



**Data Glacier**

Your Deep Learning Partner

# Exploratory Data Analysis

**G2M insight for Cab Investment firm**

**11/10/2021**

# Agenda

Executive Summary

Problem Statement

Approach

EDA

Recommendations

# Executive Summary

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
- Time series analysis for both companies
- Finding the most profitable Cab company
- Recommendations for investment

# Problem Statement

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.

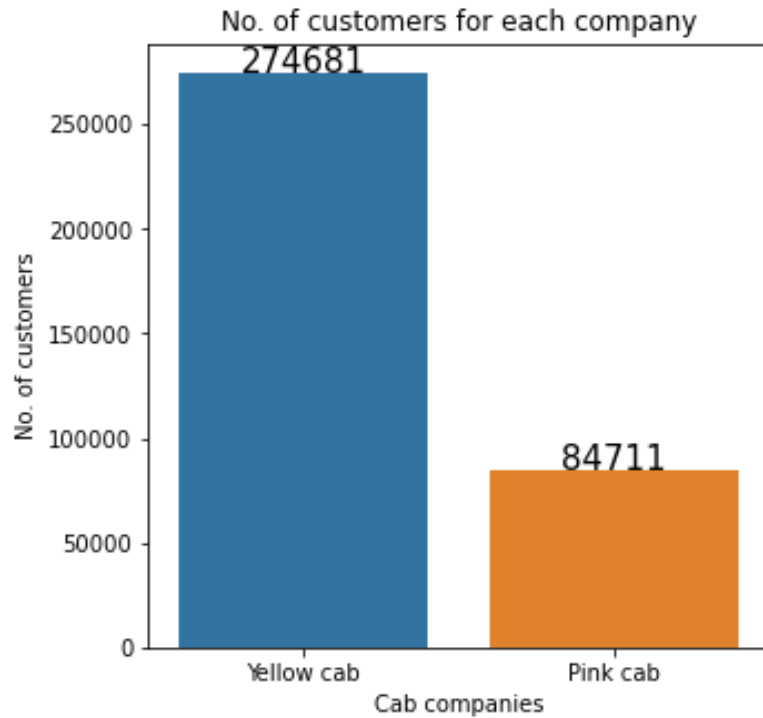
# Approach

- Data exploration:
  - 3 datasets were merged in one dataset resulting in 16 features 3 of them were derived from other features, 2 datasets also were merged to result in 5 features, so we get 21 features in total with 359,392 rows.
  - Timeframe of the data: 2016-01-31 to 2018-12-31
- 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)
  - Date including some outliers where there were some dates before 2016-1-31

# EDA

Number of customers for each company  
City Wise Cab Users Covered By Company  
Profit analysis  
Time series analysis  
Customers segmentation  
Distance travelled by each company  
Monthly rides

# Number of customers for each company



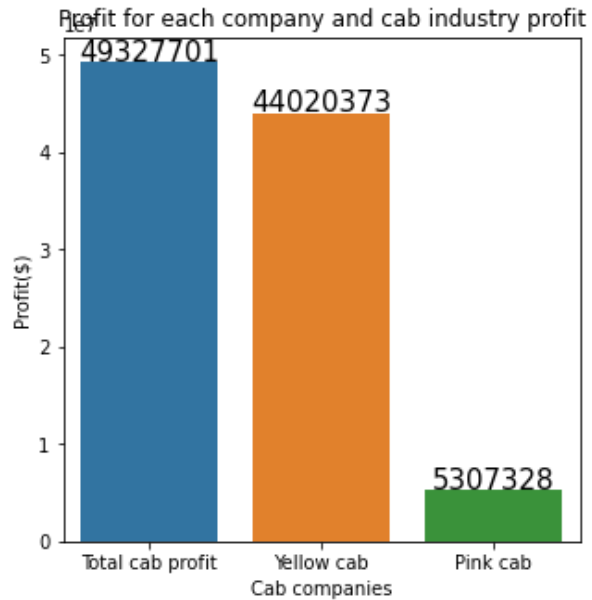
From this bar chart plot we can conclude that Yellow cab company much more customers than pink cab.



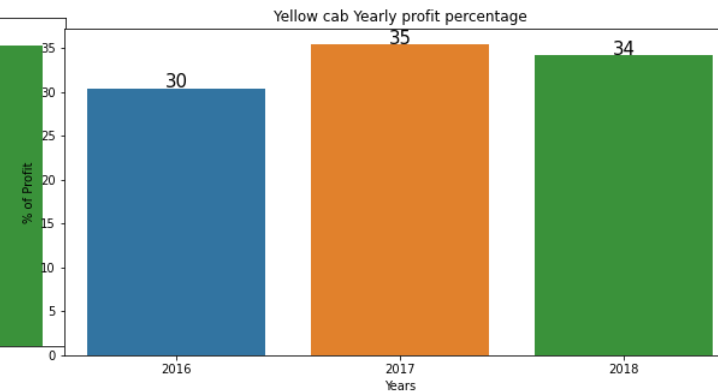
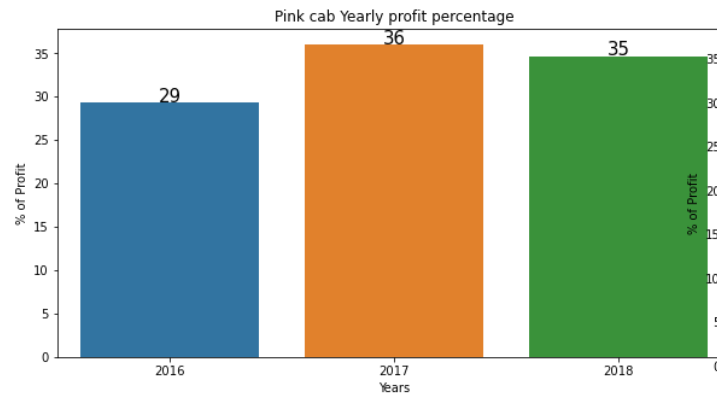
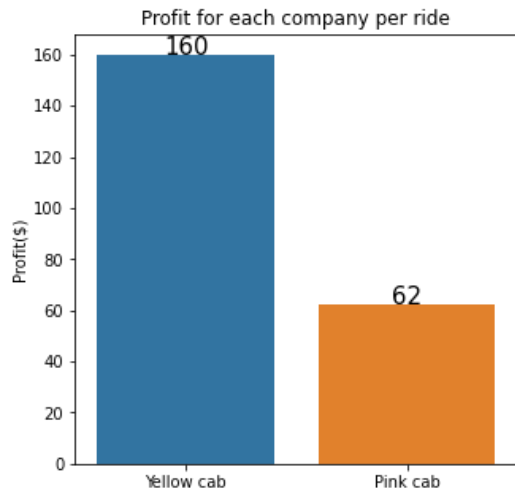
- This shows customer distribution according to states which shows that Yellow cab got more customers in most of the states except in Sam Diego, Nashville, Sacramento, Pittsburgh.
- It also shows that there are another companies in the cab industry dominating in San Francisco since there are no users belongs to both Yellow or Pink cab.



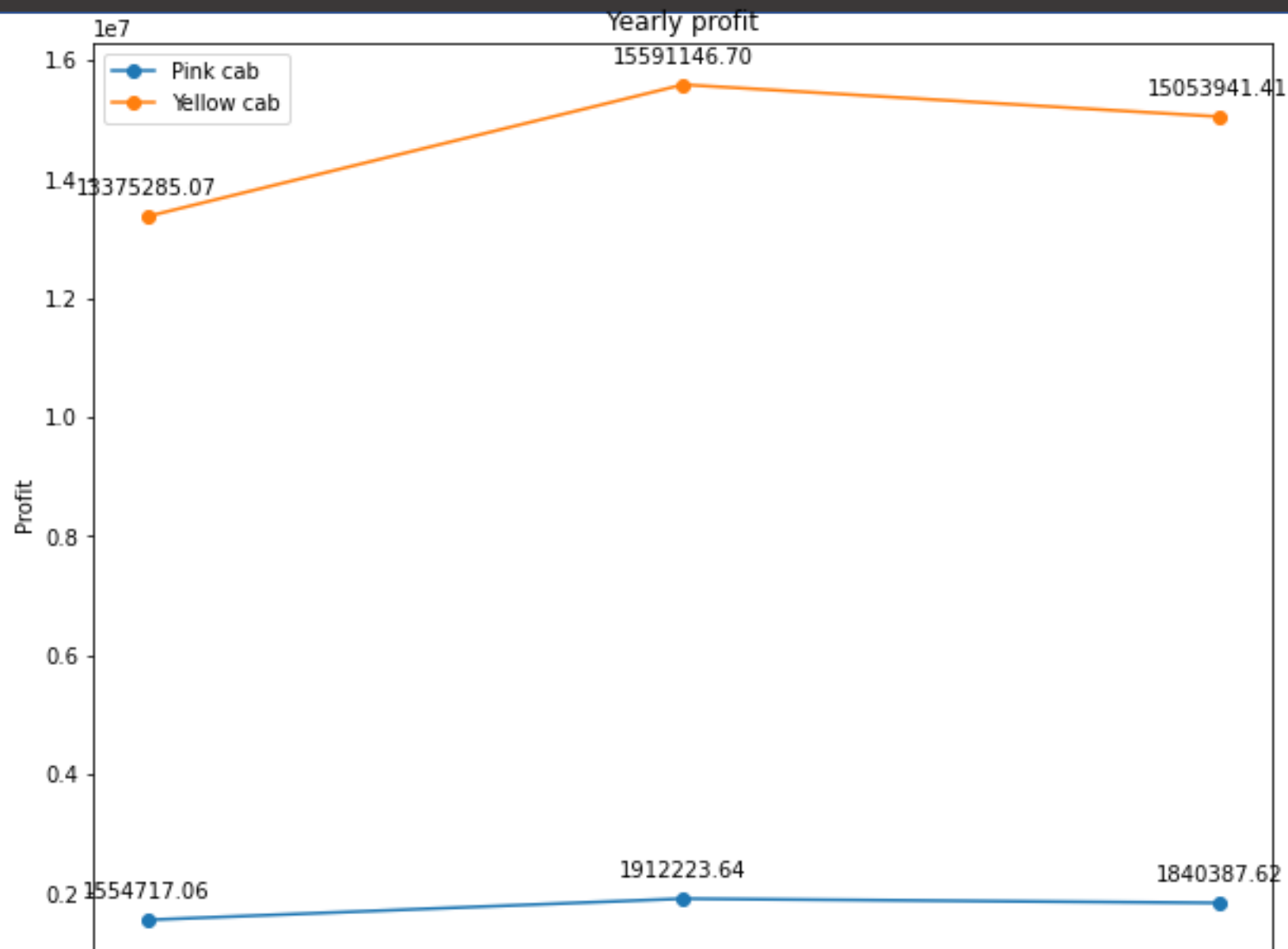
# Profit analysis



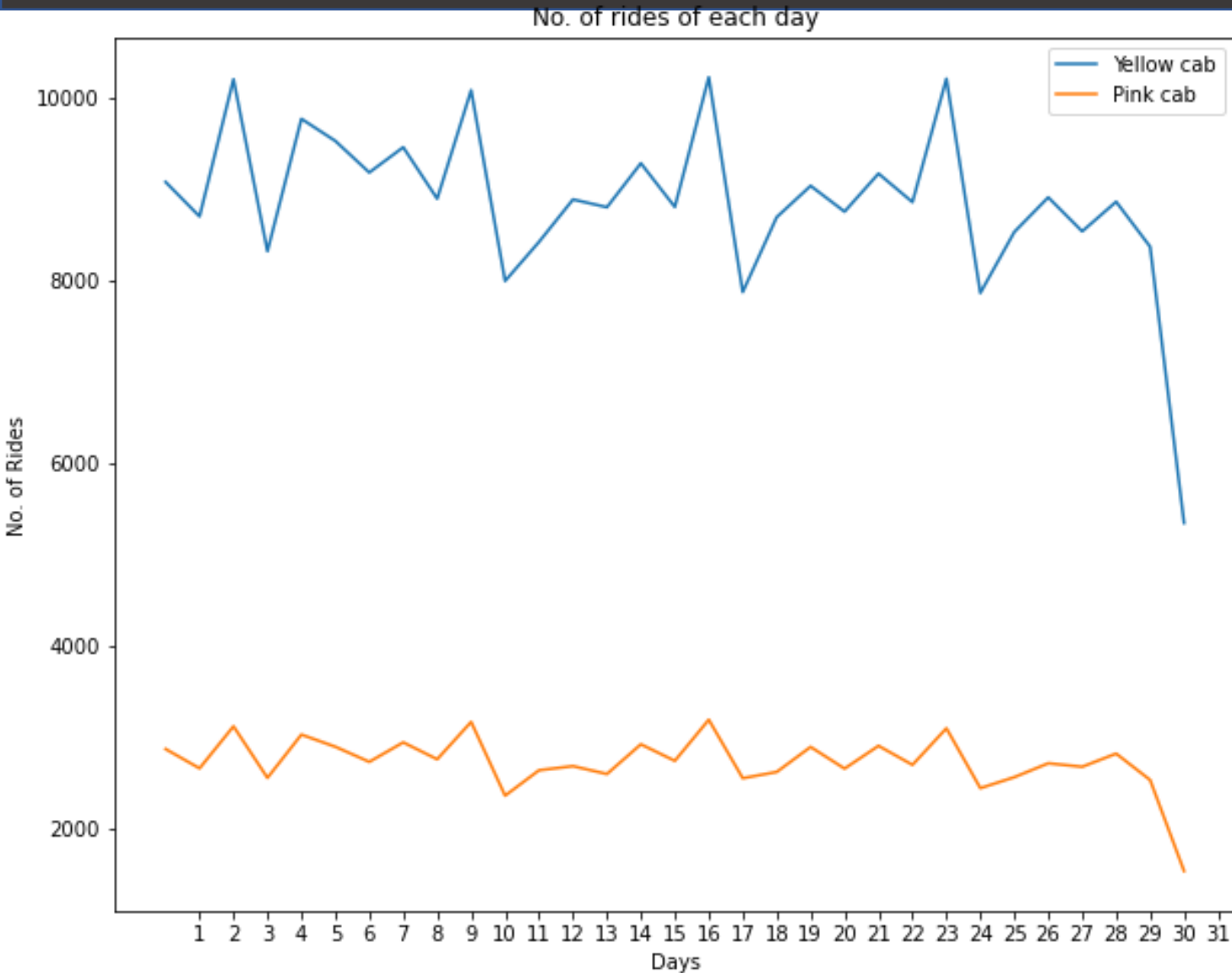
- It seems that the cab industry is profitable as we can see from the first bar where the total profit for three years is 49,327,701\$ Where most of the profit is gained by Yellow cab.
- The plot shows the 2017 was the most profitable year for both companies and 2016 was the least profitable year.



# Profit time series analysis

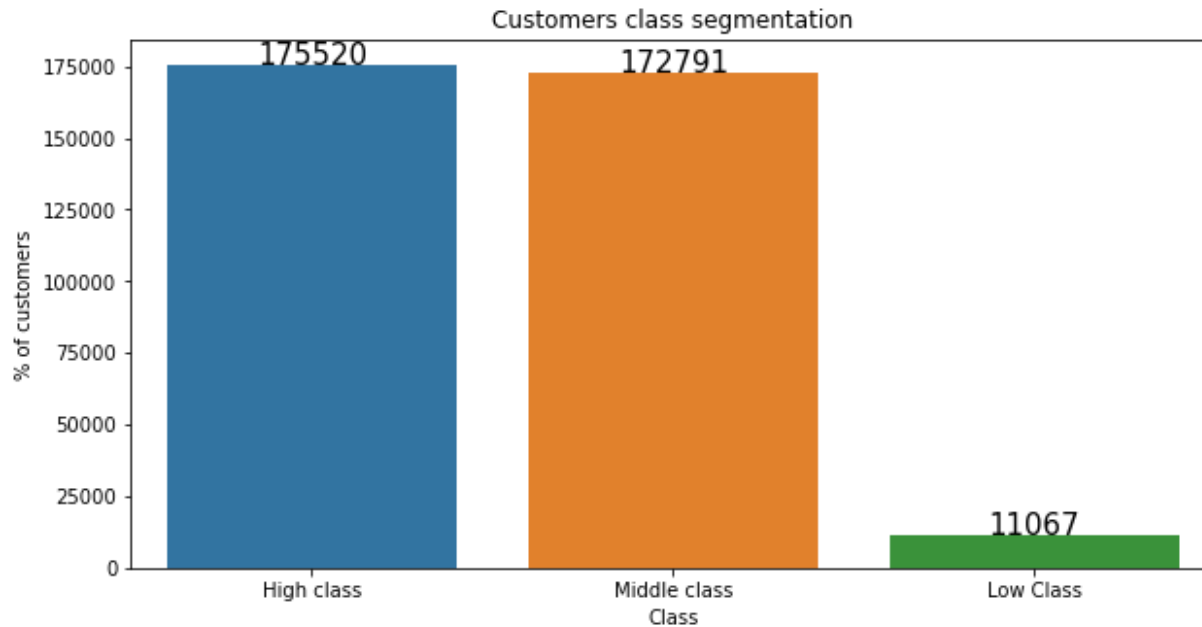


# Time series analysis



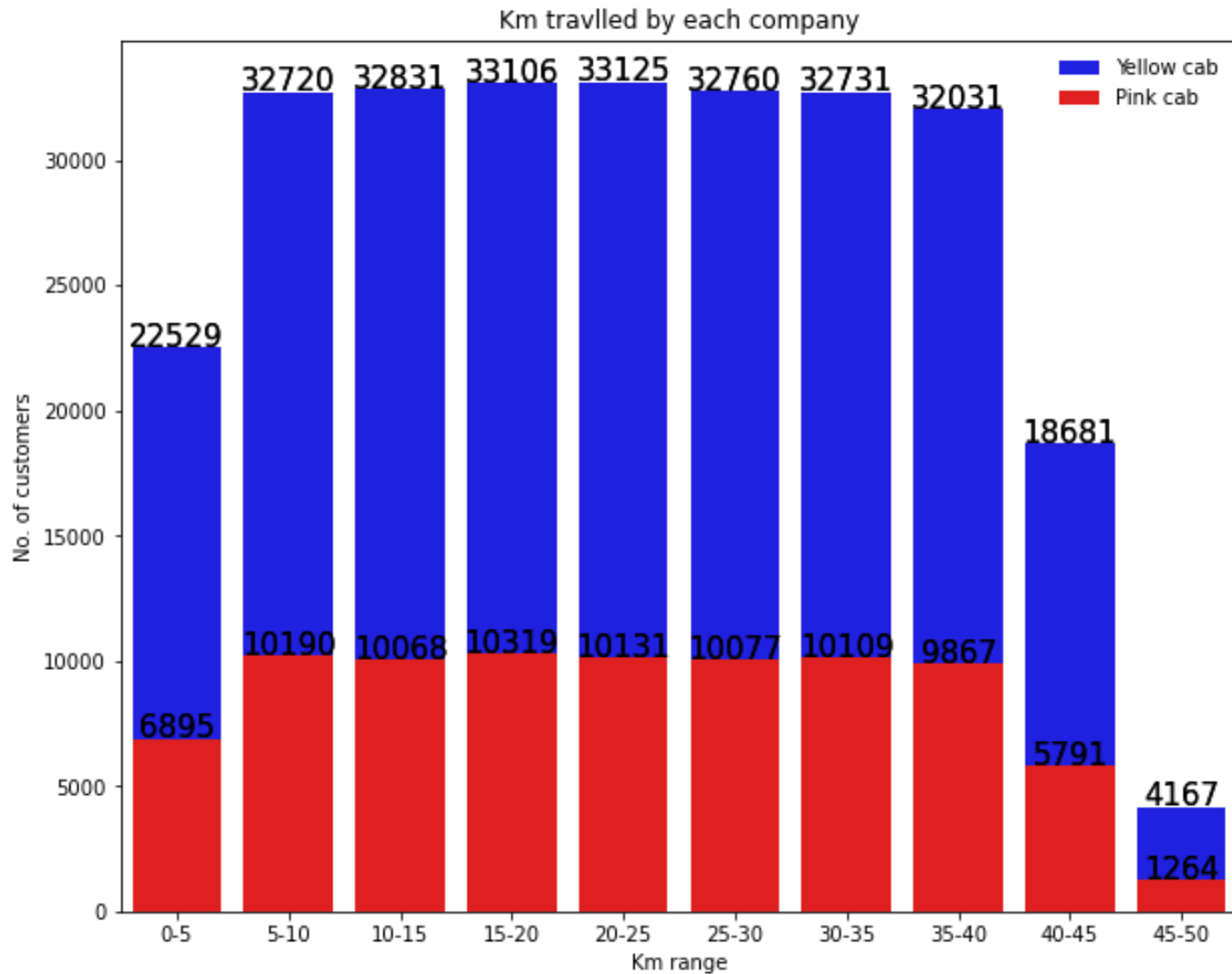
There is seasonality in number of rides which starts at the second day of each month and repeats every week.

# Customer segmentation



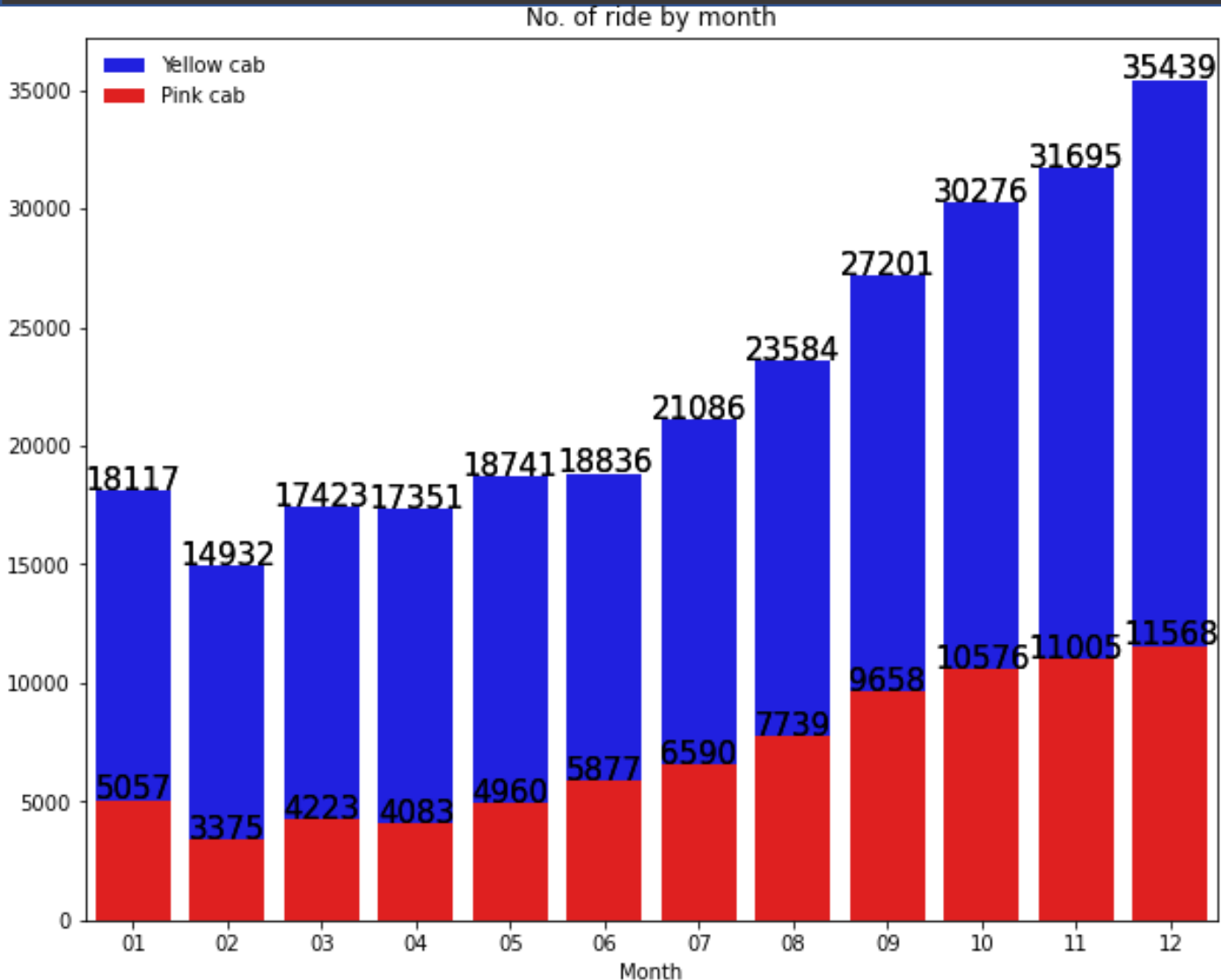
The plot shows that the majority of the cab users are from the middle and high class which reflect having more cab users in the rich states rather than poor states.

# Customer segmentation



- Yellow cab has customers almost uniform for short, medium and long Trip. Which shows yellow cab is offering better customer plan for short, medium and long Trip.
- 5-35 KM trips are contributing more In both the cabs profit
- There is a huge difference between yellow and pink cab in customer reach for short and long trip. Yellow cab has very good customer reach in this segment as well

# Monthly rides



- Number of rides increase gradually as the year starts to end which reflects more holidays affecting the number of rides positively.

# Recommendations

We have evaluated both the cab companies on following points and found Yellow cab better than Pink cab:

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

# Thank You