



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M Insight for Cab Investment Firm

September 19th 2022

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Summary

- Goal: Provide company 'XYZ' insight into the growing Cab Industry in the USA as they are looking to invest.
- The data provided contains key attributes of two cab companies, Pink Cab and Yellow Cab.
- The time period of the data is from 31/01/2016 to 31/12/2018.
- The data contains the following attributes:
 1. Cab Data: Transaction ID, Date of Transaction, Company, City, KM Travelled, Price Charged, Cost of Trip
 2. City Data: City, Population, Number of Users.
 3. Customer Data: Customer ID, Gender, Age, Income.
 4. Transaction Data: Transaction ID, Customer ID, Payment Mode.

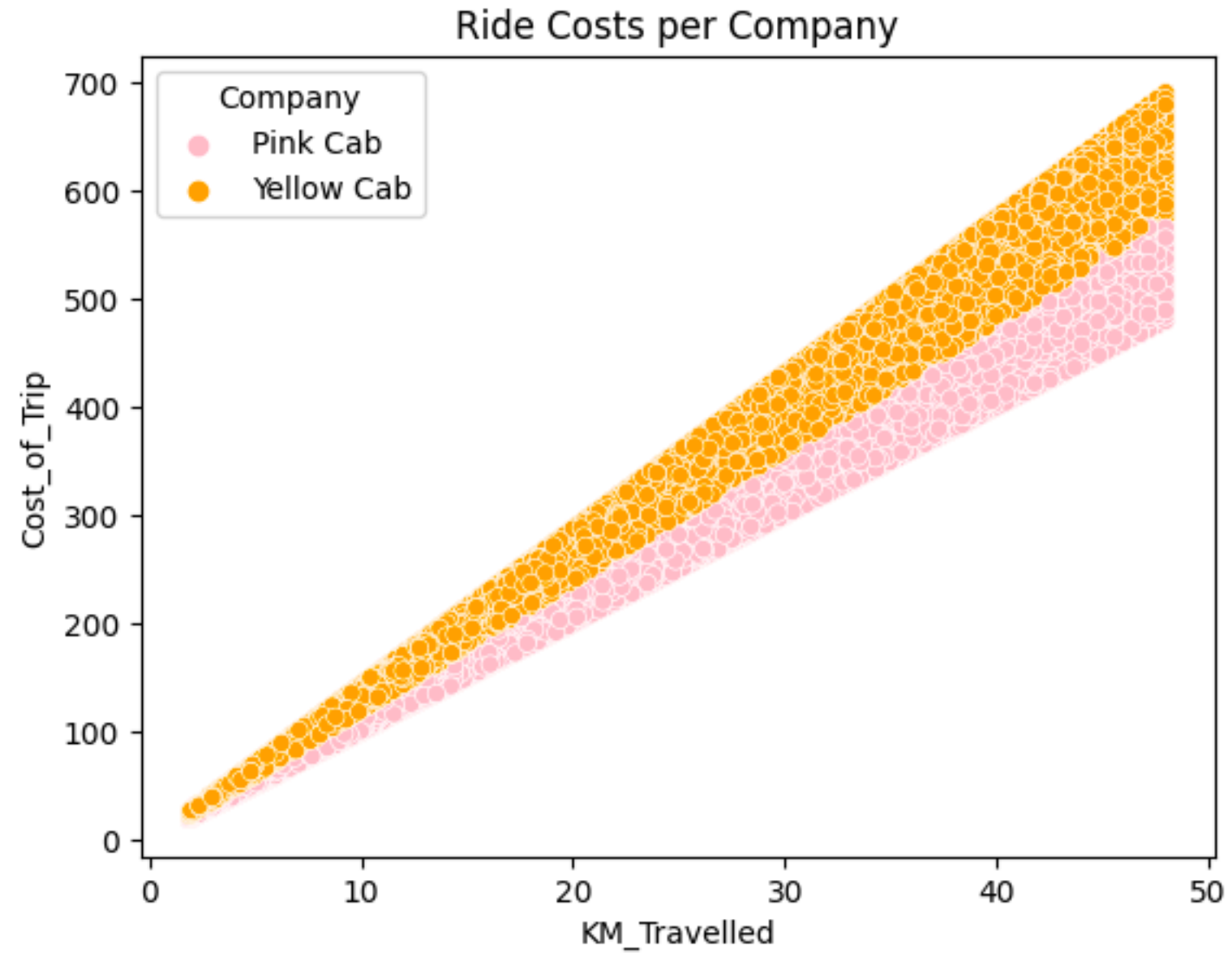
Problem Statement

- XYZ company needs to identify the right company deserving of their investment
- For this analysis, and to make sure the decision making is backed up by data, we will be looking at the following hypotheses:
 1. Identifying which company has more cab users during the fiscal year, by seasons.
 2. Finding if profit ratio and margin increases when the number of customers increase, i.e. is expansion profitable or not.
 3. Identifying customer demographics for each firm.

Approach

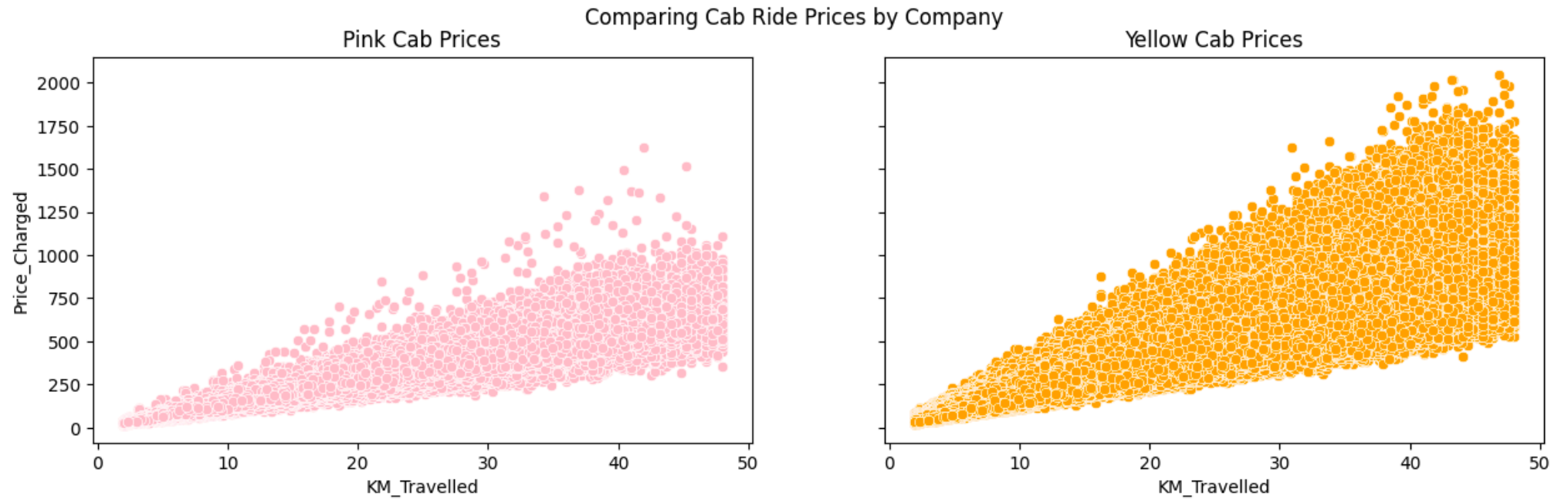
- The approach is to conduct an Exploratory Data Analysis, this requires us to use the provided 4 datasets: [Cab_Data.csv, City.csv, Customer_ID.csv, Transaction_ID.csv].
- EDA is used to take insights from the data. We have tried to find different patterns, relations, and anomalies in the data using some statistical graphs and other visualization techniques. The steps included are:
 1. Variable Identification
 2. Univariate Analysis (categorical, discrete and continuous)
 3. Bi-Variate Analysis
 4. Variable Transformation

EDA



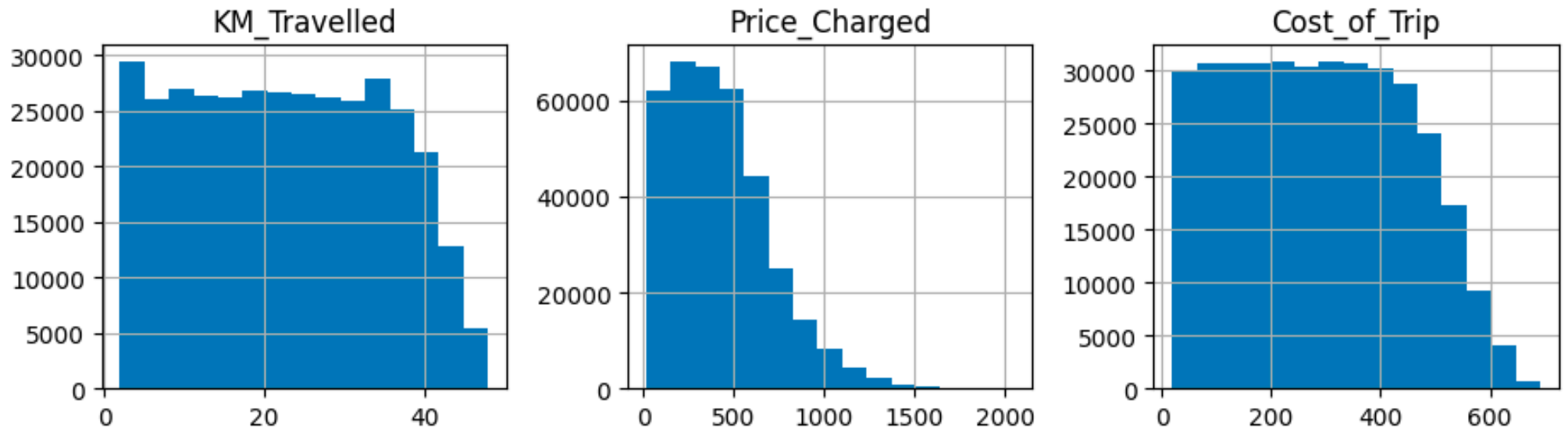
- At first glance, we can see that Yellow Cab cost are much higher as compared to Pink Cab. Let's take a deeper look.

EDA



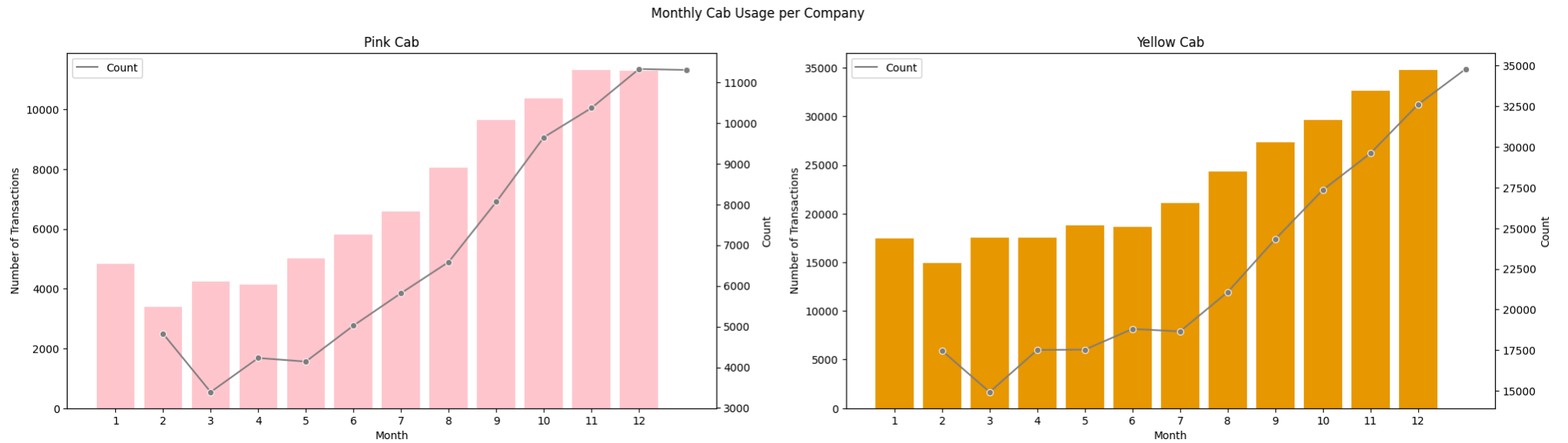
- However, we also find out that Yellow Cab prices are also significantly higher.

EDA



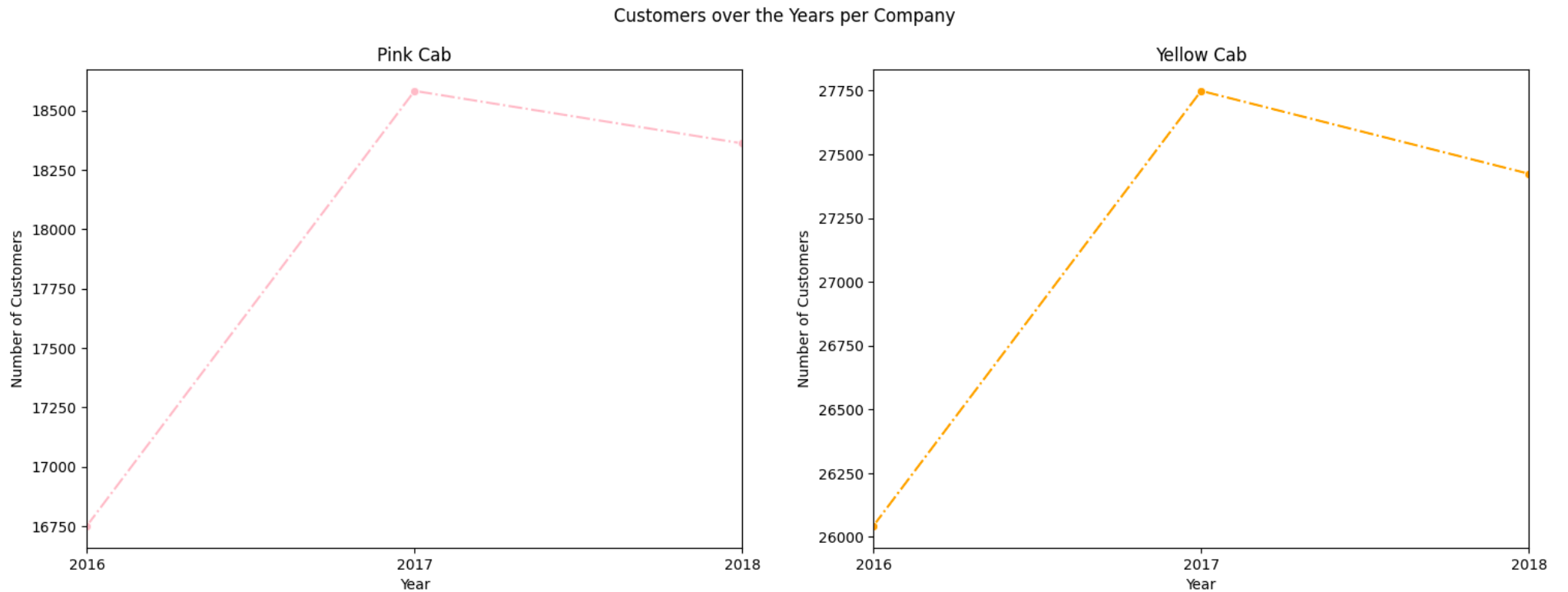
- Both Cab companies have $\text{Price} > \text{Cost}$, certainly, both are creating some profits.
- The consumer ride length also determines how much money each is bound to make.
- Both Pink and Yellow cab rides fall between the 2KM - 48KM range.

EDA



- Yellow Cab has far more transactions as per out data, in the range of nearly 35000 in the peak usage time in December, as opposed to Pink Cab that reaches a maximum of around 11000 transactions per month.

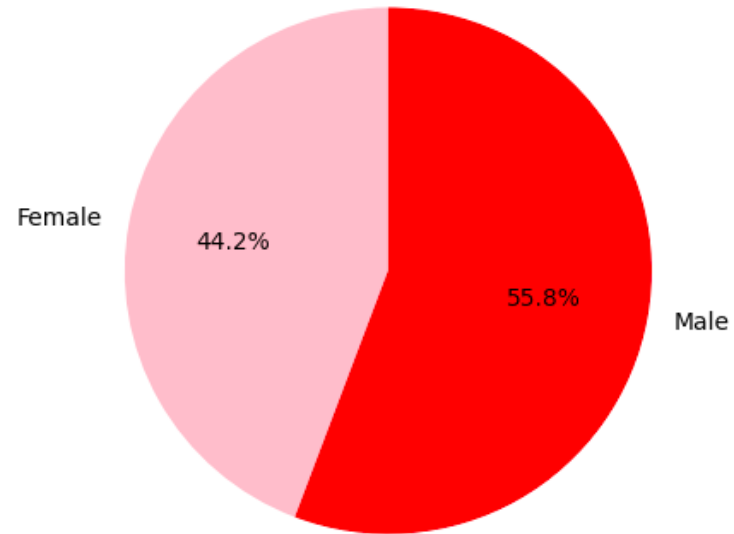
EDA



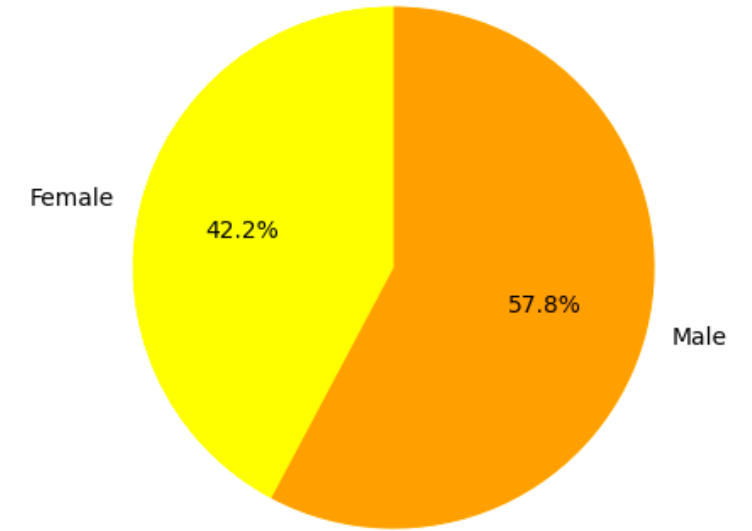
- Both cab companies had a significant jump in the number of cab users between year 2016 and 2017, both follow similar trends over time, however, as previous findings showed us, Yellow cab has more users, hitting a maximum of 27750 users in the year 2017 as opposed to 18500 by Pink Cab, nearly a 10000 user margin difference.

EDA

Gender Frequency in Pink Cab Riders

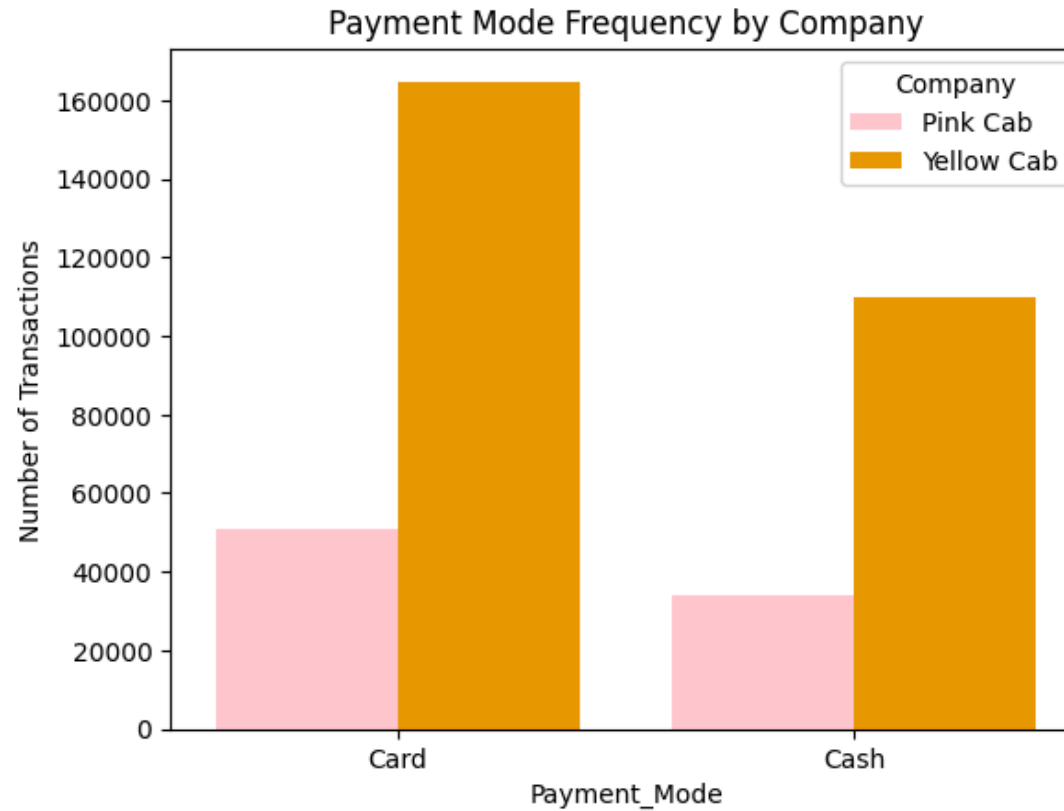


Gender Frequency in Yellow Cab Riders



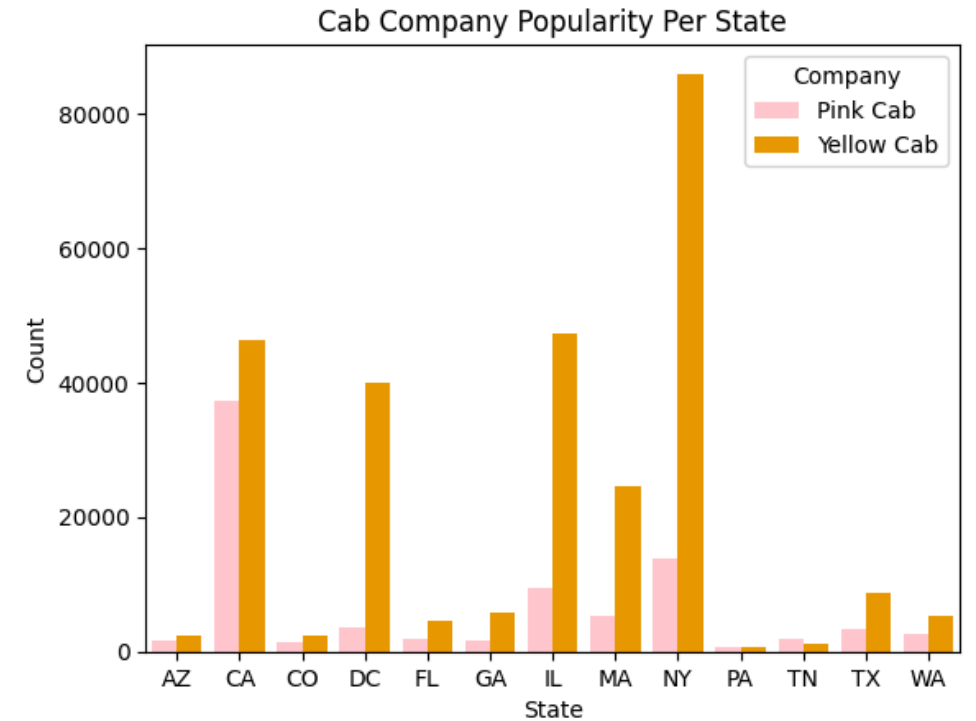
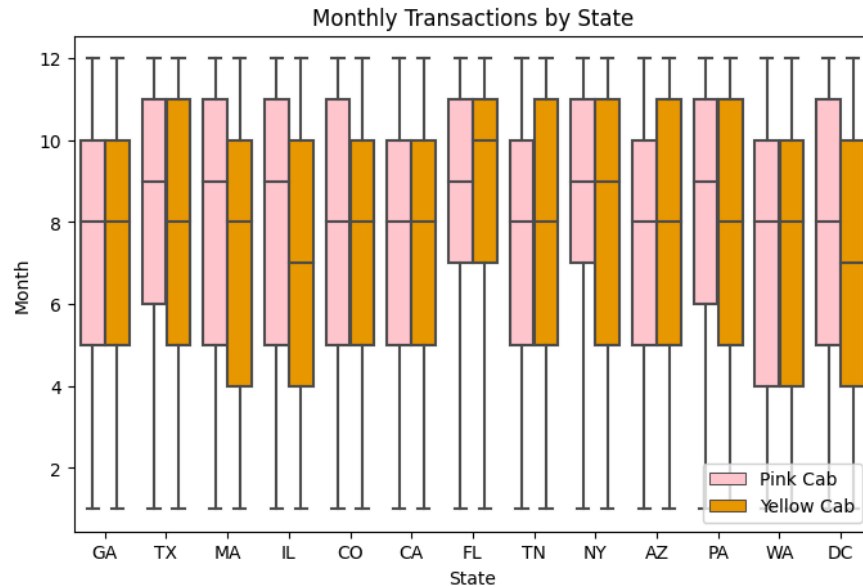
- From these users, more of the riders tend to be male than female, however, Pink Cab has slightly more female users than male. If XYZ choses to invest in Pink Cab, it can be their brand strategy to market themselves to the female crowd for potential growth.

EDA



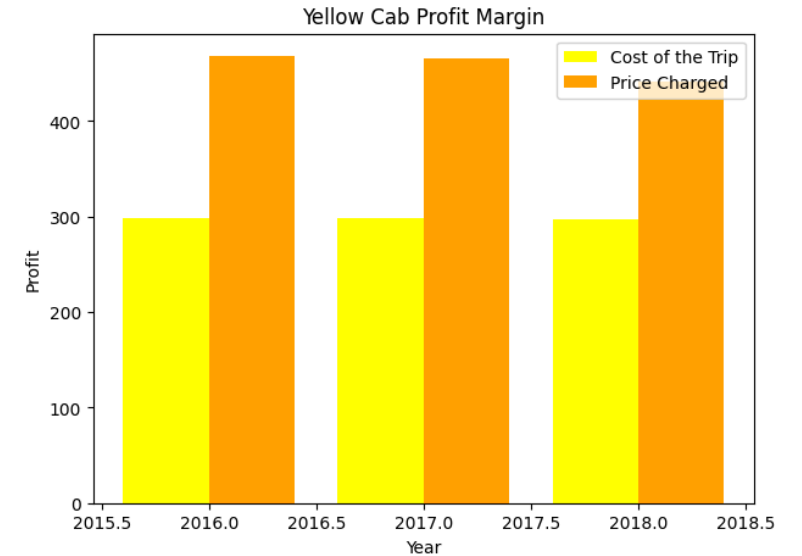
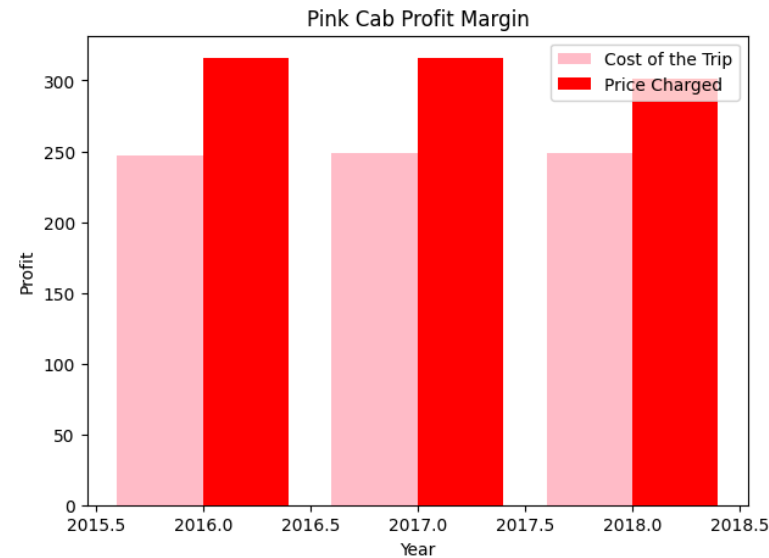
- Amongst the users of both companies, more users use Card payment methods as compared to Cash, making electronic payments a key priority.

EDA



- Although Yellow Cab has more users, we can also see this is due to it's high popularity in Metropolitan and crowded cities like New York City, Chicago, DC and California.
- No matter what state, and time of the year, Yellow Cab tends to be a popular choice for it's users.
- A key thing to note though is that Pink Cab almost meets Yellow cab with its popularity in California.

EDA SUMMARY



- From the graphs, we can see that Yellow Cab has a higher profit Margin (Price Charged - Cost of Trip) as compared to Pink Cab.

Recommendation

- From our previous insights, we also discovered that yellow cab has a far wider customer network, hence making it the better choice for XYZ's future investment.
- However, if XYZ decides otherwise, key things to take away while investing in Pink Cab is that
 1. California is a good State for growth for Pink Cab
 2. Pink Cab can be marketed towards Women and promotes as a safe cab company for female riders.

Thank You
By: Devika Chandnani