



Data Glacier

Your Deep Learning Partner

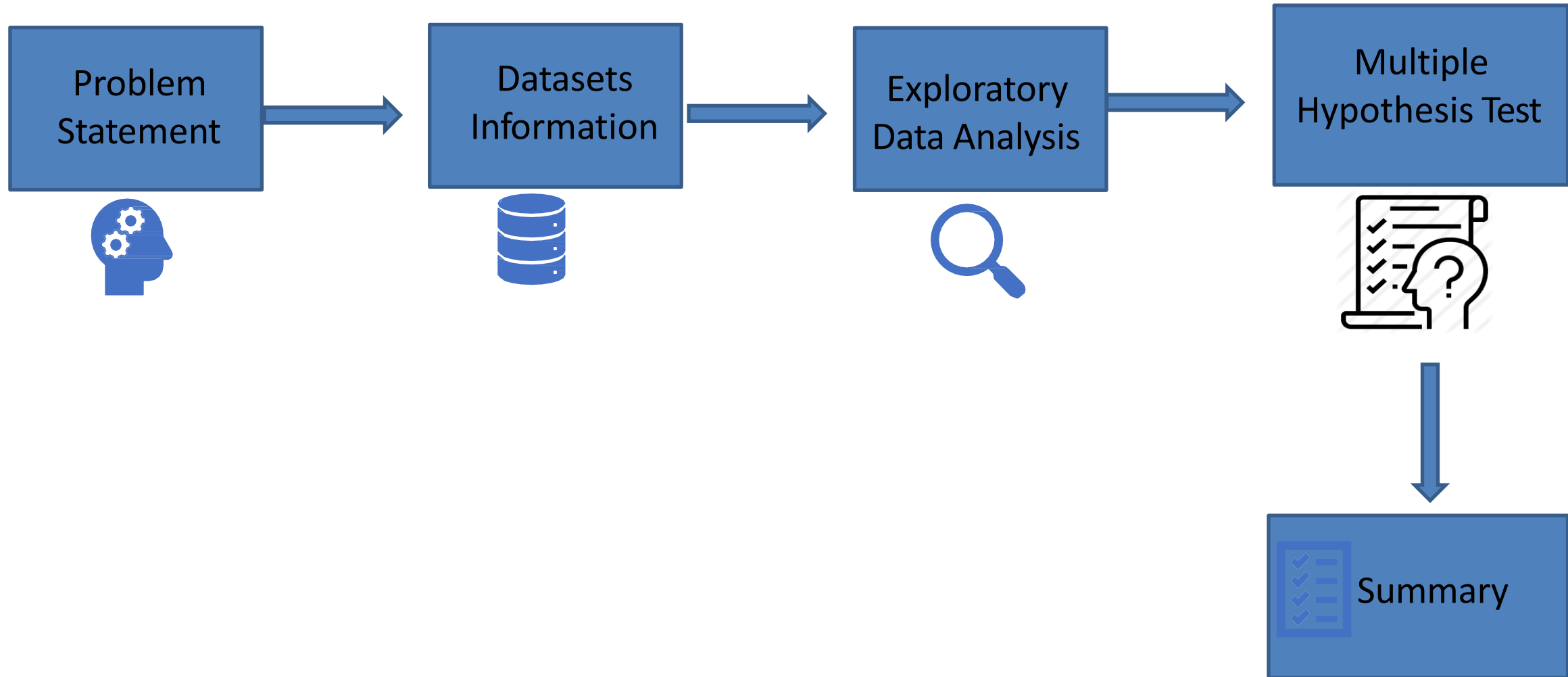
Project: G2M Insight For Cab Investment Firm

Data Science Virtual Internship

Tharun Kumar Kouda

07-DEC-2022

Outline



Problem Statement

- XYZ is a private equity 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.
- Objective: Provide actionable insights to help XYZ firm in identifying the right company for making an investment.

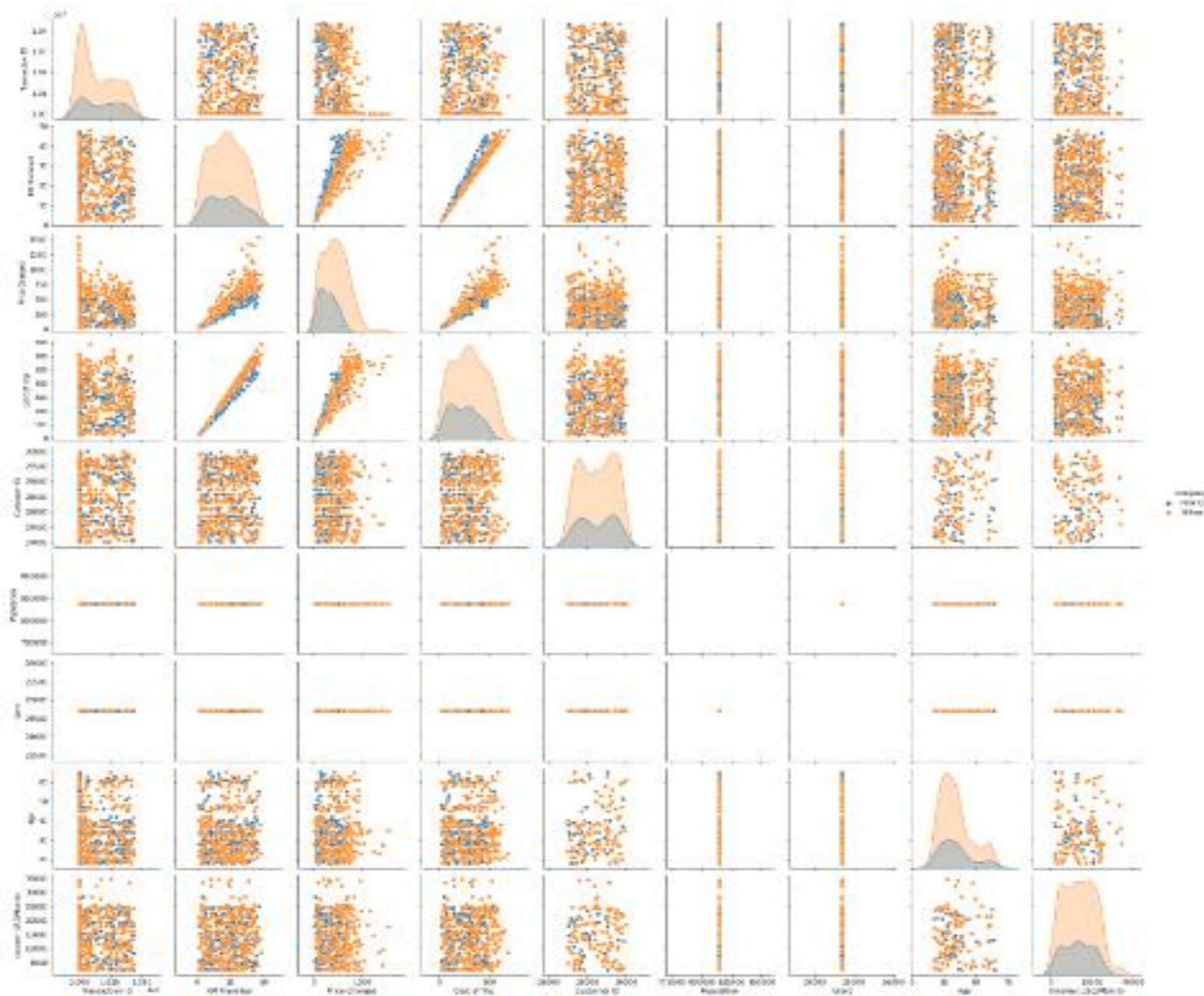
The analysis has done in following ways:

- Data Understanding and Visualization
- Finding the most users Cab Company
- Finding the cheapest Cab company for users
- Finding the most profitable Cab Company
- Multiple Hypothesis and Investigate

Data Information

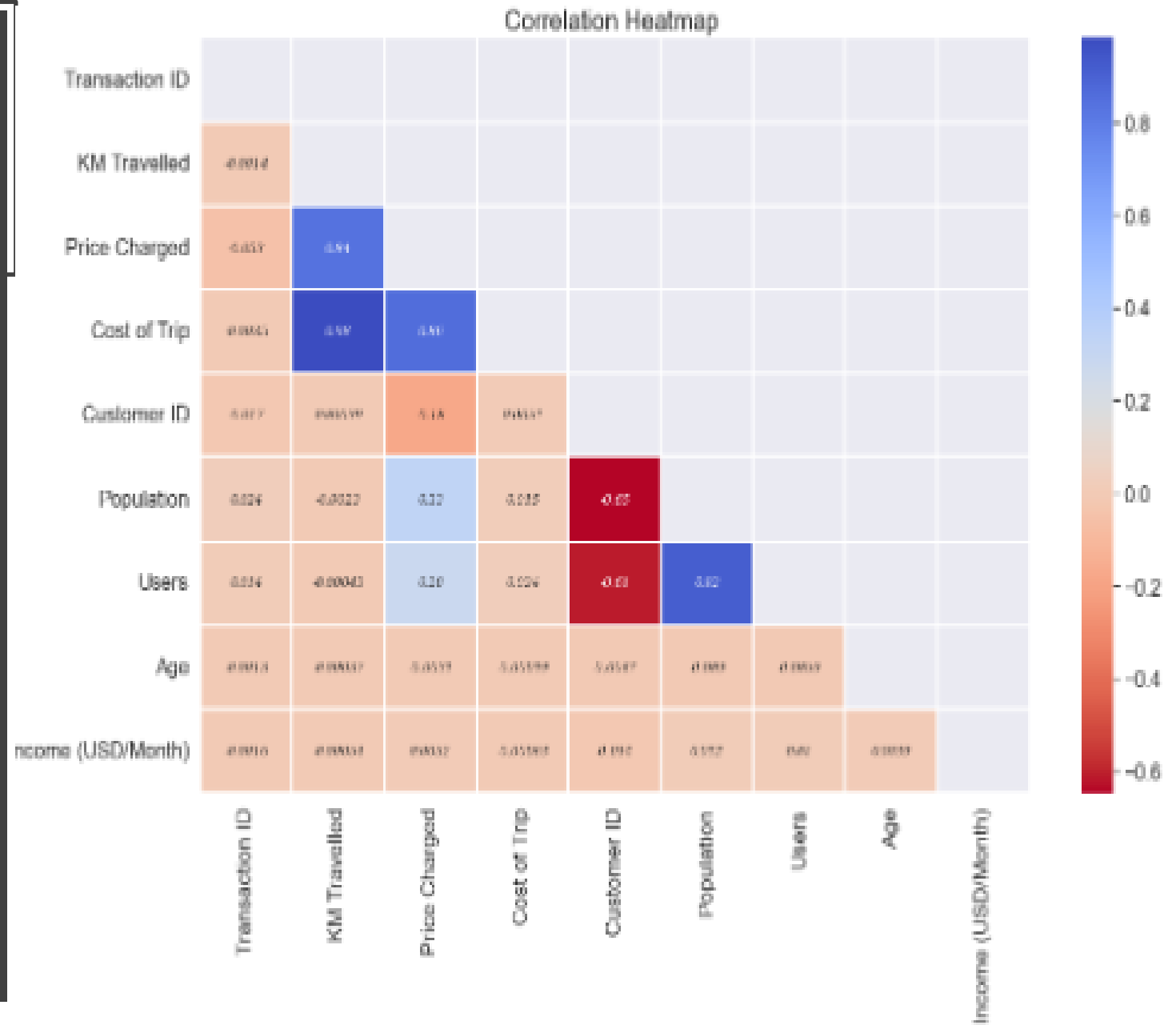
- Cab_Data.csv – this file includes details of transactions for 2 cab companies
- Customer_ID.csv – this is a mapping table that contains a unique identifier that links the customer's demographic details
- Transaction_ID.csv – this is a mapping table that contains transaction to customer mapping and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

Relationships Between Variables



Correlation Between Variables

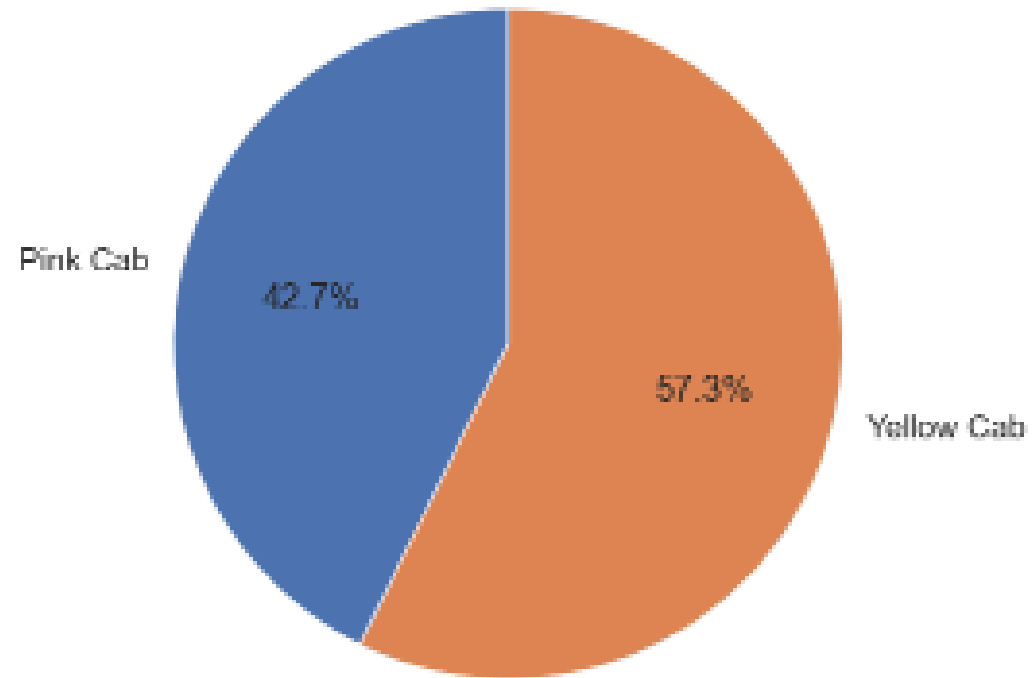
price charged - KM travelled , KM travelled - cost of Trip and Population - users are having the positive correlation which are indicated in the blue color



Which Company has more Users?

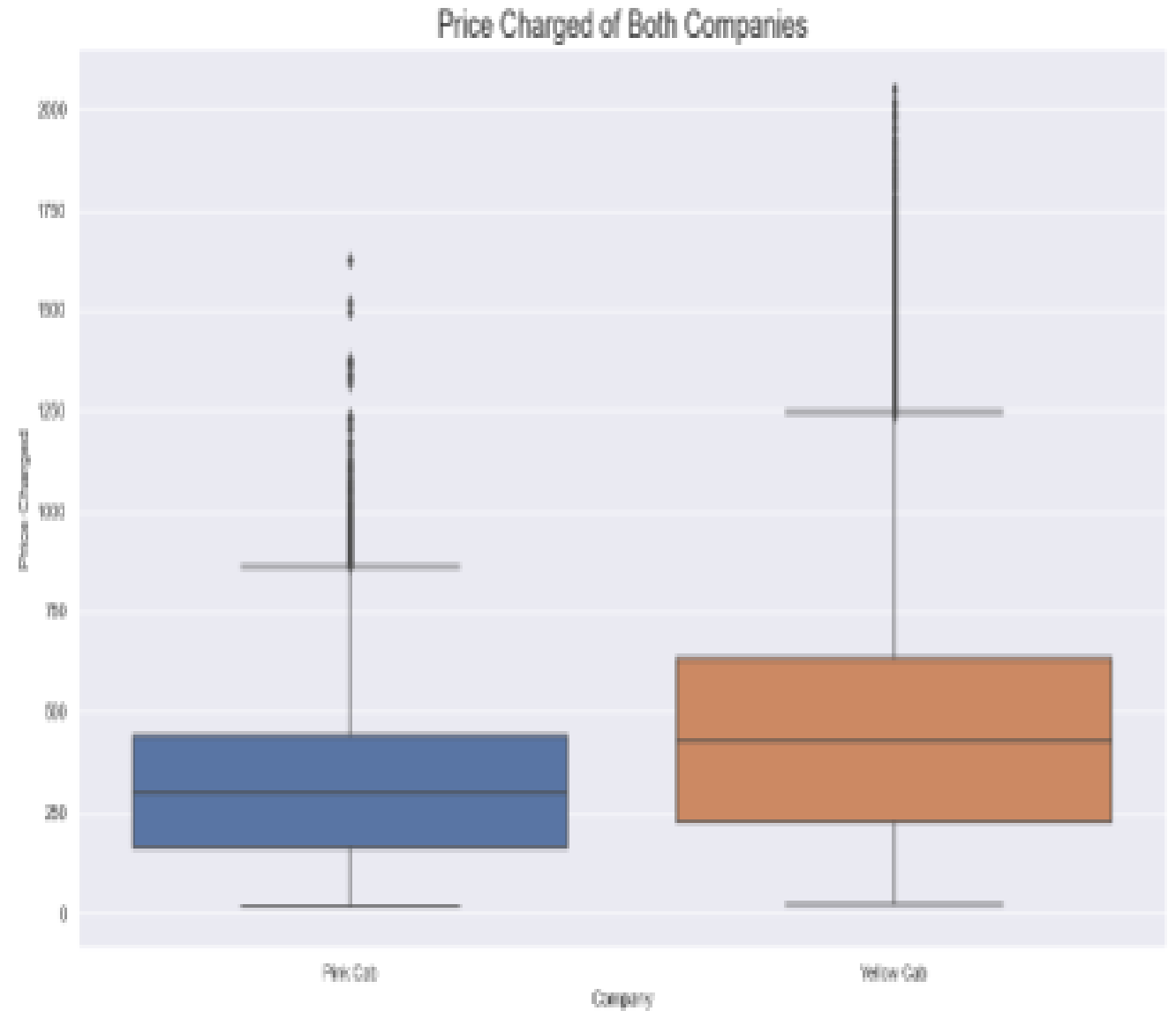
Pink cab services are having '42.7' percent and Yellow cab services are having '57.3' percent of customer shares respectively

Customer distribution between Yellow and Pink cab services.



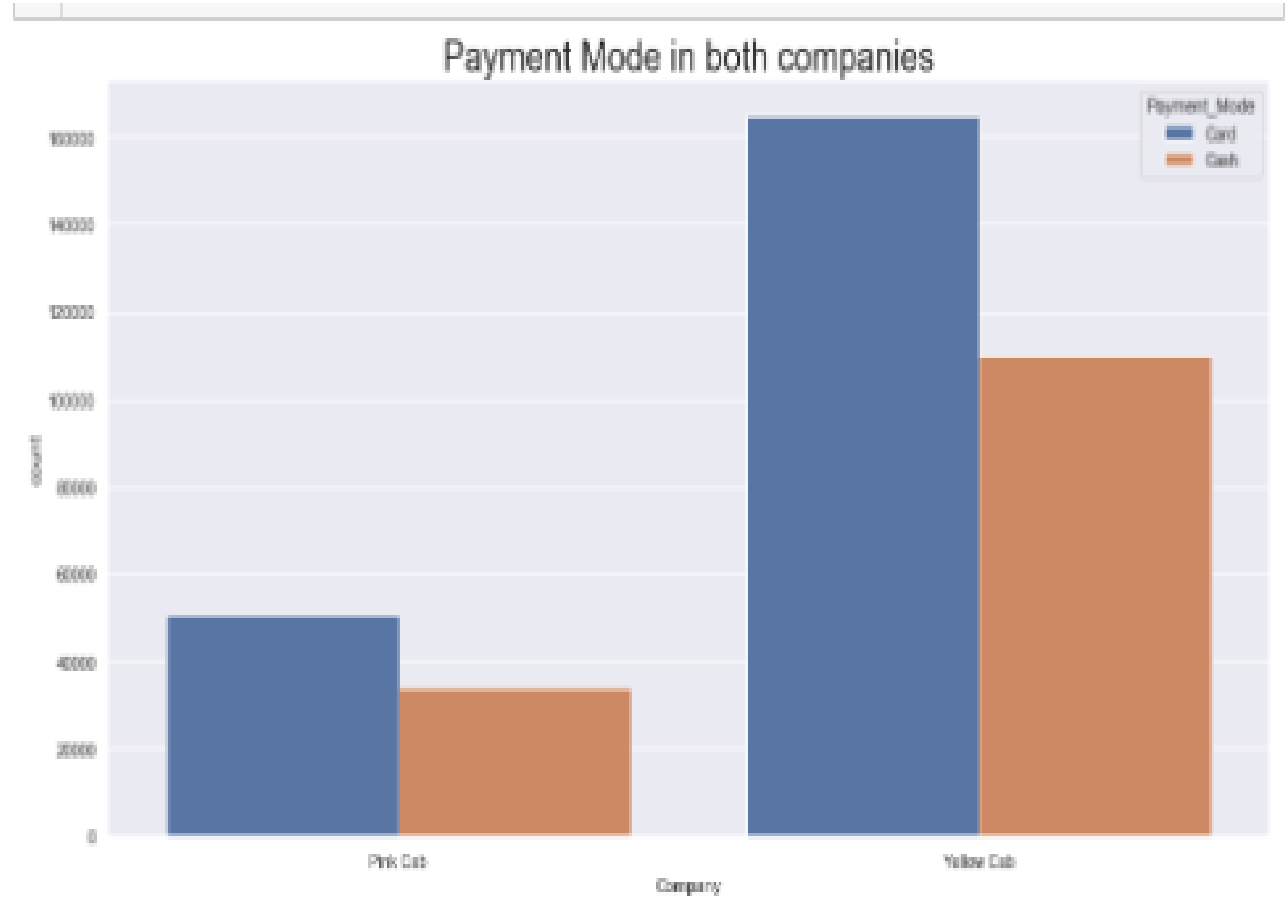
Which Company has a high price charged?

- As we can see Price Charged for Yellow Cab is highest as compared to Pink Cab



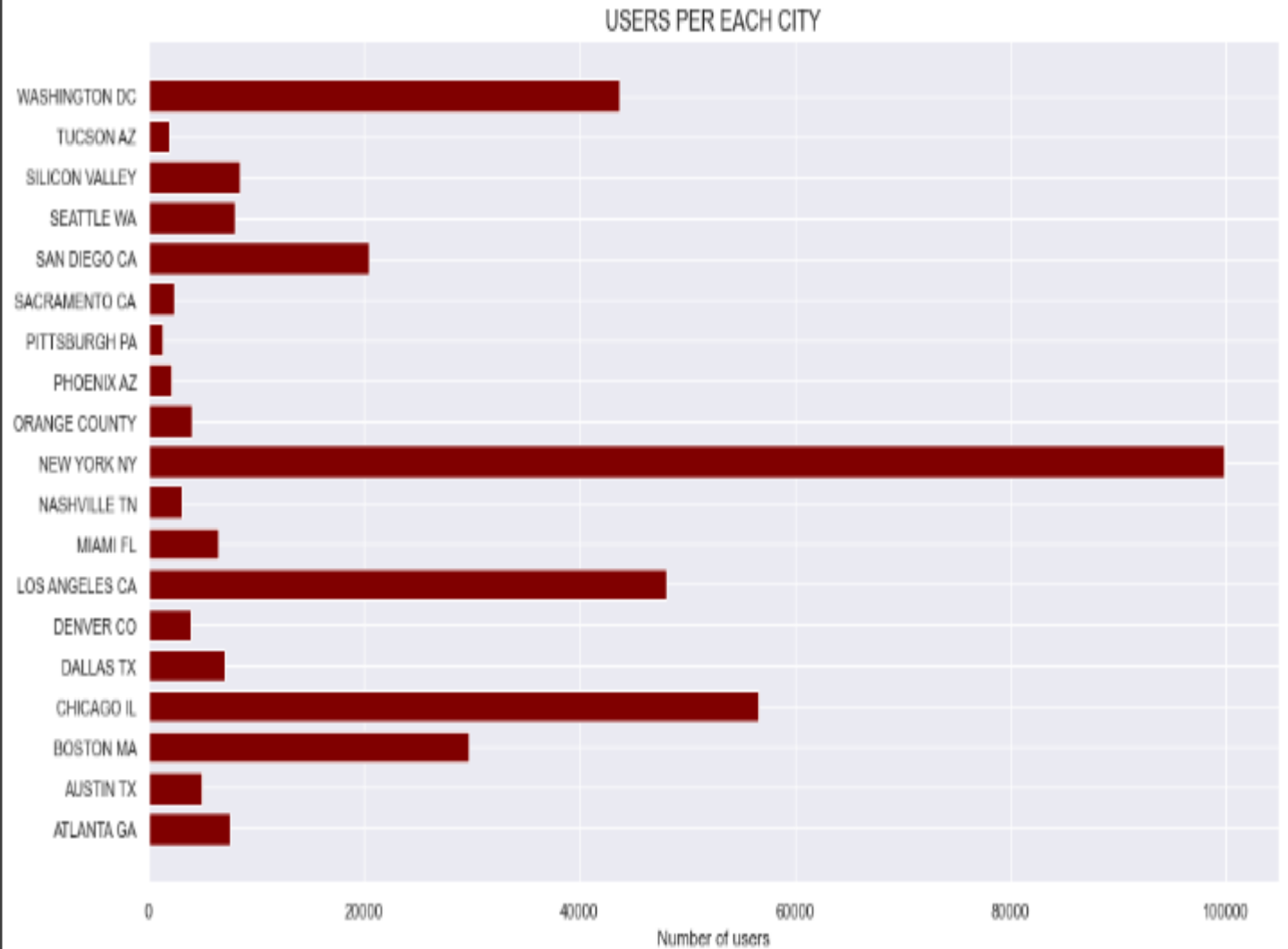
Payment Mode

Pink cab services are having '42.7' percent and Yellow cab services are having '57.3' percent of customer shares respectively



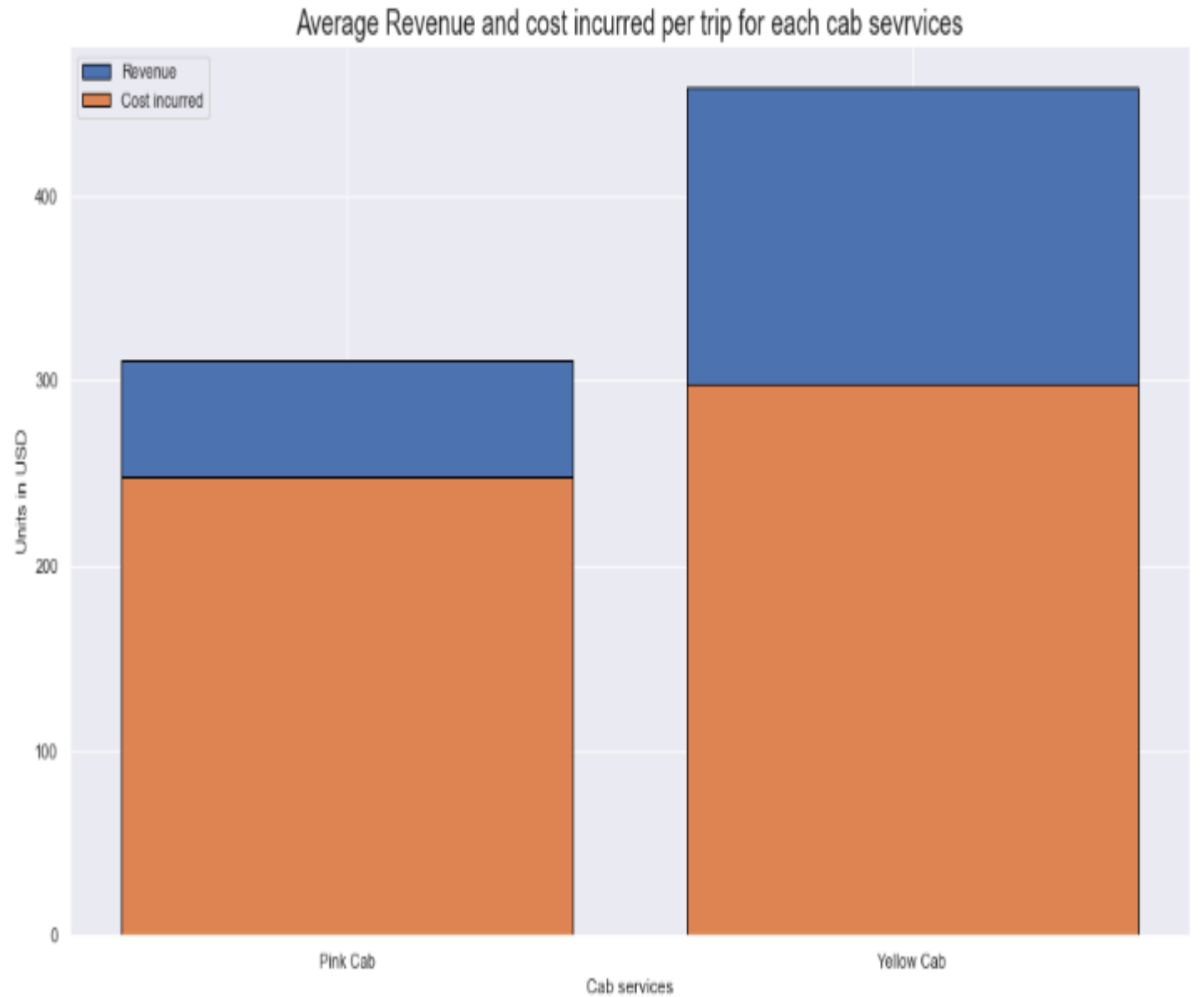
Users w.r.t Cities

- New York City has the highest Cab users with 28% followed by Chicago with 16% and Los Angeles with 13%



Profit Margin

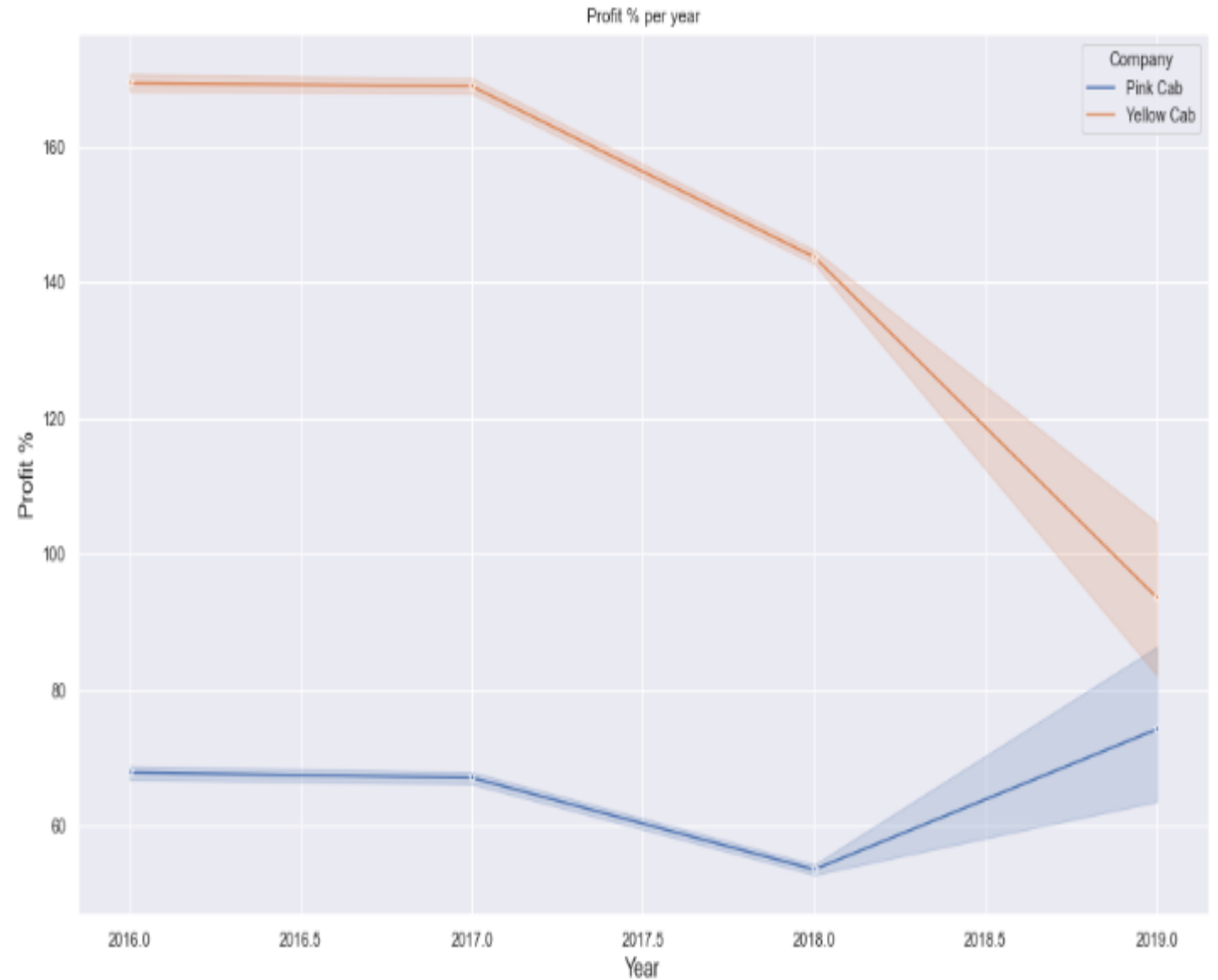
- The Yellow cab has a higher Profit Margin (Price Charged - Cost of Trip) compared to Pink cab



Profit Margin w.r.t Time

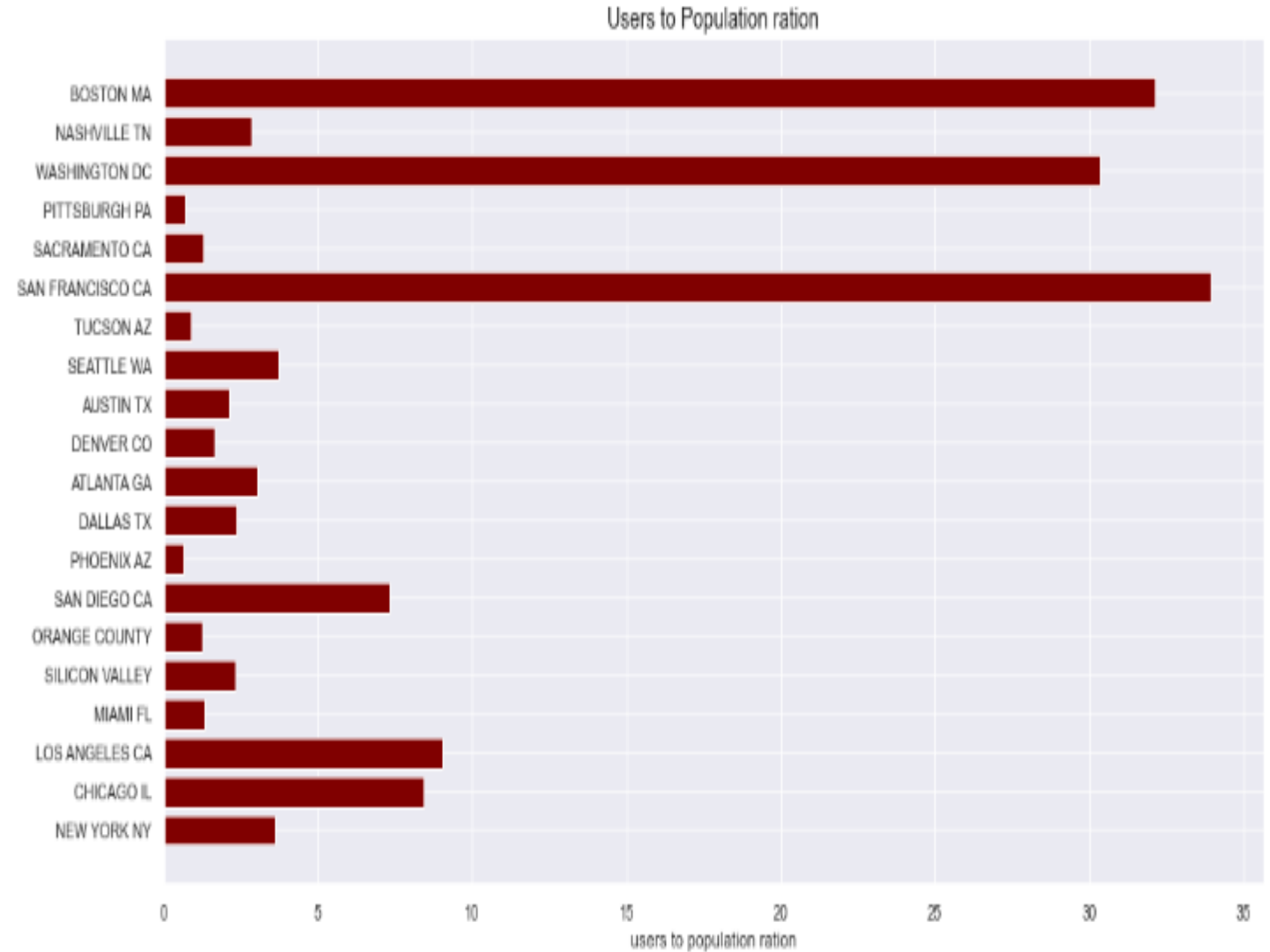
* profit of pink cab services decreased in 2018, but it raised back in 2019

* profit of Yellow cab services decreased after 2017



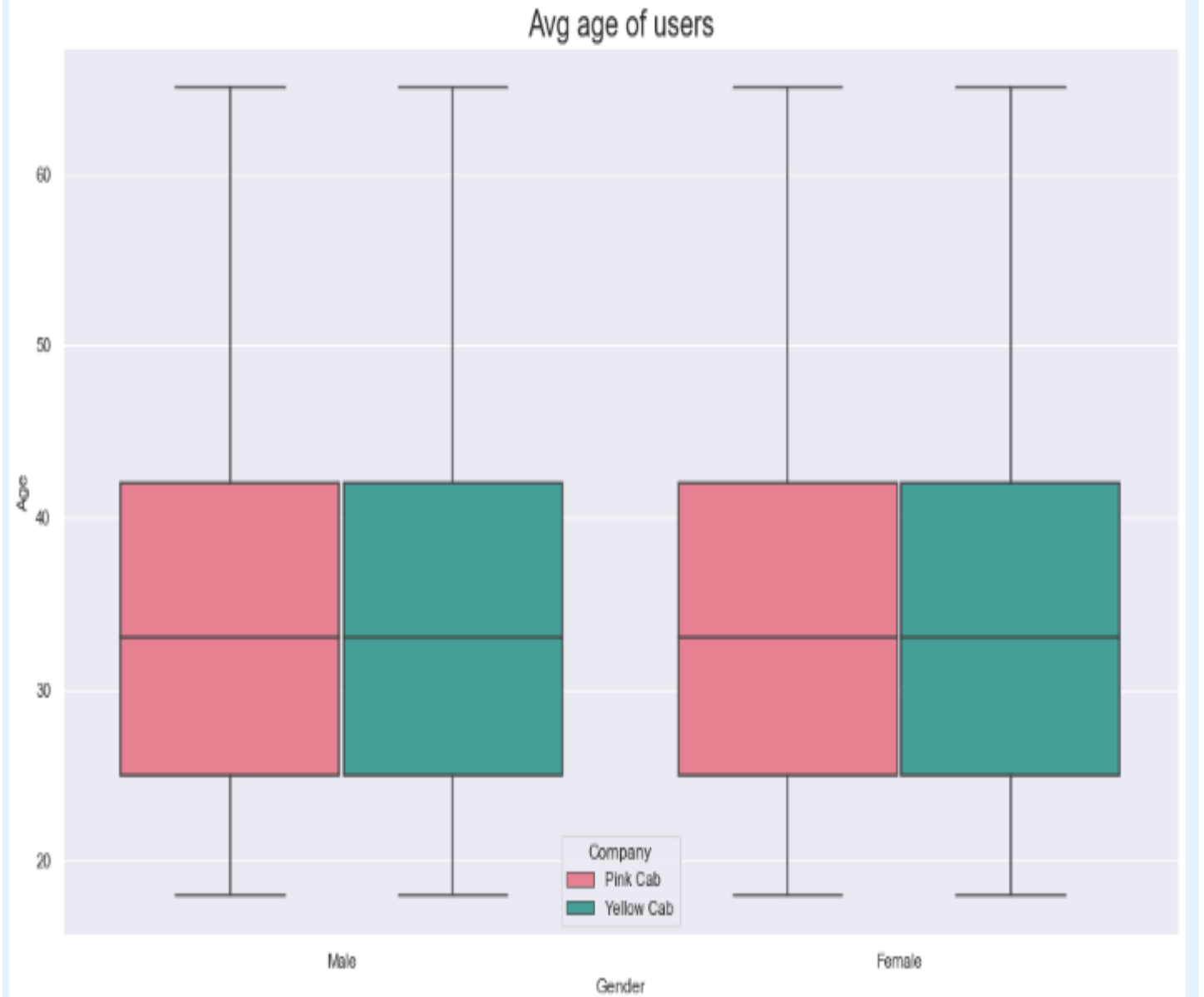
Users w.r.t Population

- As we can see in the cities of San Francisco, Washington, and Boston more than 30% of the population use cab service



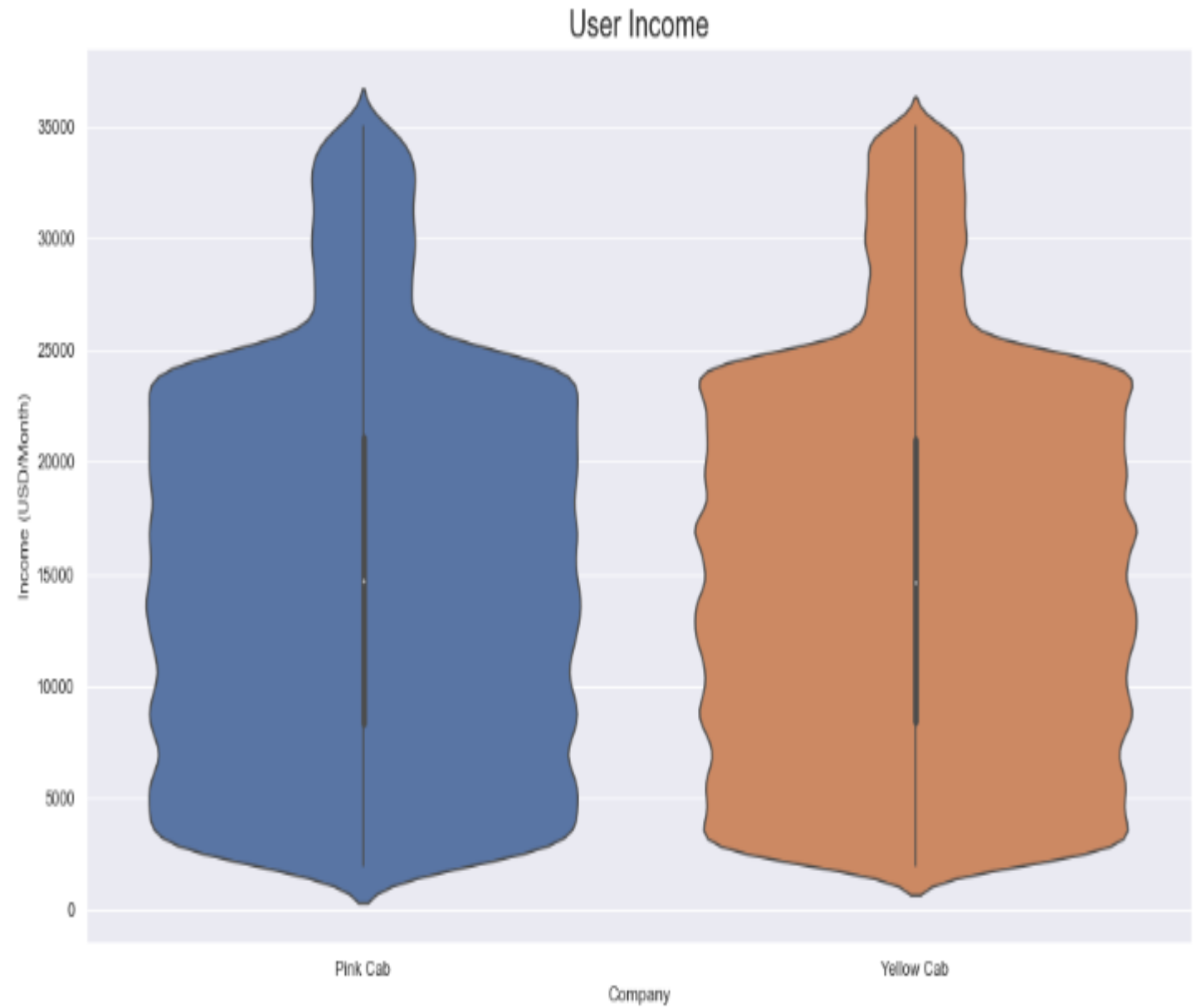
Average Age of Users

Average age of customers are 35 years for both male and female customers



Average Income of Users

Avg income of cab services is around `15k\$`



Hypothesis 1: Is there any difference in profit regarding Gender

H0 (Null hypothesis) : There is no difference regarding Gender in both cab companies.

H1 (Alternative hypothesis) : There is a difference regarding Gender in both cab companies.

Pink Cab Company

47231 37480

P value is 0.11515305900425798

We accept null hypothesis (H0) that there is no difference regarding gender for Pink Cab

Yellow Cab Company

P value is 6.060473042494144e-25

We accept alternative hypothesis (H1) that there is a difference regarding gender for Yellow Cab

Conclusion

There is no difference regarding Gender in both cab companies

Hypothesis 2: Is there any difference in Profit regarding Age

H0: There is no difference regarding Age in both cab companies.

H1: There is a difference regarding Age in both cab companies.

Pink Cab Company

P value is 0.4816748536155635

We accept null hypothesis (H0) that there is no difference regarding age for Pink Cab

Yellow Cab Company

P value is 6.328485471267631e-05

We accept alternative hypothesis (H1) that there is a difference regarding age for Yellow Cab

Conclusion

Looks like Yellow Cab company offers discounts for their customers who are older than 60 years old.

Hypothesis 3: Is there any
difference in Profit regarding Payment mode

**H0 : There is no difference regarding
Payment_Mode in both cab companies.**

**H1 : There is a difference regarding
Payment_Mode in both cab companies.**

Pink Cab Company

P value is 0.7900465828793288

We accept null hypothesis (H0) that there is no difference in payment mode for Pink Cab

Yellow Cab Company


P value is 0.2933060638298729

We accept null hypothesis (H0) that there is no difference in payment mode for Yellow Cab

Conclusion

There is no difference in payment mode for both cab companies

Conclusion



Conclusion

Yellow Cab Company is better than Pink Cab Company Because:

- Yellow cabs have a better profit margin
- Have more users than the pink cab company
- As a result more transactions

Hence better to invest in Yellow cab than pink cab company

**THANK
YOU**



Data Glacier

Your Deep Learning Partner