



Data Glacier

Your Deep Learning Partner

G2M insight for Cab Investment Firm

Data Science Virtual Internship

Name : Yashwanth Bharadwaj V

Date: 17 April 2023

Agenda

Executive Summary
Approach
EDA
Recommendations

Executive Summary

- Data 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.
- **Objective:** Provide actionable insights for the firm to identify the right company to make their investment

Approach

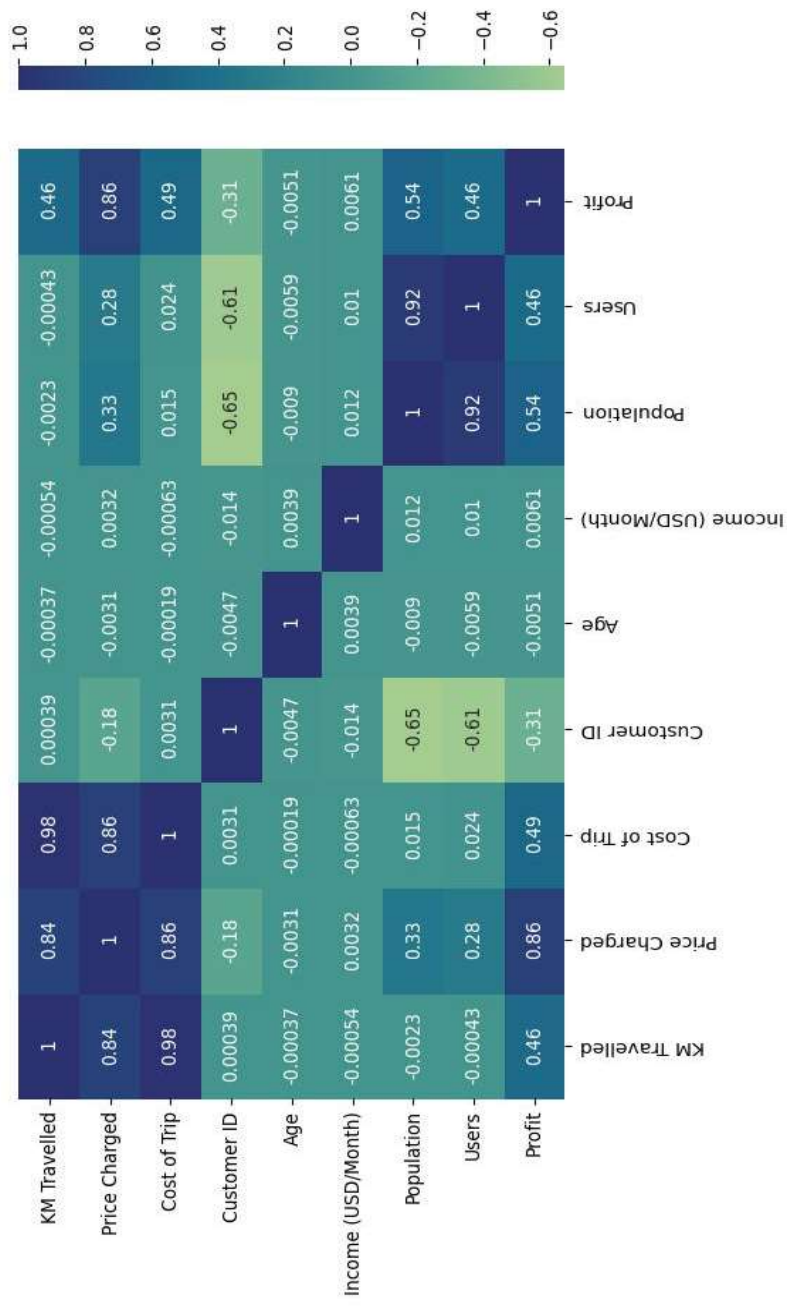
- Data was obtained through GitHub
- The .csv files were checked for any outliers and null values and were cleaned if found
- Data types of the attributes were converted to their actual types
- New attributes were created off the existing columns for easier data analysis
- All the data frames are merged based on the attribute that was same across the two merged data frames until we have a single data frame

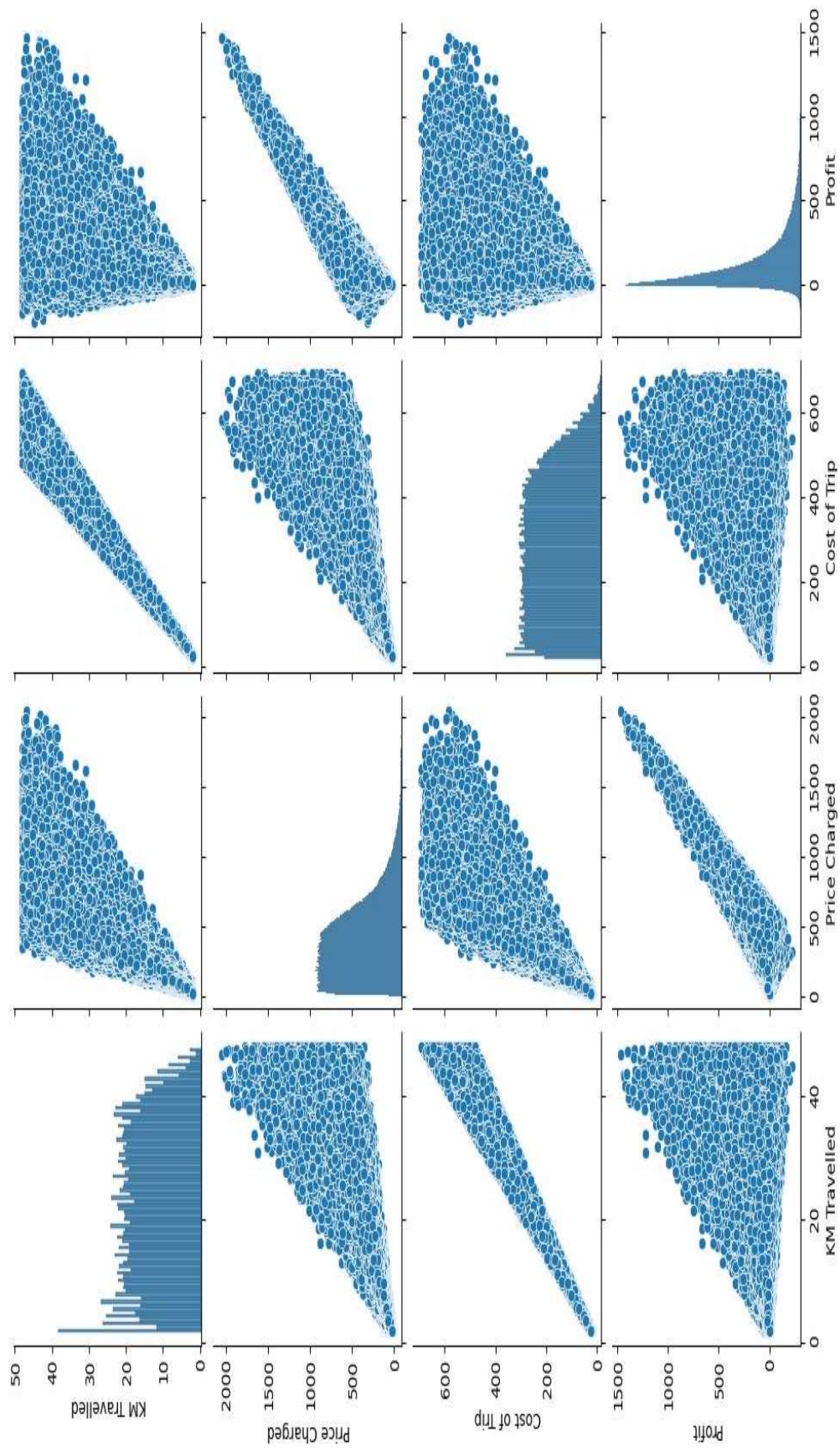
Data Information

- **Cab_Data.csv** - includes details of all the transactions of yellow and pink cab companies
- **City.csv** - includes the list of US cities, their population and the number of cab users
- **Customer_ID.csv** - contains a unique identifier that has the user information
- **Transaction_ID.csv** - contains a unique identifier that maps to customer table and has the payment mode

Correlation between Attributes

- Strong Correlation between
 - Population vs Users
 - KM Travelled vs Price Charged vs Cost of Trip

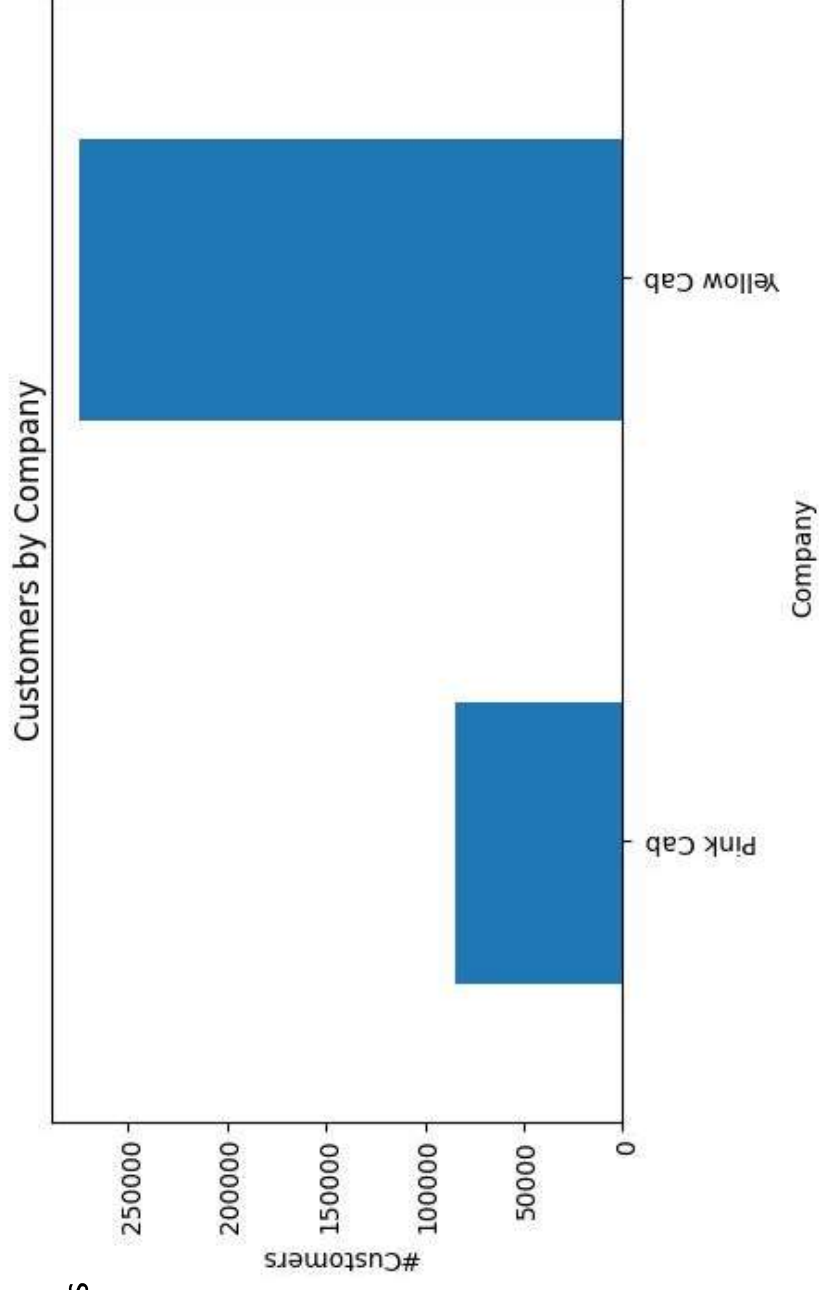




Exploratory Data Analysis

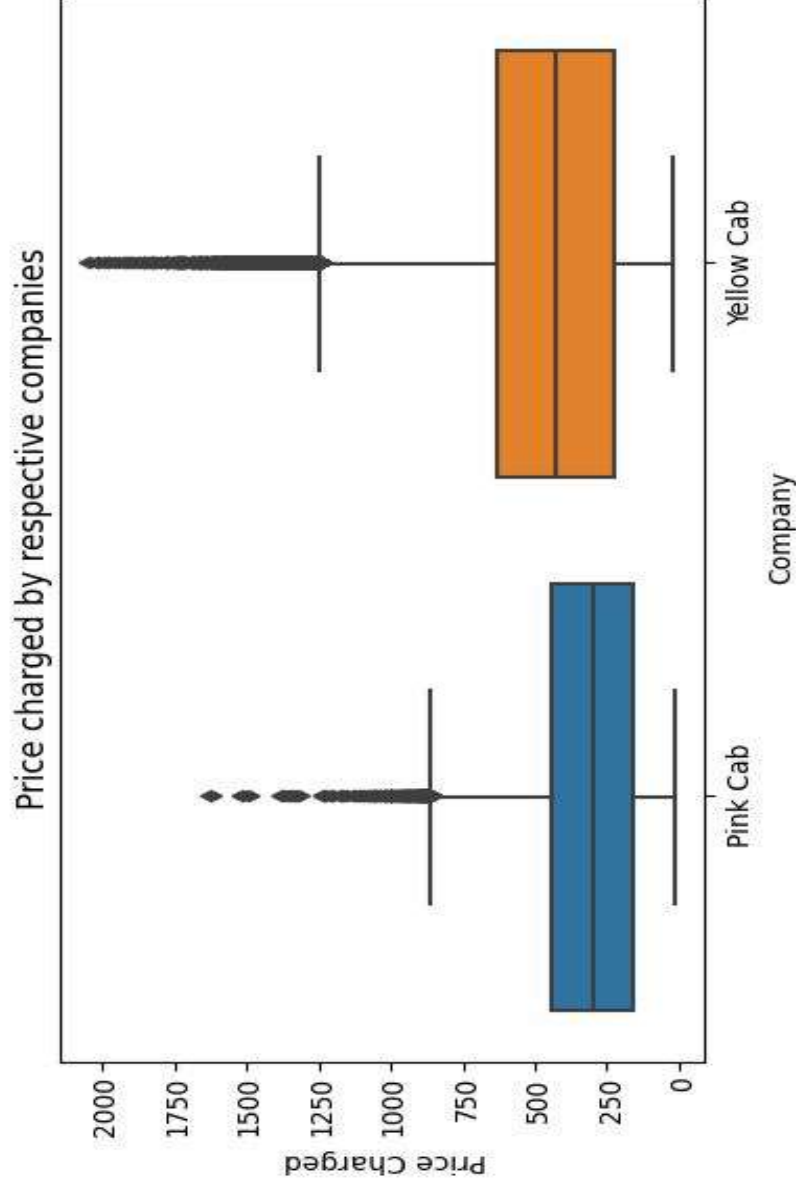
Company with most Users

- Yellow cab has more users compared to pink cab



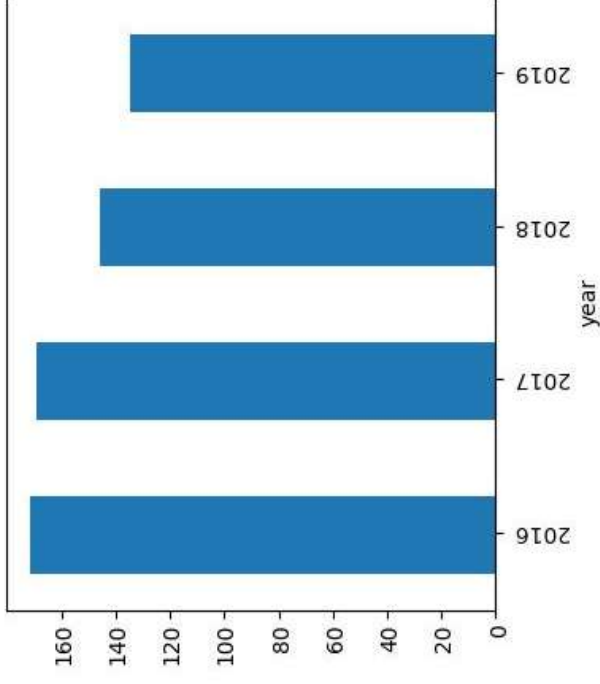
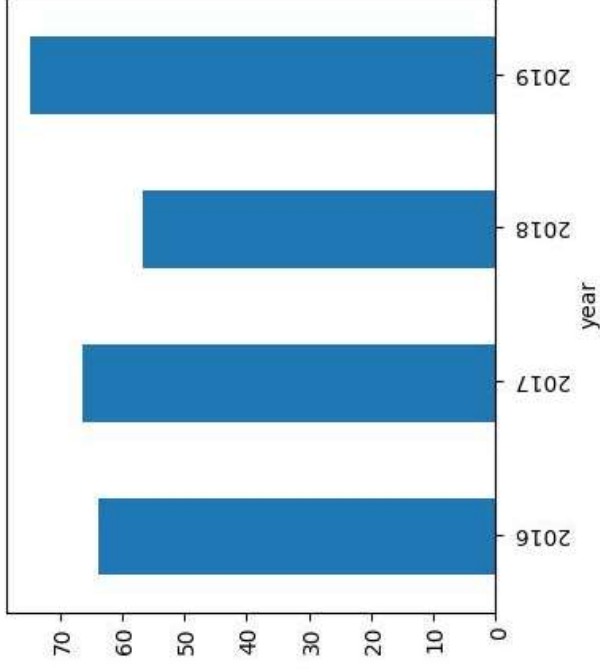
Price charged by the Companies

- Yellow cab charges more per user on average compared to the pink cab
- Despite this fact, users prefer yellow cabs as can be seen from the previous slide



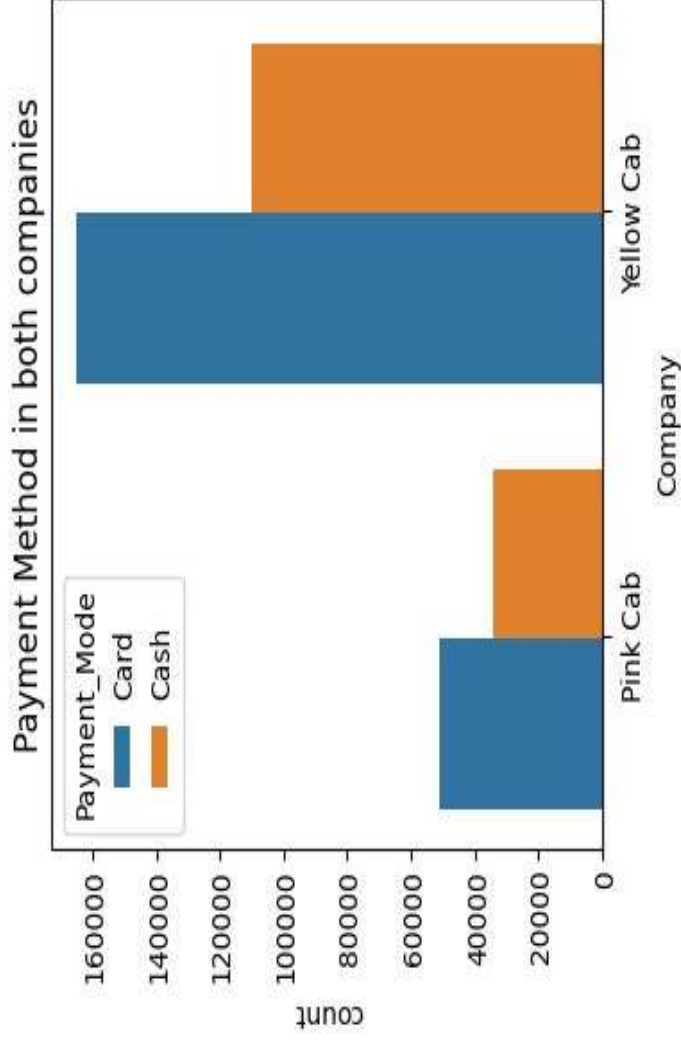
Profit Over the Years

Yellow cab company has seen a slight decrease in profits over the last 2 years but it has the highest net profit across the time period.



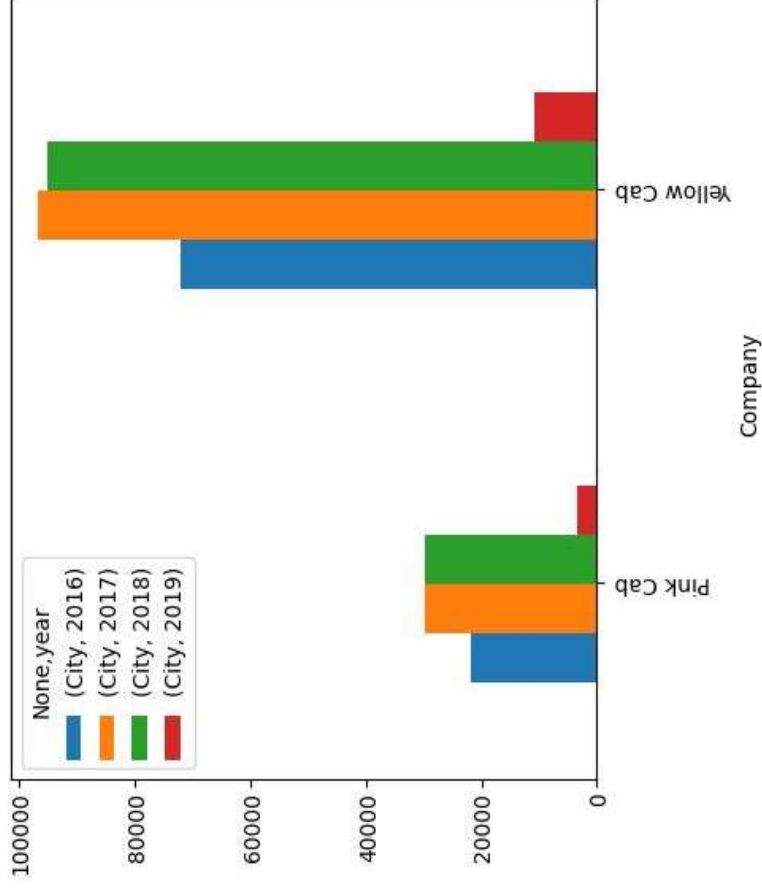
Payment Mode

- Users availing services from both the companies prefer card over cash.



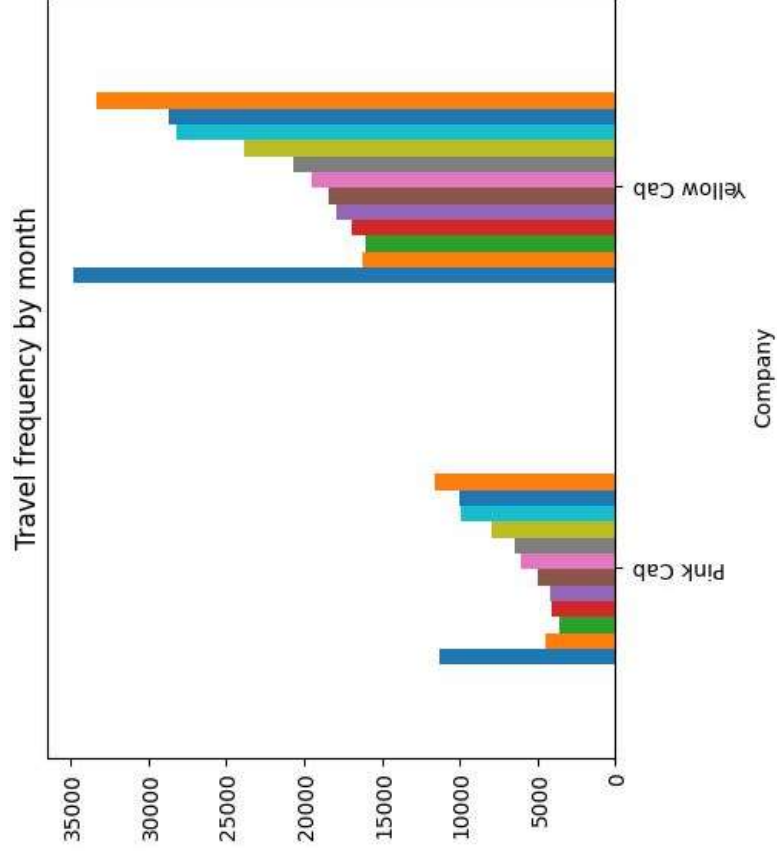
Company Users over the years

- The number of users using the cab services increased over the last 3 years
- The percentage increase of the number of users of yellow cab company is comparatively more than that of the pink cab company



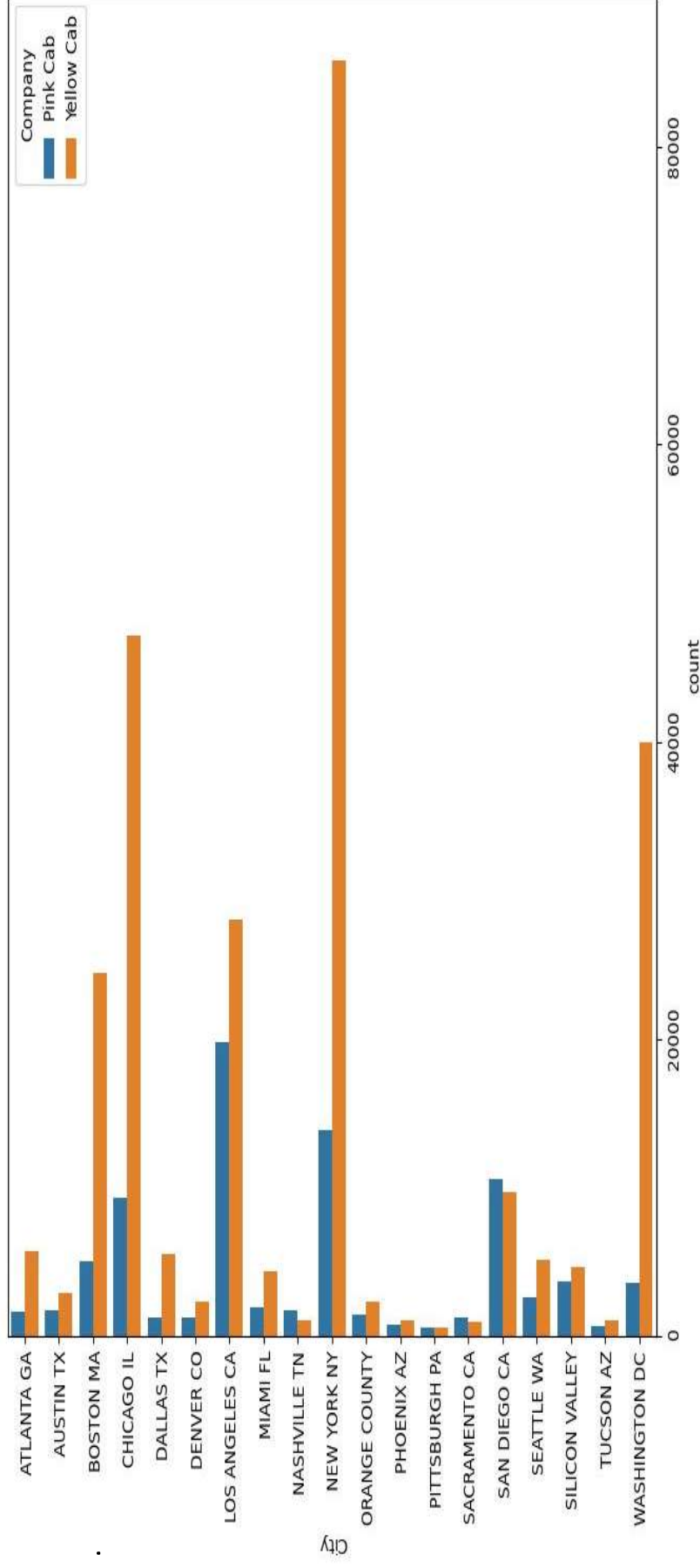
Travel Frequency by Month

- Most of the traffic for both the cab companies is during fall and January as that is the holiday season.



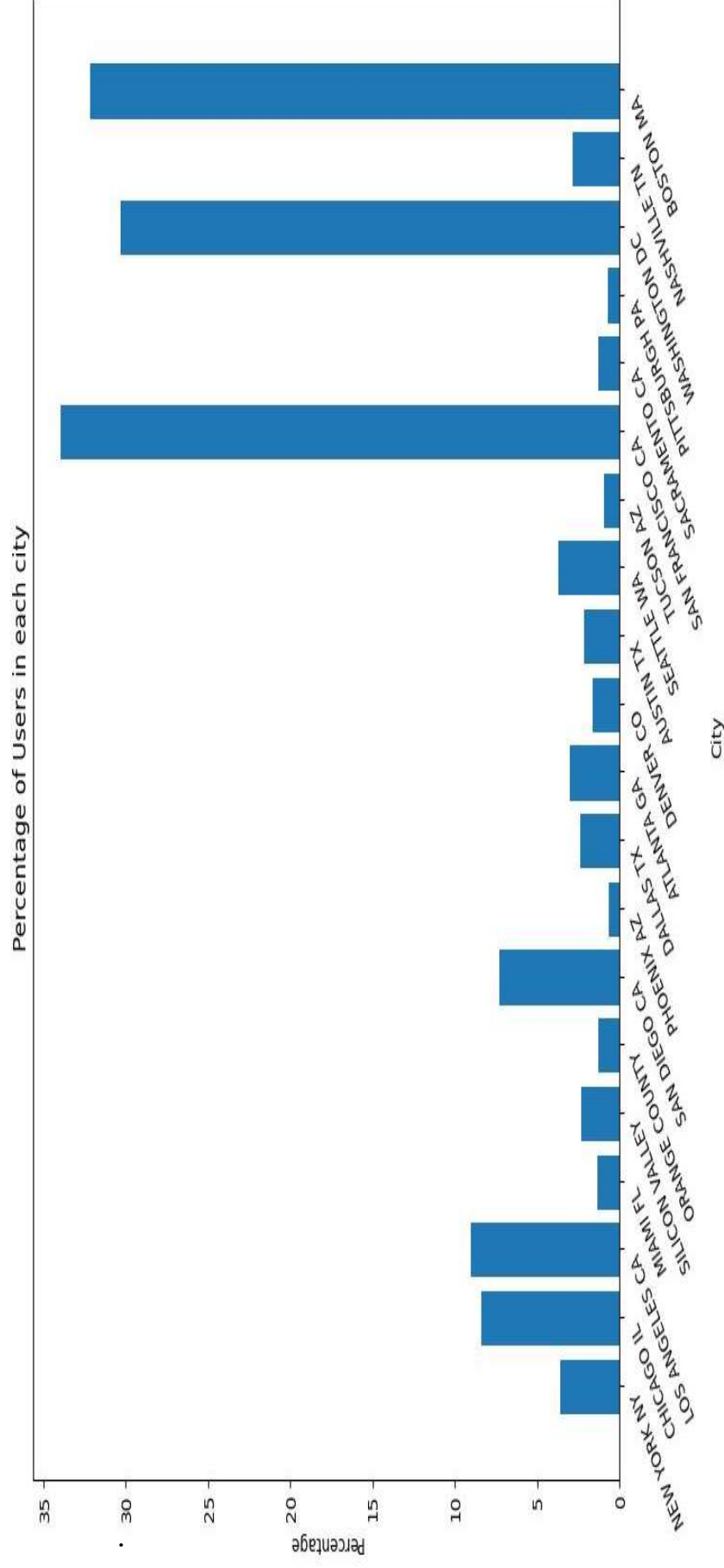
Company's User base

Most of the user base is centered around New York, Los Angeles, Chicago and Washington for both the companies and all of these cities's users prefer yellow cabs



Users percentage by city

Most users are located in Washington, Chicago, Boston and San Francisco.



Recommendations

- Yellow cab company has more users than Pink cab company
- Despite the fact that yellow cabs charge higher in average to pink cabs, users tend to prefer yellow cabs more indicating its high demand
- The annual net profit of both the companies increased YOY but yellow cabs has a huge margin over pink cabs
- Number of users availing the services of both the companies increased over the last 3 years
- Most of the user base is located in New York, Los Angeles, Chicago, Washington and Boston. Again, most users from these cities chose yellow cabs
- Percentage of unique users are located in Chicago, Washington, Boston and San Francisco. These numbers correspond to the total user base numbers from the respective cities
- New York and Los Angeles has less users and yet they are the cities that are availing more cab services and that too yellow cab services which tend to put more weight on the analysis. Thus, investing in Yellow cab company is recommended backed up by this analysis

Thank You