G2M Case Study

Virtual Internship

By:
Mohammad Shehzar Khan

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.

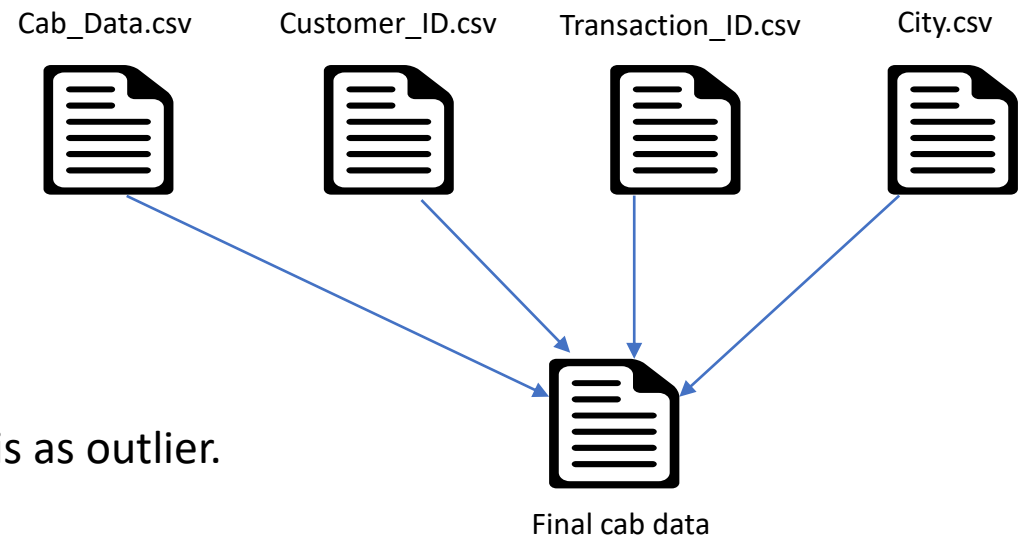
- Objective : Provide actionable insights to help XYZ firm in identifying the right company for making investment.

The analysis has been divided into four parts:

- Data Understanding
- Number of rides for each cab type
- Finding the most profitable Cab company
- Recommendations for investment

Data Exploration

- 22 Features(including 9 derived features)
- Timeframe of the data: 2016-01-31 to 2018-12-31
- Total data points :359,392



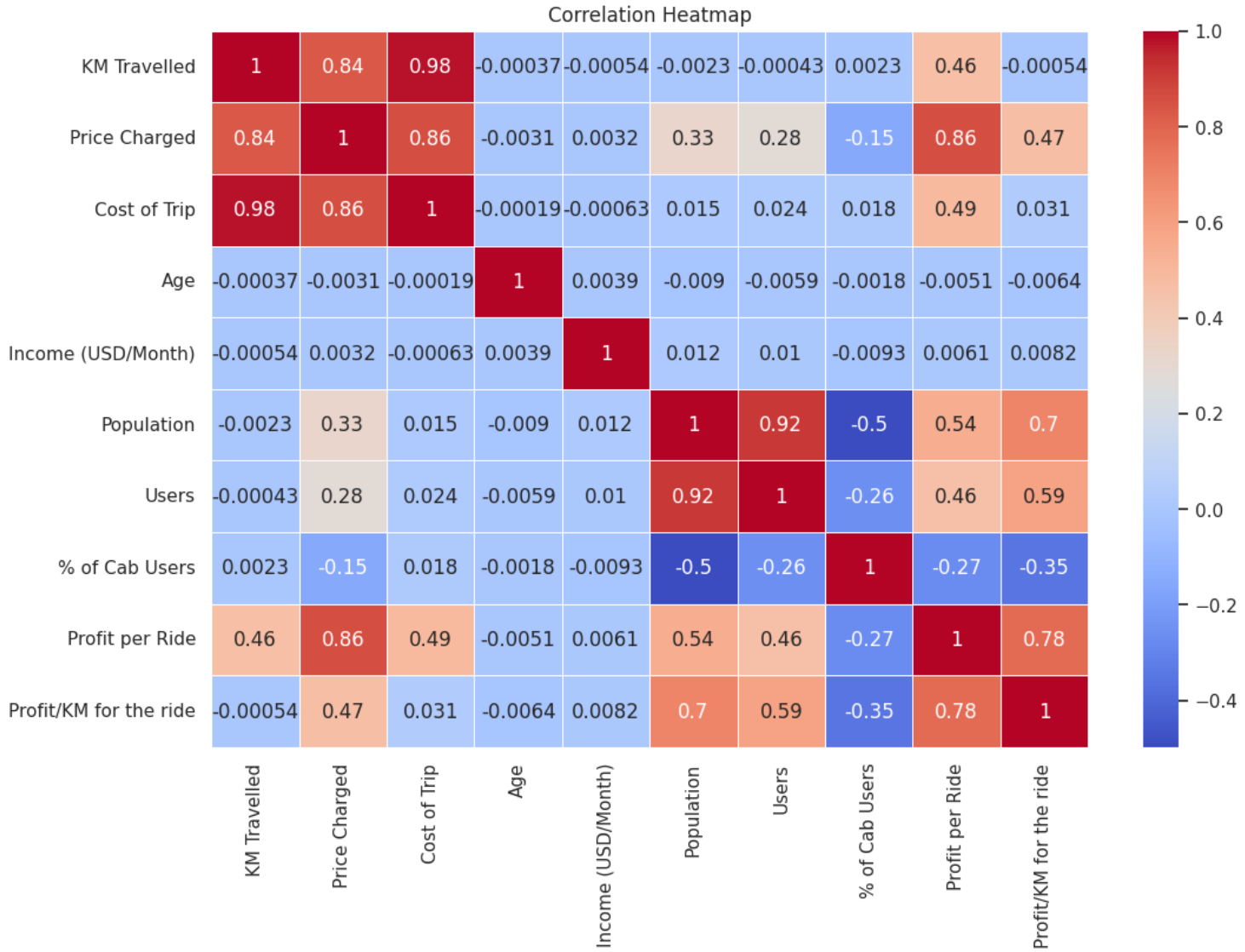
Assumptions:

- 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)

Data Exploration

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 359392 entries, 0 to 359391
Data columns (total 23 columns):
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Transaction ID          359392 non-null  int64
 1   Date of Travel          359392 non-null  datetime64[ns]
 2   Company                 359392 non-null  object
 3   City                    359392 non-null  object
 4   KM Travelled            359392 non-null  float64
 5   Price Charged           359392 non-null  float64
 6   Cost of Trip            359392 non-null  float64
 7   HighChargeFlag          359392 non-null  int64
 8   Customer ID             359392 non-null  int64
 9   Payment_Mode            359392 non-null  object
 10  Gender                  359392 non-null  object
 11  Age                     359392 non-null  int64
 12  Income (USD/Month)      359392 non-null  int64
 13  Population              359392 non-null  int64
 14  Users                   359392 non-null  int64
 15  % of Cab Users          359392 non-null  float64
 16  Profit per Ride         359392 non-null  float64
 17  Profit/KM for the ride  359392 non-null  float64
 18  Year                    359392 non-null  int64
 19  Month                   359392 non-null  int64
 20  Day                     359392 non-null  int64
 21  Year-Month              359392 non-null  period[M]
 22  Distance Category       359392 non-null  category
dtypes: category(1), datetime64[ns](1), float64(6), int64(10), object(4), period[M](1)
memory usage: 63.4+ MB
```

Relation between different variables/features



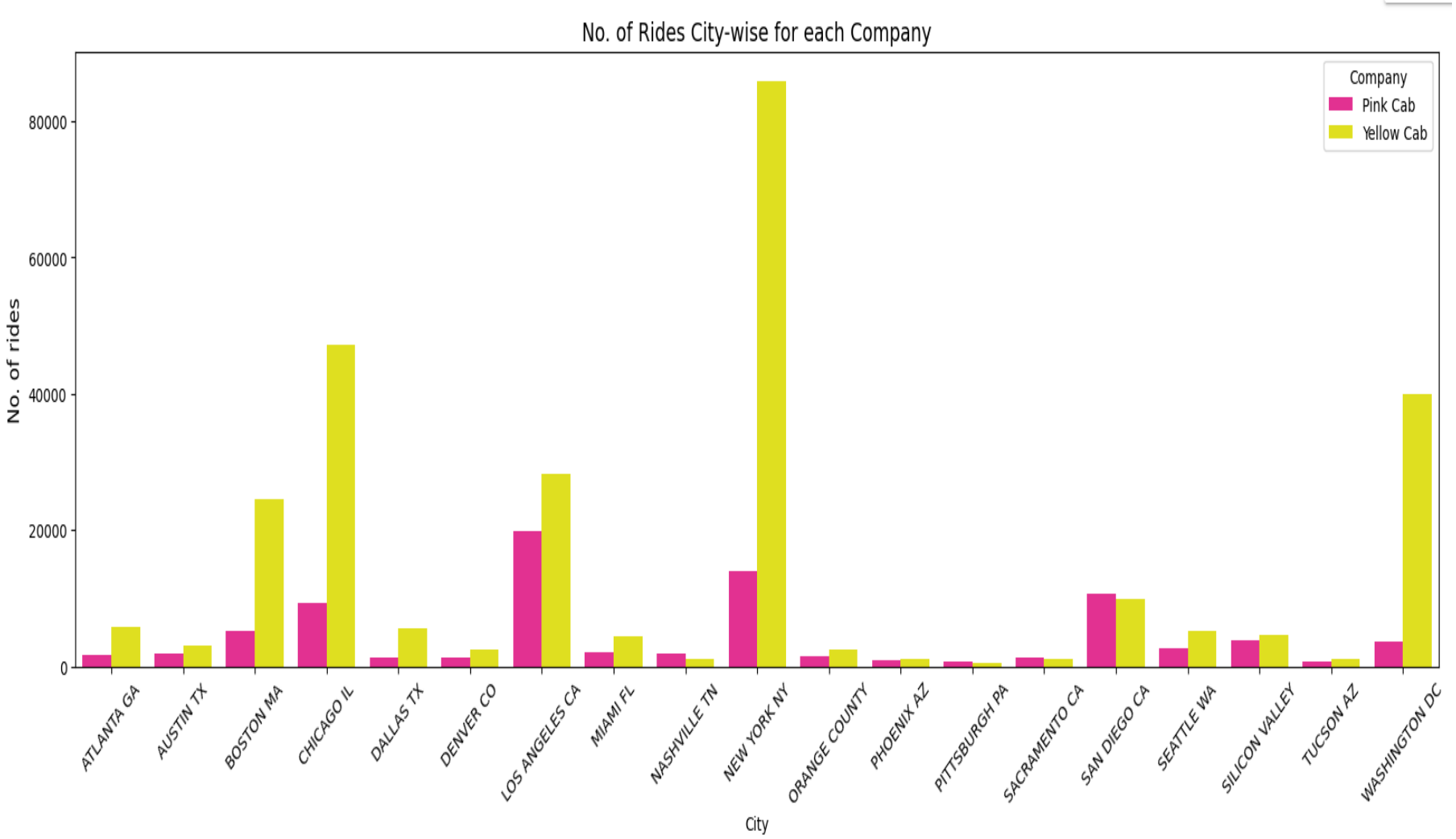
- This is a heatmap. It displays how much two features are affected by each other. The darker the color is, the strong is the relation between the particular two features.

- Here:

1. Price Charged & KM Travelled
2. Cost of Trip & KM Traveled
3. % of Cab Users & Population
4. Profit per Ride & Price Charged
5. Profit/KM for the ride & Population
6. Profit/KM for the ride & Users

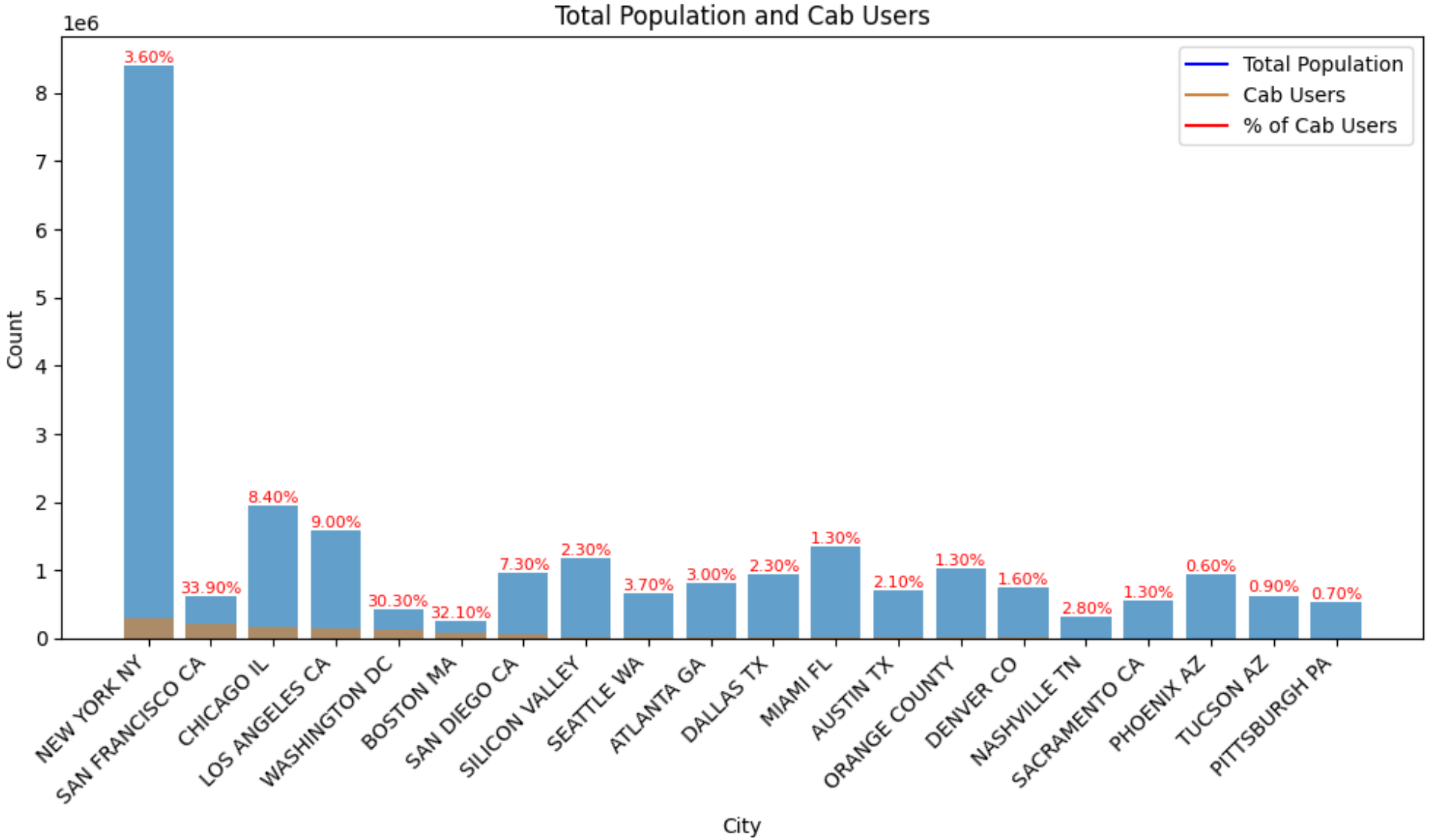
are highly correlated.

Customer Presence of Yellow and Pink cab city wise



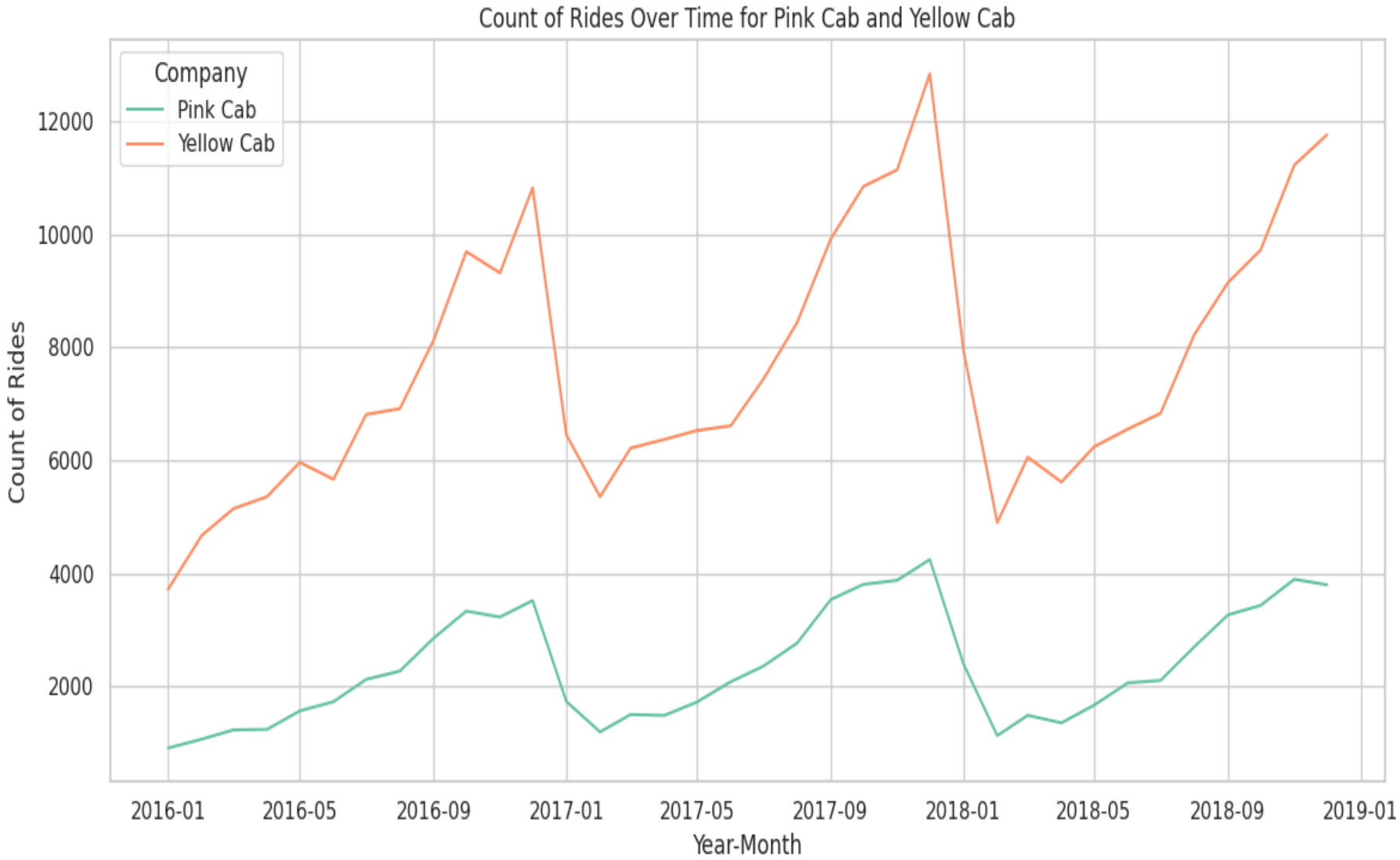
- In the cities like **New York**, **Chicago**, **Washington DC**, **Los Angeles**, and **Boston** Yellow Cabs are more preferred or popular as compared to Pink Cabs.
- In cities like **Nashville**, **Pittsburgh**, **Sacramento**, and **San Diego** the Pink Cabs are more preferred or popular as compared to Yellow Cabs.
- Overall it looks like people in bigger cities prefer Yellow Cabs over Pink Cabs.

Population and Cab Users in each City



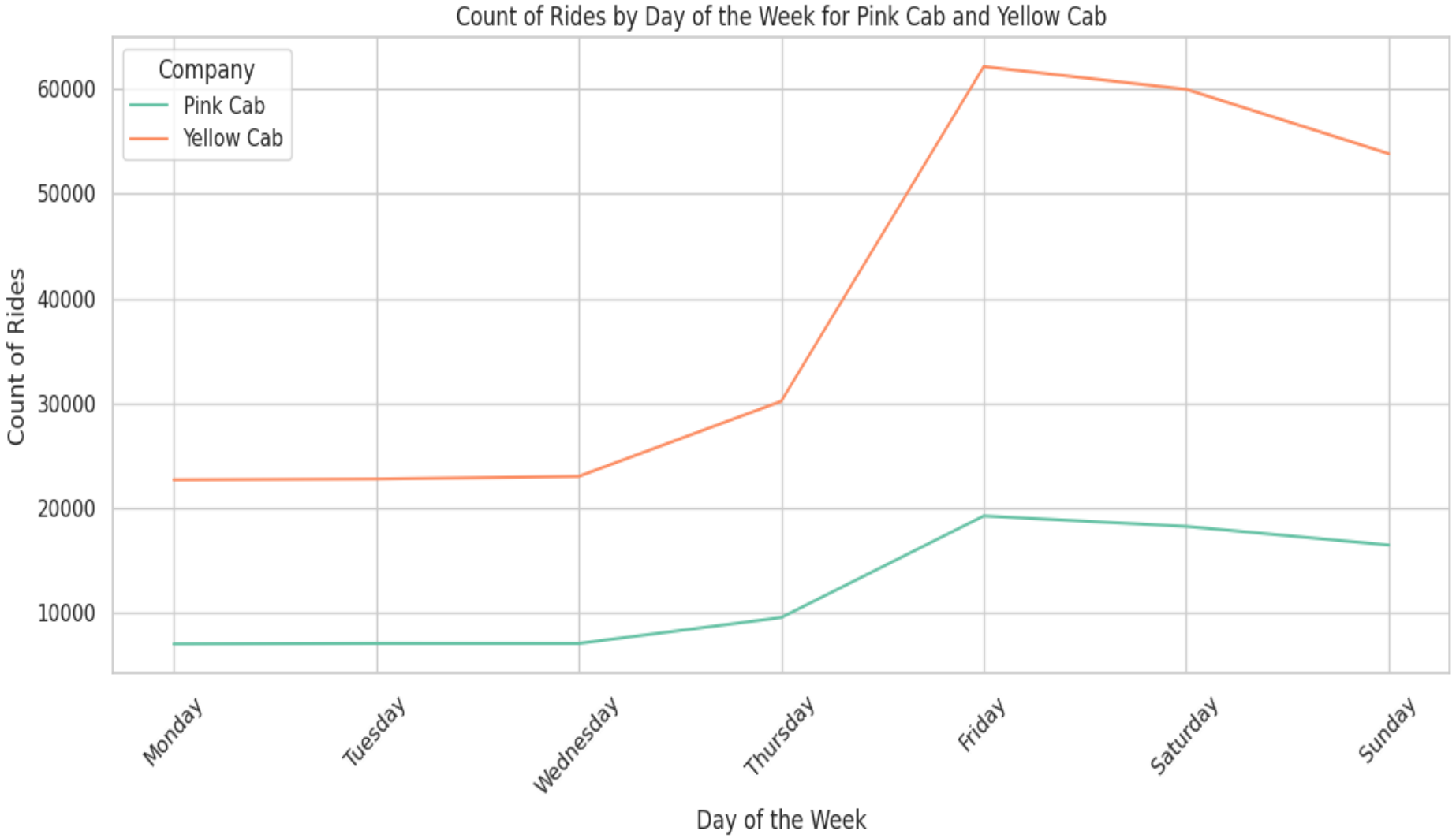
- New York city has the most number of cab users, followed by San Francisco and Chicago.
- San Francisco has the most cab users by % of total population followed by Boston and Washington DC.

Seasonality in Demand



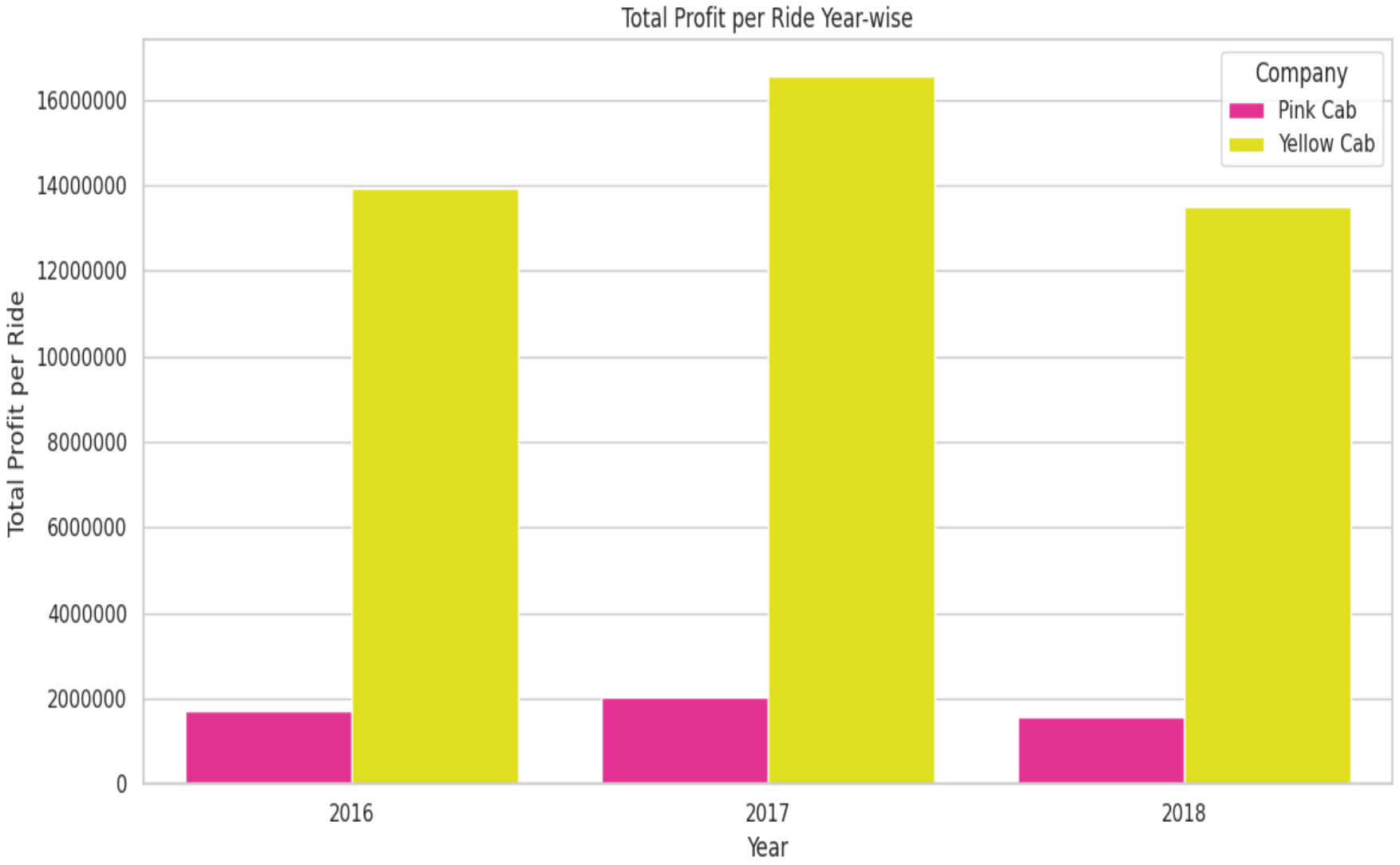
- If we look at the demand, there are some months where demand is more in each year.
- **Demand is minimum** at around **January-February** and it keeps on increasing from there and reaches to **maximum** at around the **November-December**.
- Also, the worst season for Yellow Cabs is as good as the best season of Pink Cabs.

Ride Analysis Day-Wise



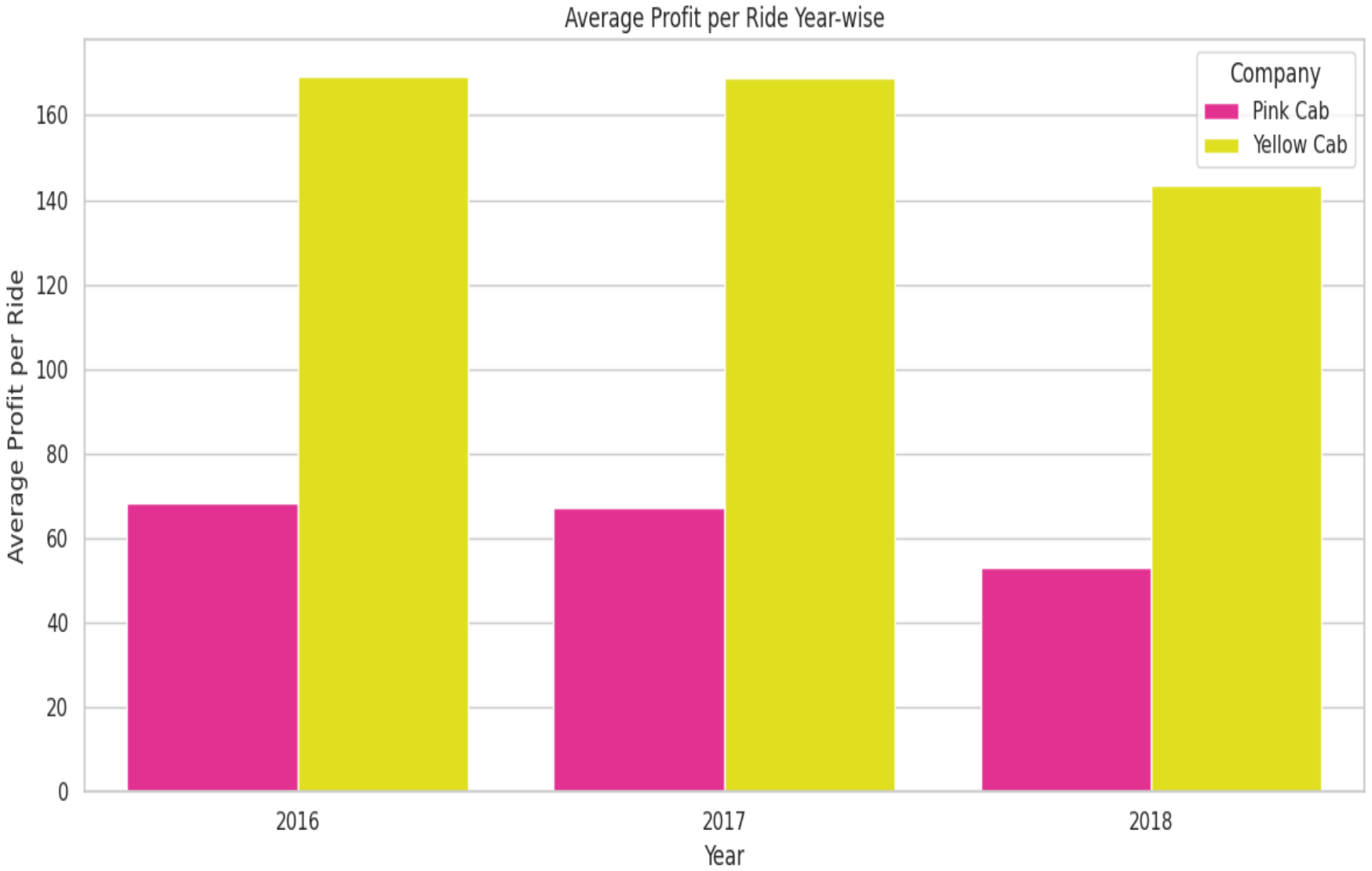
- This plot shows the total no. of rides for each day of the week.
- The no. of rides constant for Monday, Tuesday, and Wednesday.
- As we approach towards the weekend, we can see a dramatic increase in the ride count with **maximum on Friday, followed by Saturday and Sunday.**
- The Yellow Cabs see a much more relative increase in the rides in weekends as compared to Pink Cabs.

Profit Analysis (Total Profit Year-Wise)



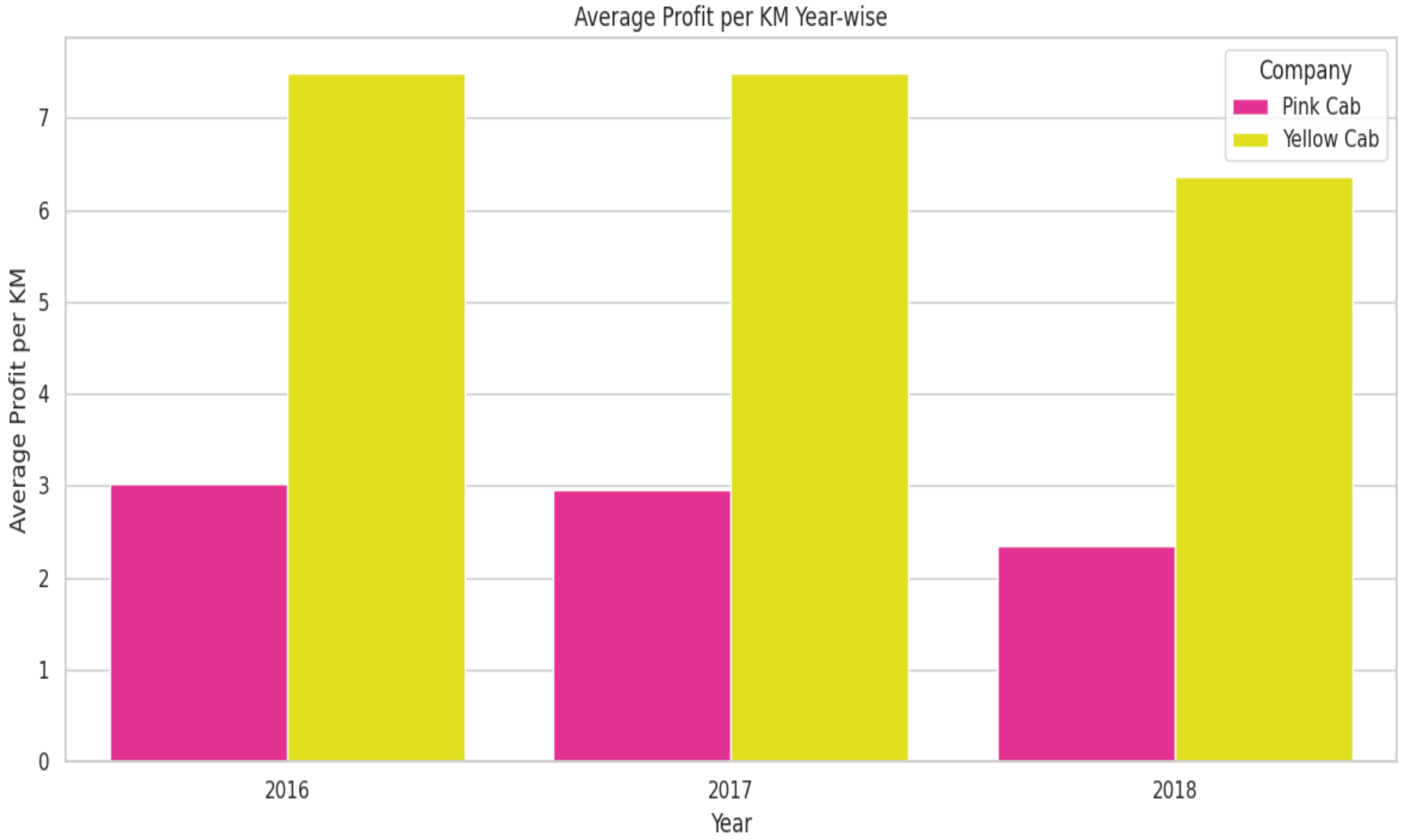
- The Total Profit of Yellow Cabs is more than 5 times the profit of Pink Cabs for years 2016, 2017 and 2018.

Profit Analysis (Average Profit Year-Wise)



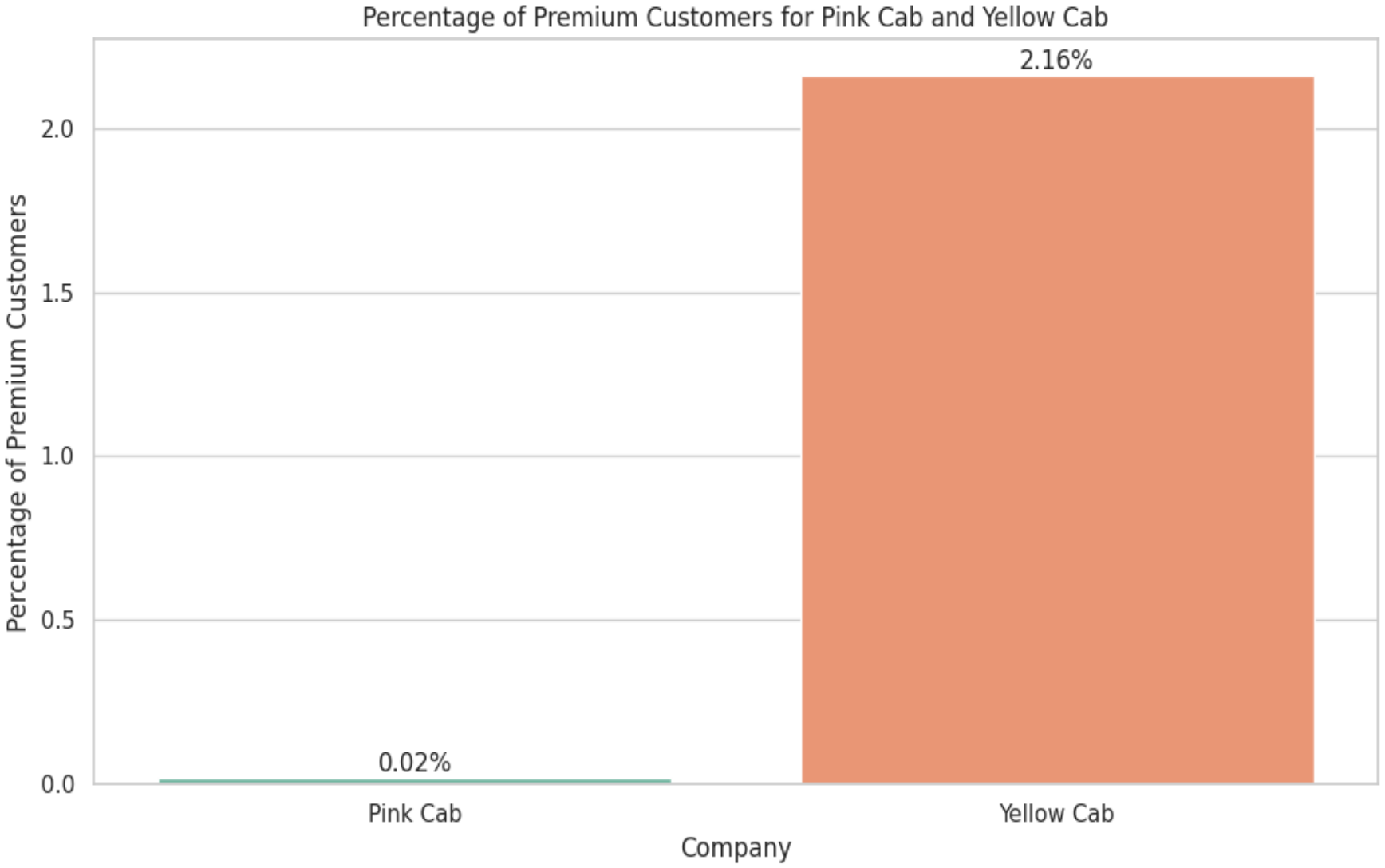
- The Average Profit per ride for the Yellow Cabs is more than twice the Pink Cabs for the years 2016, 2017, and 2018.

Profit Analysis (Average Profit/KM Year-Wise)



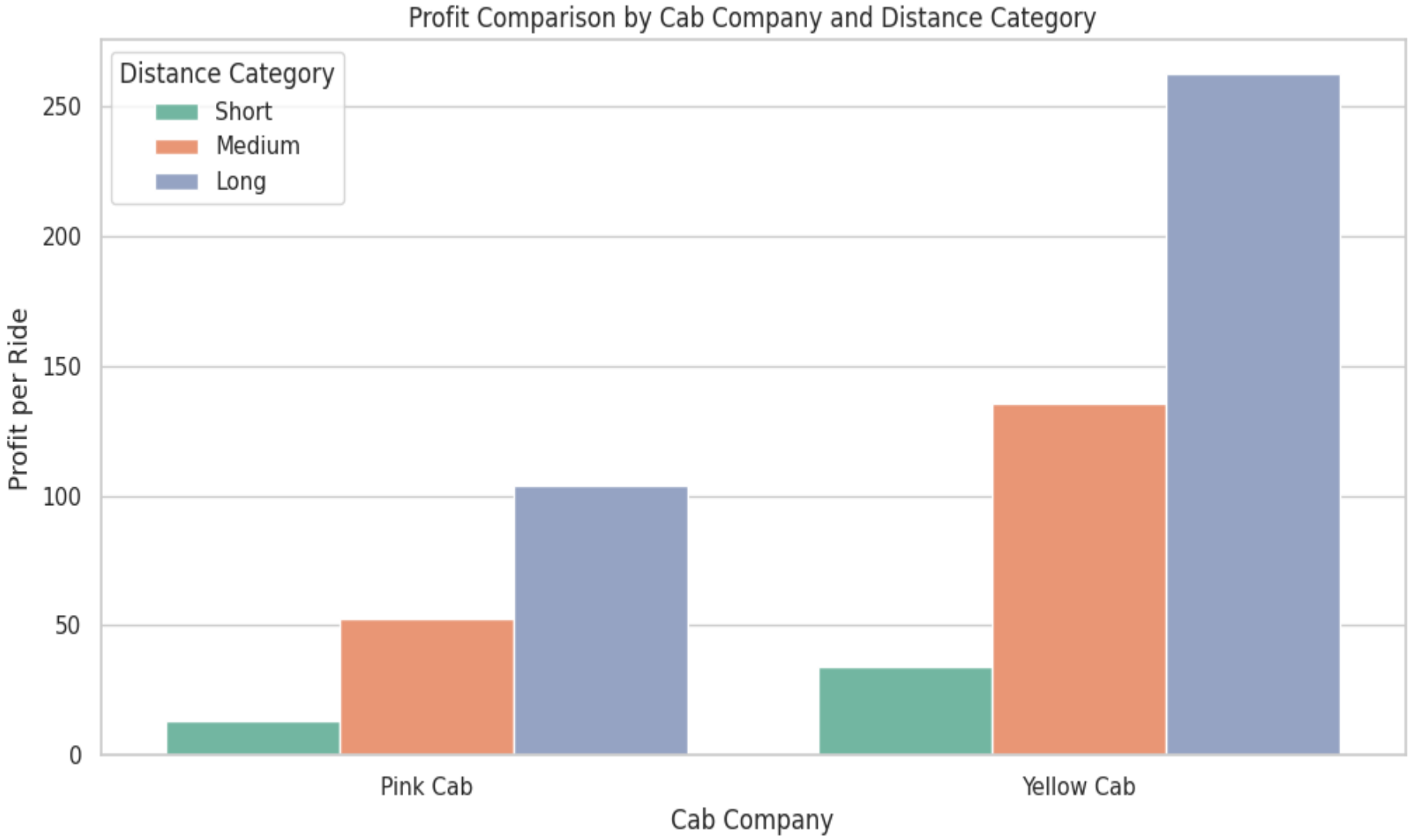
- The Average Profit per KM for the Yellow Cabs is more than twice the Pink Cabs.

Premium Customers



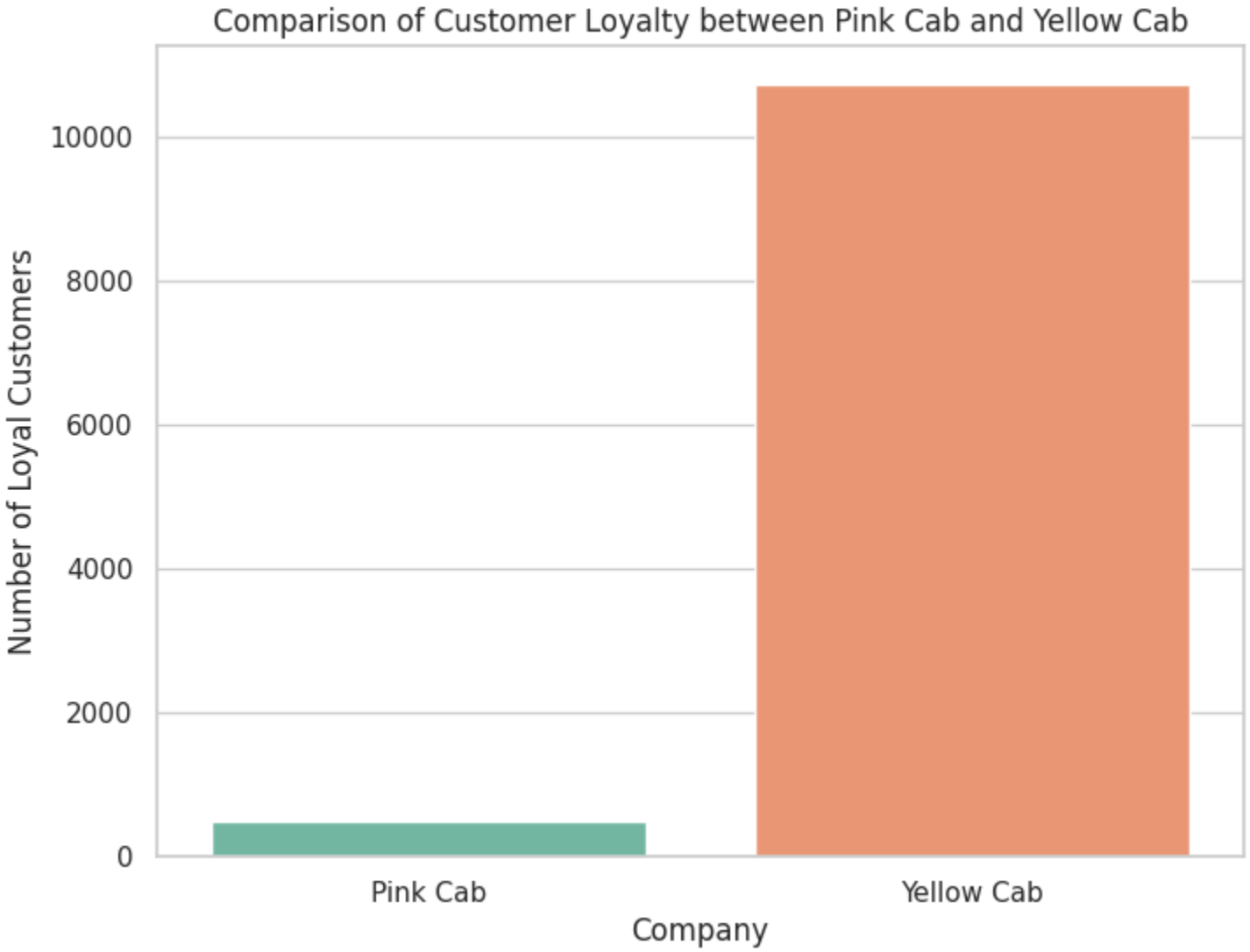
- When we looked at the Cost of the Trip for each ride, there were some rides that were unusually costly as compared to all other rides.
- Possible reasons for such expensive rides could be either Premium rides, or Overnight rides, or Long distance rides.
- In any case, it usually shows that customer is not that concerned about the expenses.
- These people are called premium customers.
- Here, we found that **0.02%** of customers who take **Pink Cab** are Premium customers and **2.16%** of customers who take **Yellow Cab** are Premium customers.

Profit Analysis by Distance



- We have assumed that a trip upto 8km is short distance trip, a trip from 8km to 30km is Medium distance trip, and a trip above 30km is a long distance trip.
- Here, we can see that as the trip distance increases, profit also increases in both the companies.
- Also, the profit from Long distance trip is more than short and medium distance combined.
- Yellow cab as always make more than double profit in every category.

Customer Loyalty Analysis



- We have assumed that any customer who uses the same cab service atleast every 4 months on average is a loyal customer to that cab service.
- Here, we have data for 36 months approx., so any customer who has 9 rides with a particular company is a loyal customer to that company.
- Although Yellow Cabs have costlier rides, still they have far more loyal customers as compared to the pink cabs.
- Pink Cabs have less than 1000 customers, whereas, Yellow Cabs have more than 10,000 loyal customers.

Recommendations

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

- **Customer Reach :** Yellow cab has higher customer reach in 15 cities by a huge margin while Pink cab has higher customer reach in 4 cities but the difference is not that much.

- **Customer Retention:** We have assumed that any user who uses the same cab company atleast 9 times in these 3 years is a Loyal/Retained Customer. Yellow Cabs has 10 times more customer retention as compared to Pink Cabs.

- **Profit:** Yellow cab's total profit is more than six times the total profit of the Pink cab for each of the three years. If we look more precisely, the average profit per KM for Yellow cabs is more than twice the average profit per KM for Pink Cabs.

- **Seasonality & Demand:** If we look at the demand over time, the demand of Yellow Cabs in worst season is more than the demand of Pink Cabs in their best season.

- **Trip Distance:** Even though the profits of Yellow cabs are almost double than Pink cabs for any length of the trip, but both the companies have less profit in short distance trips. As the distance of the trip increases, Yellow cabs profit improves drastically.

- **Customer Loyalty:** The Yellow Cabs have 10 times more loyal customers than Pink Cabs.

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

Thank You