



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

Virtual Internship

21-July-2024

Background –G2M(cab industry) case study

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

The analysis has been divided into four parts:

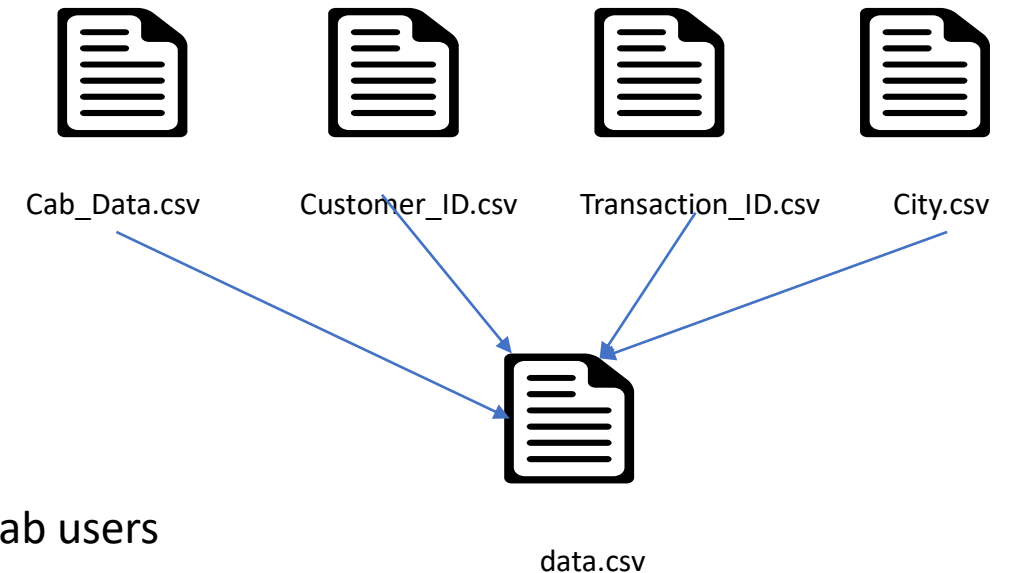
- Data Understanding
- Forecasting profit and number of rides for each cab type
- Finding the most profitable Cab company
- Recommendations for investment

Data Exploration

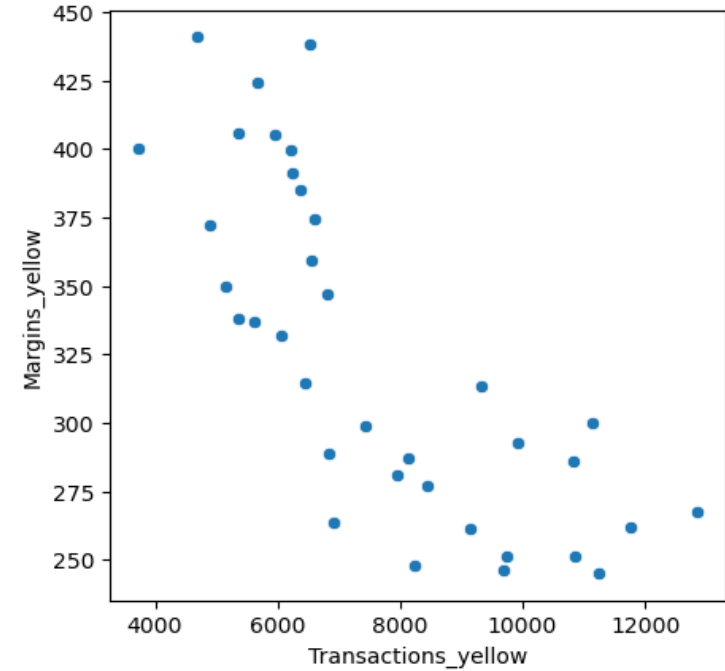
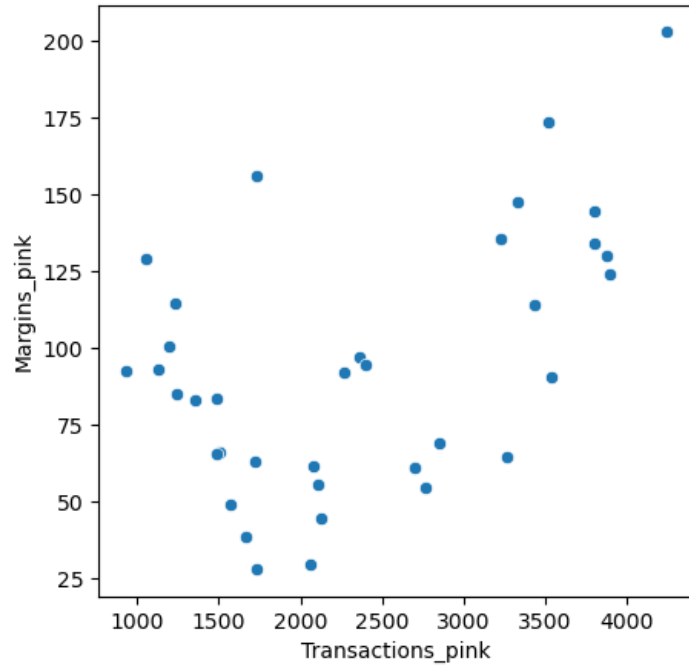
- 12 Features
- Timeframe of the data: 2016-01-01 to 2018-12-31
- Total data points : 3,59,392

Assumptions:

- Outliers are present but due to high end cars included them.
- The Users feature in the city dataset is treated as the number of cab users in the city, assuming it includes users of both Yellow and Pink cabs.
- The profit of rides is calculated by considering only the Price_Charged and Cost_of_Trip features, while keeping other factors constant.

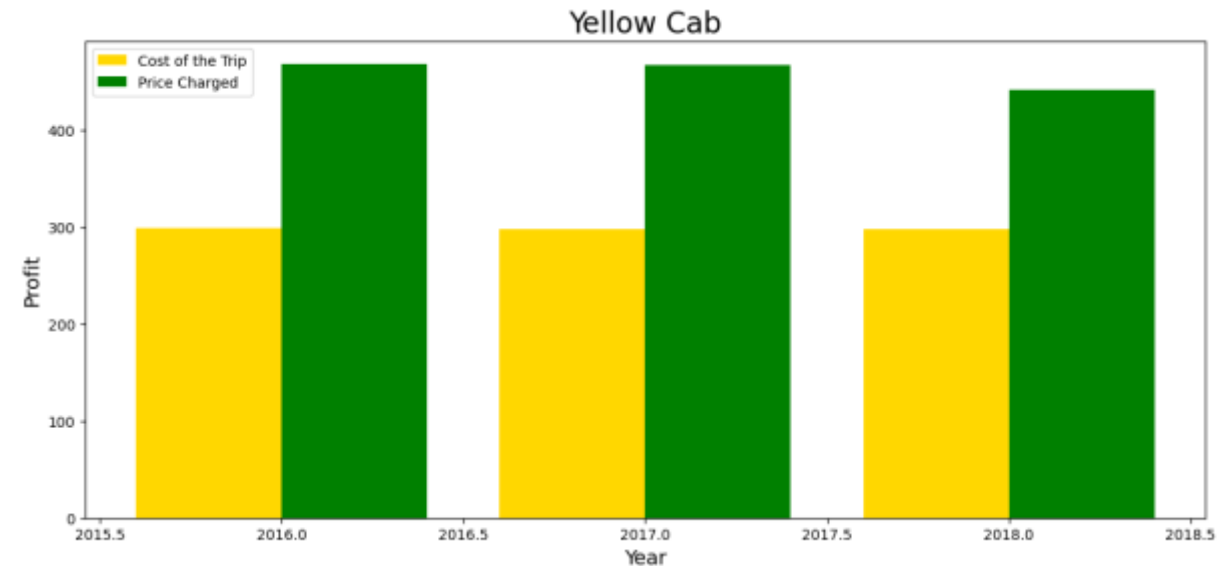
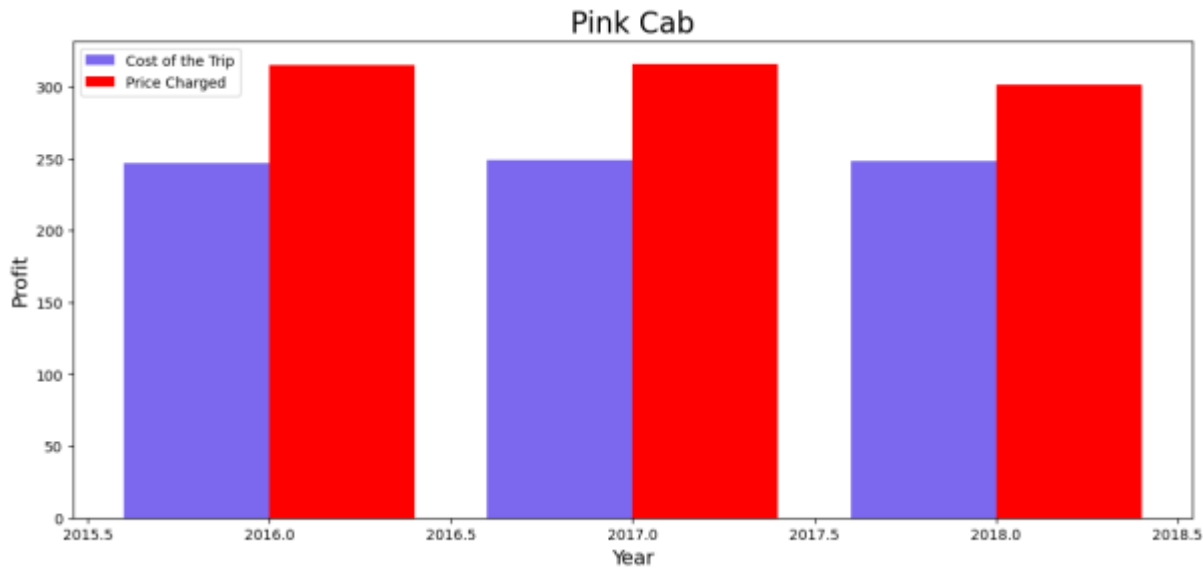


Profit Analysis



- From the above graph, it is clear that yellow cab have more profits based on transactions compared to pink cab.

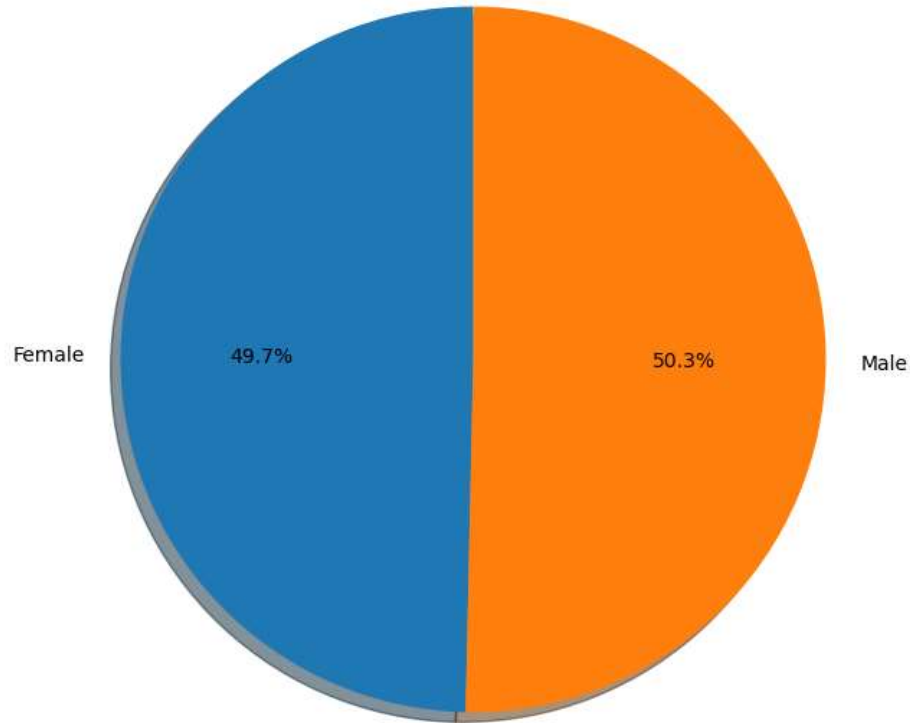
Yearly Profit Analysis



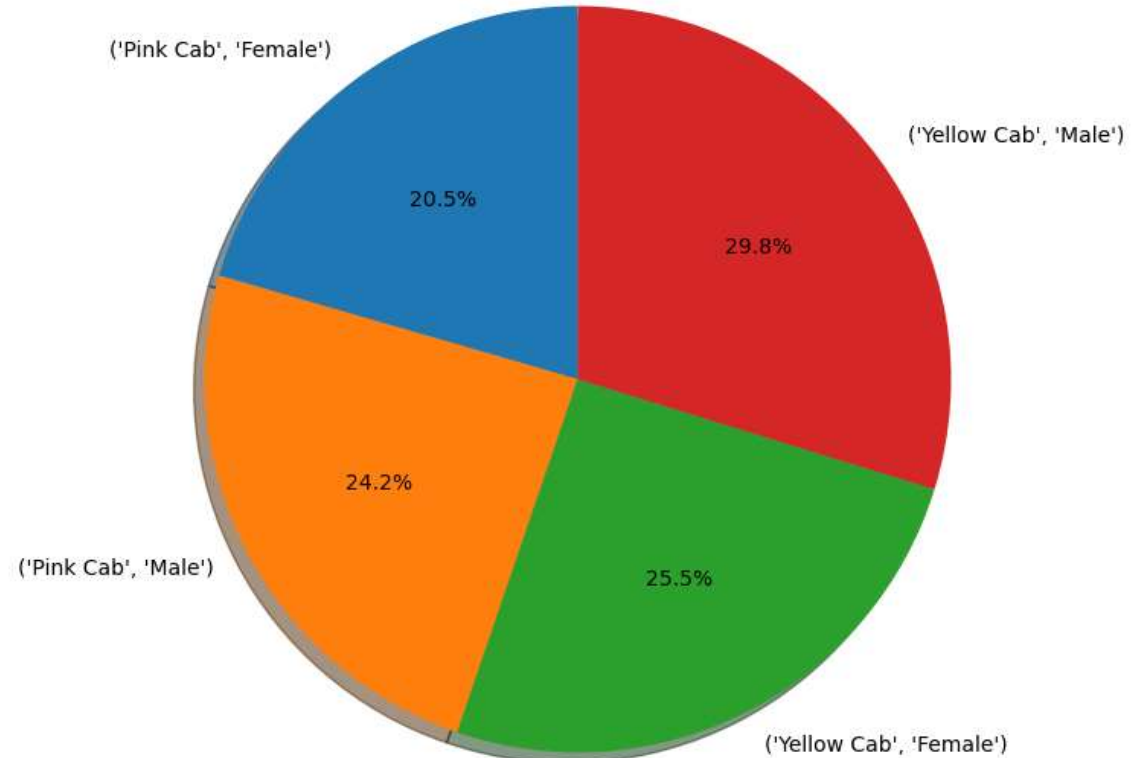
- From the above bar graphs, it is observed that yellow cab have more profits yearly based on considering the “Cost of the trip” and “Price charged” compared to the pink cab profits yearly.

Profit and customer base Analysis Gender wise

Price Charged per Gender (Yellow Cab)

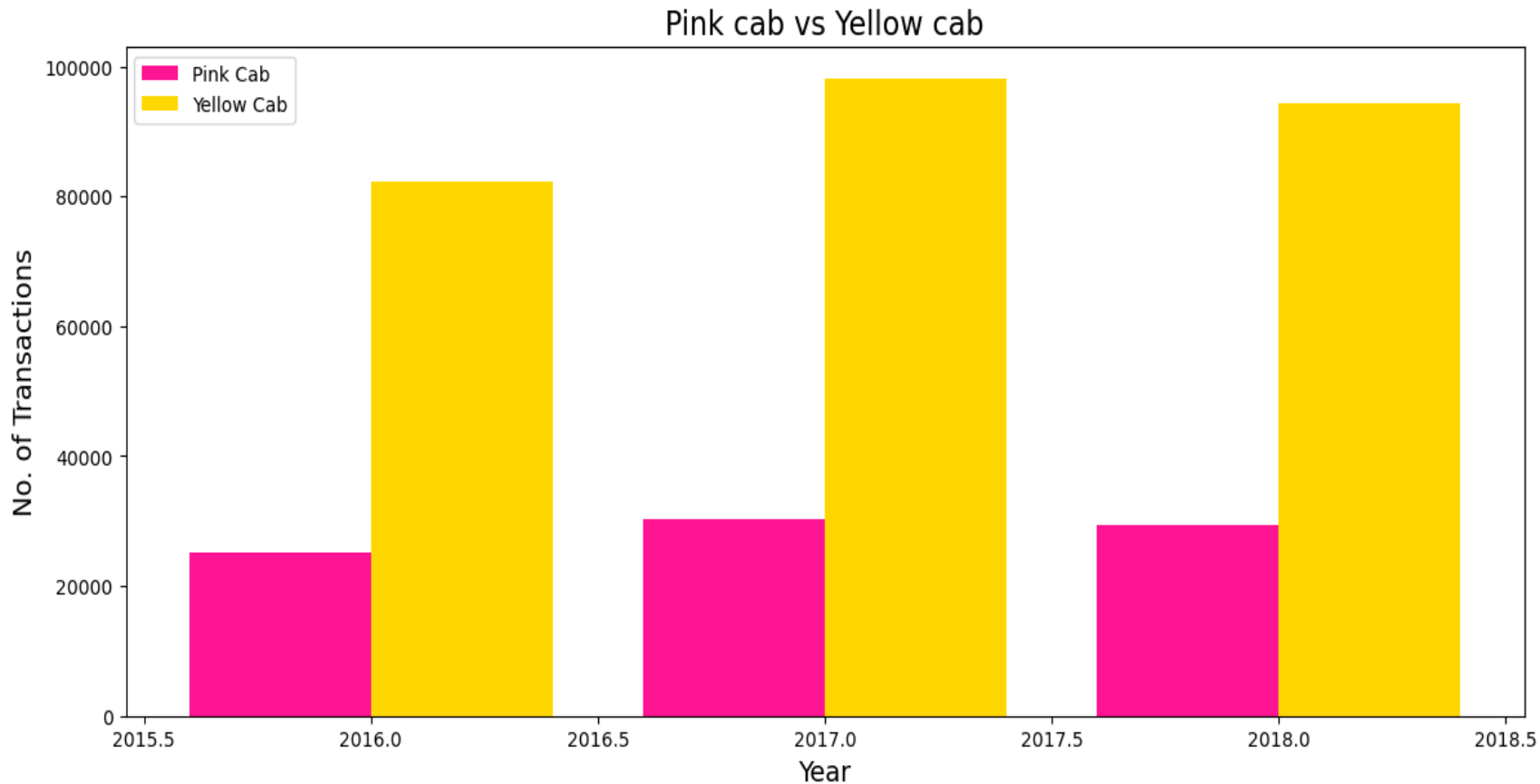


Customer share for Gender per cab



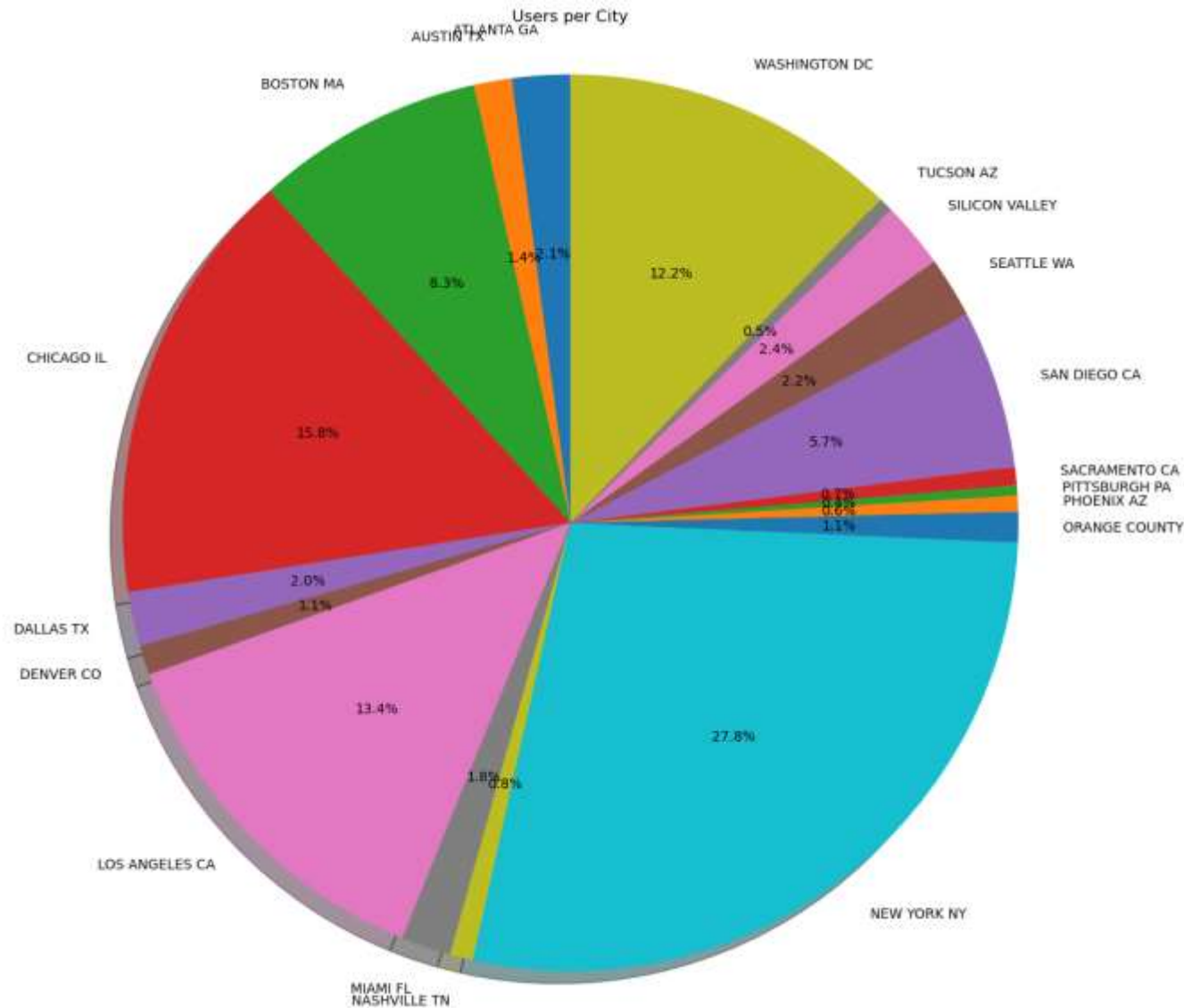
- From the above Pie chart, it is observed that there is almost equal distribution of gender in the profit and customer base of both the cabs.

Customer Retention



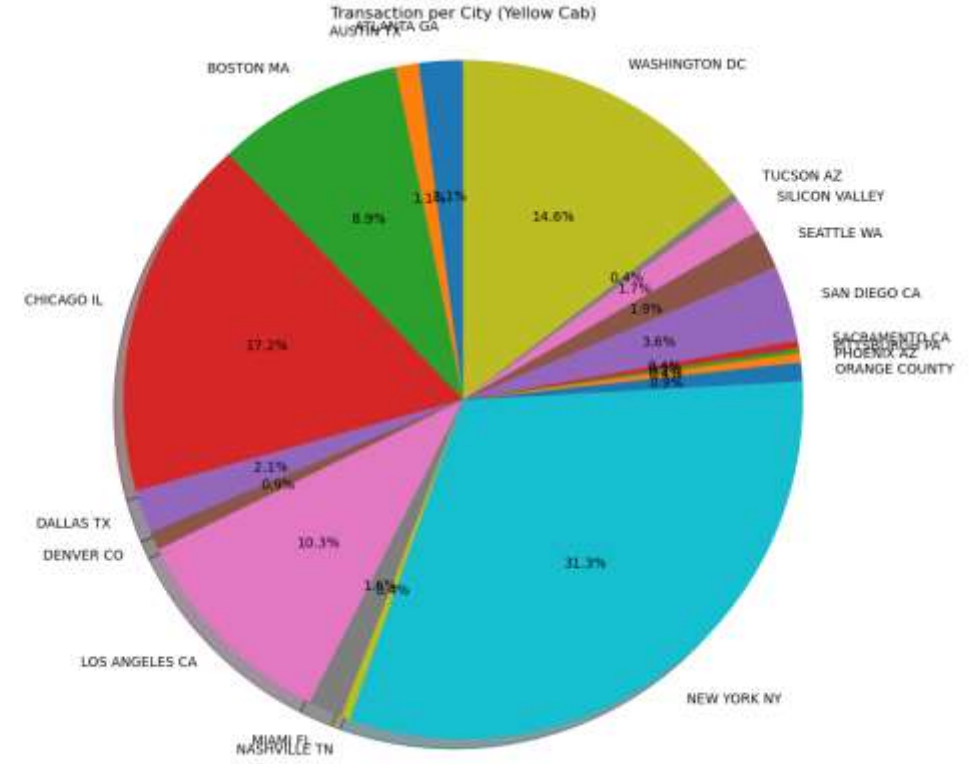
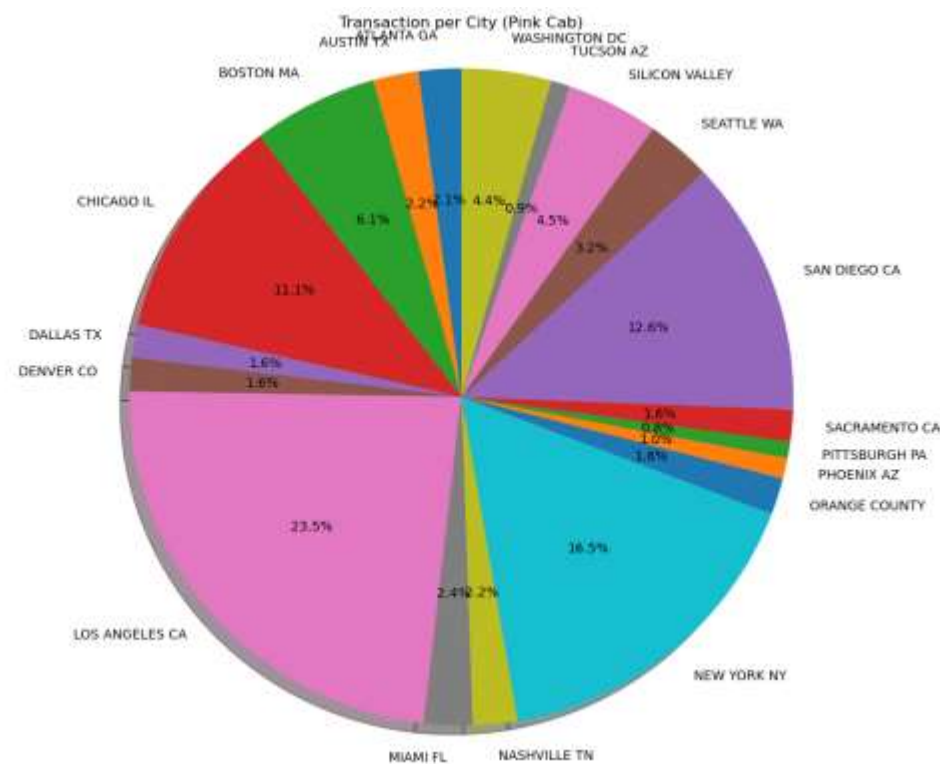
- From the graph, it is clear that yellow cab constantly maintained their customers with good profits.
- So, we can see that Yellow cab is able to retain their customers well as compared to Pink cab.

City Wise Cab Users Covered By Company



- This pie chart shows the number of users covered by Yellow and Pink cab in the city against all cab users present in the city.

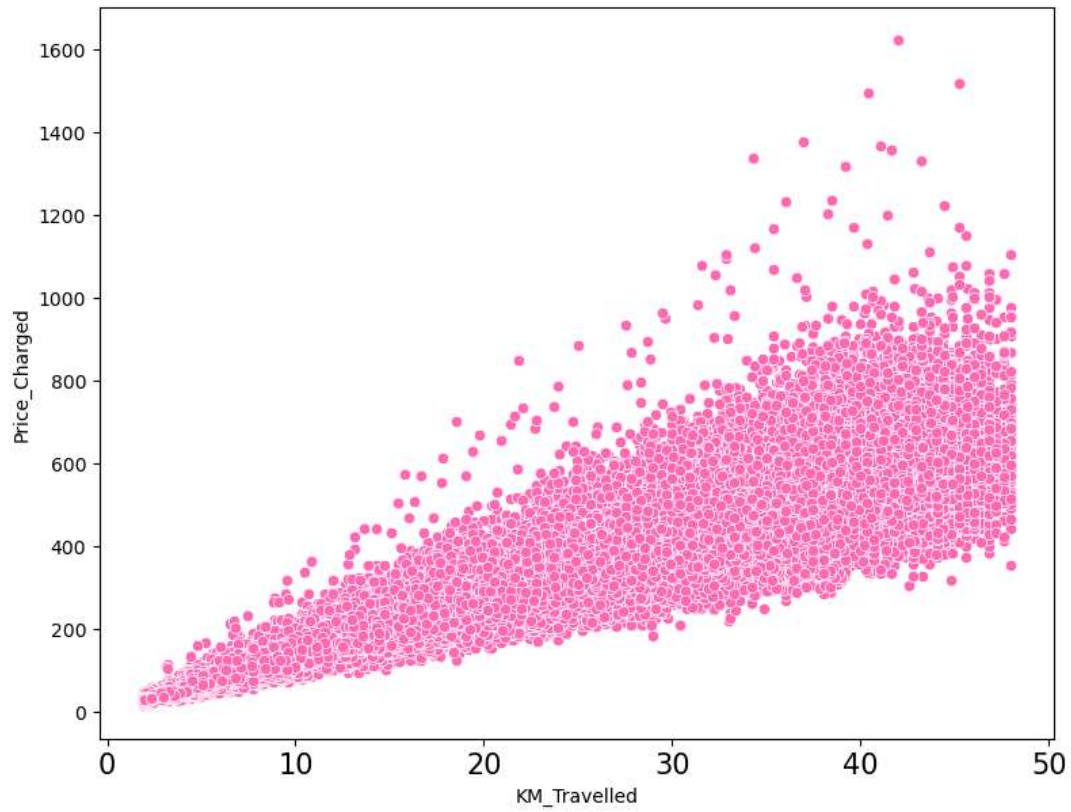
Customer Presence of Yellow and Pink cab city wise based on Transactions



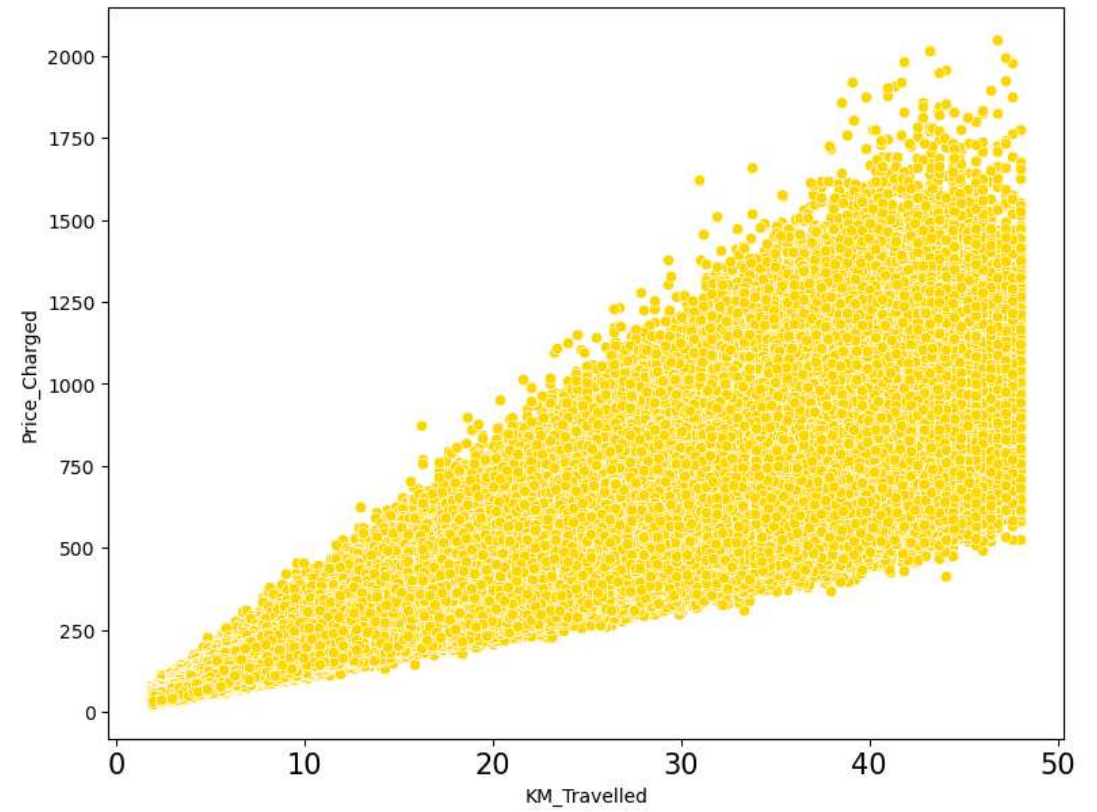
- From above pie chart, Yellow Cab has highest transactions in New York City which has the highest Cab Users of 28%.
- Also, transactions for Pink Cab is more in Los Angeles City.

Customer analysis based on ride distance and Price charged

Price vs Distance in Pink Cab



Price vs Distance in Yellow Cab



Recommendations

After evaluating both cab companies, it was found that Yellow Cab outperforms Pink Cab based on the following criteria:

1. **Customers:** Yellow Cab has a wider customer reach across most cities compared to Pink Cab, which has a higher reach in fewer cities. Additionally, Yellow Cab is more successful in attracting other cab users than Pink Cab.
2. **Customer Retention:** Analysis indicates that Yellow Cab excels significantly over Pink Cab in retaining customers.
3. **Age group:** Yellow Cab attracts customers across all age groups and is equally popular among the 60+ age group as it is among the 18-25 age group.
4. **Average Profit per KM:** Yellow Cab's average profit per kilometer is nearly three times that of Pink Cab.
5. **Income-wise Reach:** While both cabs are popular among high and medium income classes, Yellow Cab performs better in serving all three income groups (low, medium, and high).

As a conclusion, based on above points , Yellow cab company is recommended for investment.

Thank You



Data Glacier

Your Deep Learning Partner