



G2M insight for Cab Investment firm

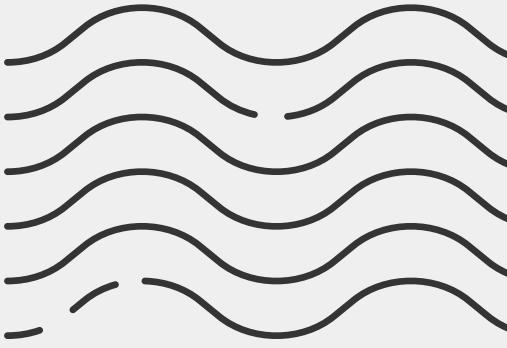
Created by **Lujain Saad**





Agenda

- Executive Summary
- Problem Statement
- Approach
- EDA
- EDA Summary
- Hypothesis Testing
- Building Models
- Recommendations





Description:

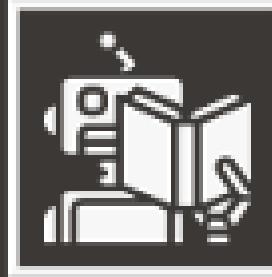
- XYZ is a private equity firm in US. Due to remarkable growth in the Cab Industry in last few years and multiple key players in the market, it is planning for an investment in Cab industry.
- Provide actionable insights to help XYZ firm in identifying the right company for making investment.

Cab Companies:

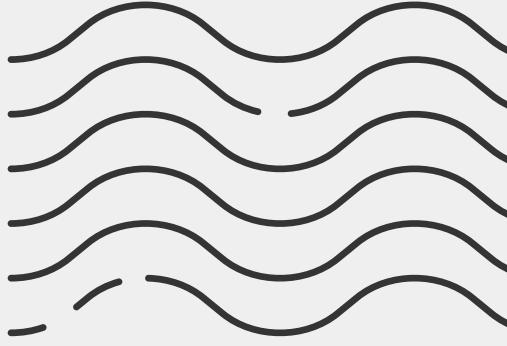
- 1) Yellow Cab
- 2) Pink Cab

The Analysis include :

- 1) Data Understanding,
- 2) Data Visualization,
- 3) Creating multiple hypothesis,
- 4) Building models and finding the best fit model based on Accuracy.



Data Preparation:



There are 4 datasets:

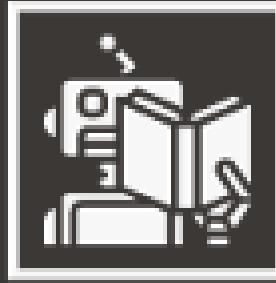
- **Cab_Data.csv** – this file includes details of transaction for 2 cab companies.
- **Customer_ID.csv** – this is a mapping table that contains a unique identifier which links the customer's demographic details.
- **Transaction_ID.csv** – this is a mapping table that contains transaction to customer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



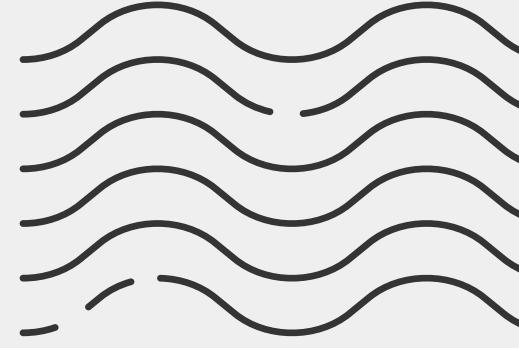


EXPLORATORY DATA ANALYSIS

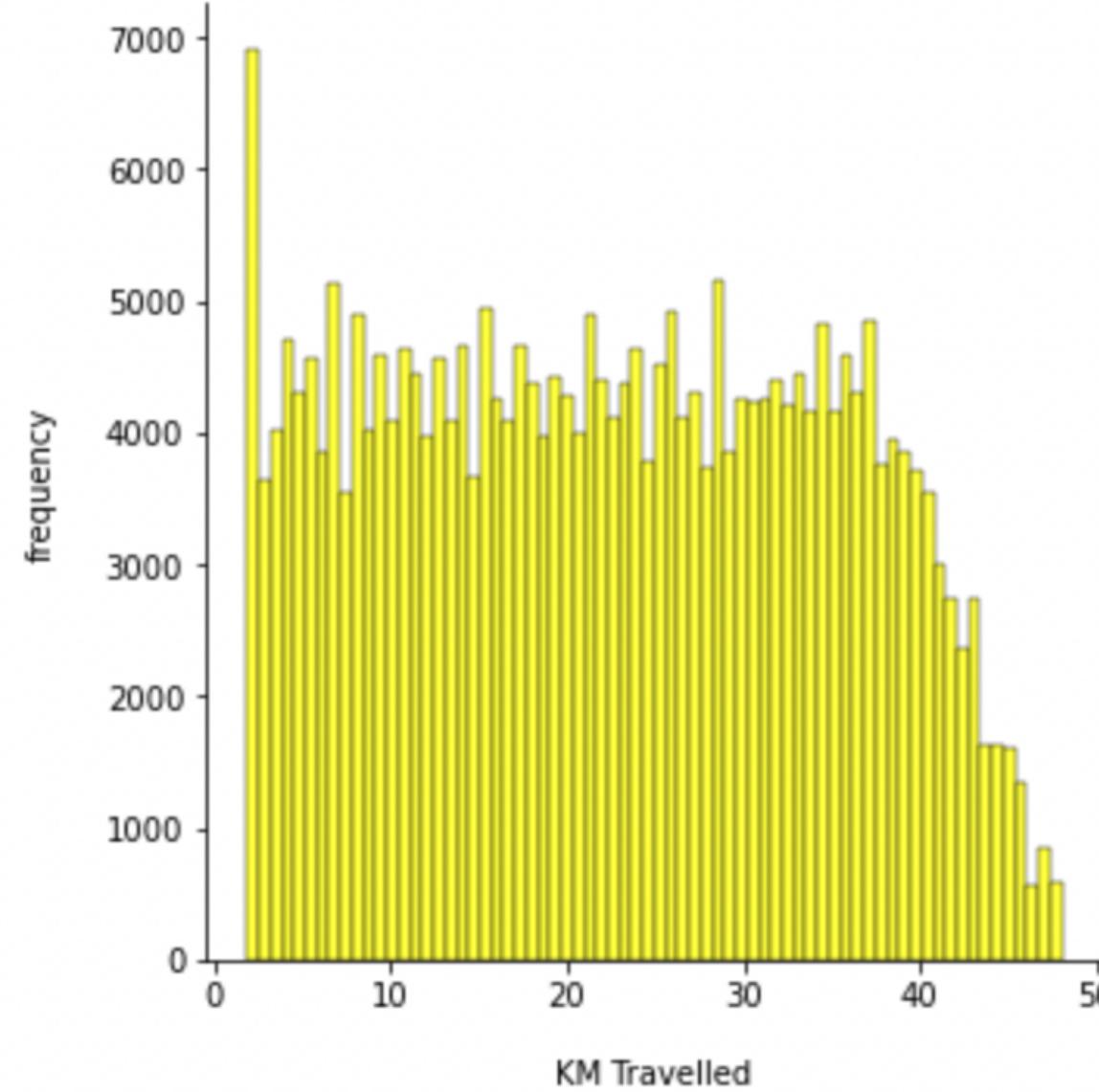




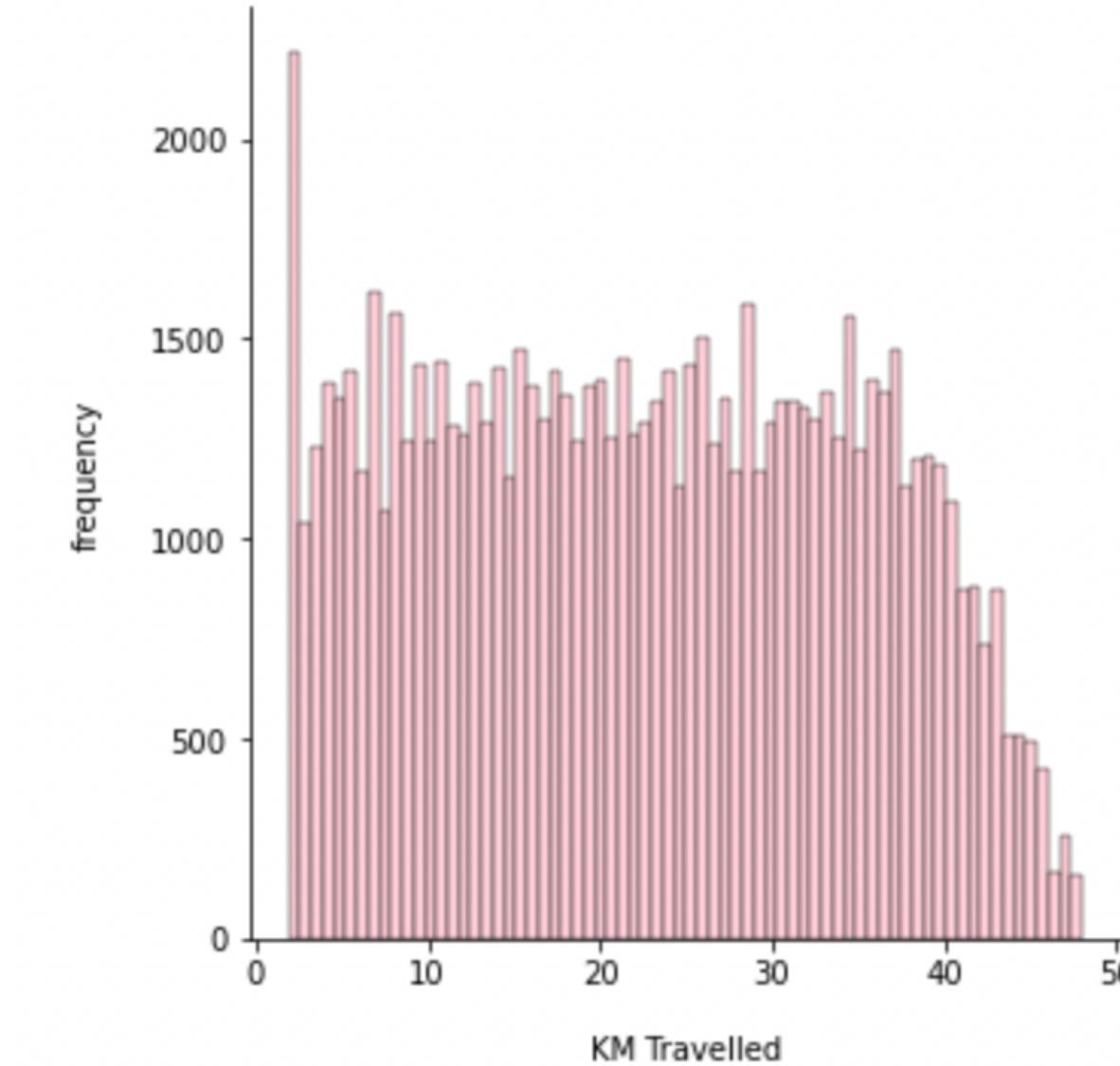
Distribution of KM Travelled for both Cabs:



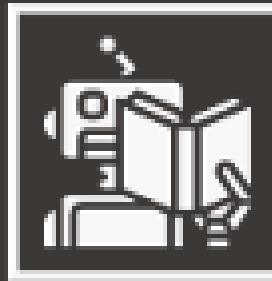
Distribution of KM Travelled (Yellow Cab)



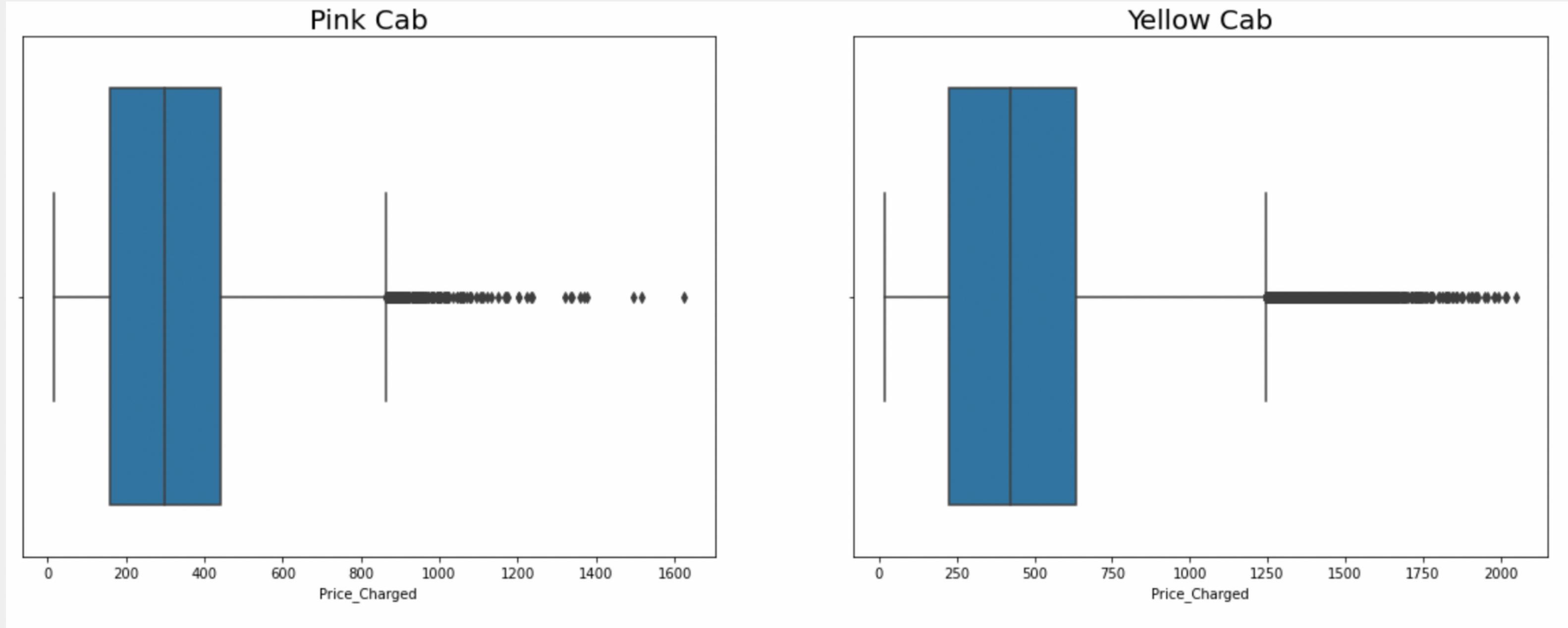
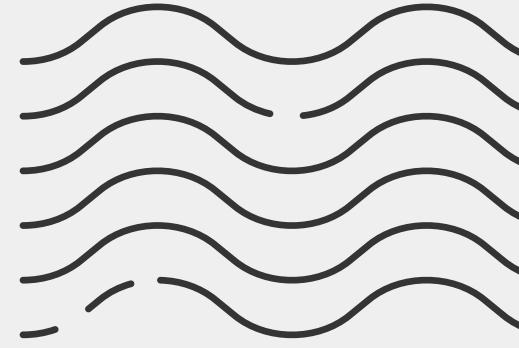
Distribution of KM Travelled (Pink Cab)



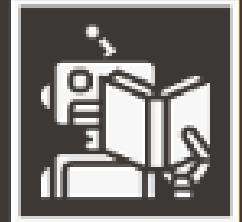
Most of the rides are in the range of approximately 2 to 48 KM.



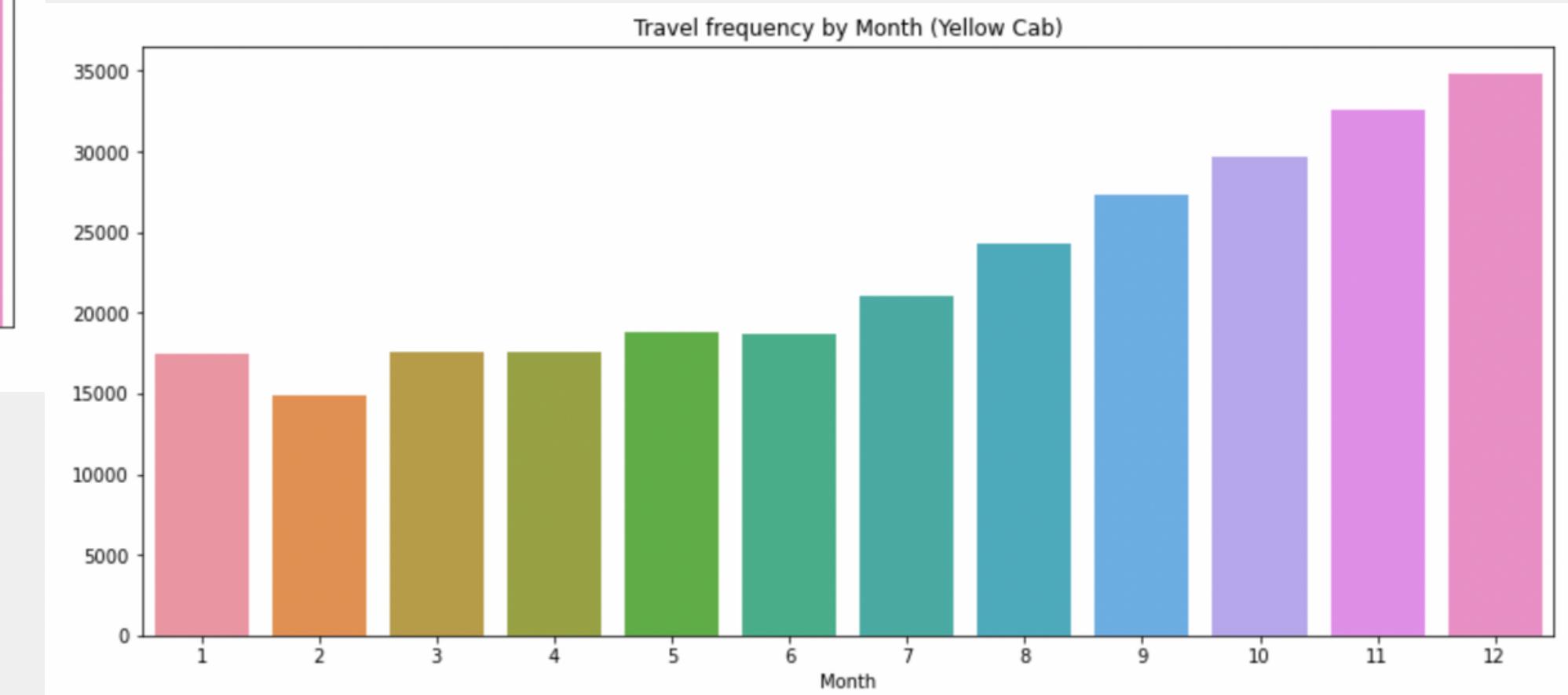
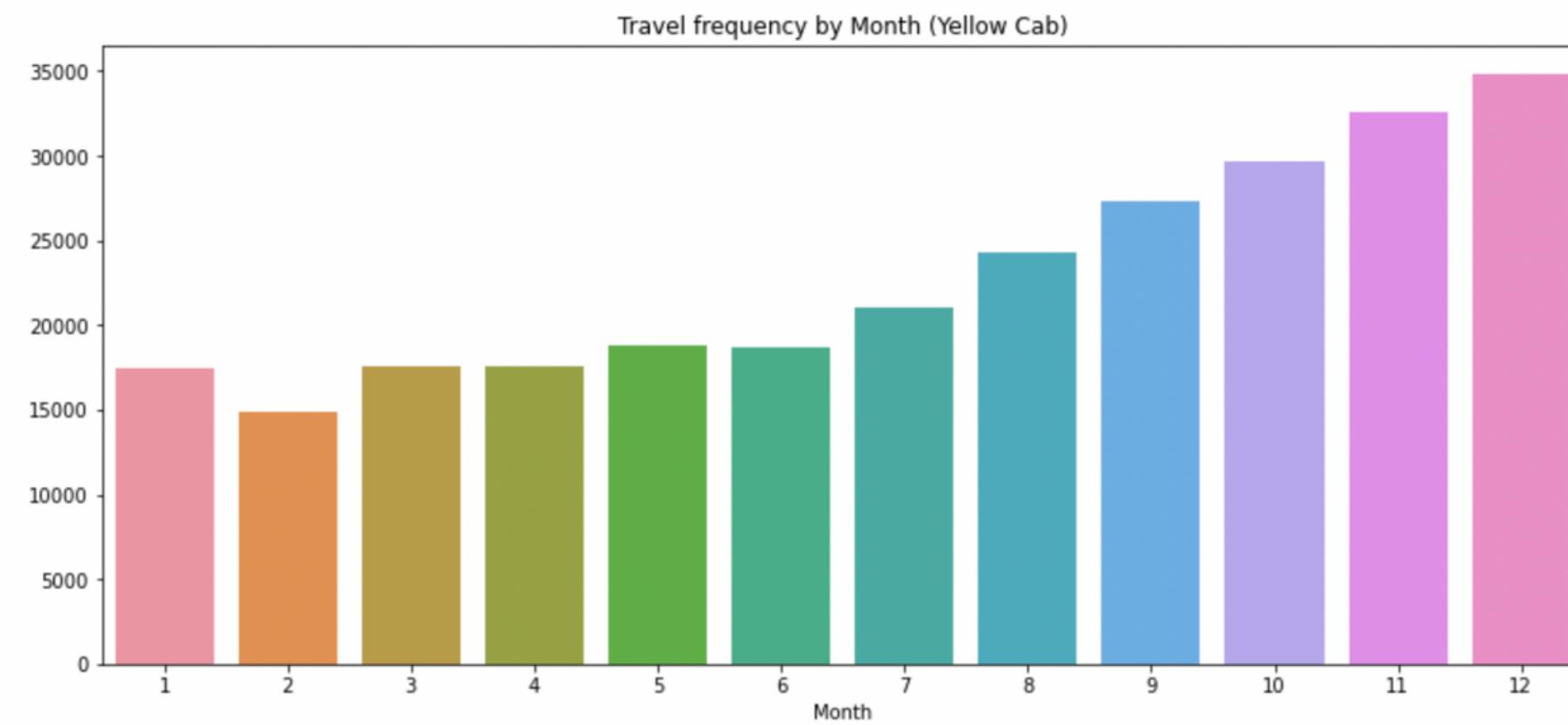
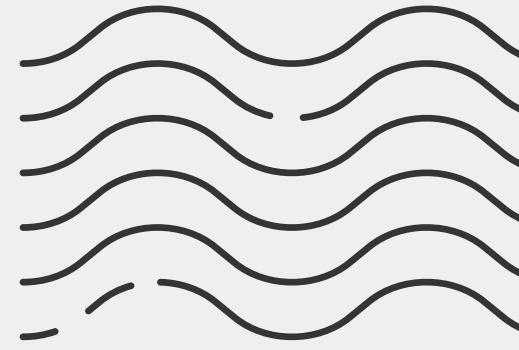
Distribution of Price Charged for both Cabs:



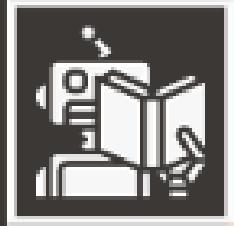
- The Price Charge range for Yellow cab is more than the Pink cab.
- The outliers are due to use of high-end cars.



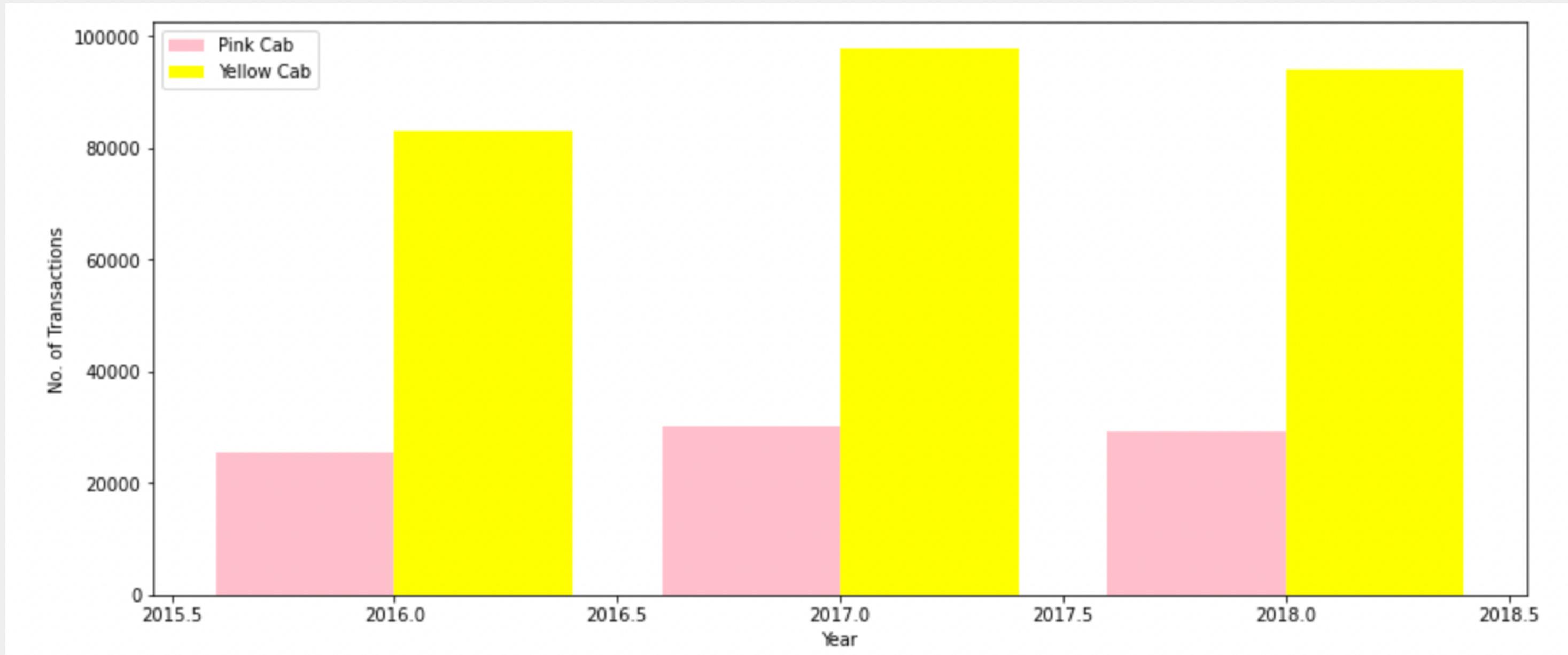
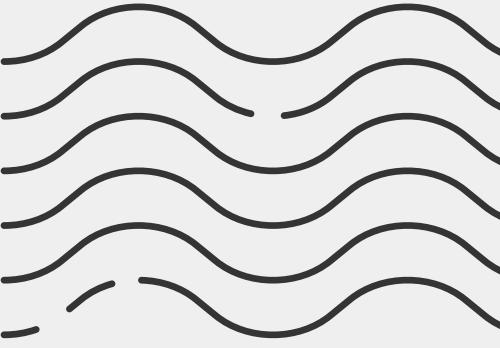
Distribution of KM Travelled for both Cabs:



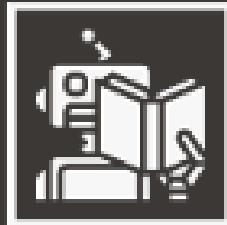
Yellow Cab has higher travels (35000) in the month of December which is the holiday season compared to Pink Cab (11000).



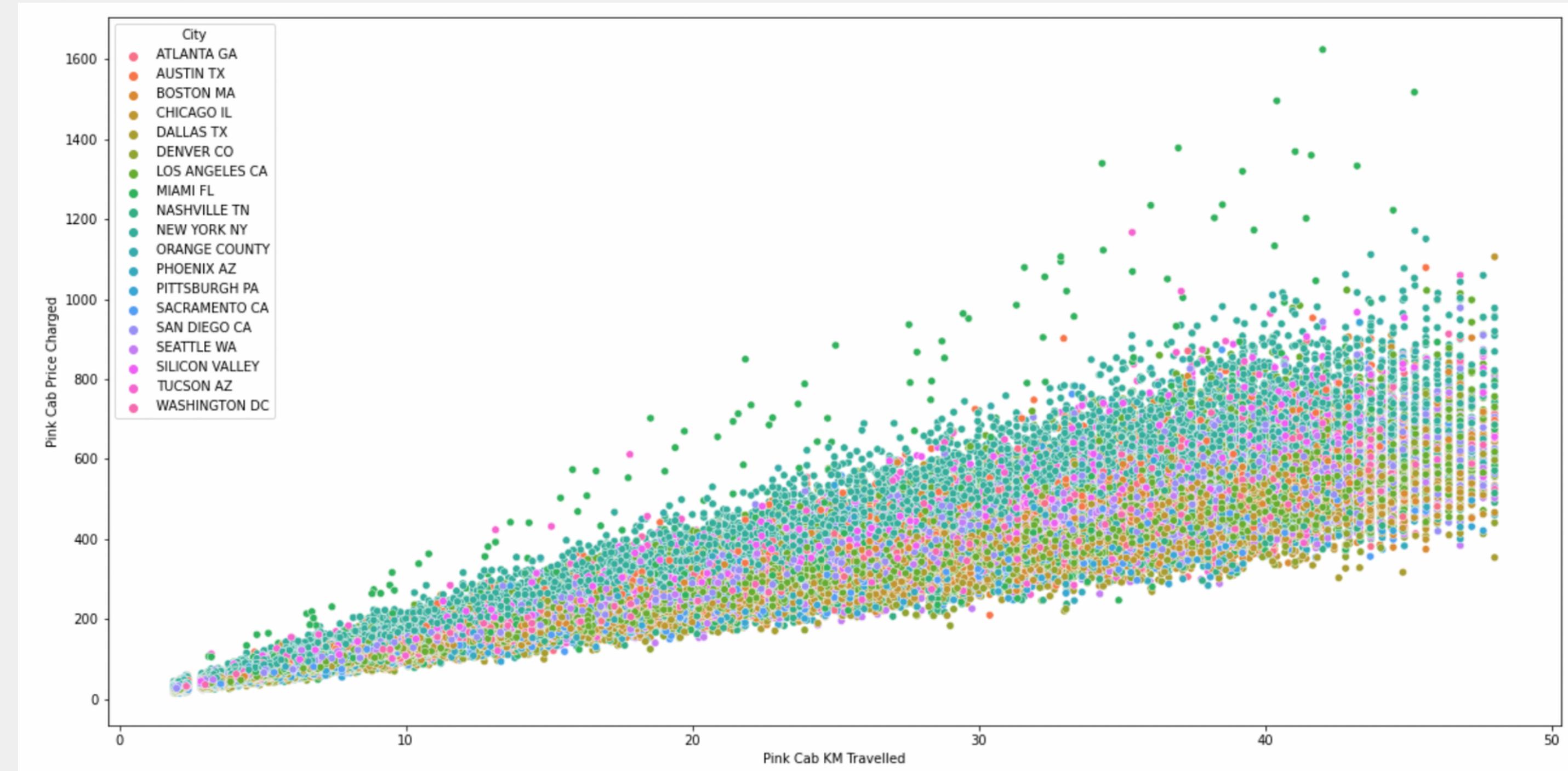
Transaction per Year for both Cabs:



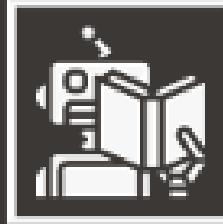
From the graph it shows that on yearly basis no. of transactions for Yellow cab is higher than Pink cab.



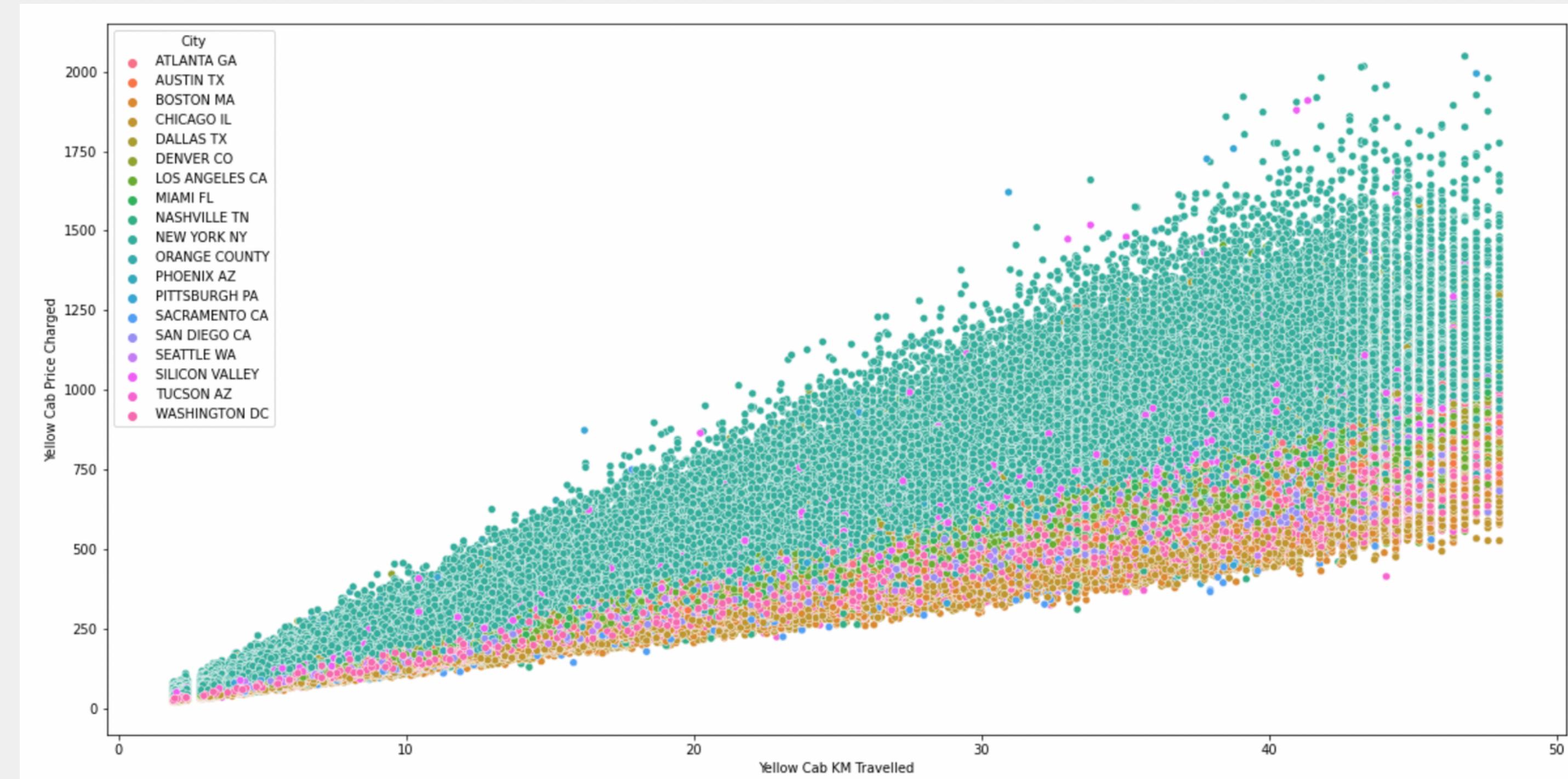
Pink Cab: Price Charged per KM per City



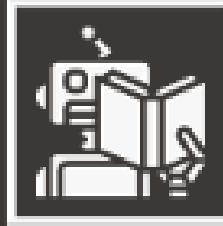
For Pink cab all the cities have the same increase in price charge with increase in distance



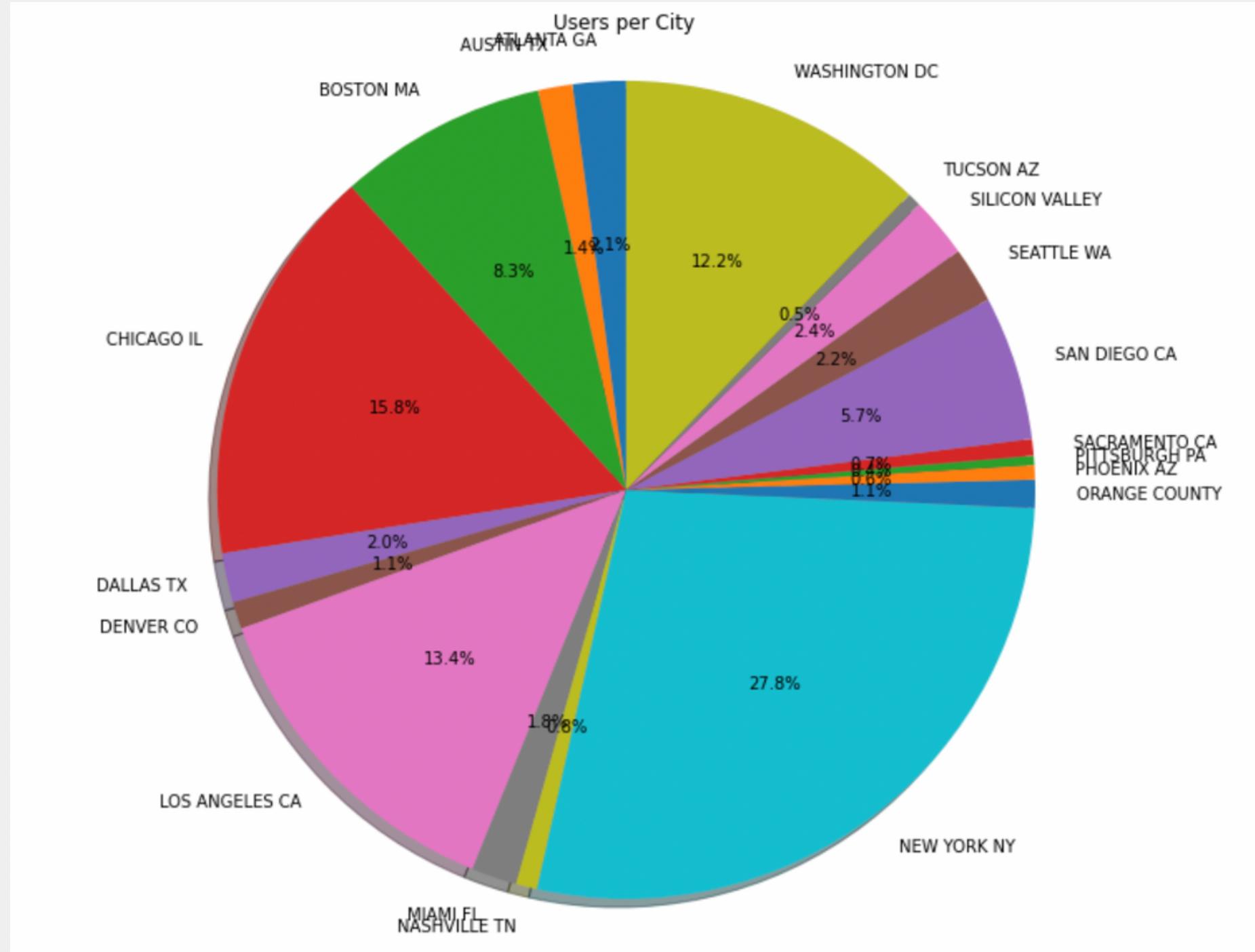
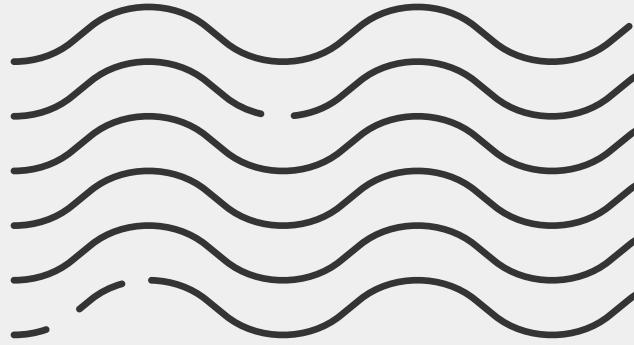
Yellow Cab: Price Charged per KM per City



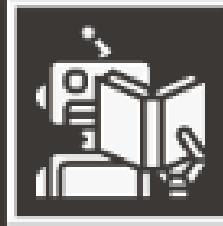
In New York City the Price charged for Yellow Cab is more in comparison to the other cities



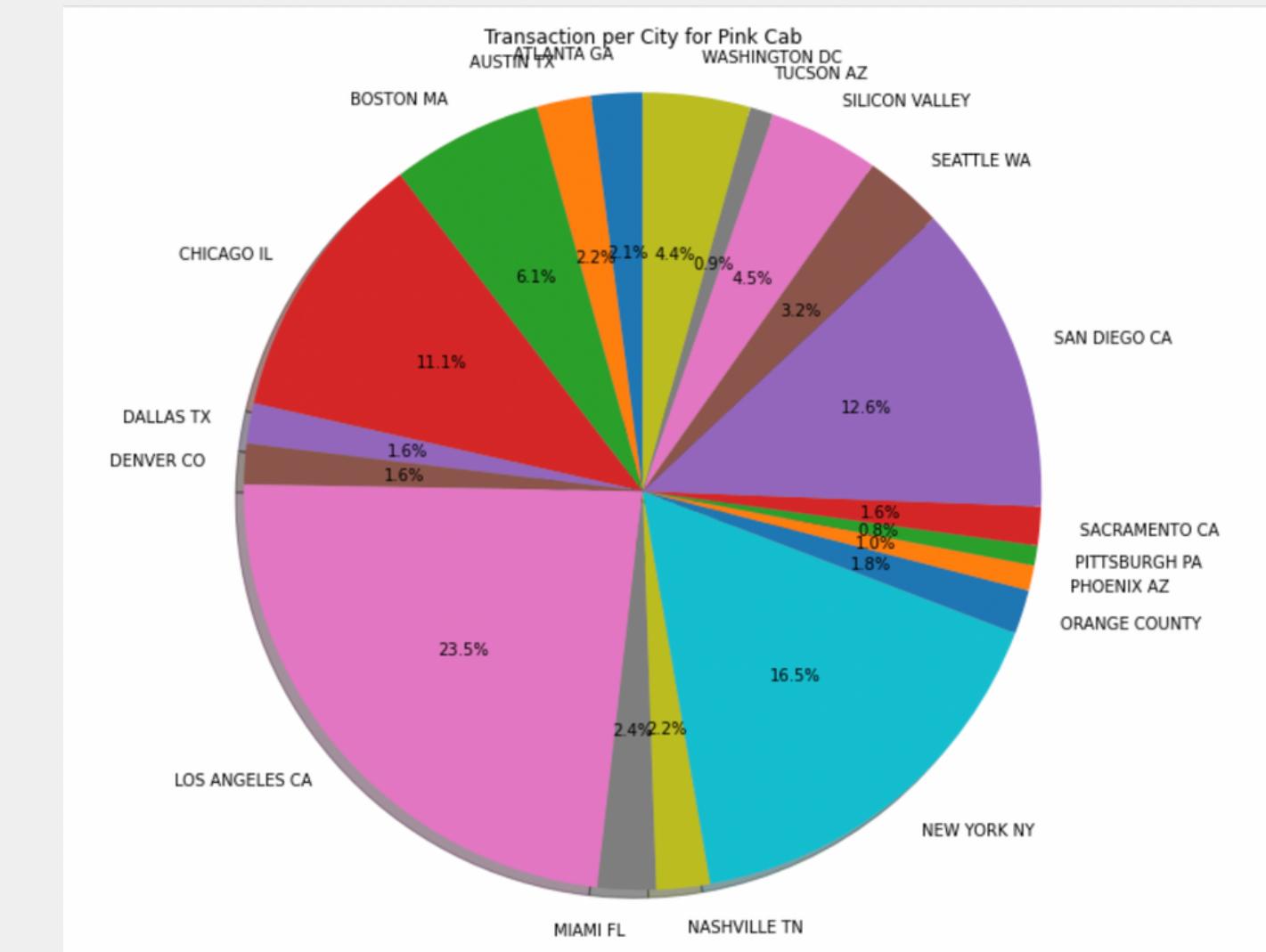
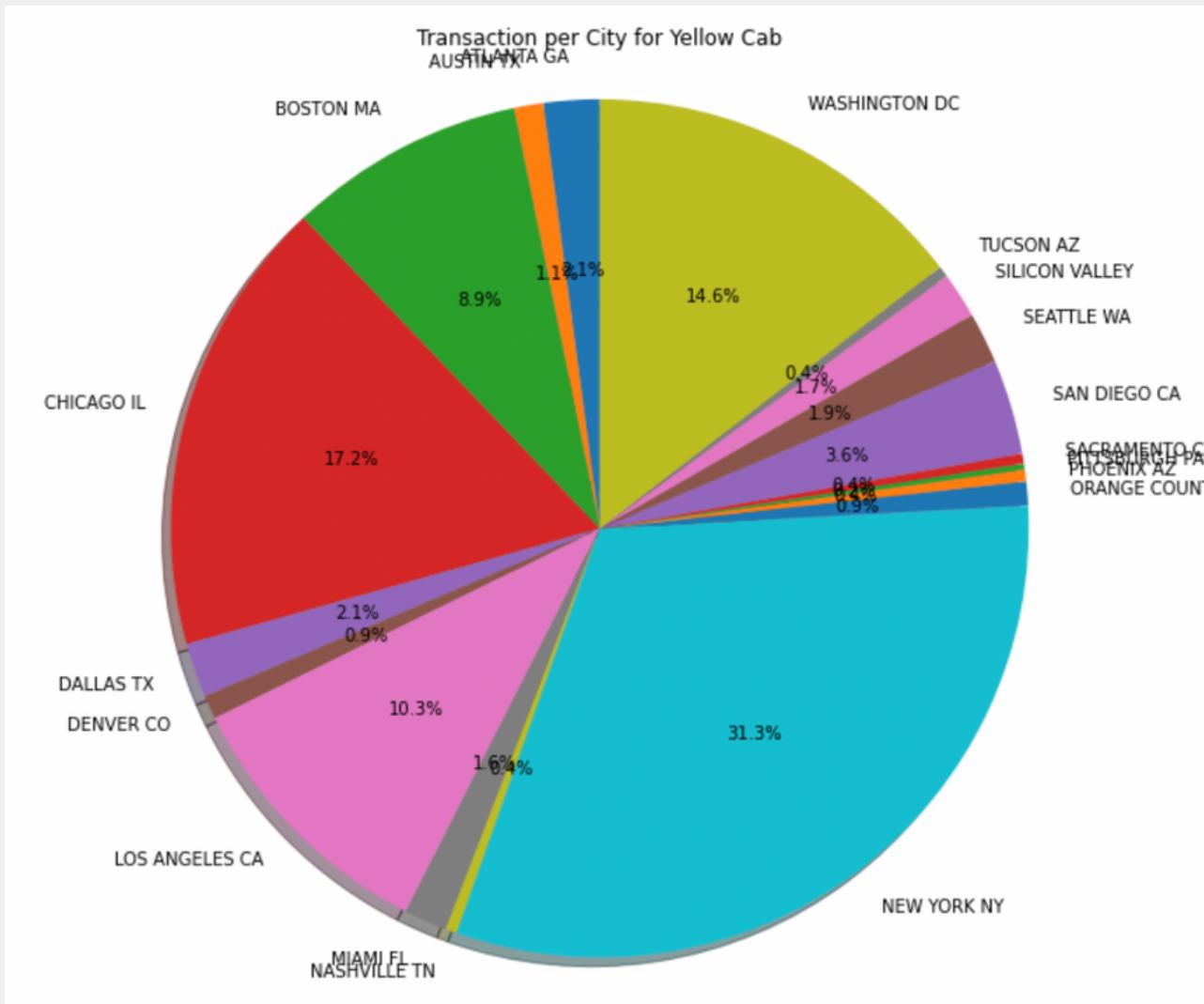
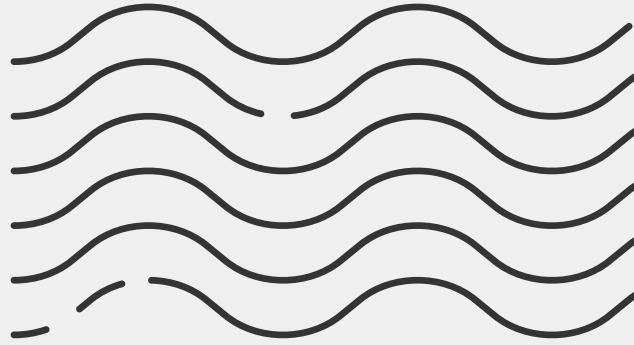
Cab Users per City:



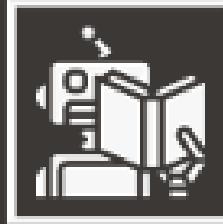
New York City has the highest Cab users with 27.8% followed by Chicago with 15.8% and Los Angeles with 13.4%



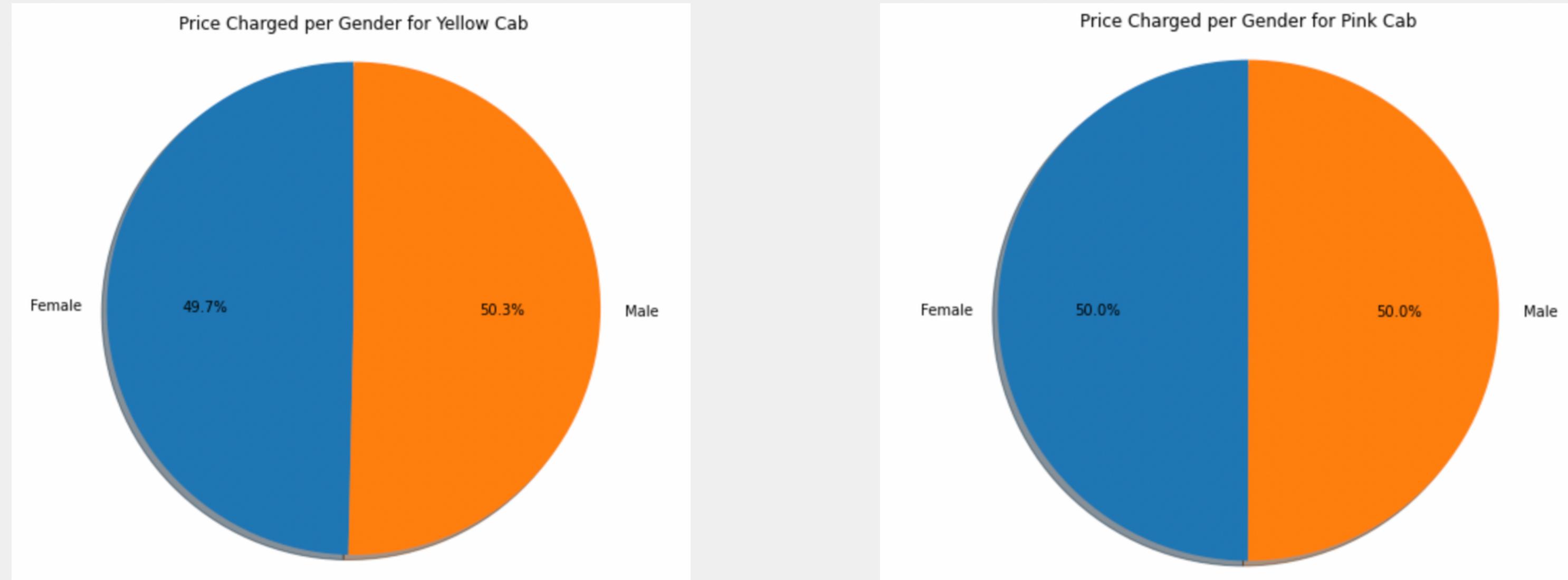
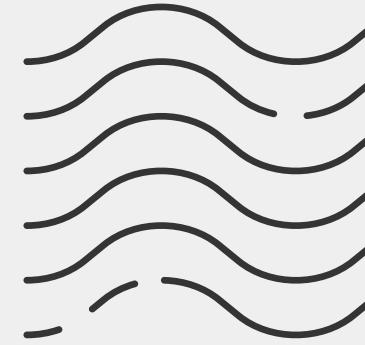
Transaction per City for both Cabs:



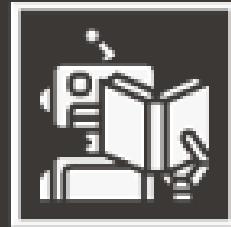
- Transaction for Yellow Cab is highest in New York City(31%) and New York City has the highest Cab Users of 28% as per the previous slide.
- Transaction for Pink Cab is highest in Los Angeles City.



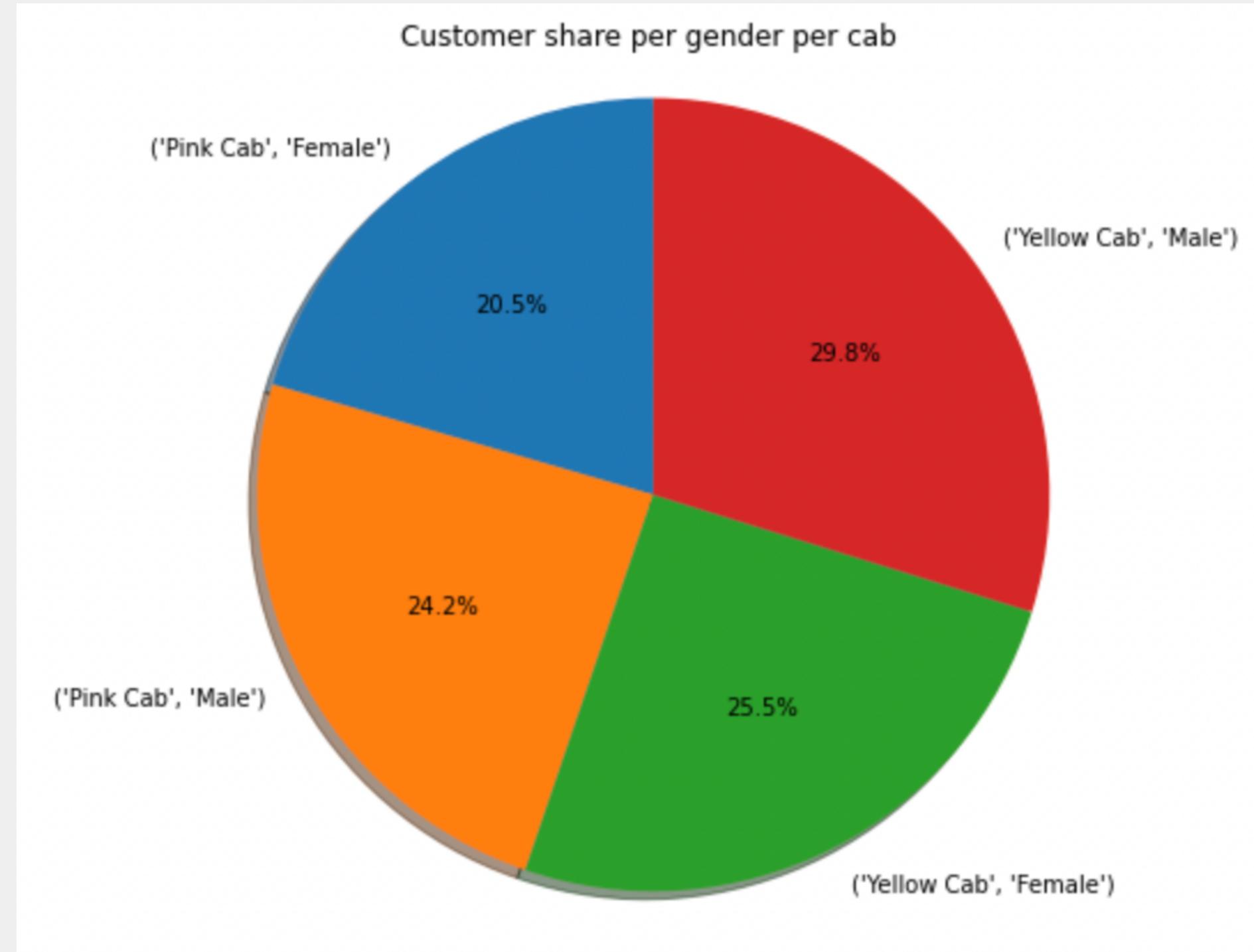
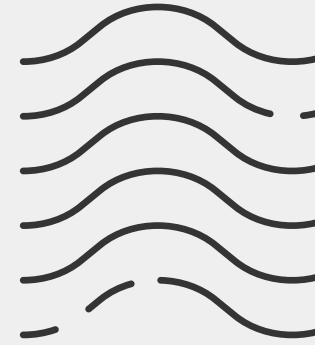
Price Charged per Gender for both Cabs:



- Yellow Cab charge less from Female Customers whereas Pink Cab charges same for both Male and Female Customers.



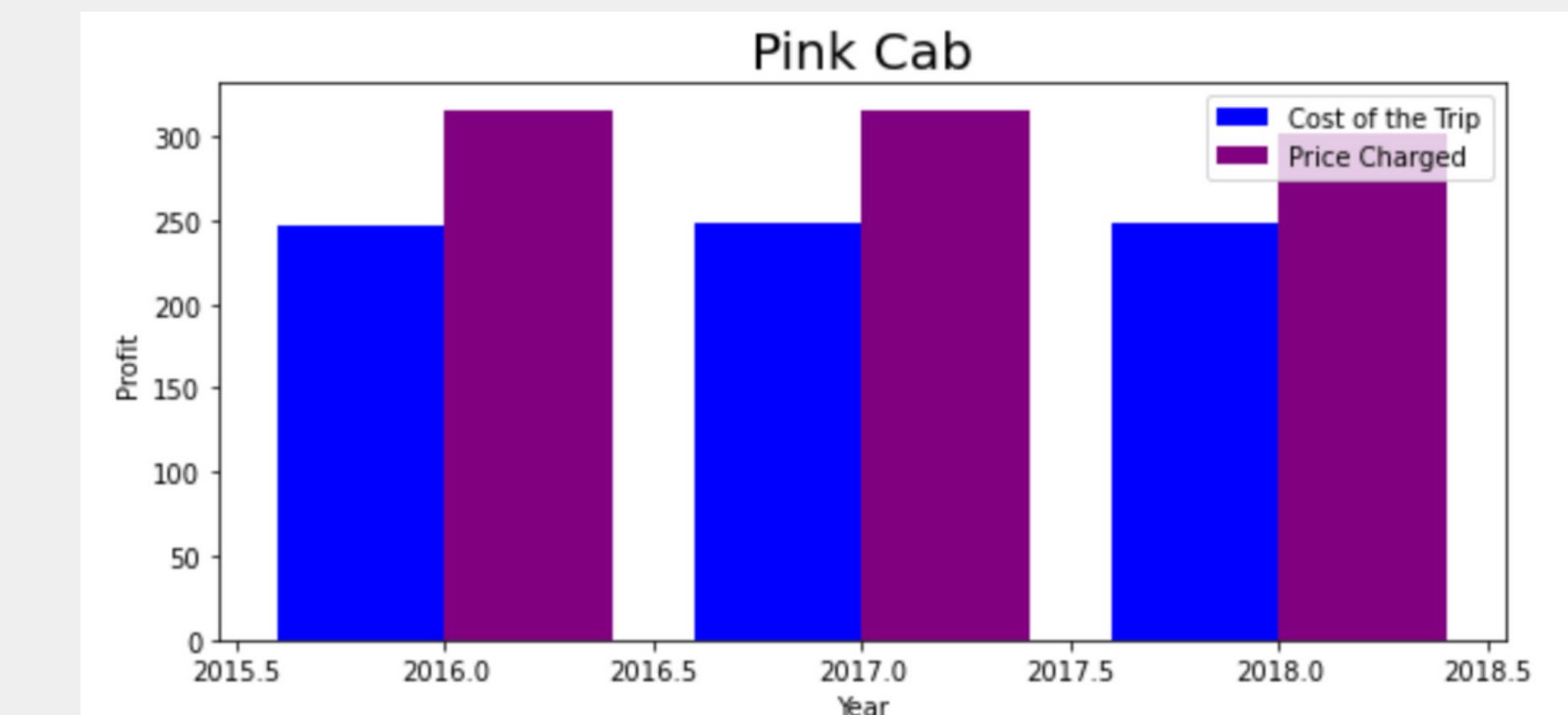
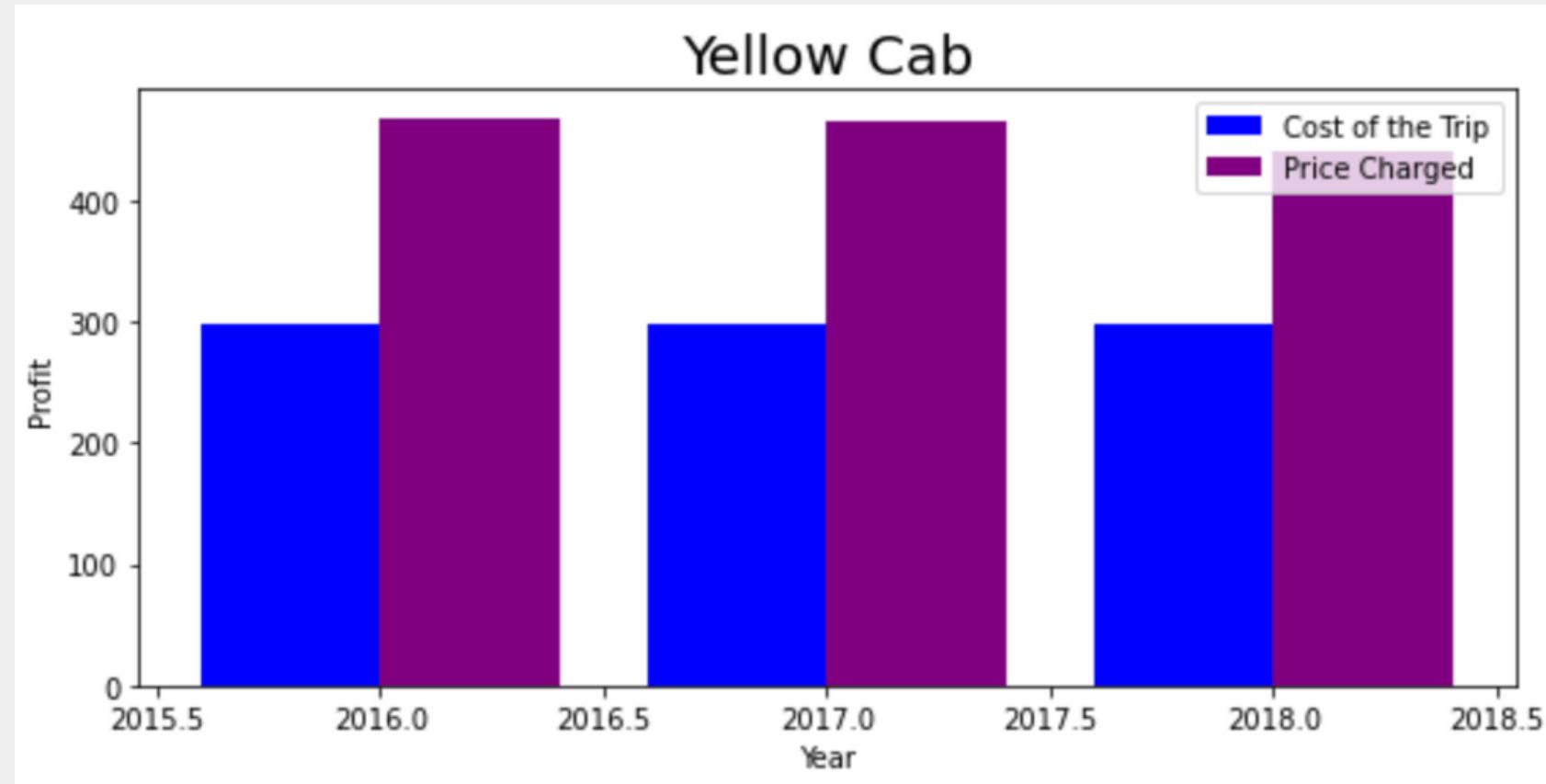
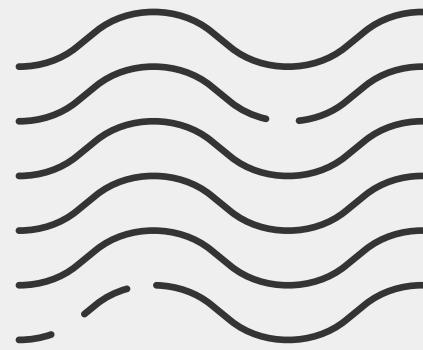
Customer Share per Gender for both Cabs:



Female Customers in Yellow Cab(25.5%) is higher compared to Pink Cab (20.5%)



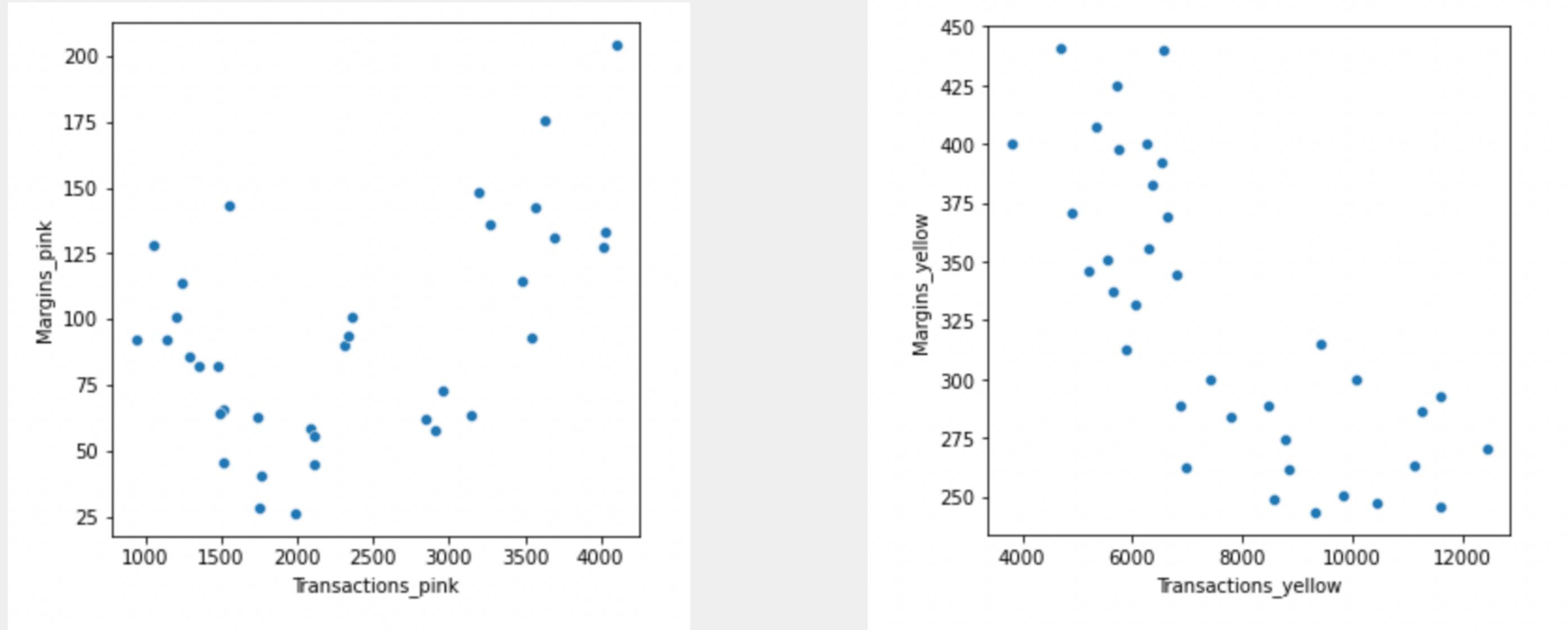
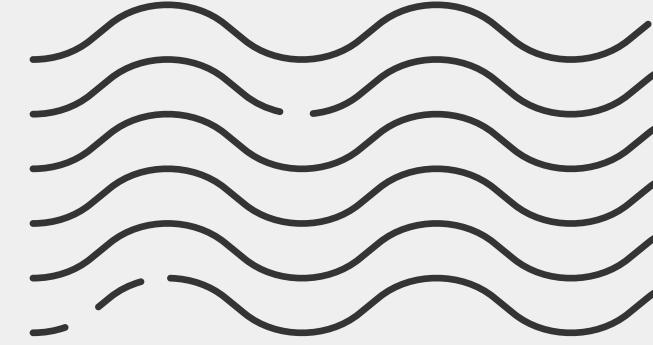
Profit Margin per year for both Cabs:



From the Graphs, it shows that the Yellow cab has a higher Profit Margin (Price Charged - Cost of Trip) compared to Pink cab.



Margins per Transactions:



- Margins: Price Charged – Cost of Trip
- Pink Cabs increase margins with increase in number of Transactions.
- Yellow Cab decrease Margins with the increase in Transaction.



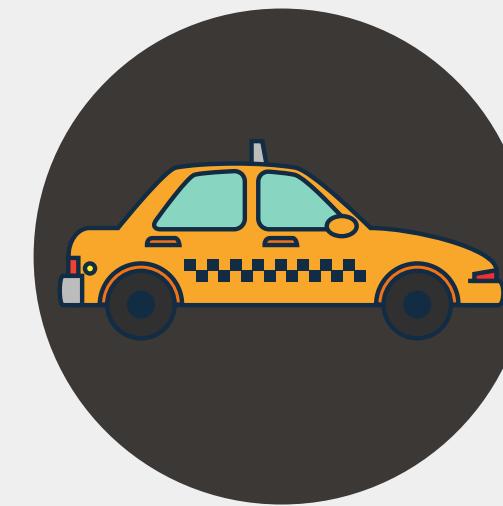
EDA SUMMARY





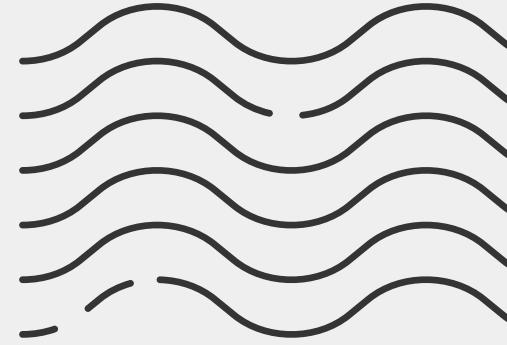
Pink Cab

XYZ Company



Yellow Cab

XYZ Company



- Rides are in the range of approximately 2 to 48 KM.
- Price Charge range from 150 to 450 dollars.
- In December which is the holiday season, no. of travels was around 11000.
- Transaction per year:
 - 2016: 20000 – 40000
 - 2017: 20000 – 40000
 - 2018: 20000 – 40000
- All the cities have the same increase in price charge with increase in distance.

- Rides are in the range of approximately 2 to 48 KM.
- Price Charge range from 250 to 600 dollars.
- In December which is the holiday season, no. of travels was around 35000.
- Transaction per year:
 - 2016: 80000 – 100000
 - 2017: 80000 – 100000
 - 2018: 80000 – 100000
- In New York City the Price charged for Yellow Cab is more in comparison to the other cities.



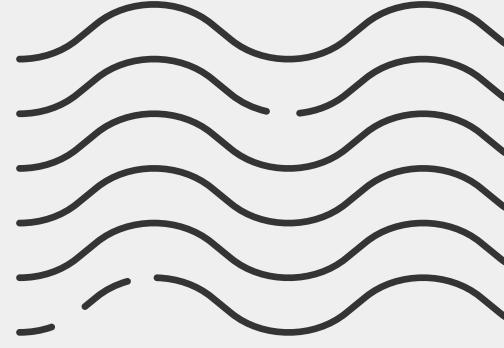
Pink Cab

XYZ Company



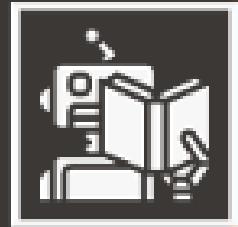
Yellow Cab

XYZ Company



- Pink Cab charges same for both Male and Female Customers.
- Female customers are around 20.5% out of the total Customers.
- Profit Margin is low each year (2016-2018) compared to Yellow Cab.
- Pink Cabs increase margins with increase in number of Transactions.

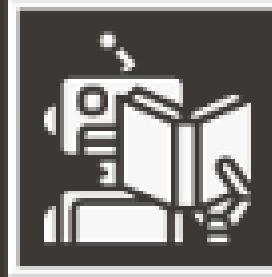
- Yellow Cab charge less from Female Customers.
- Female customers are around 25.5% out of the total Customers.
- Profit Margin is high each year (2016-2018) compared to Pink Cab.
- Yellow Cab decrease Margins with the increase in Transaction.



Correlation:



As per the graph, there is a positive correlation between Margin & Price Charged



Recommendation:

- **Transaction per year:** For Yellow Cab Transaction per year from 2016 to 2018 is almost double than Pink Cab.
- **Margin per Gender:** For Yellow Cab there is difference in Margin between Male and Female Customers due to which Female Customer percentage is higher in Yellow Cab in comparison to Pink Cab.
- **Profit Margin:** For Yellow Cab the Profit Margin is higher per year from 2016 to 2018 in comparison to Pink Cab.
- **Margin per Age:** In Yellow Cab there is difference in Margin for people older than 50 yrs, whereas in Pink Cab there is no difference in Margin of all age group.
- Yellow Cab decreases Margins with the increase in Transaction, hence for Yellow Cab the travel frequency during the Month of December which is the holiday season is 3 times more than Pink Cab.
- Customers for Yellow Cab is highest in New York City which has the highest Cab Users of 28%.

On the basis of the above points, Yellow Cab is recommended for investment.



Thank you

