



**Data Glacier**

Your Deep Learning Partner

# G2M Case Study

Virtual Internship

21 - Sep - 2022

# 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

## ❖ Utilized 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

# Data Exploration

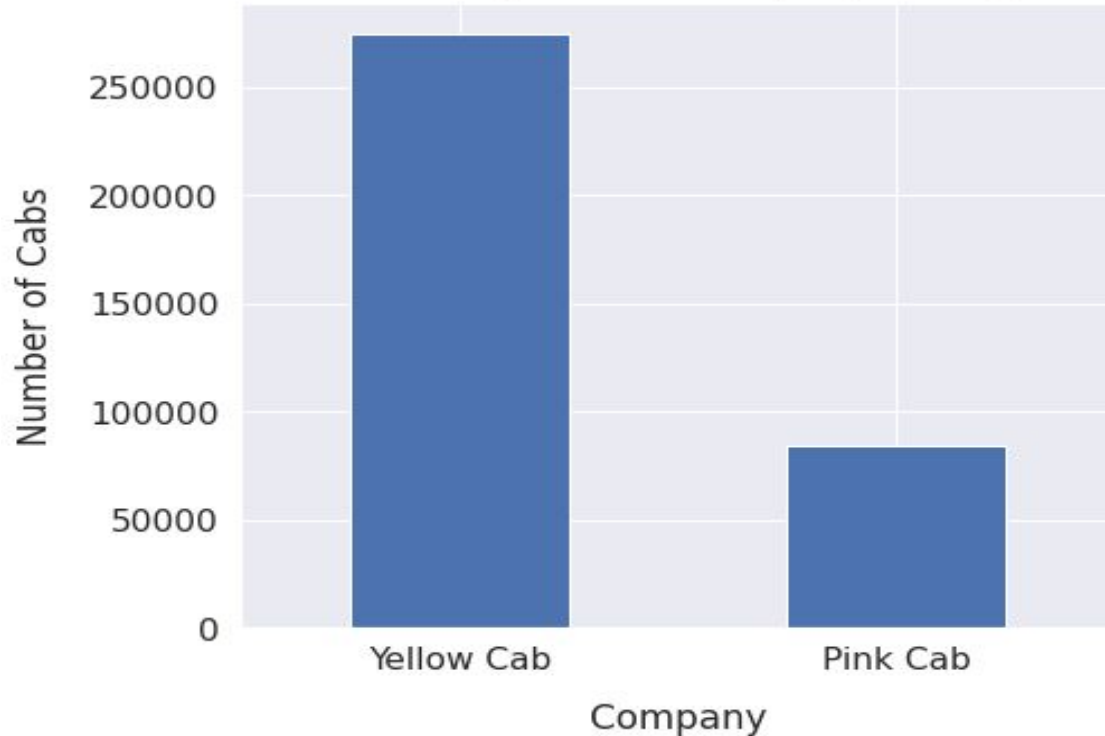
- Cab Data:
  - It has 359392 rows and 7 columns.
  - It consists of 7 features and they are : 'Transaction ID', 'Date of Travel', 'Company', 'City', 'KM Travelled', 'Price Charged', 'Cost of Trip'.
- City Data:
  - It has 20 rows and 3 columns.
  - It consists of 3 features namely, 'City', 'Population', 'Users'.
- Customer\_ID:
  - It consists of 49171 rows and 4 columns.
  - The 4 features are 'Customer ID', 'Gender', 'Age', 'Income (USD/Month)'.
- Transaction\_ID:
  - It consists of 440098 rows and 3 columns.
  - The 3 features are 'Transaction ID', 'Customer ID', 'Payment\_Mode'

# Hypothesis

- Which company has maximum cab users at a particular time period?
- Is there any relation between the age and the cabs utilized by them?
- Are companies having a good business?
- Which company performs better?
- Where to invest in?
-

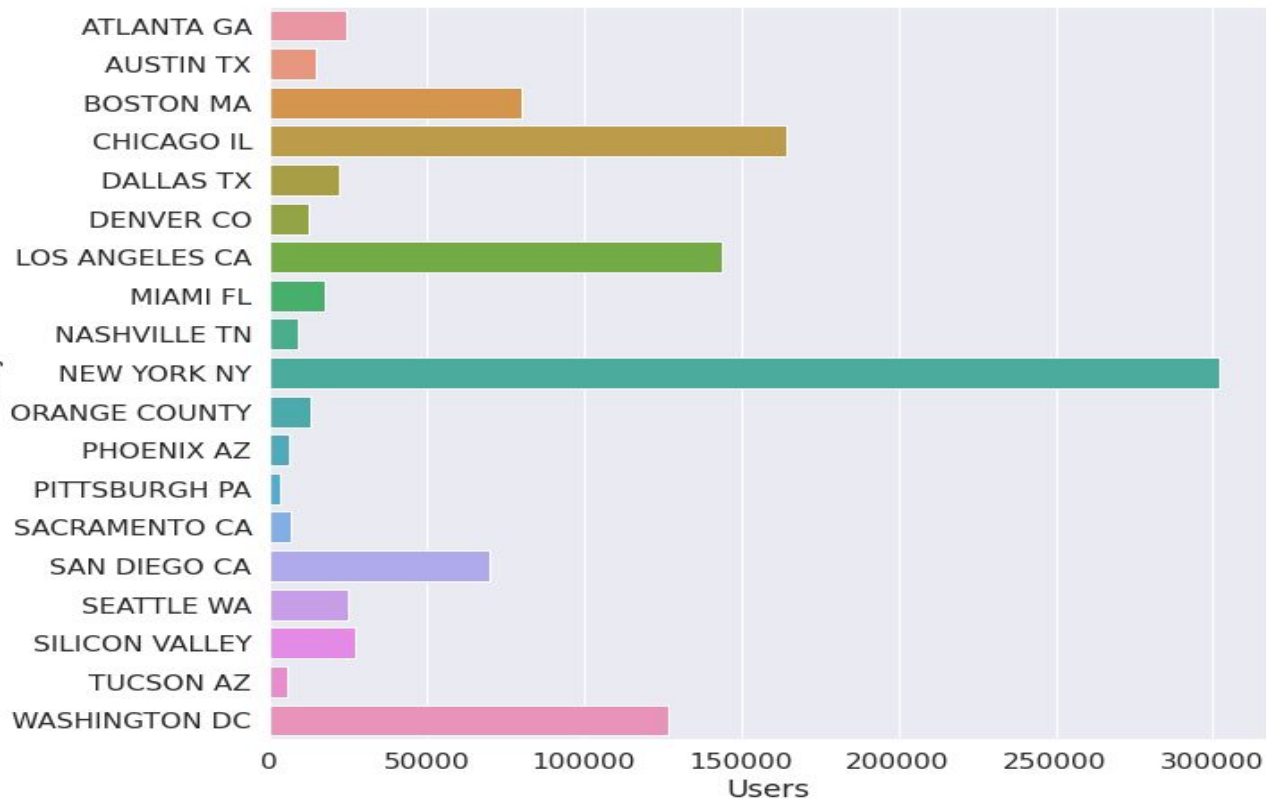
# Cabs per company

Number of cabs of a particular company running in the cities



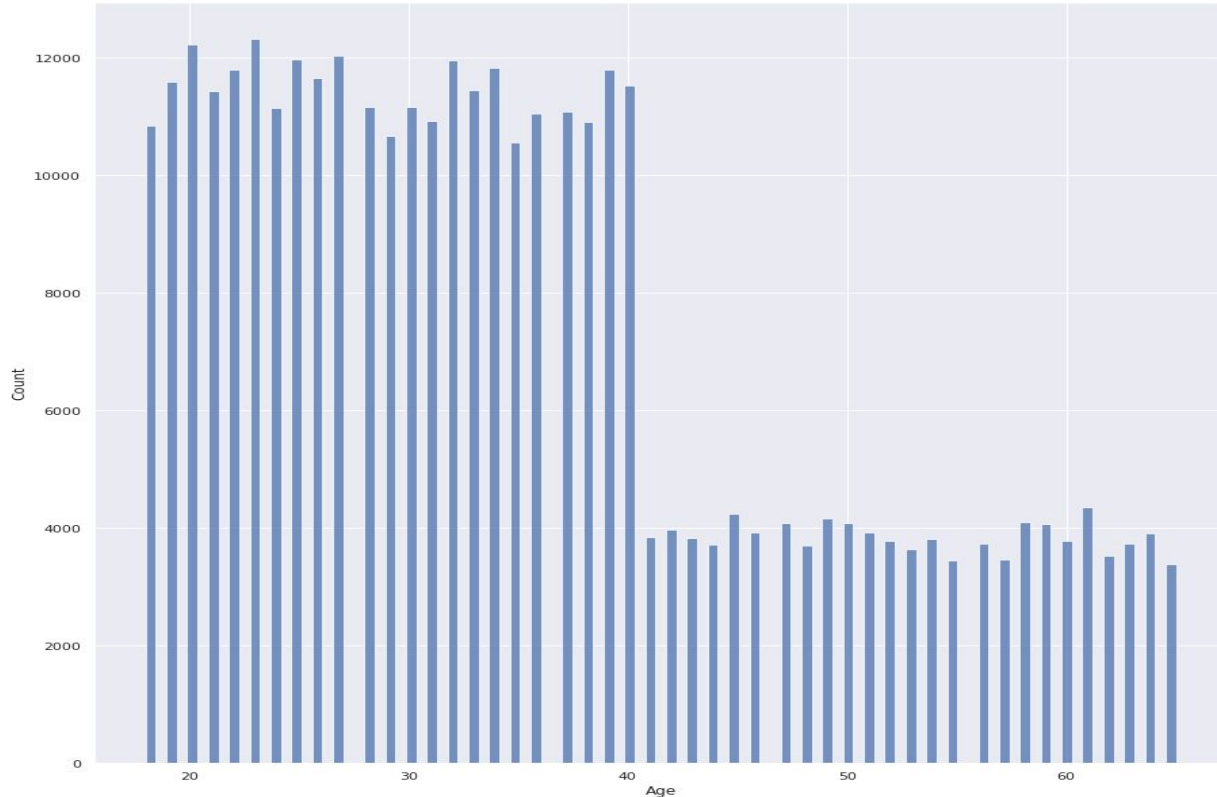
- More yellow cabs running in the cities compared to the pink cabs

# Users vs City



- New York has most number of users which is more than 300,000 followed by Chicago and Los Angeles.

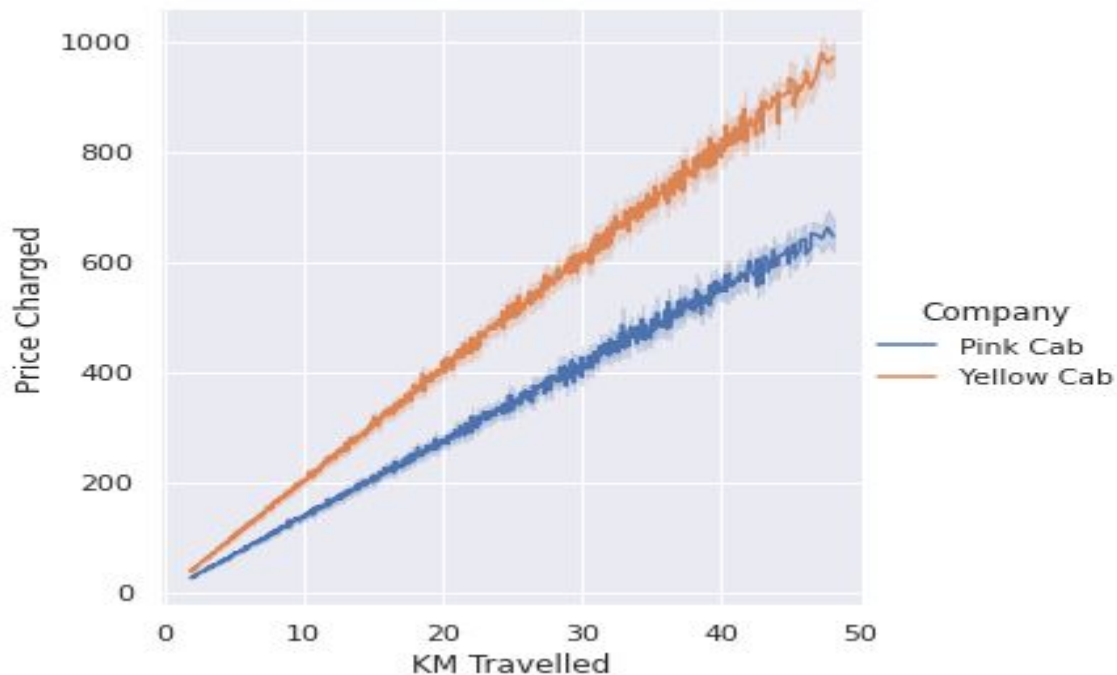
# Cab requirement analysis as per the age



- Significant decrease in the number of users opting for the cab service after the age of 40.

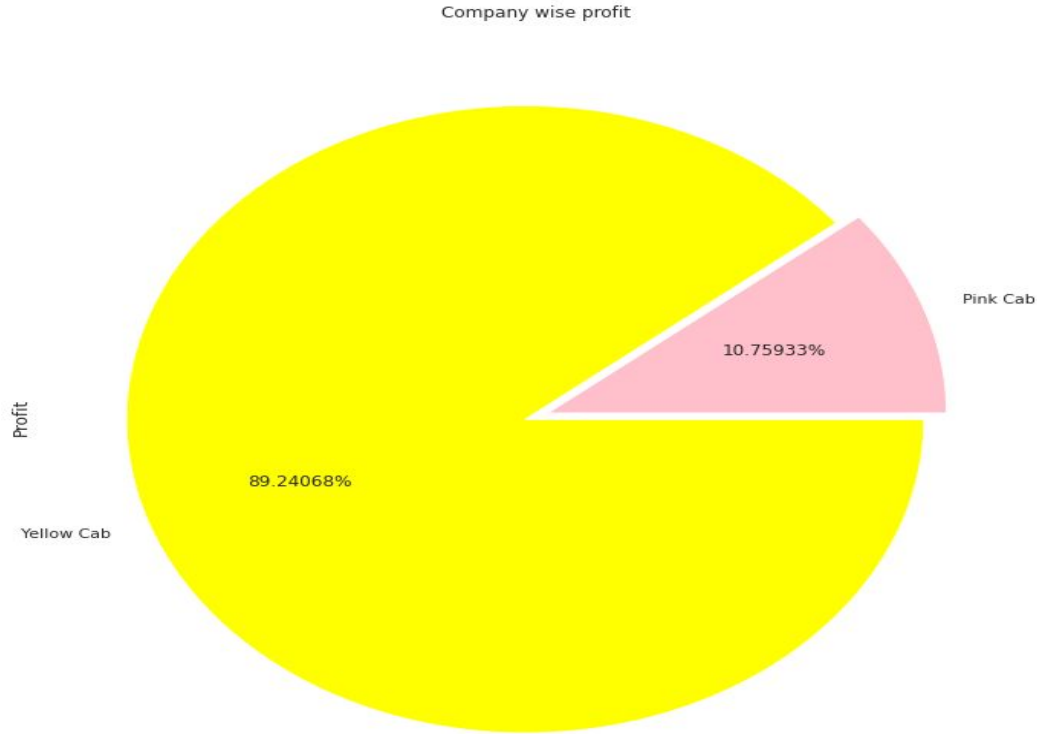


# Price vs KM Travelled



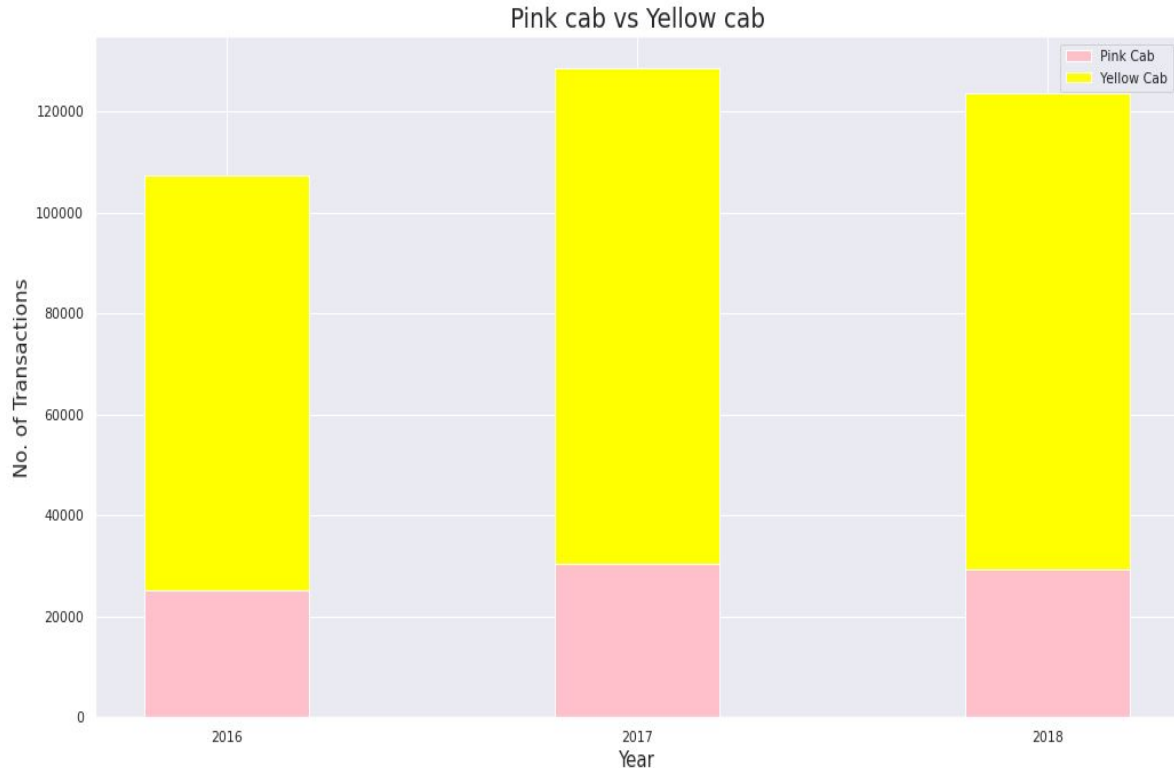
- Linear relation between KM travelled and Prices charged.
- Yellow cab prices higher than pink cab company.

# Profit share



- Yellow cab company has almost 90% of the profit share.
- Pink has a mere 10% of the profit share.

# Yearly transaction analysis



- Yellow cab has more customers than pink cab services in all the three years.

# Yearly profit analysis



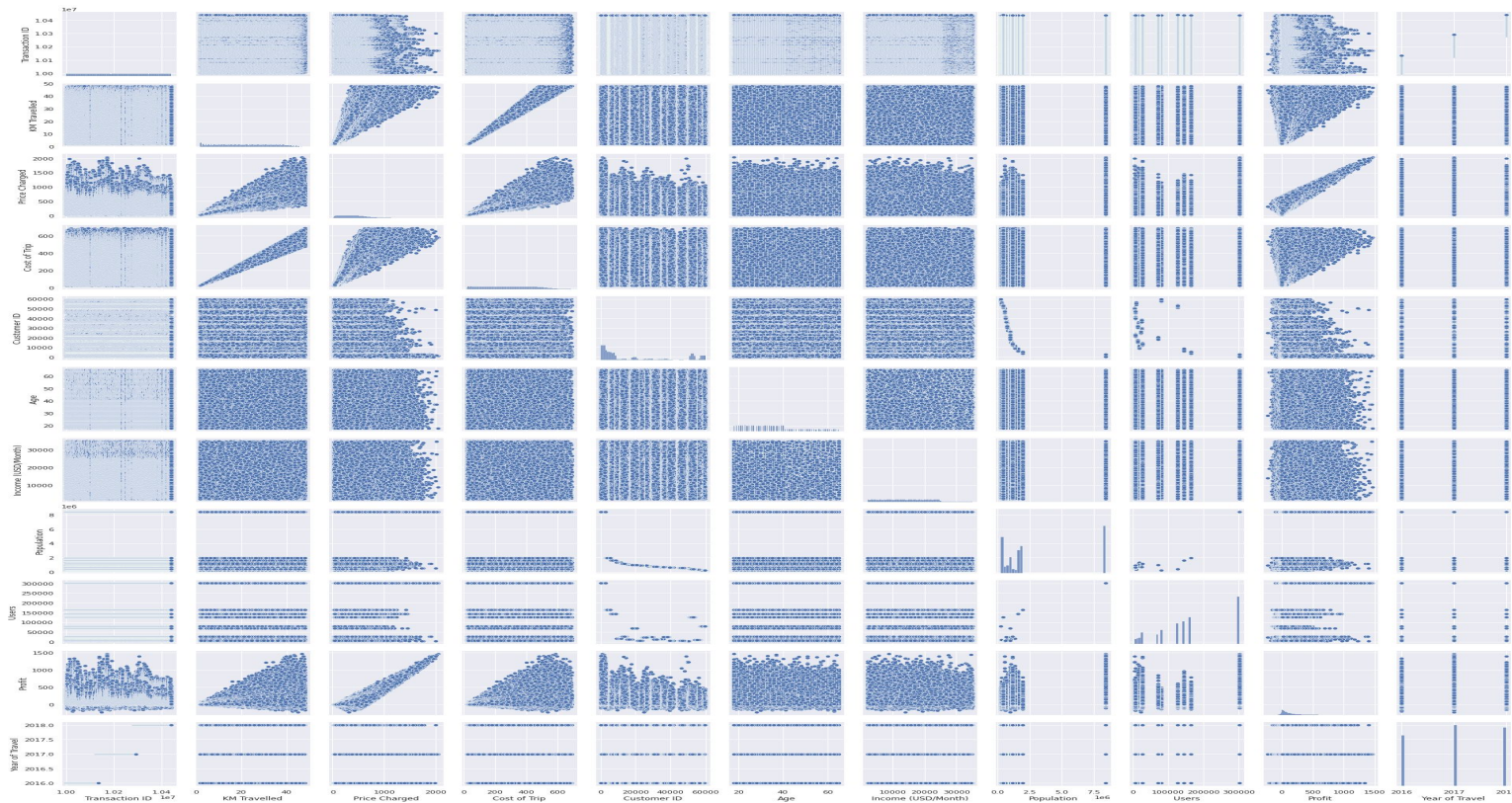
- Decline in the profits for both the cab services.
- Yellow cab has a steeper profit decline.

# KM Travelled vs Profit



- The profit margin for the Yellow cab services is more for the KM travelled than the pink services.

# Relation between all features



# Answers to the Hypothesis

- Which company has maximum cab users at a particular time period? - Yellow Cab Services have significantly higher users.
- Is there any relation between the age and the cabs utilized by them? - The age groups above 40 have a significantly lower interest in the cab services.
- Are companies having a good business? - Both the companies have a decrement of profit during the three year period.
- Which company performs better? Depending on the number of users and the profit margin of two services, Yellow cab services perform better in the market.
- Where to invest in? - It is advised to invest in the Yellow Cab services due to a larger market monopoly and profit margins.

# Recommendation

- It is observed that there are more yellow cabs running in the cities.
- New York city has the highest number of cab users.
- There are more transactions occurring for the yellow cab company compared to the pink cab company.
- Age groups below 40 prefer cabs than the ones above the age of 40.
- More prices are charged by the yellow company for the KMs travelled than its counterpart.
- The profit percent share of the yellow cab is almost 90% and that of the pink is 10%.
- There is a dip in the profit margins for both the companies.
- It is observed that more the distance travelled more is the profit for the companies.
- There are outliers present in the dataset for prices charged and the profit.

Depending on the analysis from the graphs I would recommend to invest on the Yellow Cab Services.



# Thank You



**Data Glacier**

your Deep Learning Partner