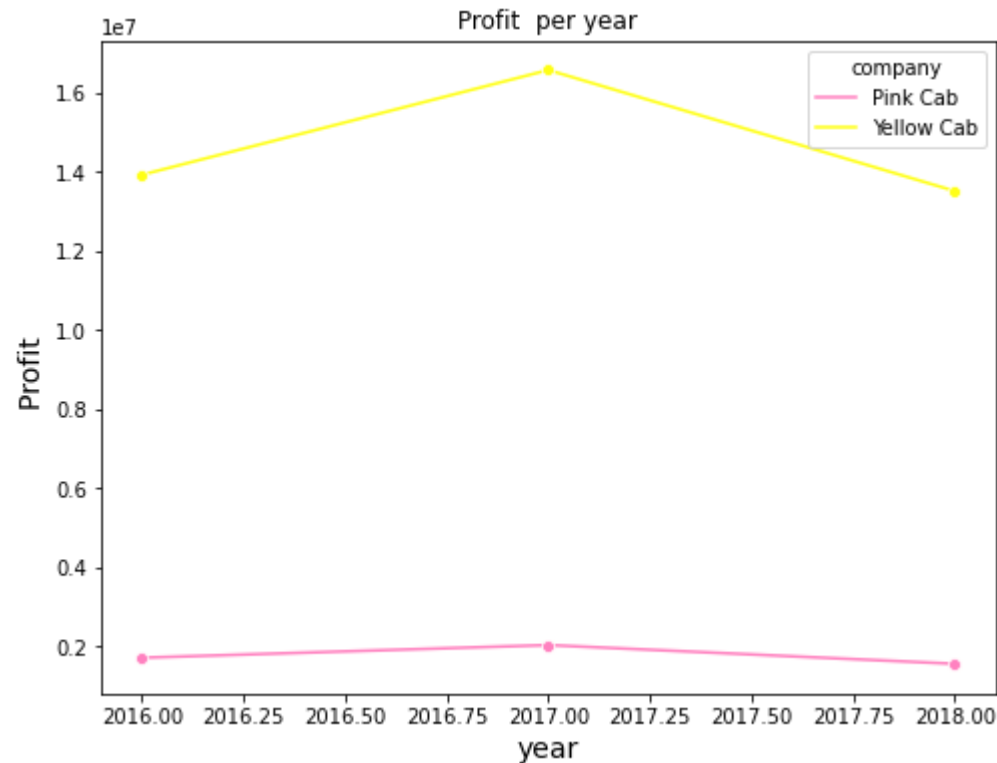
Average profit per kilometer travelled and Customer

Average Net Profit per KM and per Customer between 2016 and 2018



For each customer riding with Yellow Cab, each customer brought 169.35, 168.82 and 143.42 dollars from 2016 to 2018 respectively whereas this number is 68.32, 67.07 and 53.23 dollars respectively.

Profit repartition per years and months



Yellow in general has higher profit compared to Pink Cab in every month. Profit from both companies starts to increase from August to December.

The profit line from Yellow not stable compared to Pink Cab, where their profit line does not change dramatically.

Hypothesis thesis Analysis

First Hypothesis is to see if there is an influence /difference regarding ages for both yellow and pink cab.

Second hypothesis is to see if gender has an impact on profit for both company.

Conclusion:

For Yellow cab there is a difference regarding age but for pink company there is no difference.

Yellow Cab: There is difference in Profit between Male and Female customers

Pink Cab: There is no difference in Profit between Male and Female customers

For Yellow Cab, gender really matter in generating profit. Male Customers tend to brings more profit compared to Pink Cab. However, there are no gender different in Profit for Male and Female customers. The reason it could be because Pink Cab charged the same amount of money regrading its gender.

Conclusion and recommendations for investment

After analyzing and treating carefully our dataset, I highly recommend to invest in Yellow Cab rather than the pink Cab.

Profit Analysis Aspect :

The Yellow cab made the higher profit than the Pink Cab in most of cities from 2016 to 2018.

Price Charged (has a strong positive correlation with distance) is more correlated with the profit features, represent the higher earning for the Yellow Cab Company than the Pink Cab Company.

Customer Analysis Aspect :

The Yellow Cab Company has more reputation and is the top choice for customers in many cities and the most popular cities in US.

We see that the number of transaction, number of customers per city is always higher for the YellowCab Company than the Pink Cab Company

Thank You



Data Glacier

Your Deep Learning Partner