



Data Glacier

Your Deep Learning Partner

Go-to-Market insight for Cab Investment firm

Virtual Internship
LISUM11
(1st July - 30th Sept)

Background:

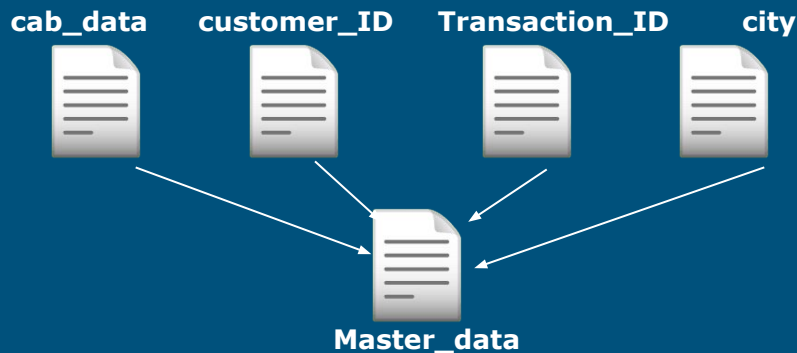
The Client:

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 to help XYZ firm in identifying the right company for making investment.

About data:



Cab_Data.csv - includes details of transaction for 2(Yellow, Pink) cab companies.

Customer_ID.csv – includes mapping table that contains a unique identifier which links the customer's demographic details.

Transaction_ID.csv – includes mapping table that contains transaction to customer mapping and payment mode.

City.csv – includes list of US cities, their population and number of cab users.

Time period of data is from **31/01/2016** to **31/12/2018**

Total number of features: **17 features**

Total number of records: **359392**

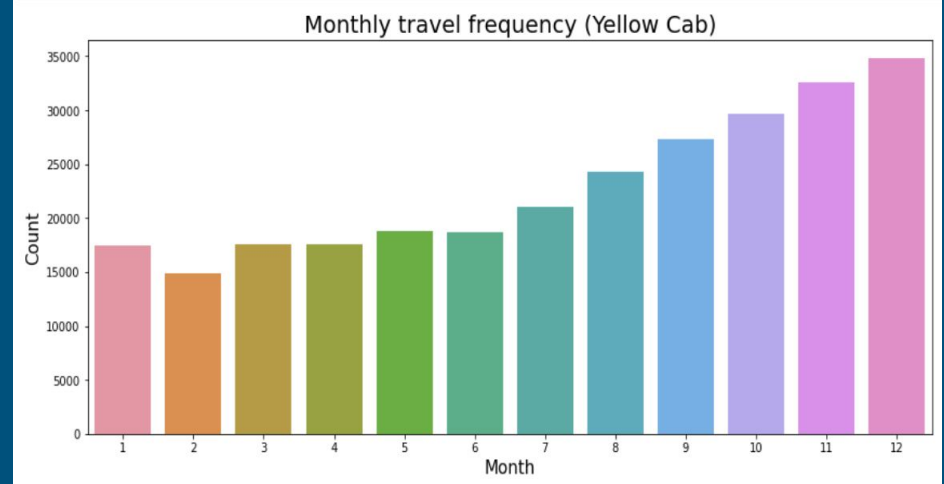
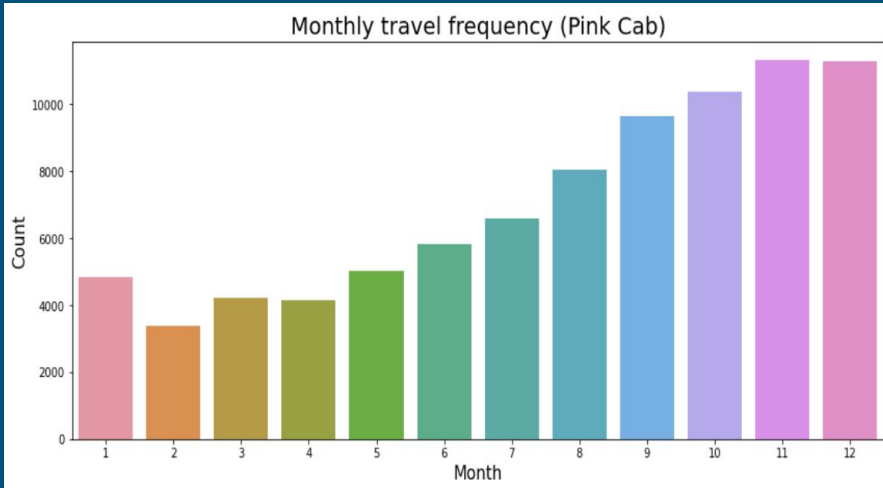


Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis *(EDA)*

Monthly travel frequency analysis:



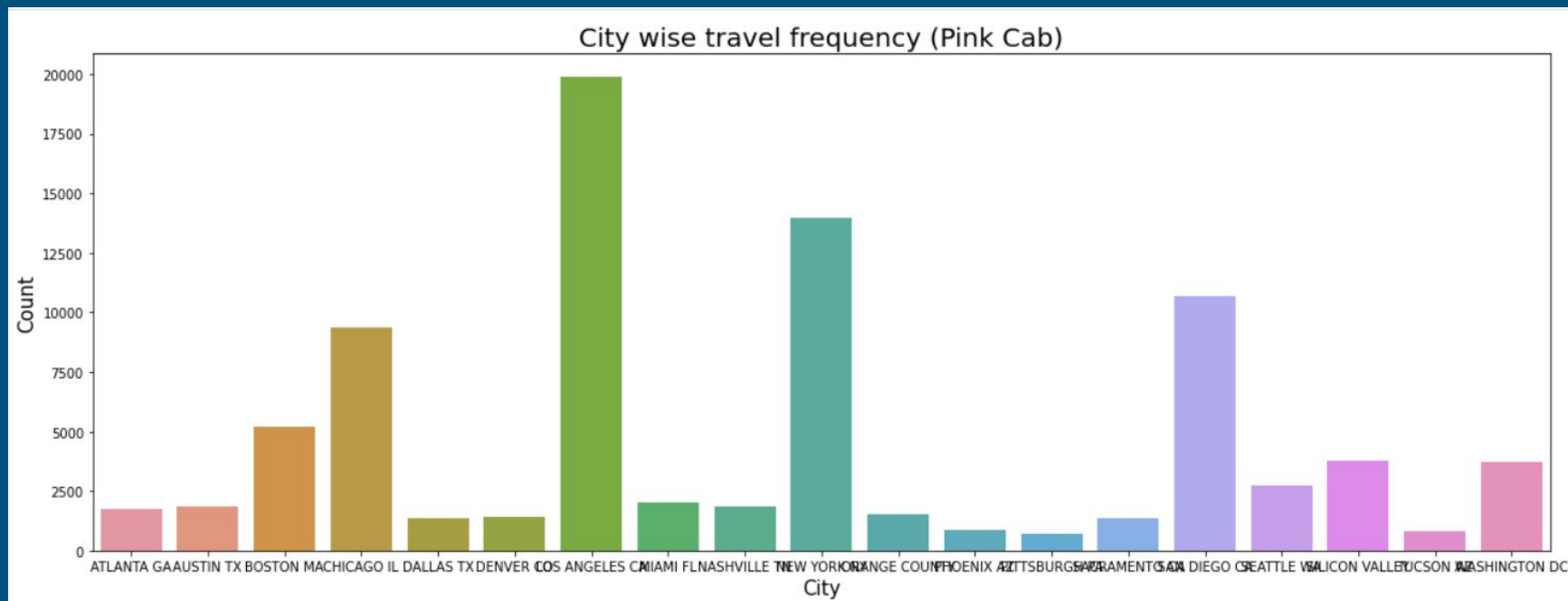
Observation:

Yellow cab has the highest number of travels and for the month of December, highest number of travels for the cab companies can be seen respectively.

City-wise travel frequency:

Pink cab:

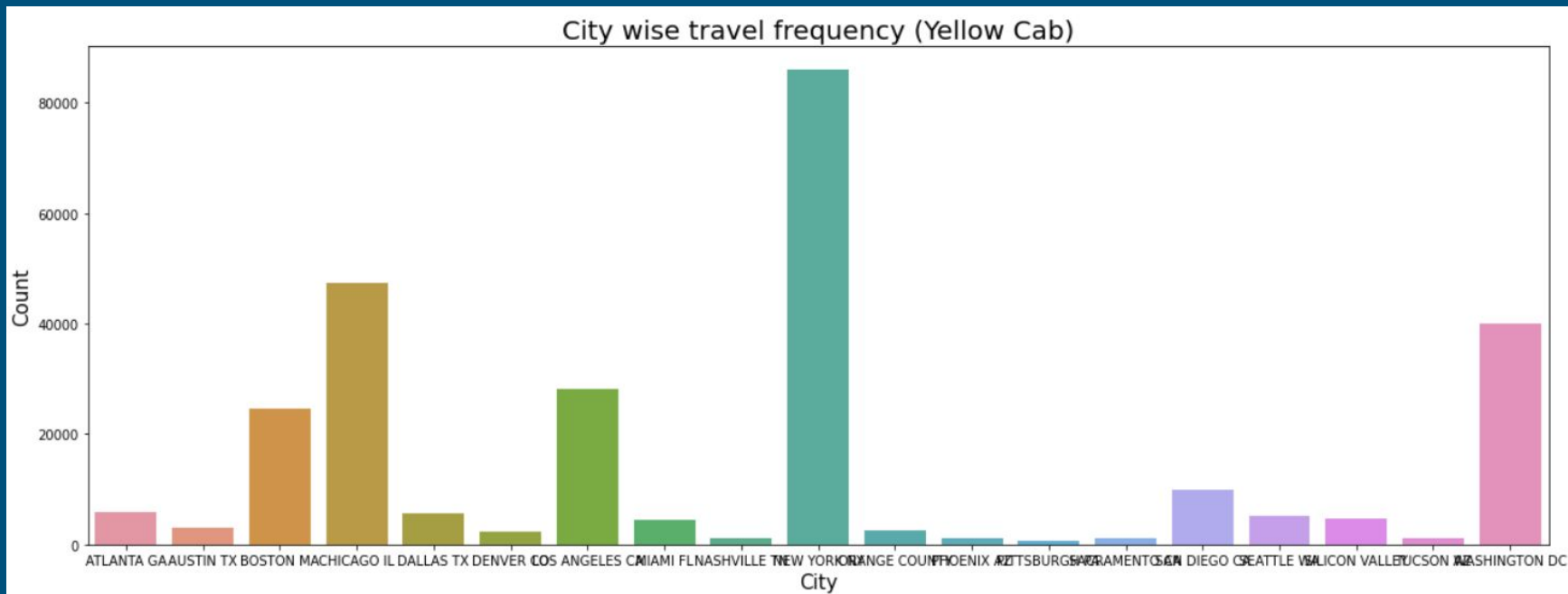
Los Angeles has the maximum count of 18000



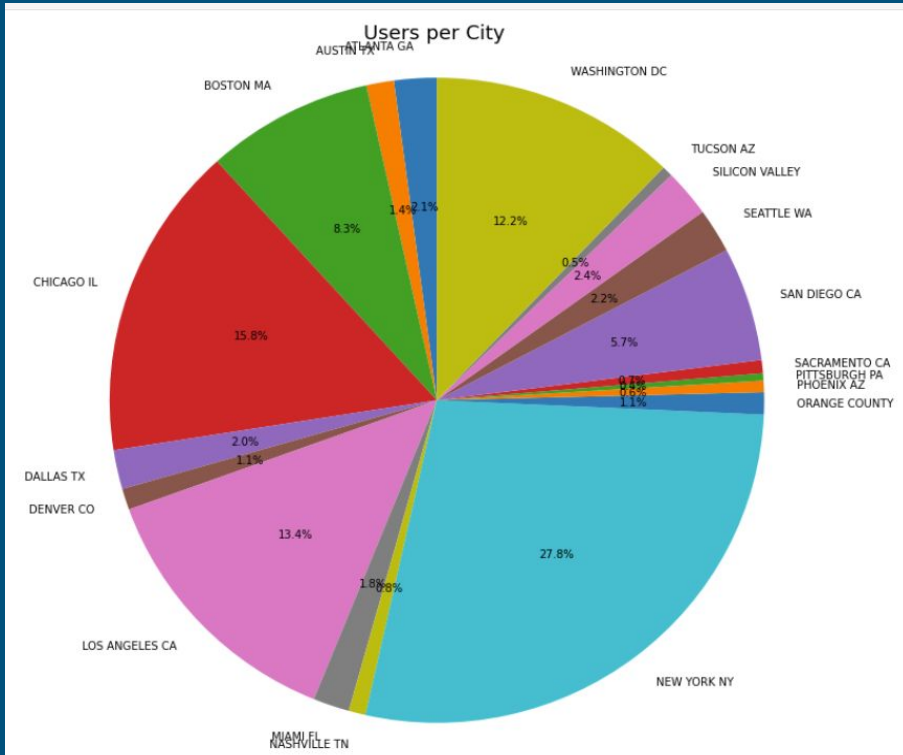
City-wise travel frequency:

Yellow cab:

New York City has the maximum count of 9000



Users per city:

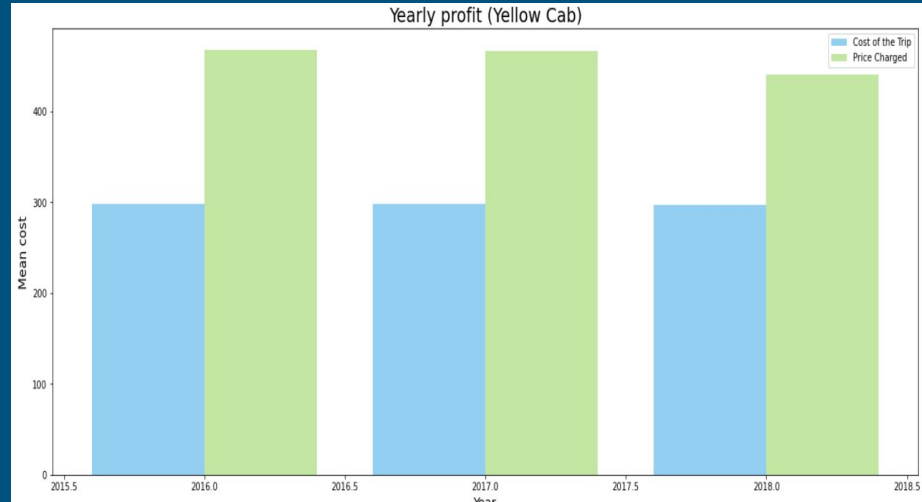
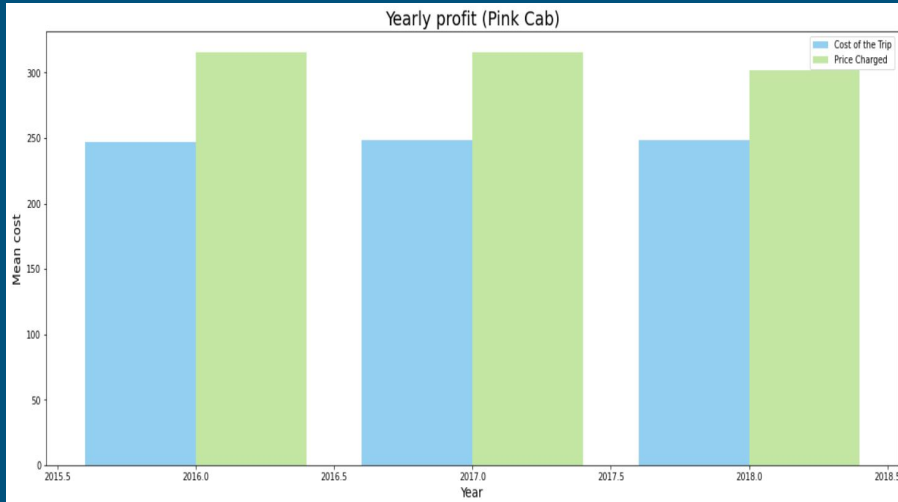


Observation:

New York city has the highest number of cab users.

From our previous analysis, Yellow cab has the maximum travel count in New York. Therefore, Yellow cabs are the most used in NYC which also produces highest profit.

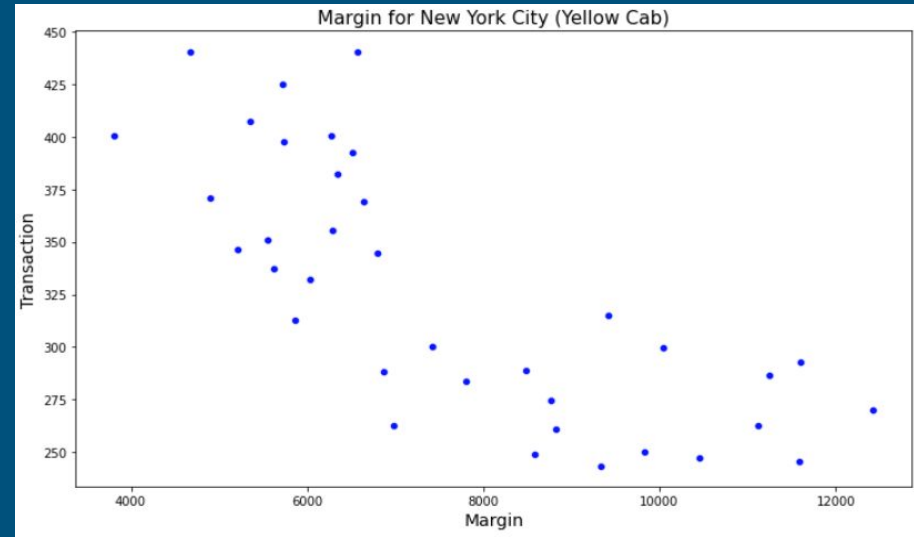
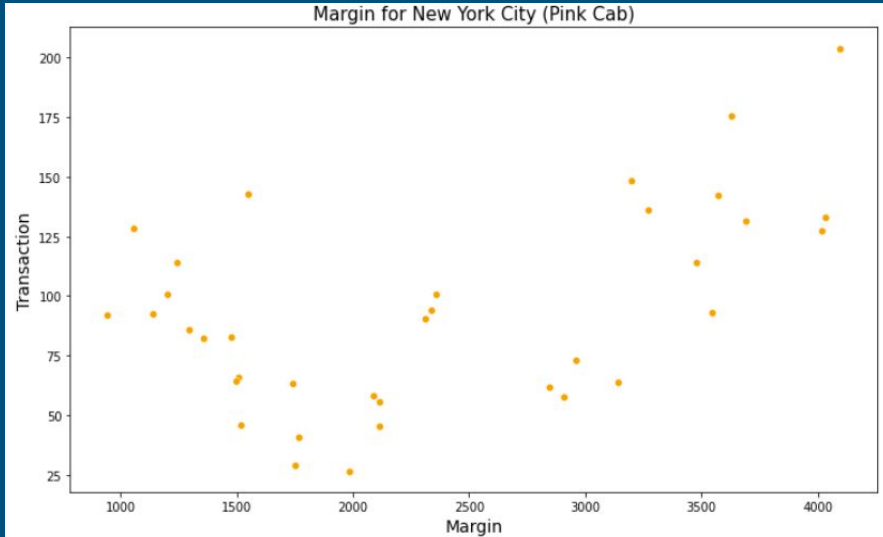
Profit Analysis:



Observation:

Most profit percentage for the cab companies can be seen in the financial year 2016. Overall, **Yellow cab** company has the highest profits in each financial year. Now, let's see profit for NYC for both cab companies.

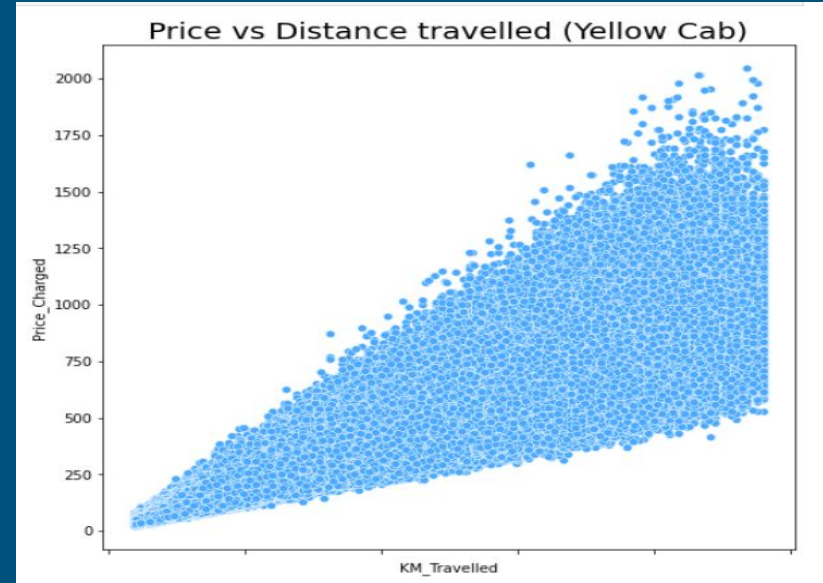
Margin for New York City:



Observation:

Yellow cab has most users/transactions in the New York City but from the above graph, we can see that the margin for Yellow cab decreases with the increase in number of transaction.

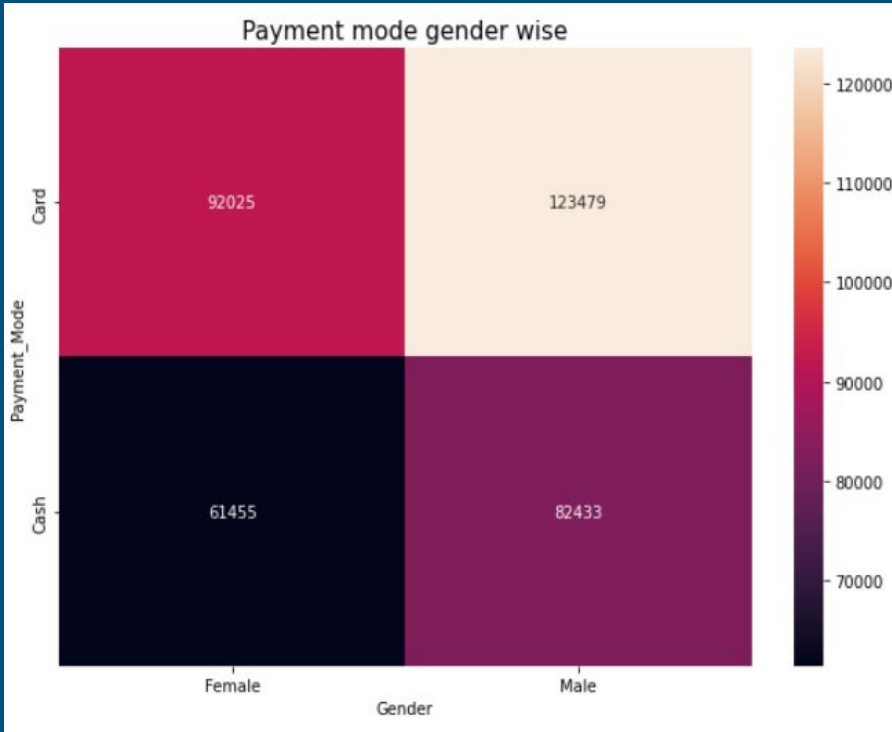
Price vs distance travelled analysis:



Observation:

We can observe that for both the cab companies prices increases with the increase in distance. That is, there is a positive correlation for both the cab companies. Yellow Cab charges more in comparison to Pink Cab.

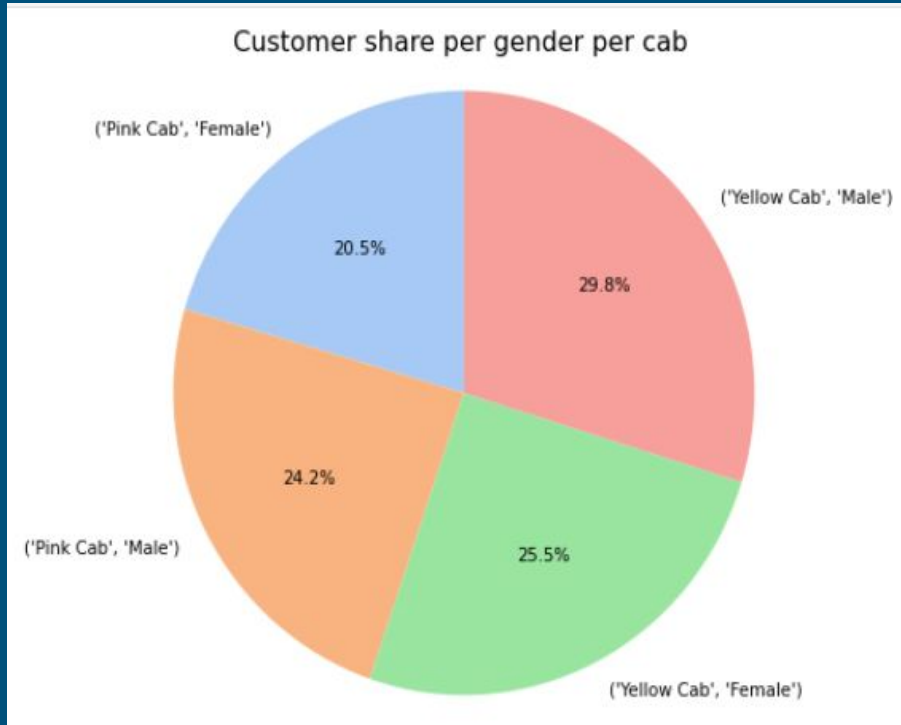
Payment mode:



Observation:

A high amount of both males and females prefer to do payment by card.

Gender-wise analysis:



Observation:

Yellow cab has more female customers than Pink cab.

Recommendations:

From EDA and various data visualisation in our presentation such as travel frequency, yearly profit, cities travelled, margin ratio, gender analysis the Yellow Cab company beats Pink Cab Company in every data analysis. Thus, I recommend XYZ firm to **invest in Yellow Cab Company**.

Some analysis and business recommendation:

- **New York** state is where the **highest investment** should be done because this state has more users, transaction, more profit margin and distance travelled.
- The male:female ratio is good in both the cab companies therefore, both cabs are safe enough.
- Yellow cab has more **female users** therefore it is safe and discounts can be made available to them during women's day, mother's day, breast cancer month etc.
- Yellow cab is quite popular among **senior citizens** (60+) therefore discounts can be made available to them to have more profits.
- As most payments are done by card therefore **various card options** can be made available.
- Since for Yellow cab, price vs distance travelled has a profit correlation then bonus points or discount can be given for distance travelled after a specific value. This will make users to book cab for longer distances.



Data Glacier

Your Deep Learning Partner

Thank you :)