G2M CASE STUDY

27/06/2021

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

BACKGROUNG G2M CASE STUDY

XYZ is a private equity firm in the USA. 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 in identifying the right company to invest with.

Cab companies:

Pink Cab

Yellow Cab

The analysis includes :

Data understanding

Data visualization

Finding the most profitable company

Recommendation.

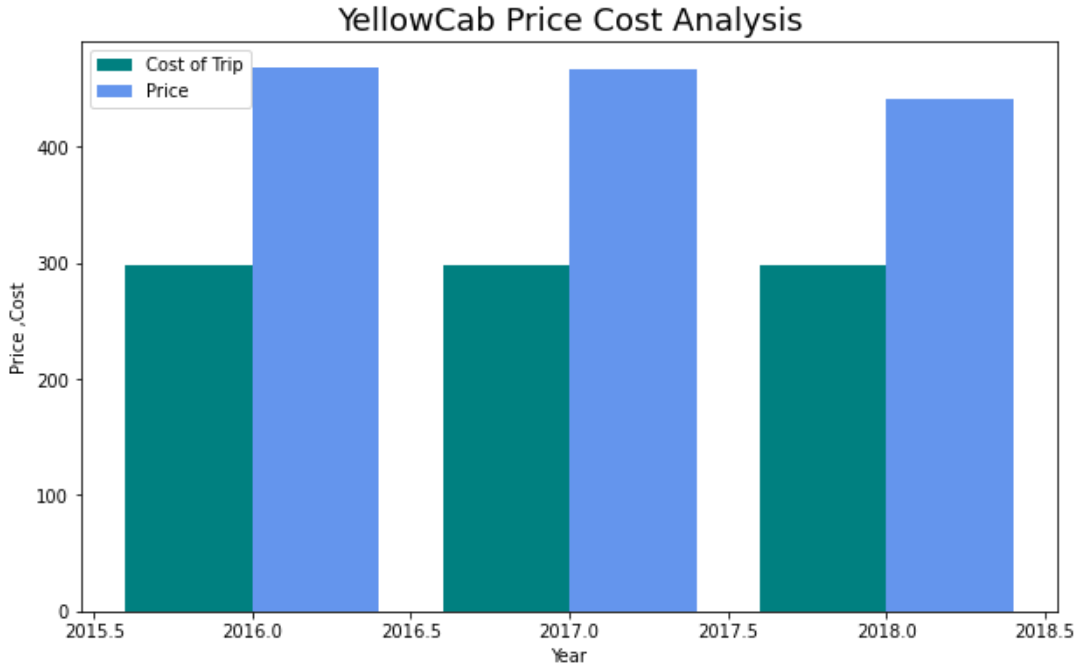
DATA EXPLORATION

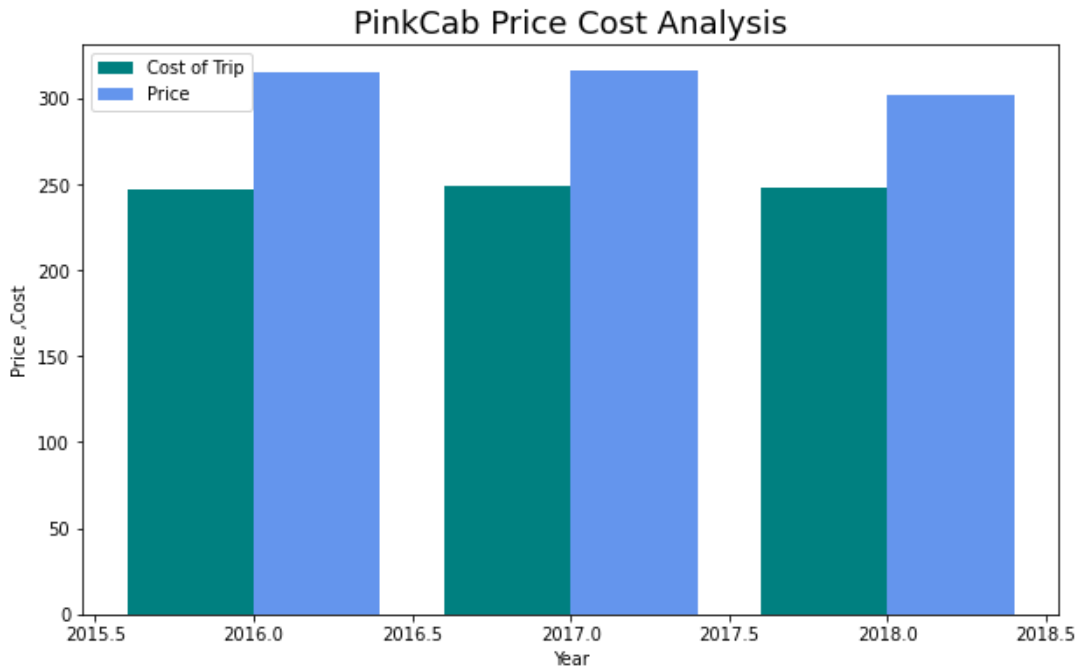
There are four datasets:

- **Cab_Data** – This file includes details of transaction for the 2 cab companies.
- **Customer_ID.csv**- This is a mapping table that contains a unique identifier which links the customer 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

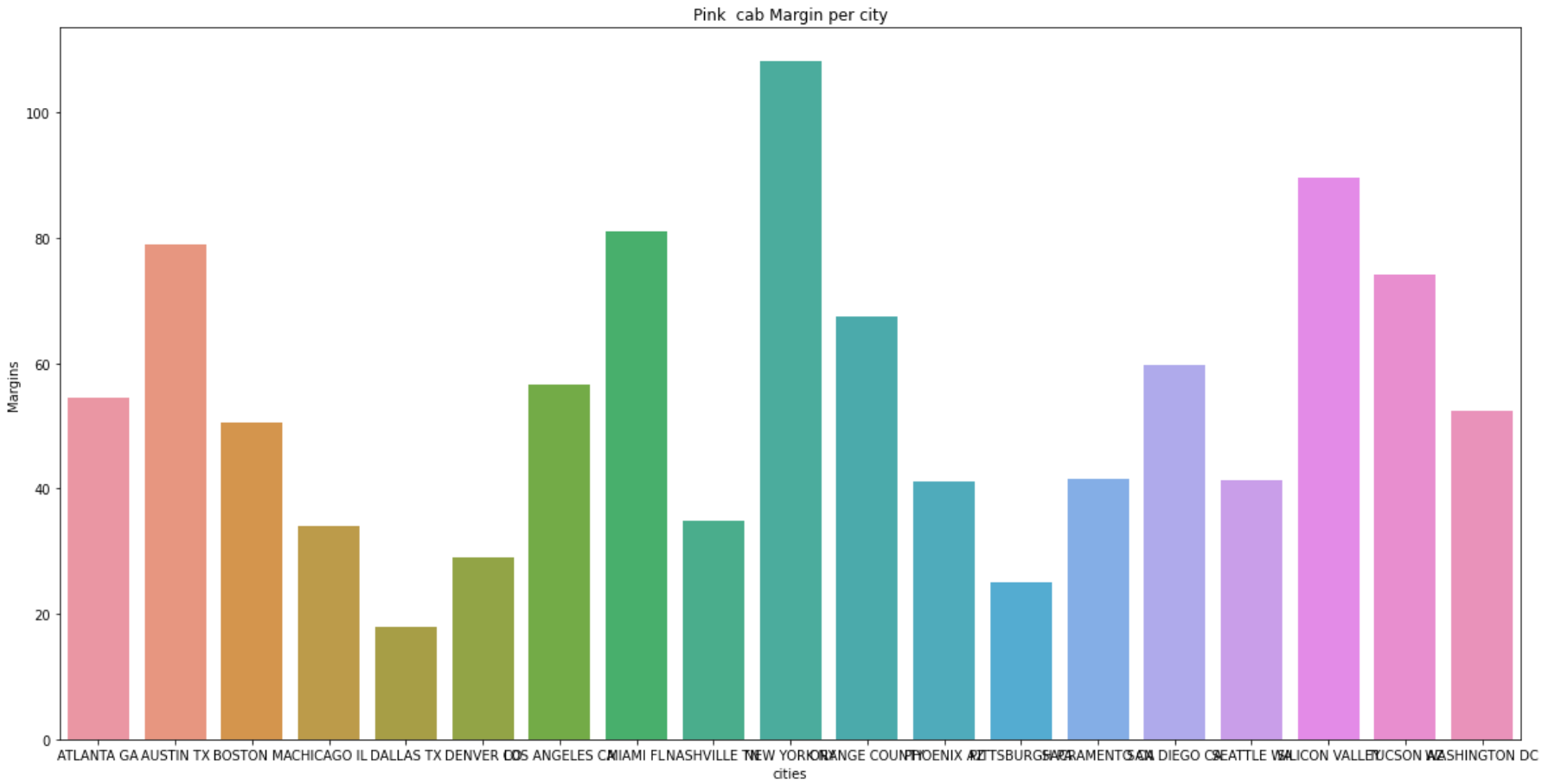
PRICE COST ANALYSIS



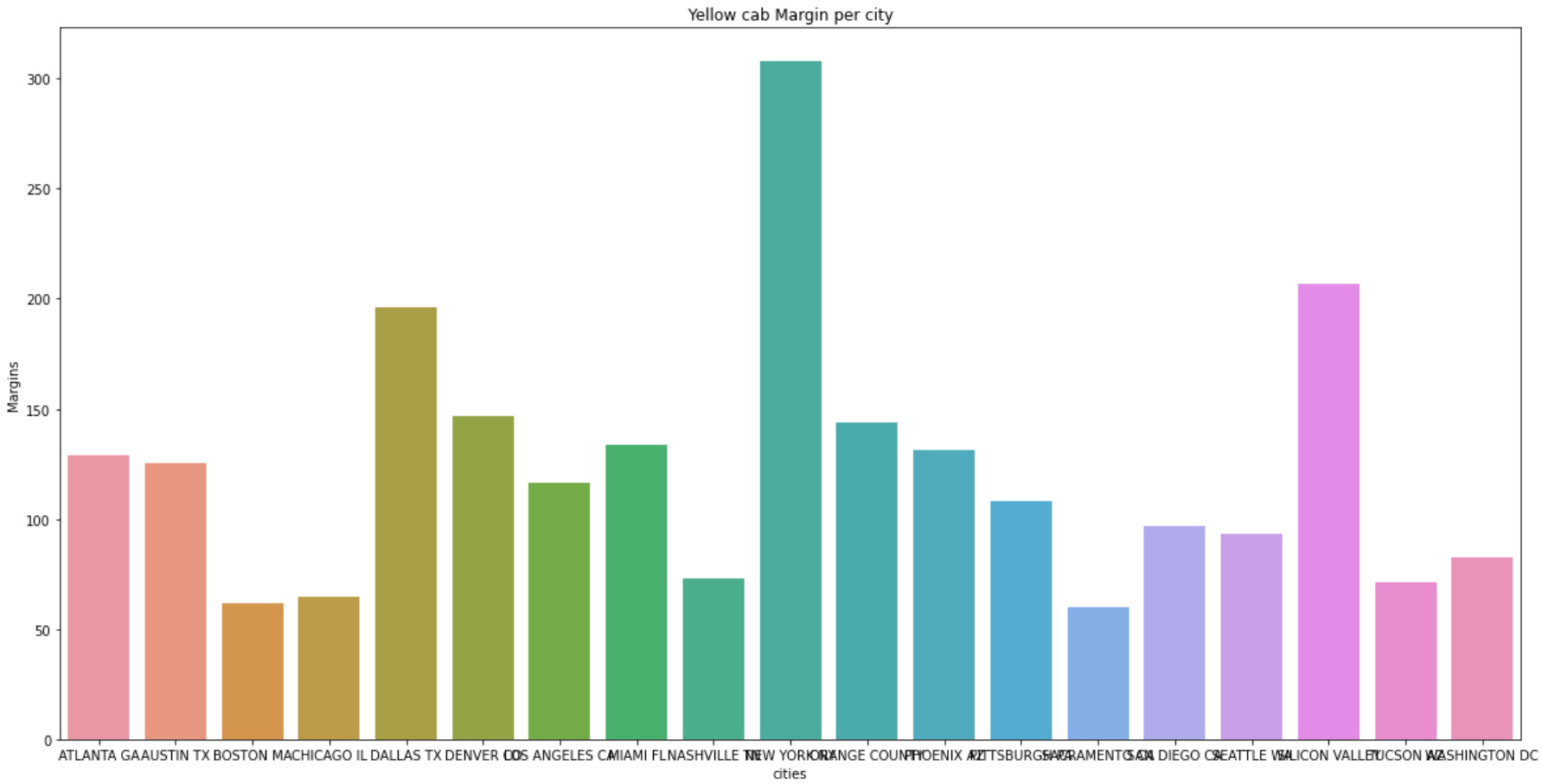


- From the above it is clear that Yellow cab has higher margins compared to Pink cab.
- Yellow cab also charges higher prices than Pink cab

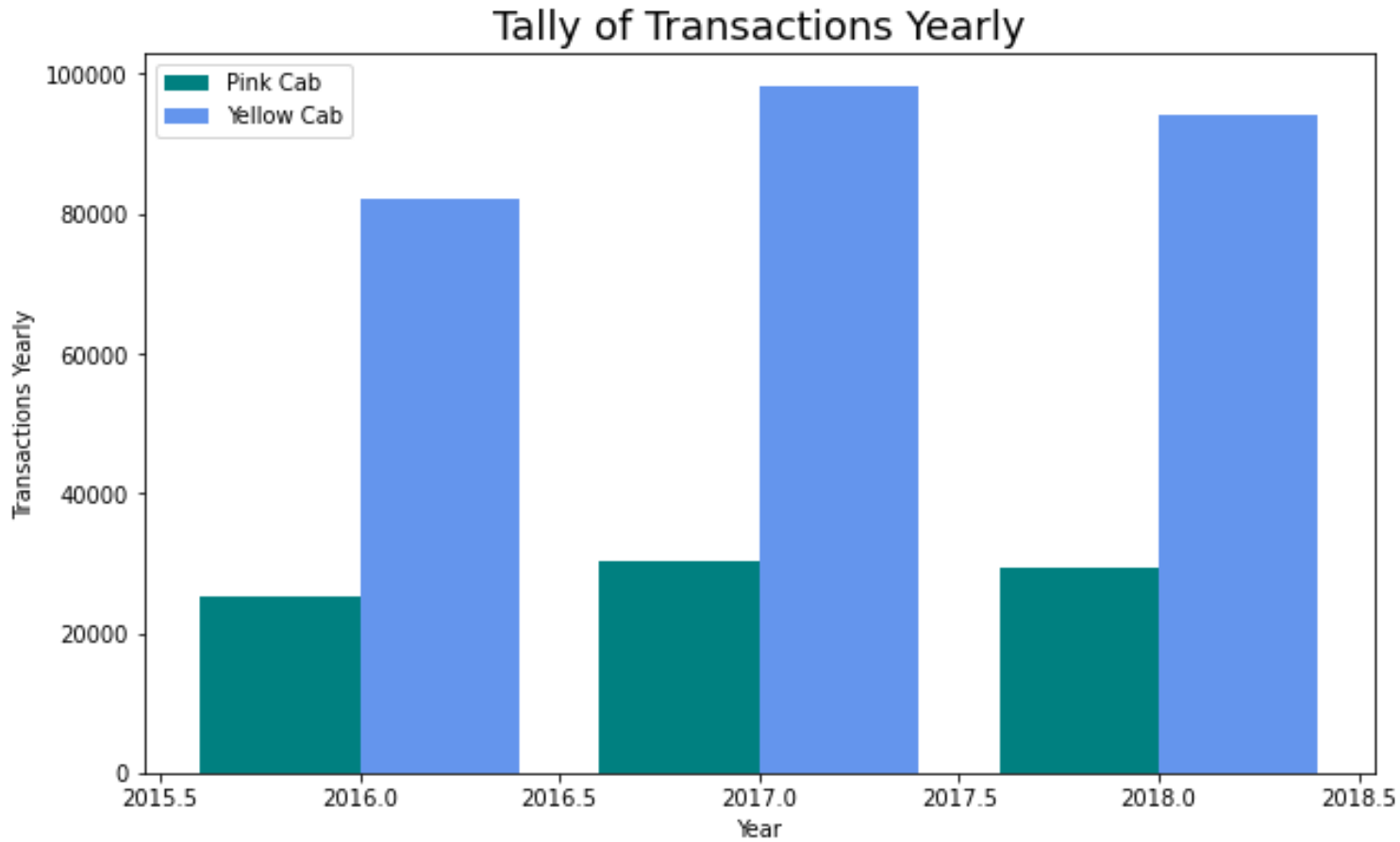
MARGIN PER CITY : PINK CAB



MARGIN PER CITY : YELLOW CAB

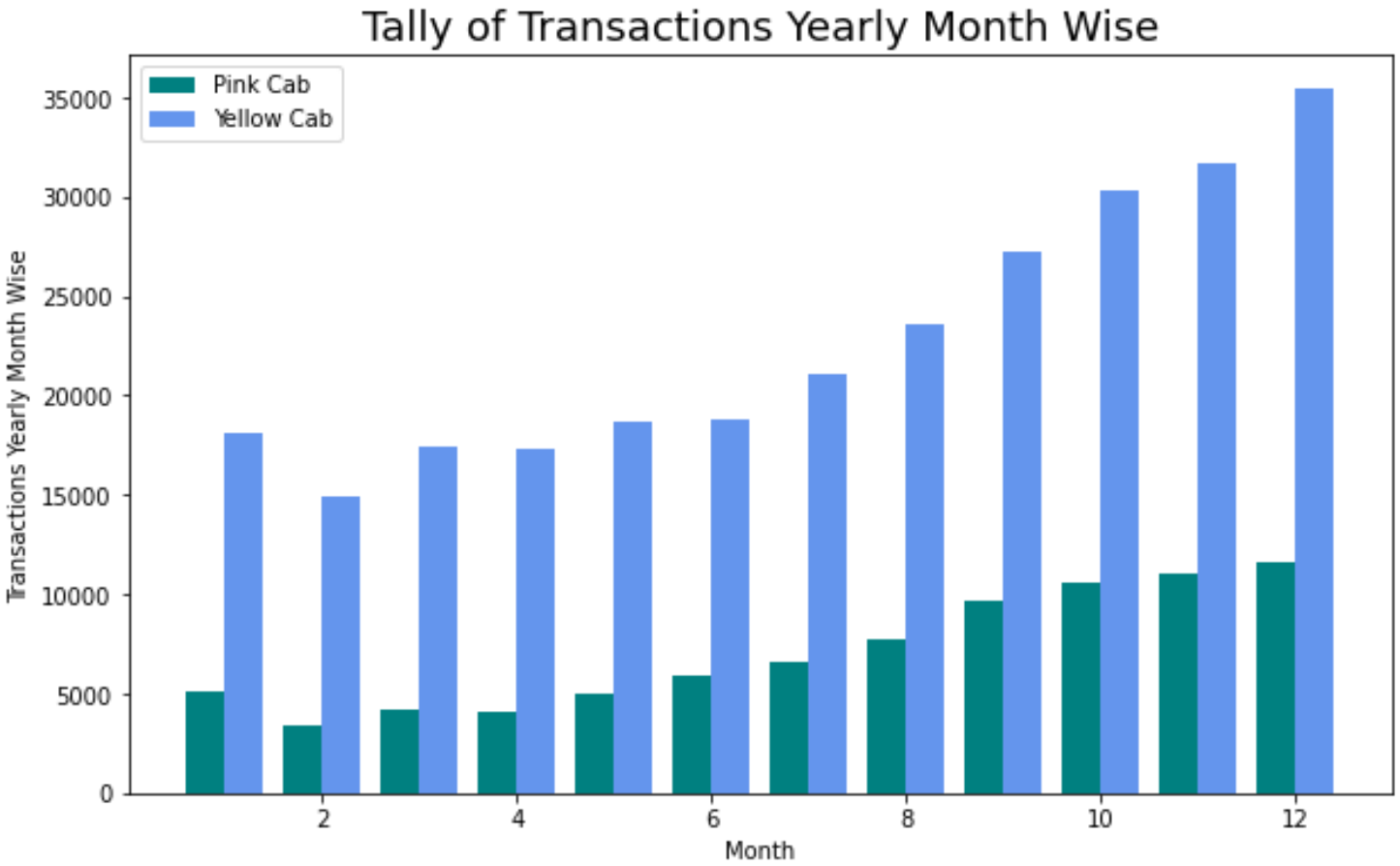


TRANSACTIONS YEARLY



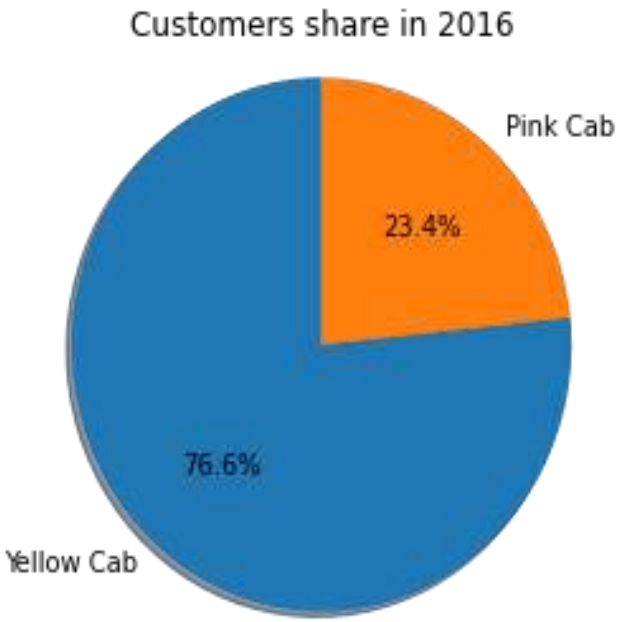
- Yellow cab has almost twice as much transactions as Pink cab

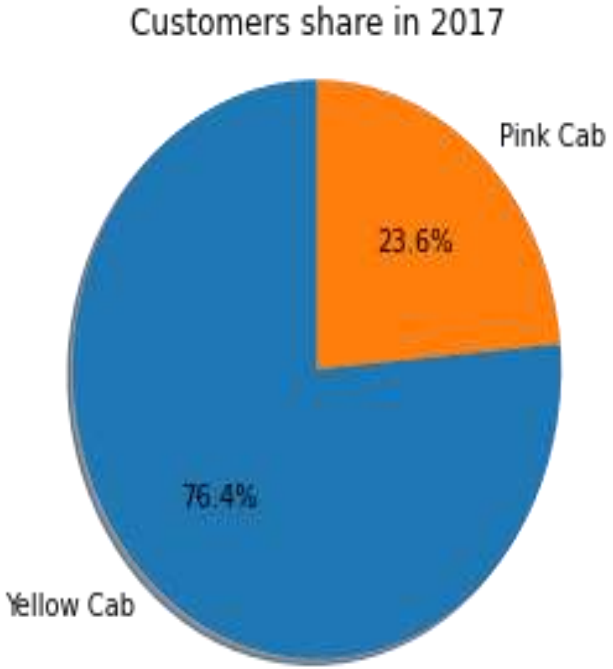
TRANSACTIONS PER MONTH

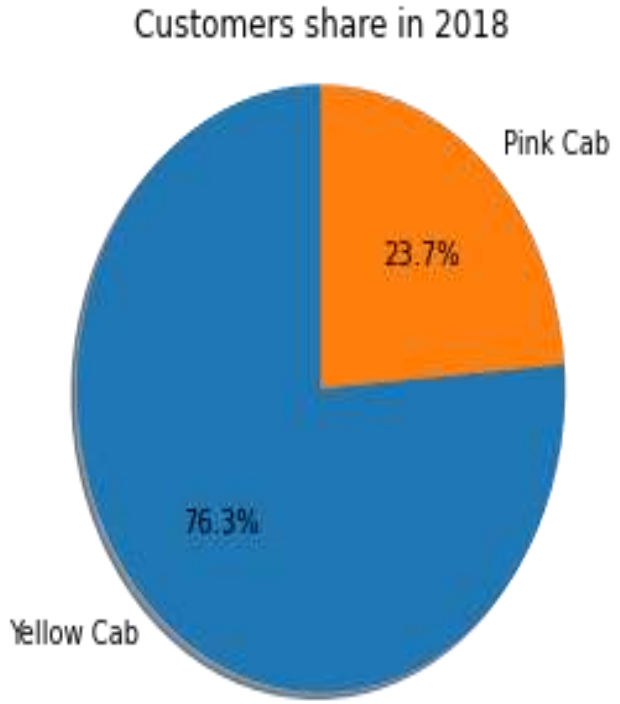


Transactions peak towards the end of the year and are highest in December which is the holiday season

CUSTOMER SHARE

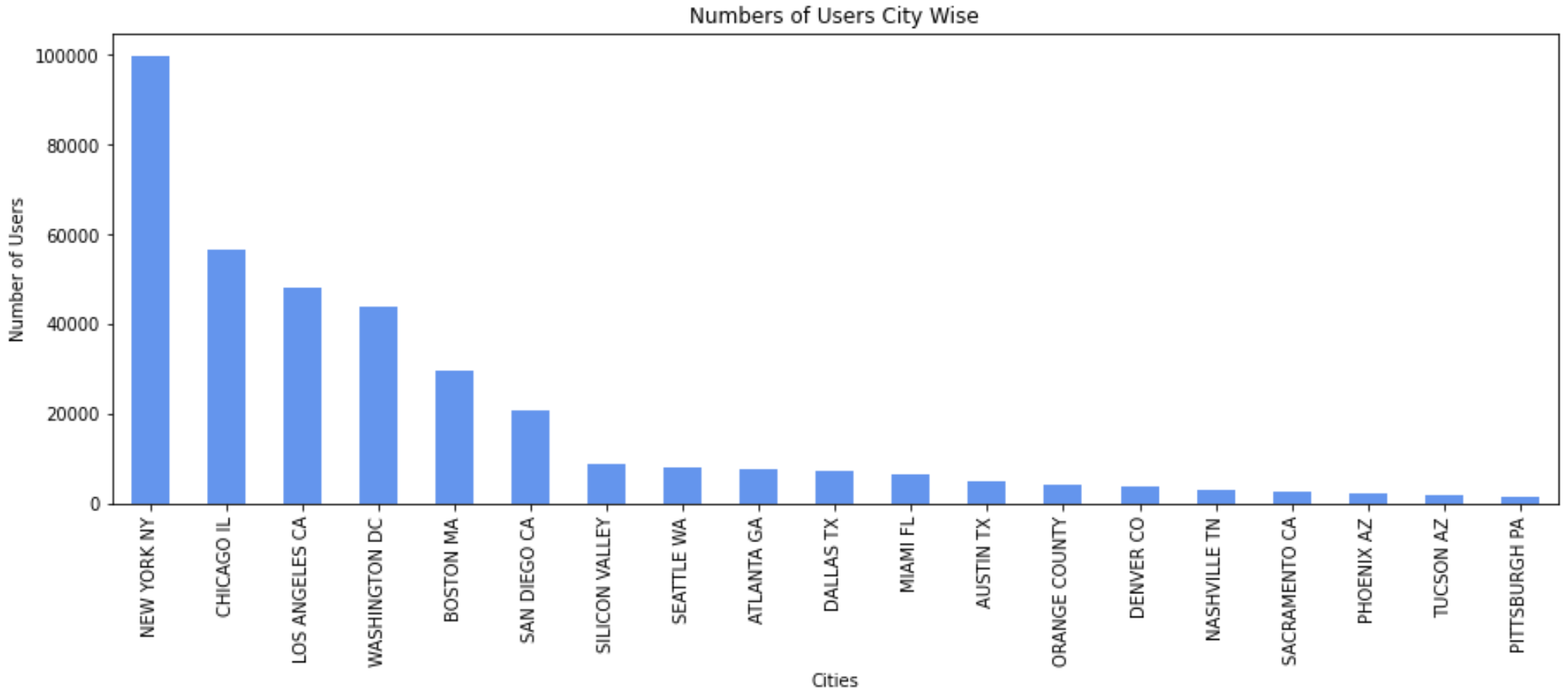






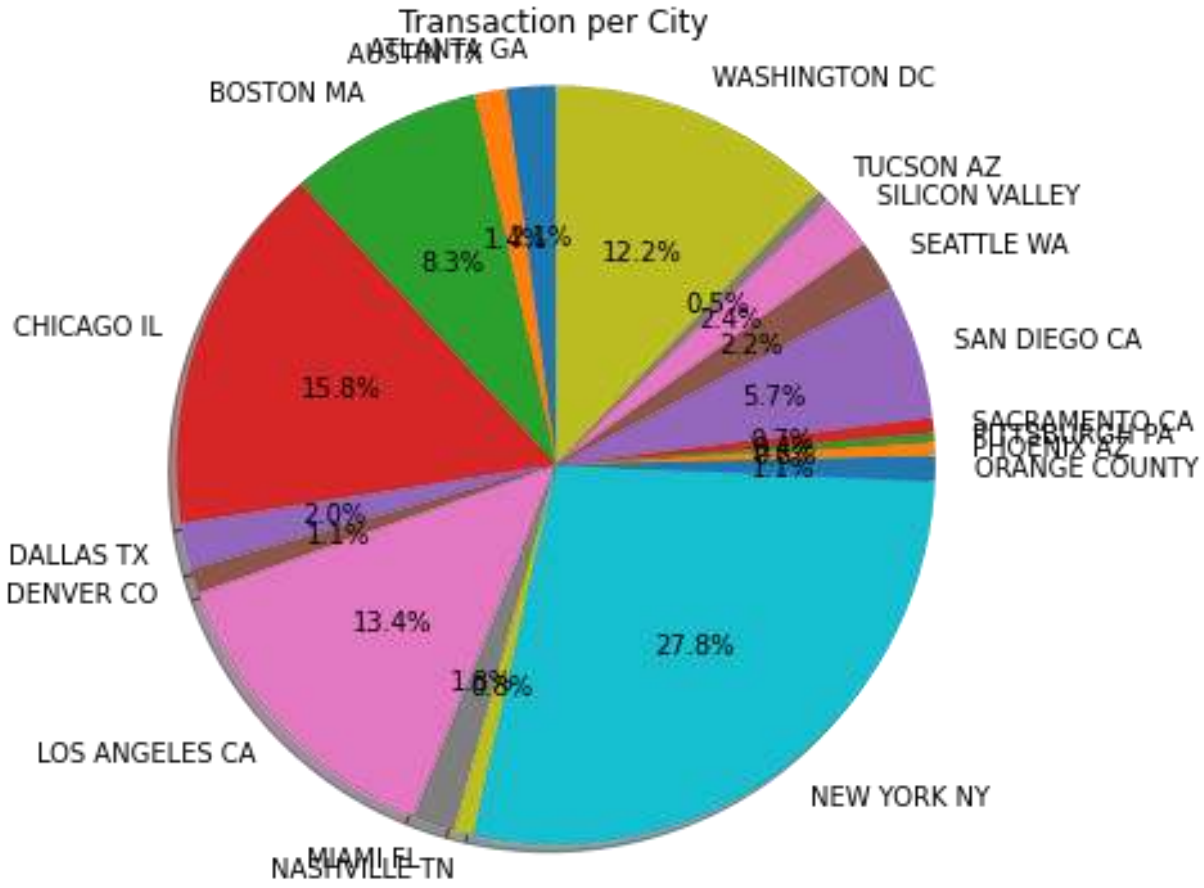
- Yellow cab has had three times as much customers from 2016 to 2018 as pink cab

NUMBER OF USERS PER CITY

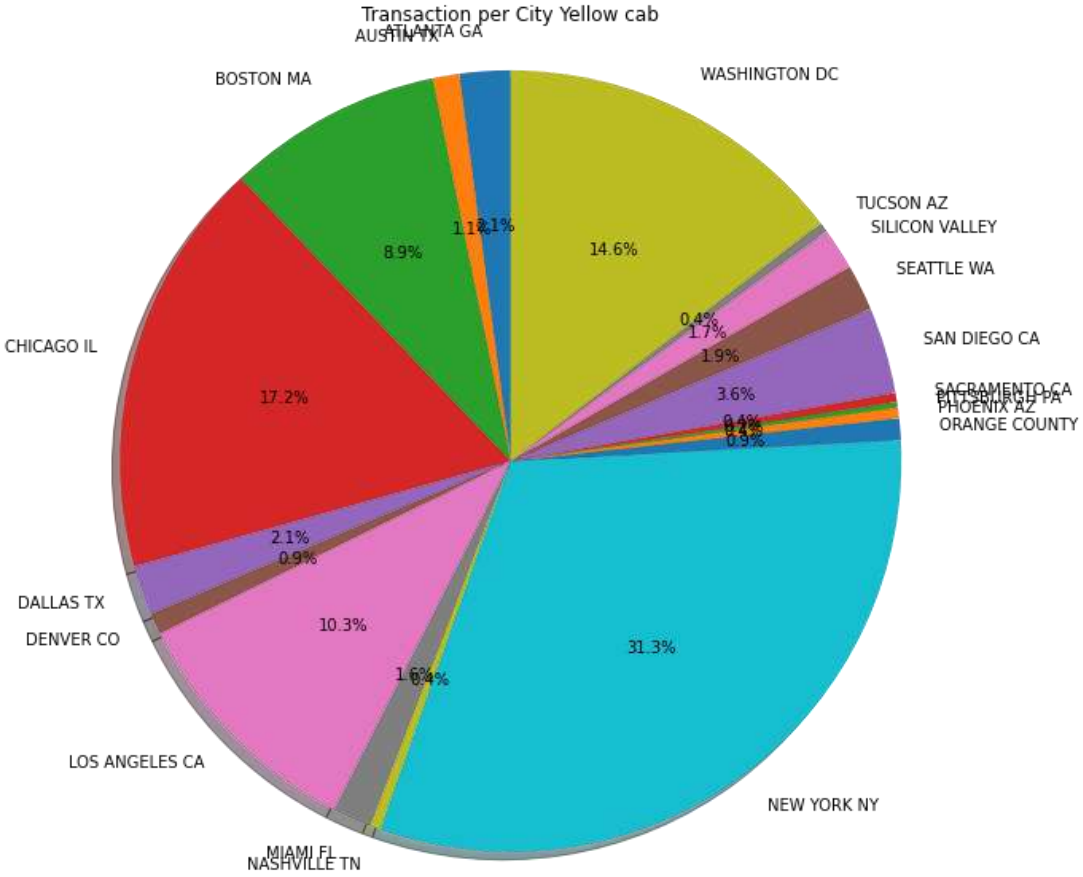


- From the above graph ,New York City has the highest number of cab users

TRANSACTION PER CITY



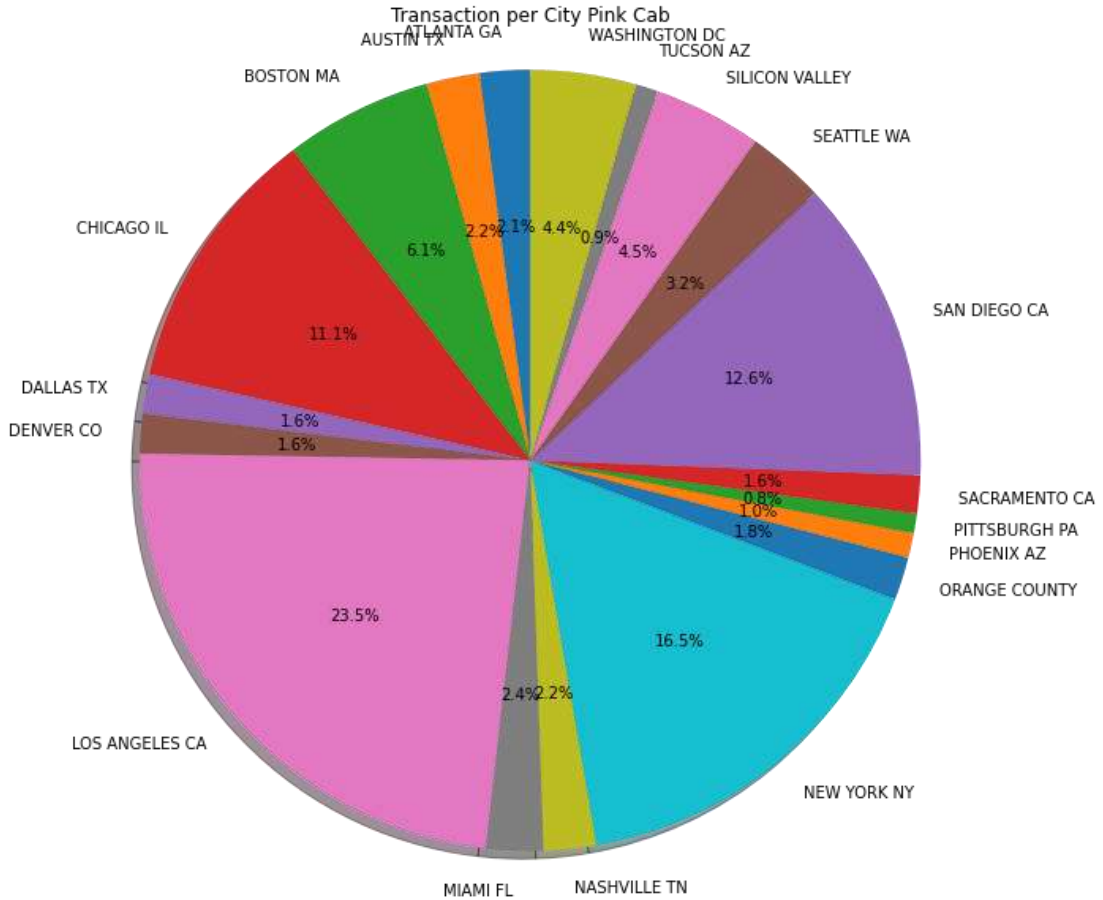
TRANSACTION PER CITY YELLOW CAB



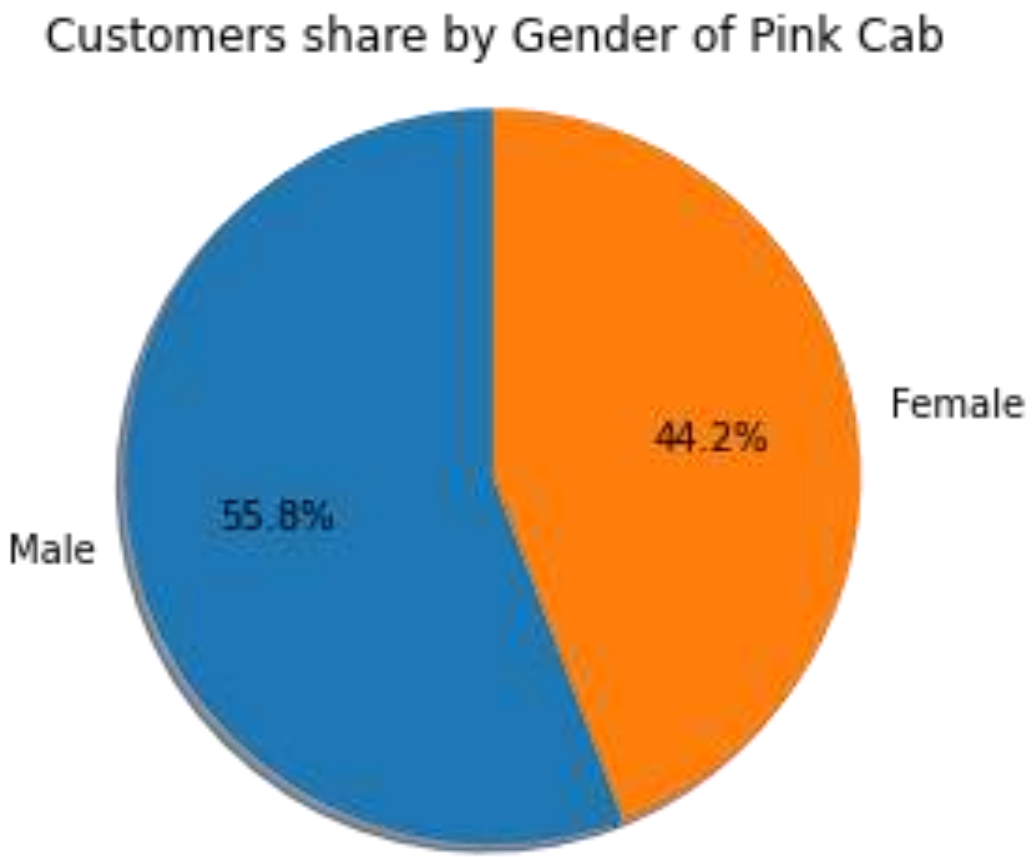
- Yellow cab has more transactions in New York city which has the highest number of cab users

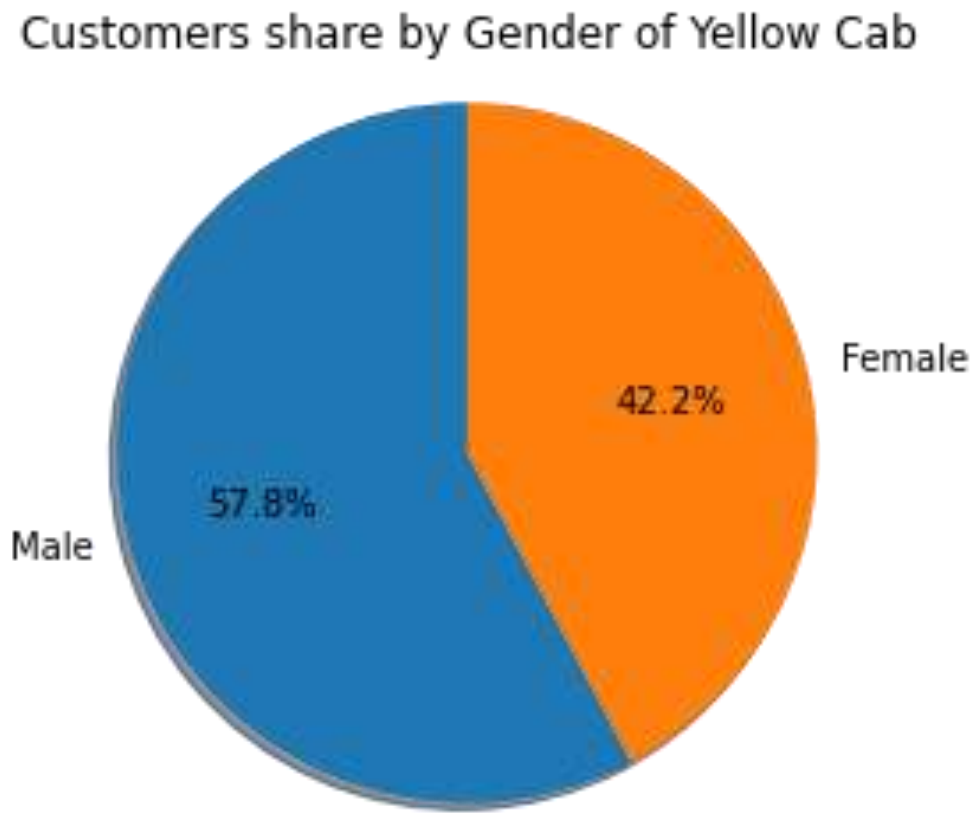
TRANSACTION PER CITY PINK CAB

Pink cab has a higher number of users in Los Angeles at 23.5% compared to New York City which has the highest number of Cab users

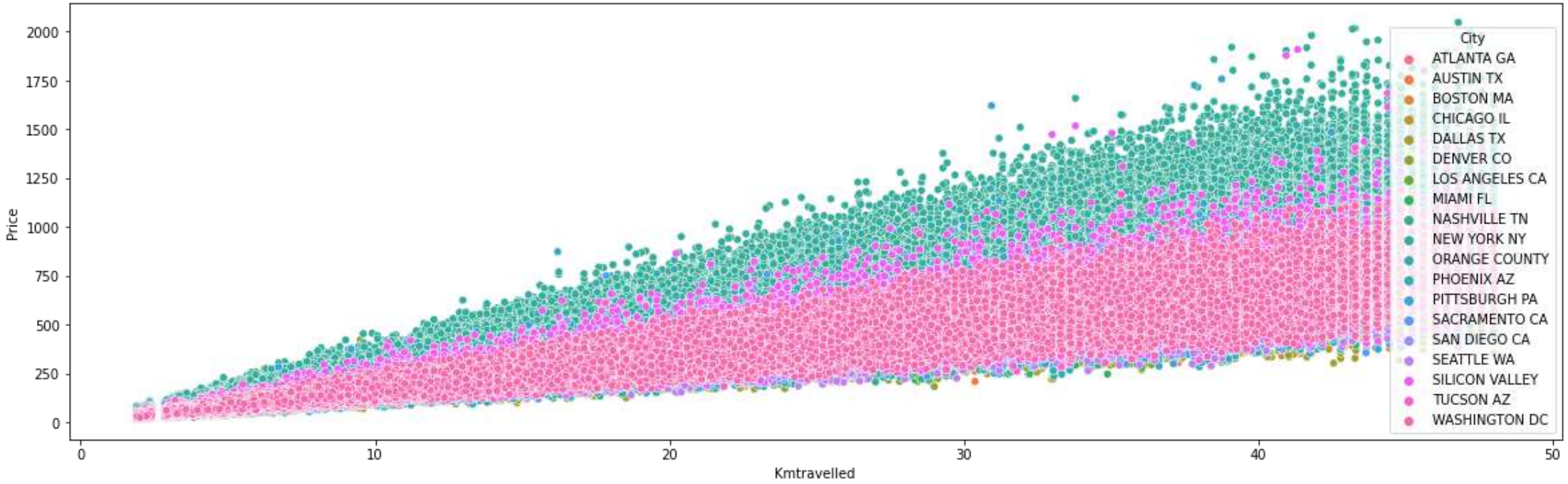


CUSTOMER SHARE



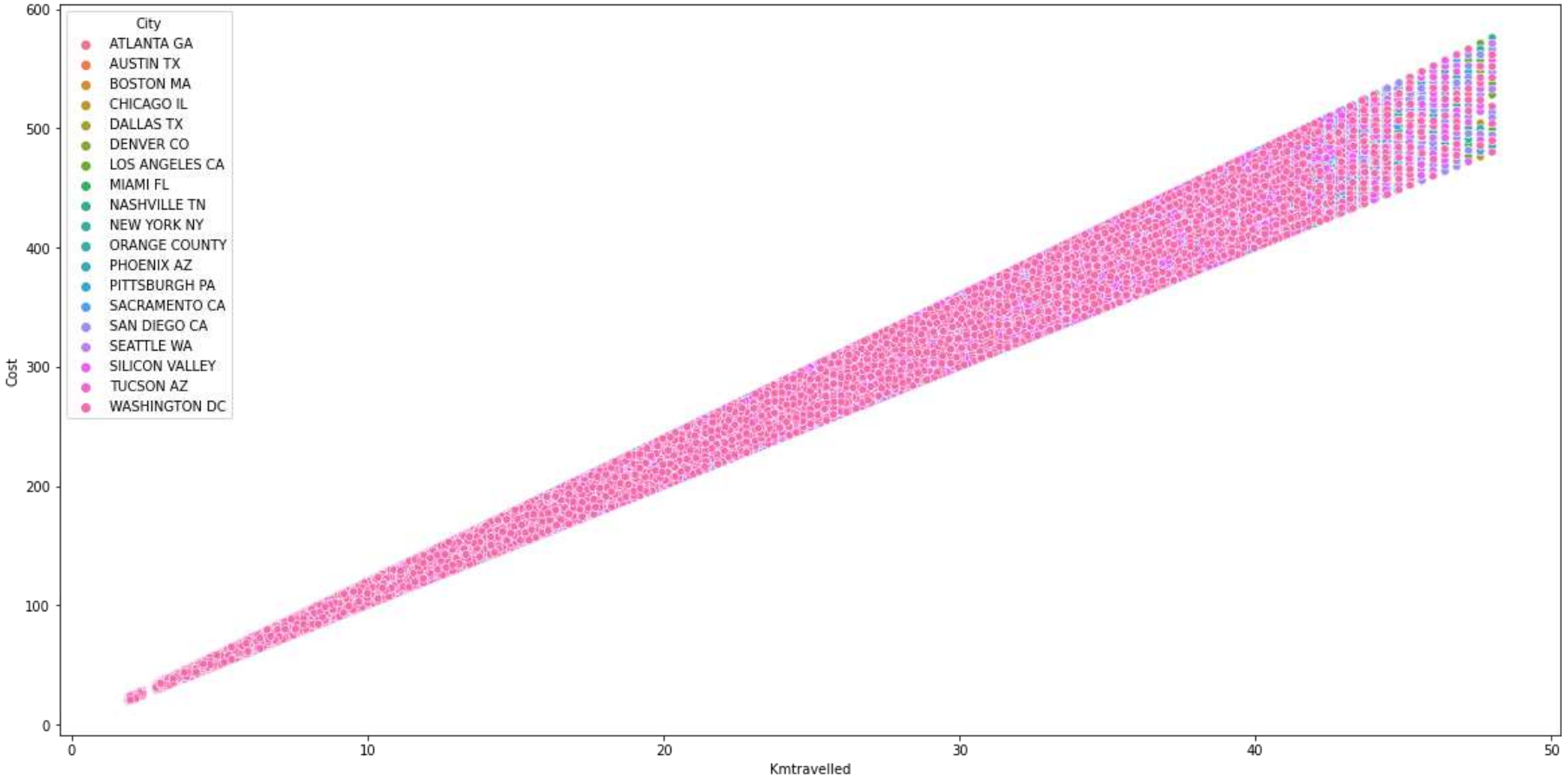


PRICE CHARGED PER KM TRAVELLED



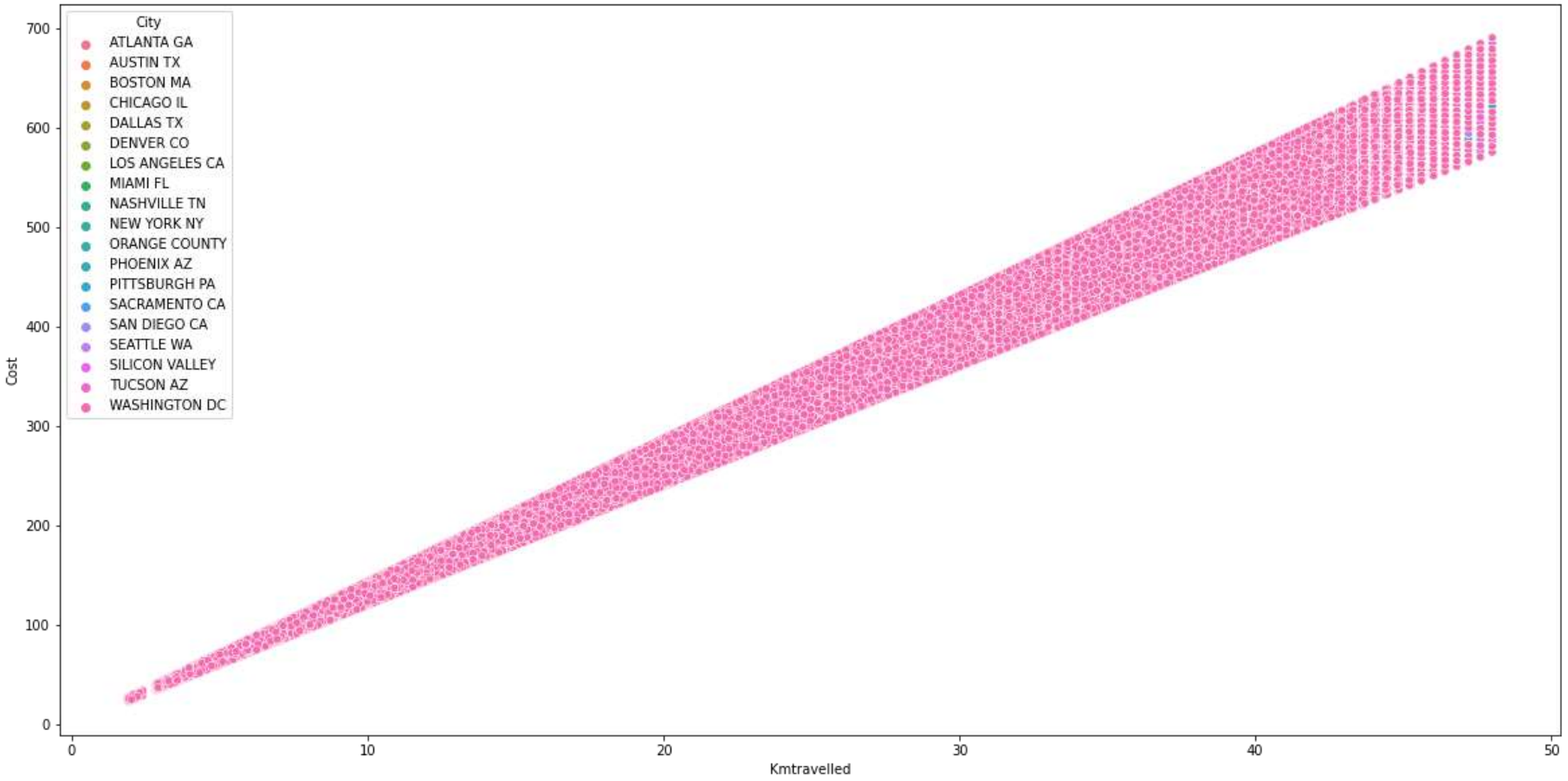
- New York City charges the highest prices in comparison with other cities

PRICE CHARGED PER KM TRAVELLED PINK CAB



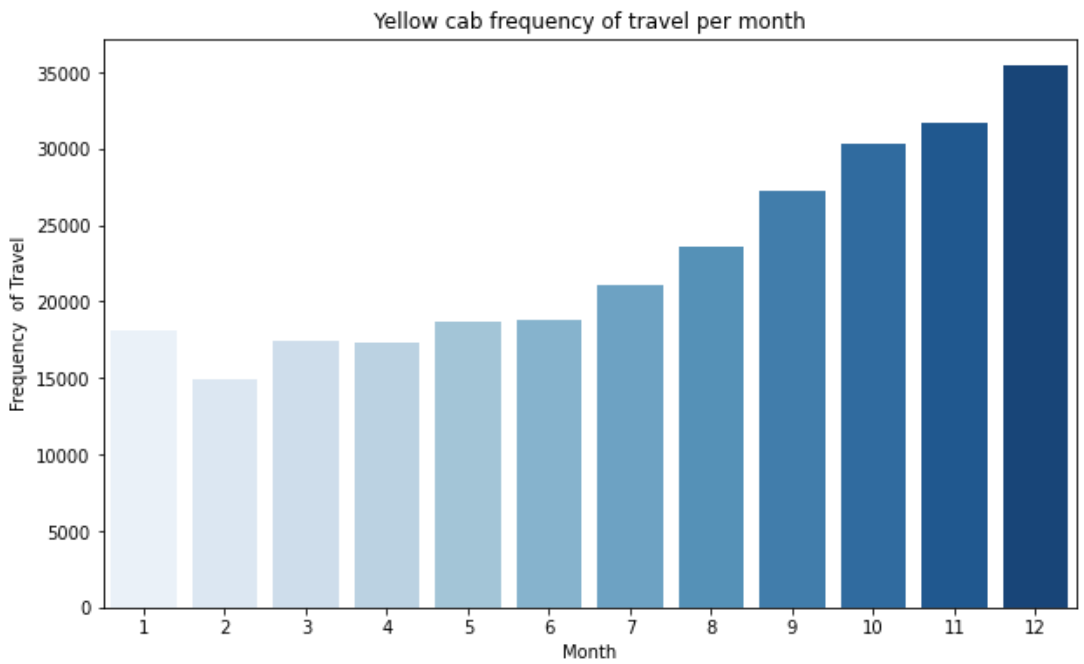
- Price charged increases linearly with distance

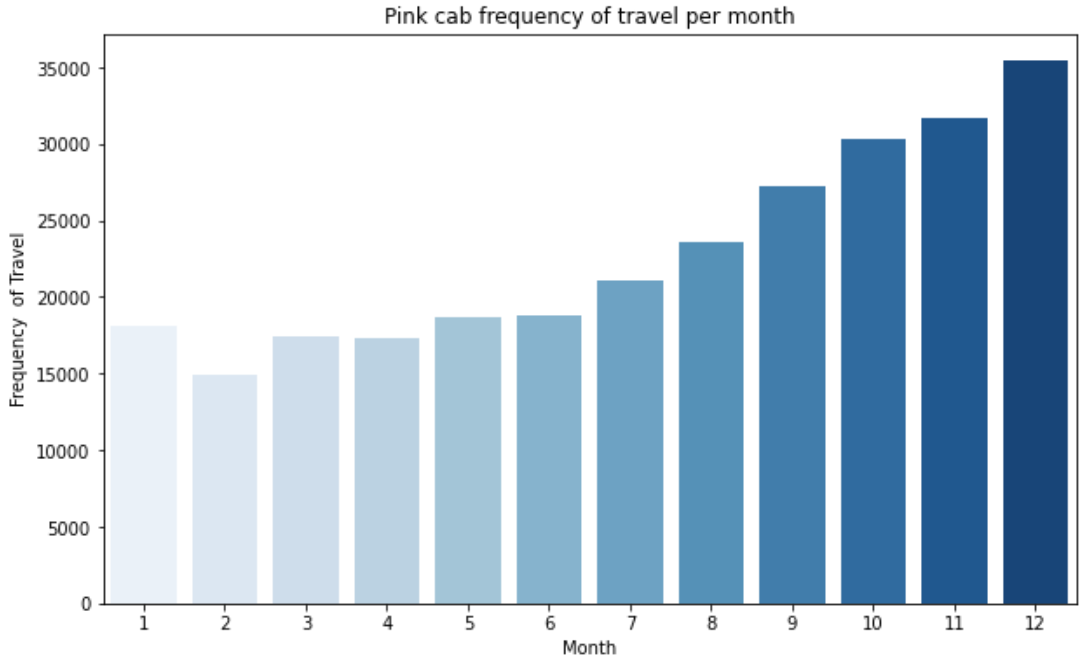
PRICE CHARGED PER KM TRAVELLED YELLOW CAB



- Prices charged increases with increase in distance for yellow cab.

TRAVEL FREQUENCY PER MONTH





- There’s a lot of travel in the month of December which is the holiday season.

HYPOTHESIS TESTING

HYPOTHESIS TESTING

HYPOTHESIS 1: There is no difference in margin given the mode of payment.

- Yellow Cabs : There is no difference in margin whether paid in cash or card.

```
print('P value is ', p_value)
```

```
We accept null hypothesis that theres no difference
P value is  0.29330606382987284
```

- Pink Cab: There is no difference in margin given the mode of payment.

```
print('P value is ', p_value)
```

```
We accept null hypothesis that theres no difference
P value is  0.7900465828793288
```

HYPOTHESIS : COMPANY MARGIN REMAINS THE SAME FOR ALL GENDER

- Yellow Cab: Gender does affect the company margin.

```
print('P value is ', p_value)
```

```
We accept alternate hypothesis that theres a difference
P value is  6.060473042494144e-25
```

- Pink Cab: Gender does not affect the company profit.

```
print('P value is ', p_value)
```

```
We accept null hypothesis that theres no effect
P value is  0.5146654429411317
```

HYPOTHESIS: Gender has no effect on distance travelled.

- Yellow Cab : There is no difference in distance travelled for both gender.

```
print('P value is ', p_value)
```

```
We accept null hypothesis that theres no effect
P value is  0.6164626165258722
```

- Pink Cab: There is no difference in distance travelled for both gender.

```
print('P value is ', p_value)
```

```
We accept null hypothesis that theres no effect
P value is  0.08738489267415496
```

RECOMMENDATION

- Margin: Yellow cab has had higher margins(price charged- cost of trip) compared to pink cab from 2016 to 2018
- Transactions : Yellow cab has had almost twice as much transactions as pink cab.
- Customer share : Yellow has a higher number of customers from 2016 through to 2018.
- Customers of Yellow cab in New York city which has the highest number of cab users is higher than that of pink cab.

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

Thank You