DATA GLACIER VIRTUAL INTERNSHIP CASE STUDY

NAME:RAHUL MAMBROL SATHEESAN

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

- 24 Features(including 9 derived features)
- Timeframe of the data: 2016-01-31 to 2018-12-31
- Total data points :355,032

Assumptions:

- Outliers are present in Price_Charged feature but due to
  unavailability of trip duration details ,we are not treating this as outlier.

- Profit of rides are calculated keeping other factors constant and only
  Price_Charged and Cost_of_Trip features used to calculate profit.

- Users feature of city dataset is treated as number of cab users in the city.
  we have assumed that this can be other cab users as well(including Yellow and
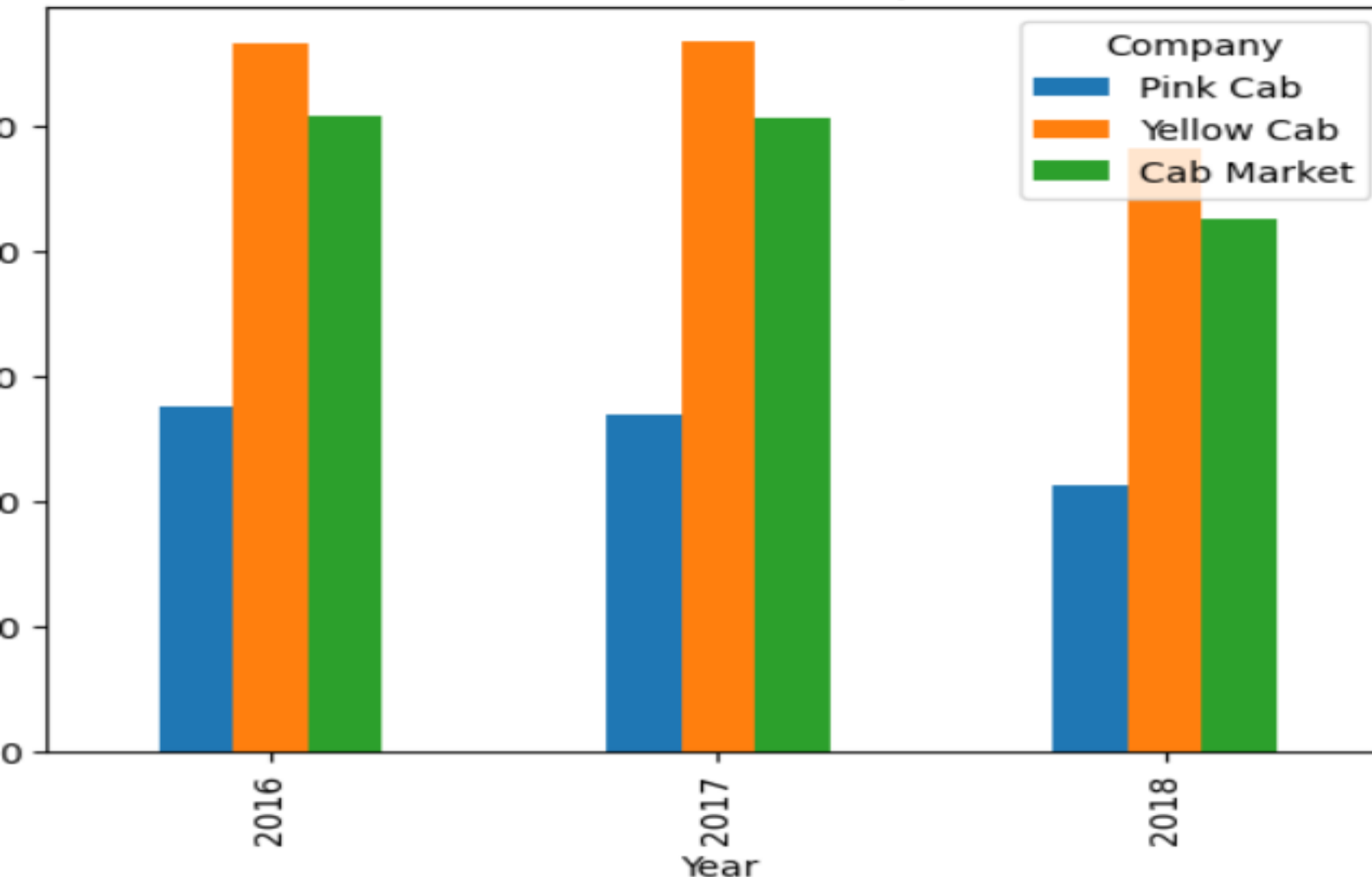 Pink cab)

Profit Analysis
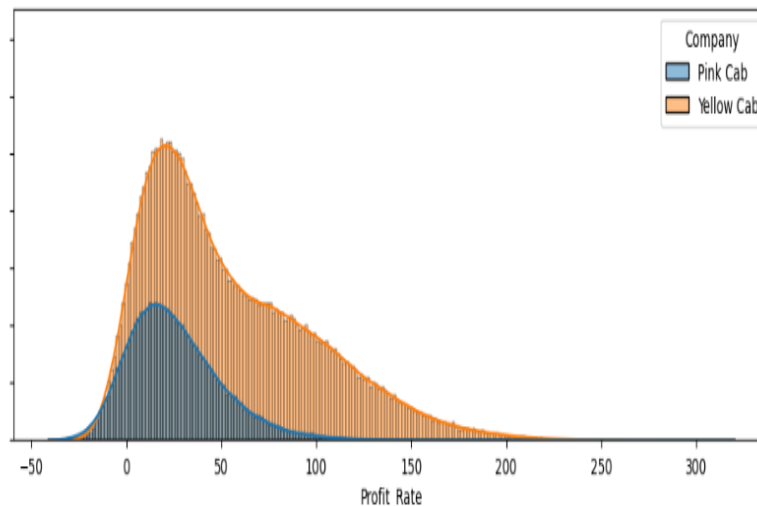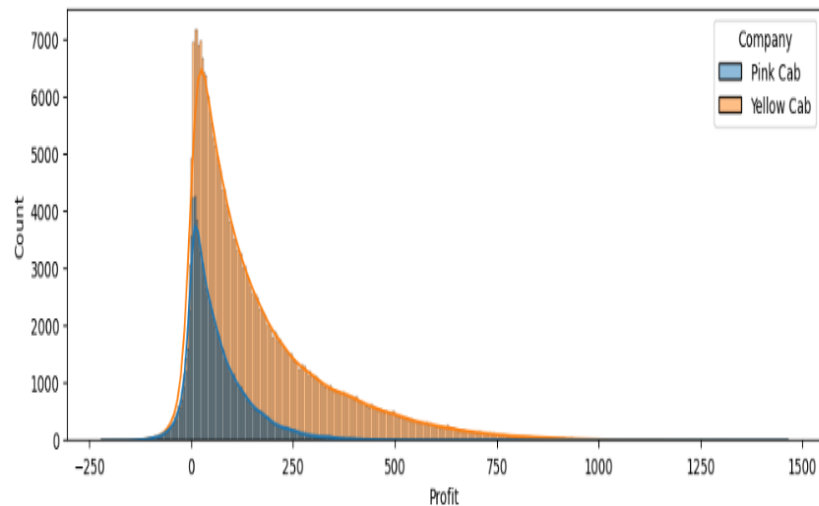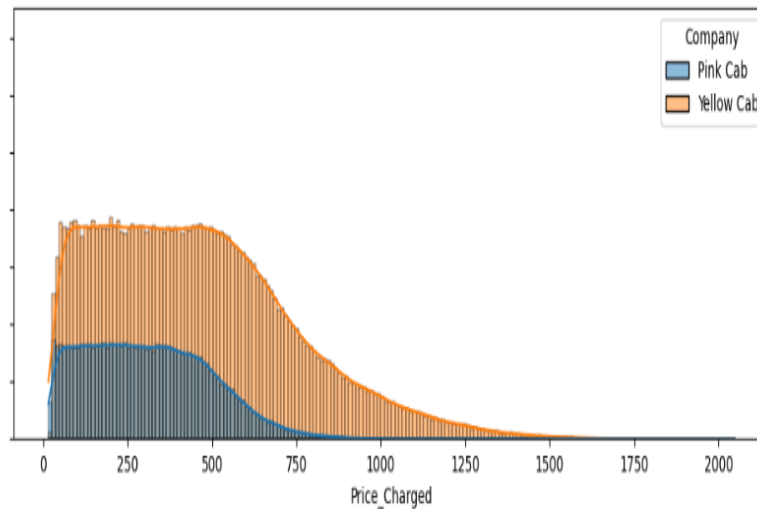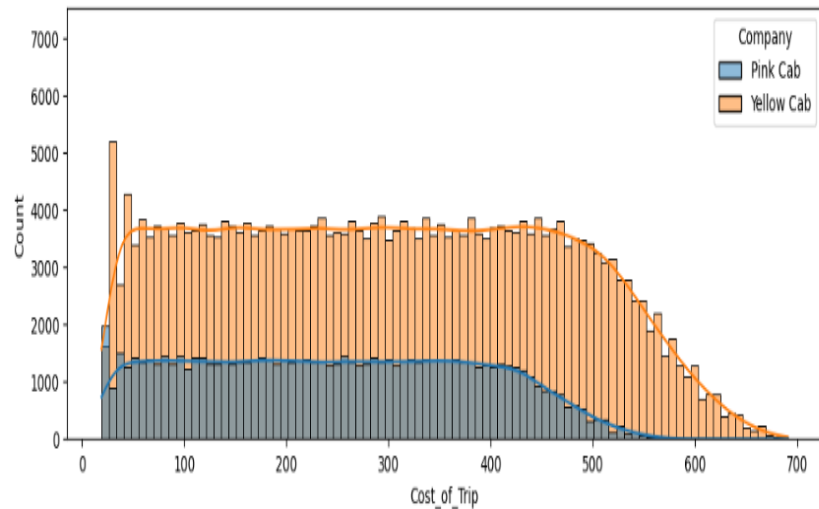
We can see from the last three figures that distributions of profits are all right skewed. But Yellow Cab has a visible bump to the right of peak in the distribution of Profit Rate and Profit per KM, which will probably lead to higher total profit rate. So we will check company's profit rate next.

Yearly Profit Analysis

Annual Profit Rate Comparison

From the last figure and table, we can see that Yellow Cab does have higher yearly profit rate (48% - 56%) than Pink Cab(21%-27%) in each of the recorded year. And one of the reasons is Yellow Cab have more high-profit-rate rides proportionally than Pink Cab. From this point of view, we think Yellow Cab is more efficient, but we still need to mine deeper into the data for an overall investment recommendation

Market share across cities

Market Share Comparison across Cities
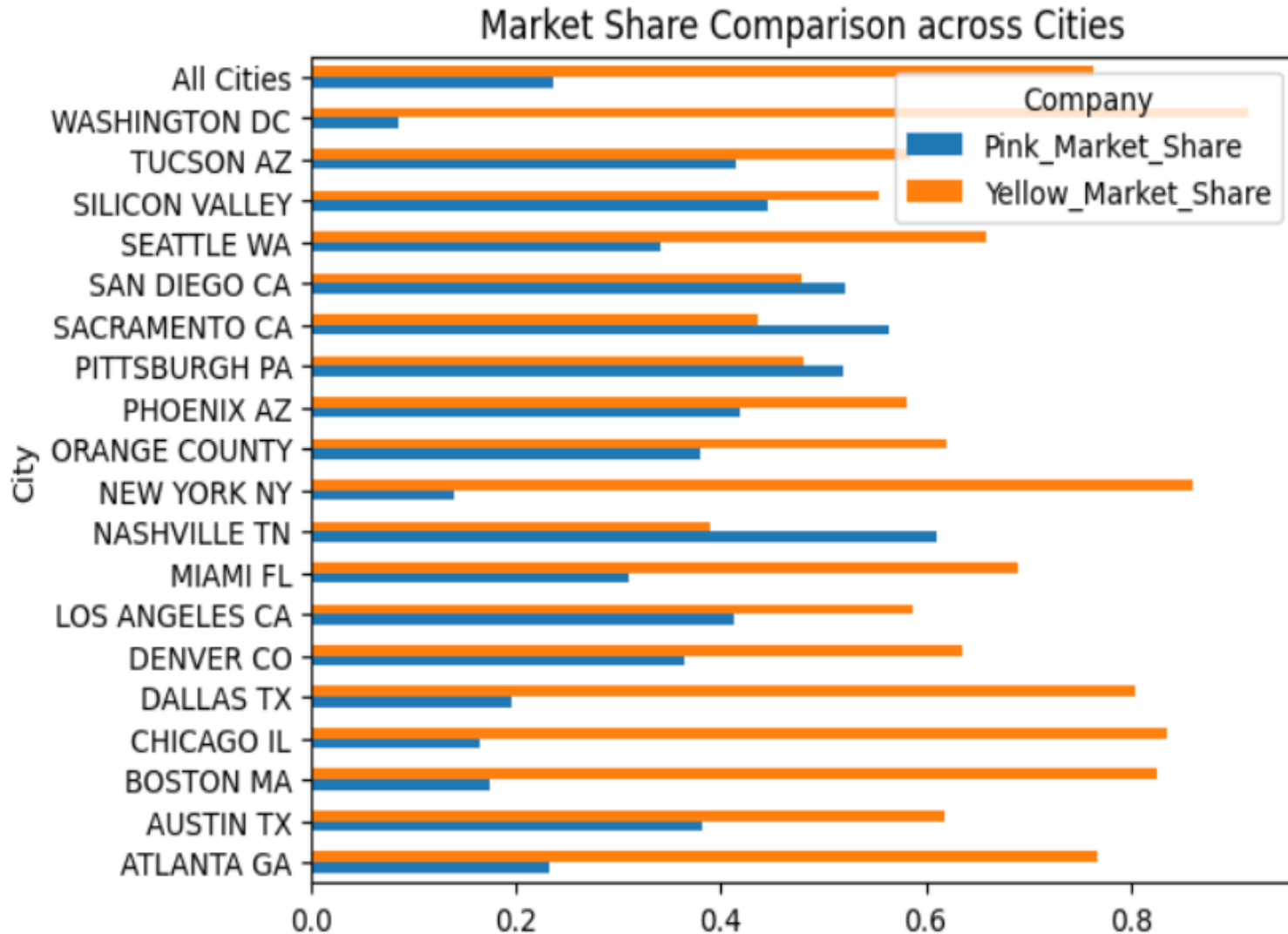
Assume Pink and Yellow Cap are the only cab companies in USA. Yellow Cap's market share is three times greater than that of Pink Cap nationally. And Yellow Cap's advantage exists in most cities. As market share is highly related to the number of cabs owned by each company, this can be seen a stable advantage of Yellow Cab.

Boxplot Distribution
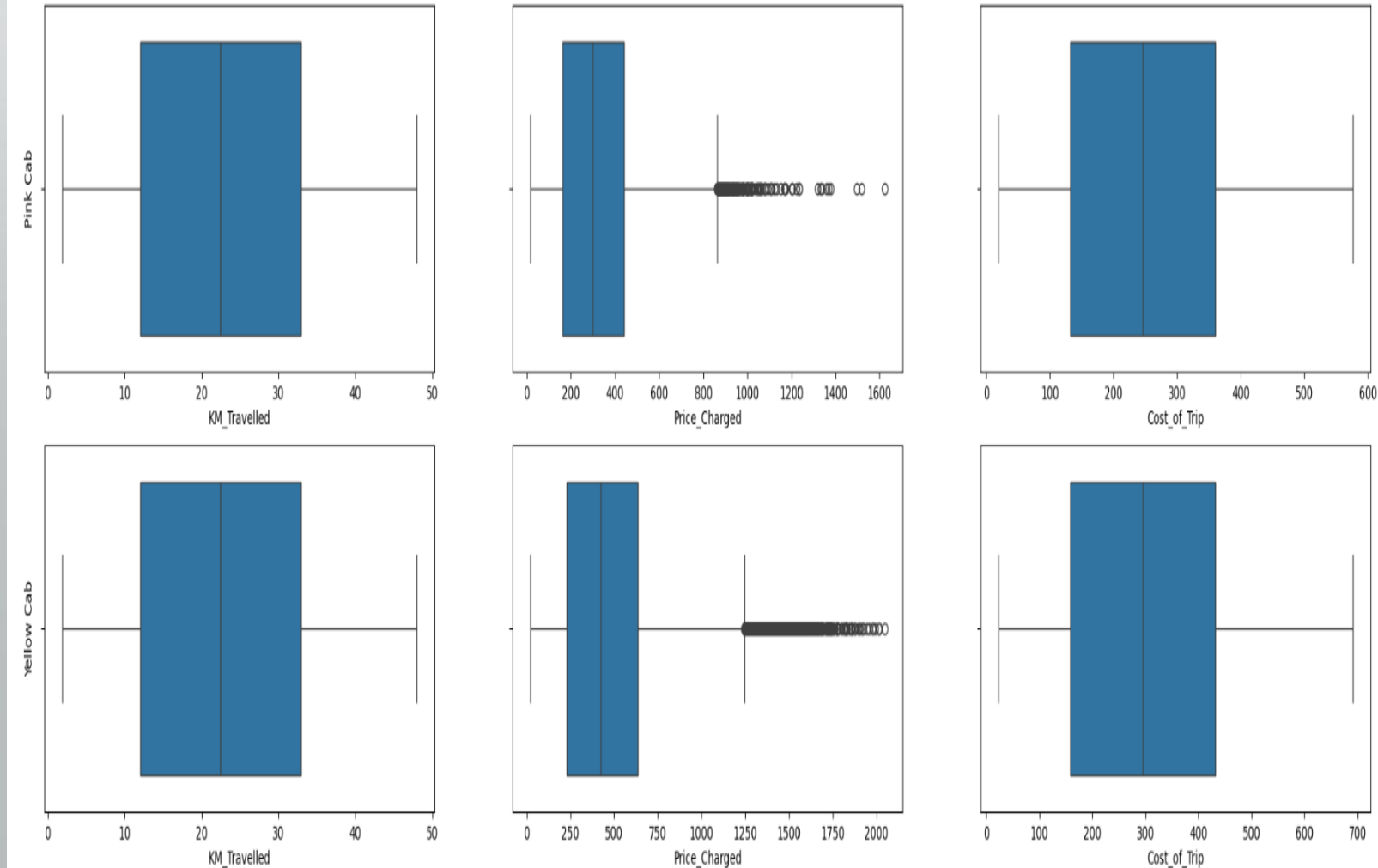
The distributions of KM_Traveled and Cost_of_Trip are similar for both companies. But Yellow Cab have greater medium in Price_Charged and higher proportion of rides with higher price

Recommendations

- **Customer Reach** : Yellow cab has higher customer reach in 25 cities while Pink cab has higher customer reach in 4 cities. We have also observed that Yellow cab is doing good in covering other cab users as compared to Pink cab.

- **Customer Retention:** We have analyzed this in two segments : at least 5 drive and at least 10 drive with the same cab company. And we found that Yellow cab is doing far better than Pink cab in both these segments.

- **Age wise Reach :** Yellow cab has customer in all age group and it's been observed that it's even popular in 60+ age group as equally as its in 18-25 age group.

- **Average Profit per KM:** Yellow cab's average profit per KM is almost three times the average profit per KM of the Pink cab.

- **Income wise Reach :**Both the cabs are very popular in high and medium income class but here also Yellow cab is performing better than Pink cab in offering their services to all the three income class group (low, medium and high)

- **Ride count and Profit Forecasting :** Both the companies are facing loss in the profit and no. of ride. Yellow cab's forecasted profit loss is around 1.83% while Pink cab's loss in 3.1%.Pink cab is facing more loss even when its forecasted no of ride loss is lesser than Yellow cab.

On the basis of above point , we will recommend Yellow cab for investment.