Data Glacier
Your Deep Learning Partner

Exploratory Data Analysis
G2M CASE STUDY
Virtual Internship

20th November 2023

Agenda

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Data Glacier
Your Deep Learning Partner

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.

Project Delivery

- You've received several datasets containing information about two cab companies. Each file represents distinct aspects of customer profiles. XYZ is seeking your actionable insights to assist them in determining the ideal company for their investment.

Approach

The analysis has been done by using below pointers

- Data Exploration and understanding

- Analysis to find the dominant cab company in terms of profitability, customer reach, customer retention, city wise cab users etc.

- Recommendations for investment

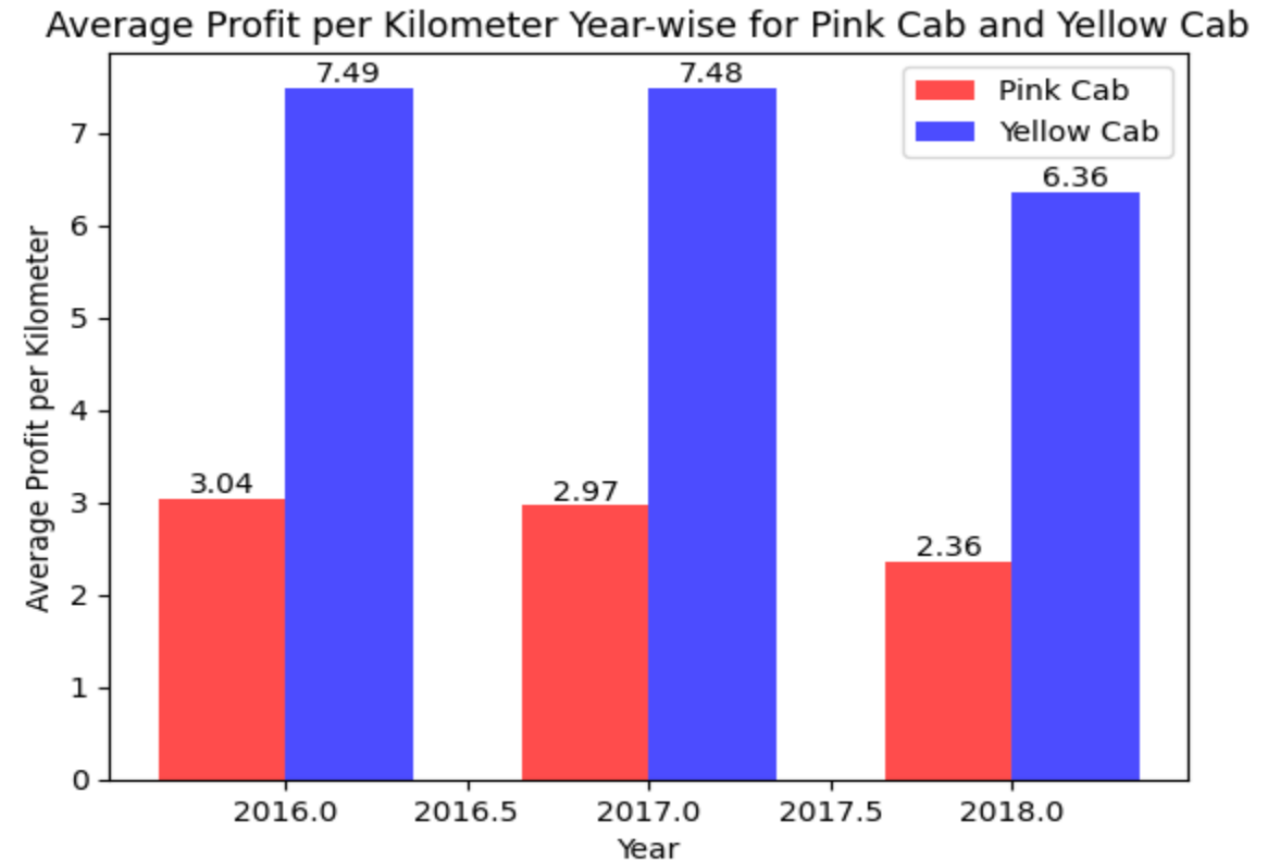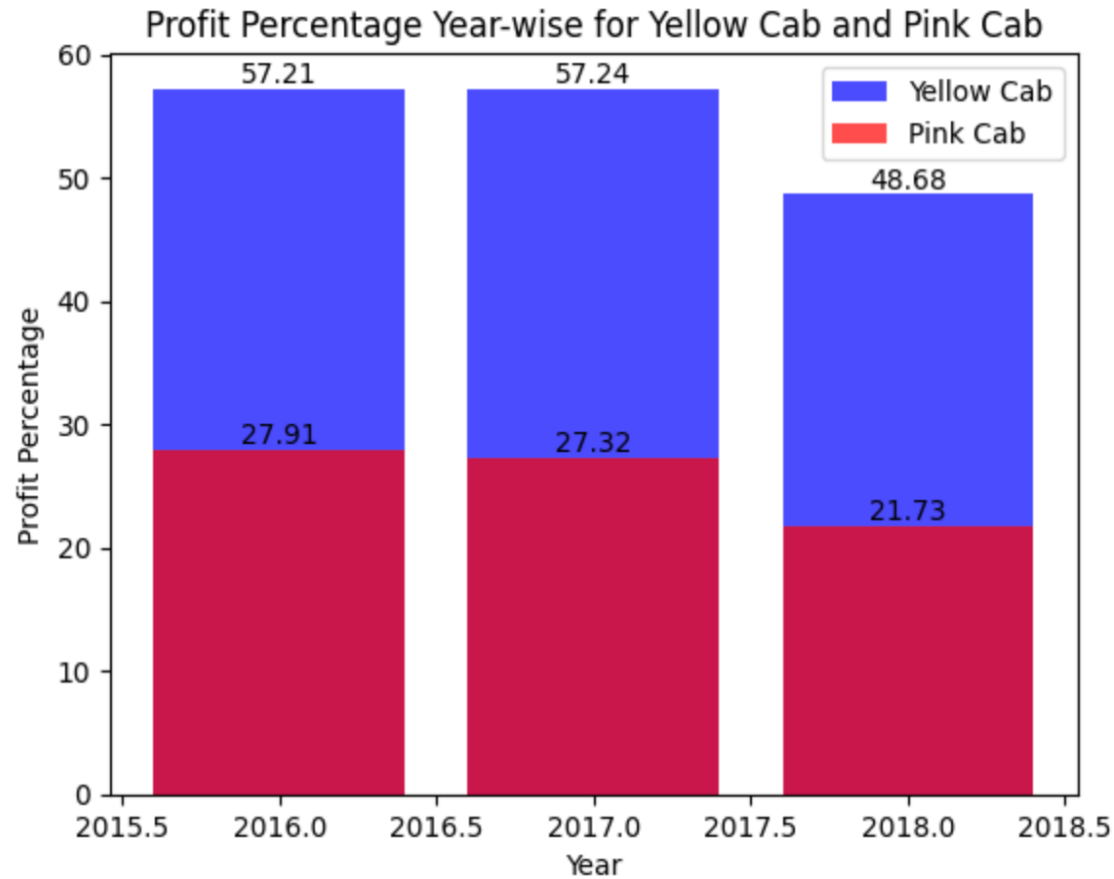# Cab Users In Each City

Percentage of Users Using Cabs in Reference to Population for Each City
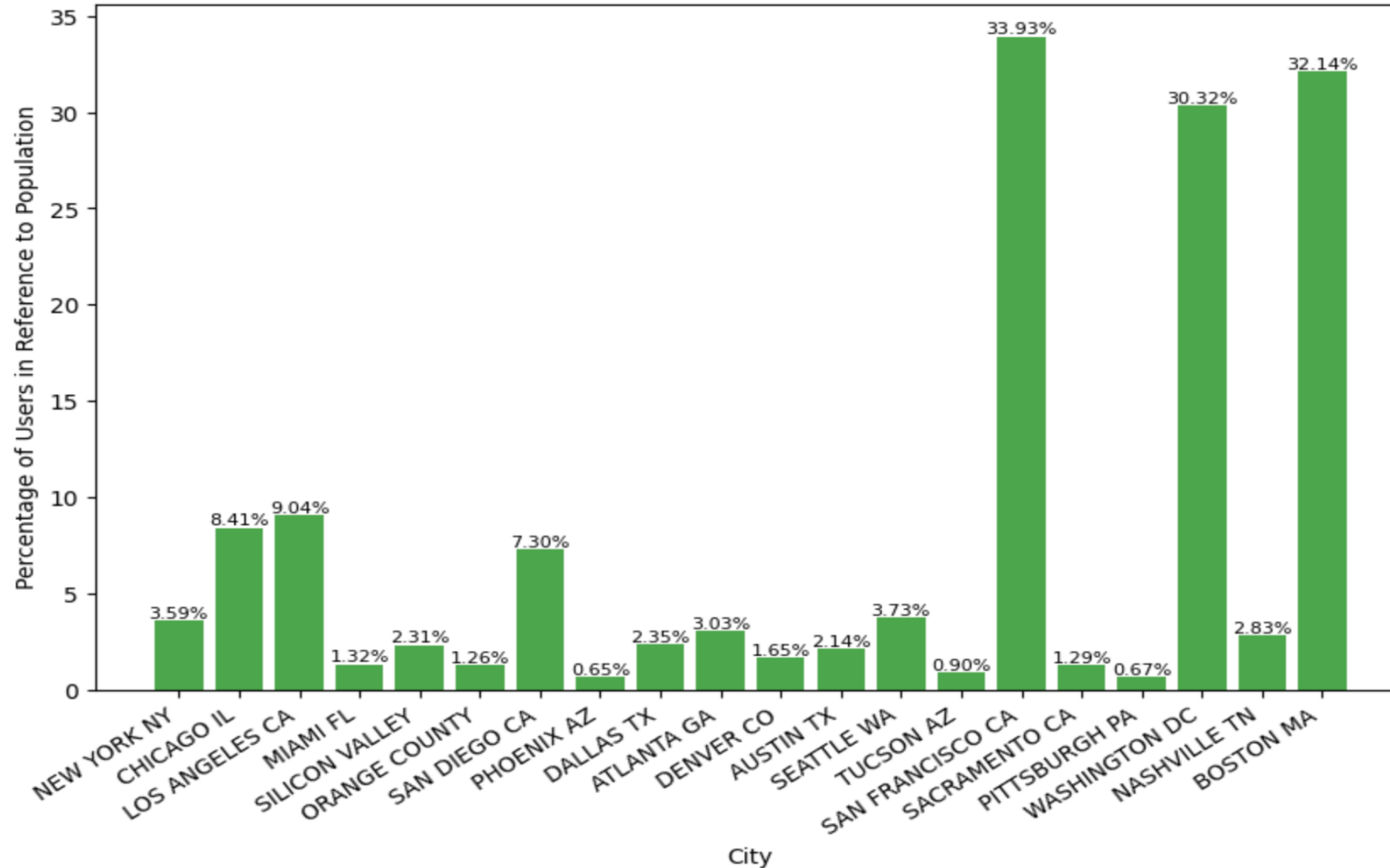
SanFrancisco CA, Washington DC and Boston has the highest numbers of users

# Profit Analysis – Average Profit Per KM

Profit Percentage Year-wise for Yellow Cab and Pink Cab

Average Profit per Kilometer Year-wise for Pink Cab and Yellow Cab

- Yellow cab has higher profit % for each year
- Yellow cab has more than two times profit margin per km than pink cab company

Profit Analysis – Cash and Card Transactions

Profit Comparison Between Payment Modes for Pink Cab and Yellow Cab (2016-2018)
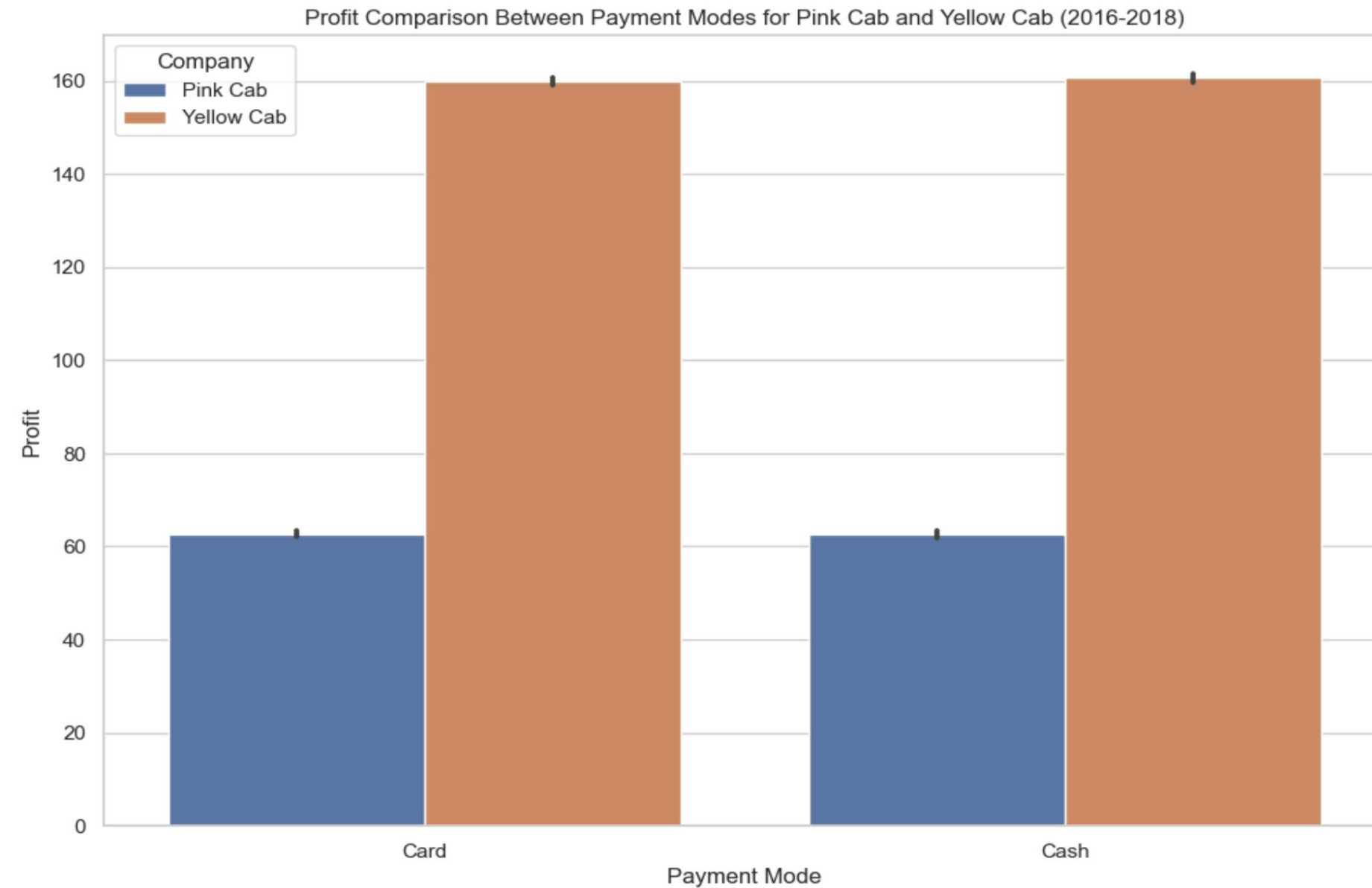
There is almost equal transactions by Yellow and Pink Cab companies in terms of cash or card

Profit and Customer base Analysis Age Wise

Profit Distribution by Age Group for Pink Cab

Profit Distribution by Age Group for Yellow Cab
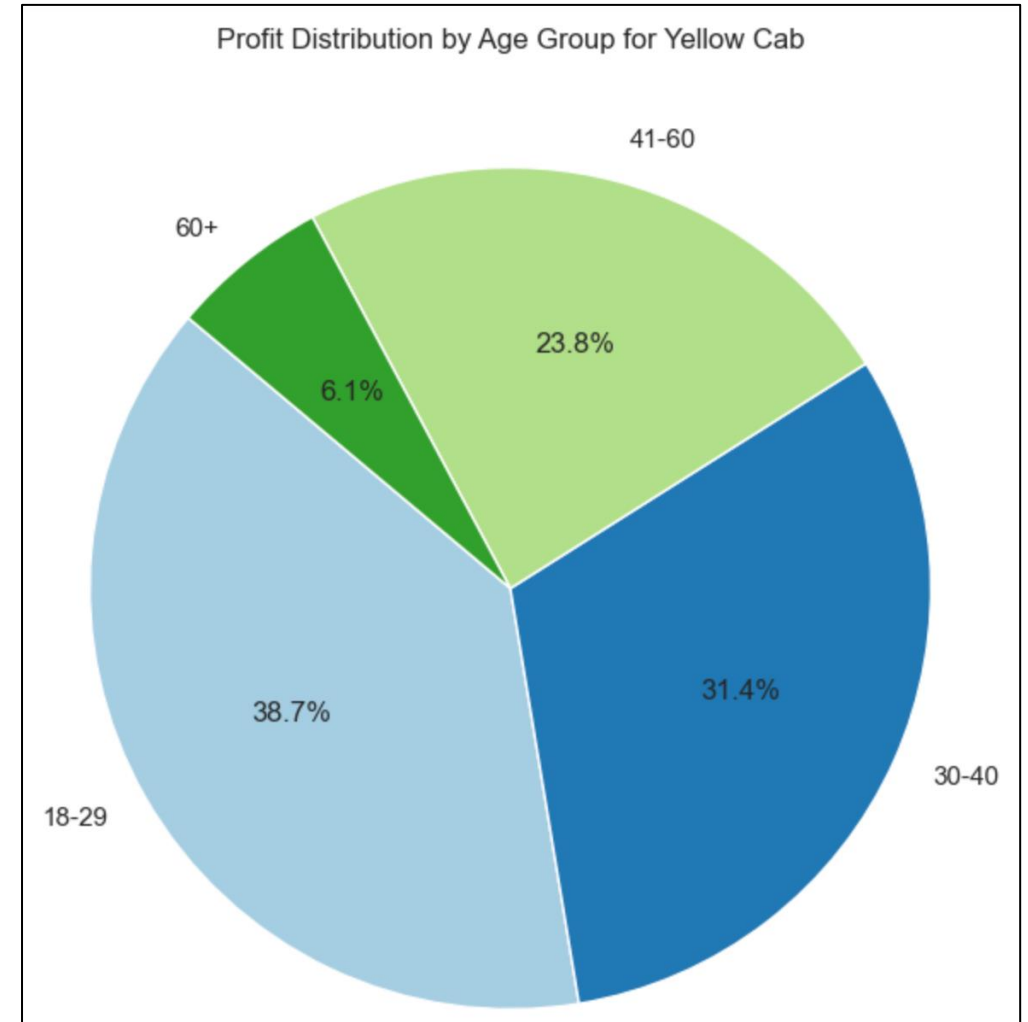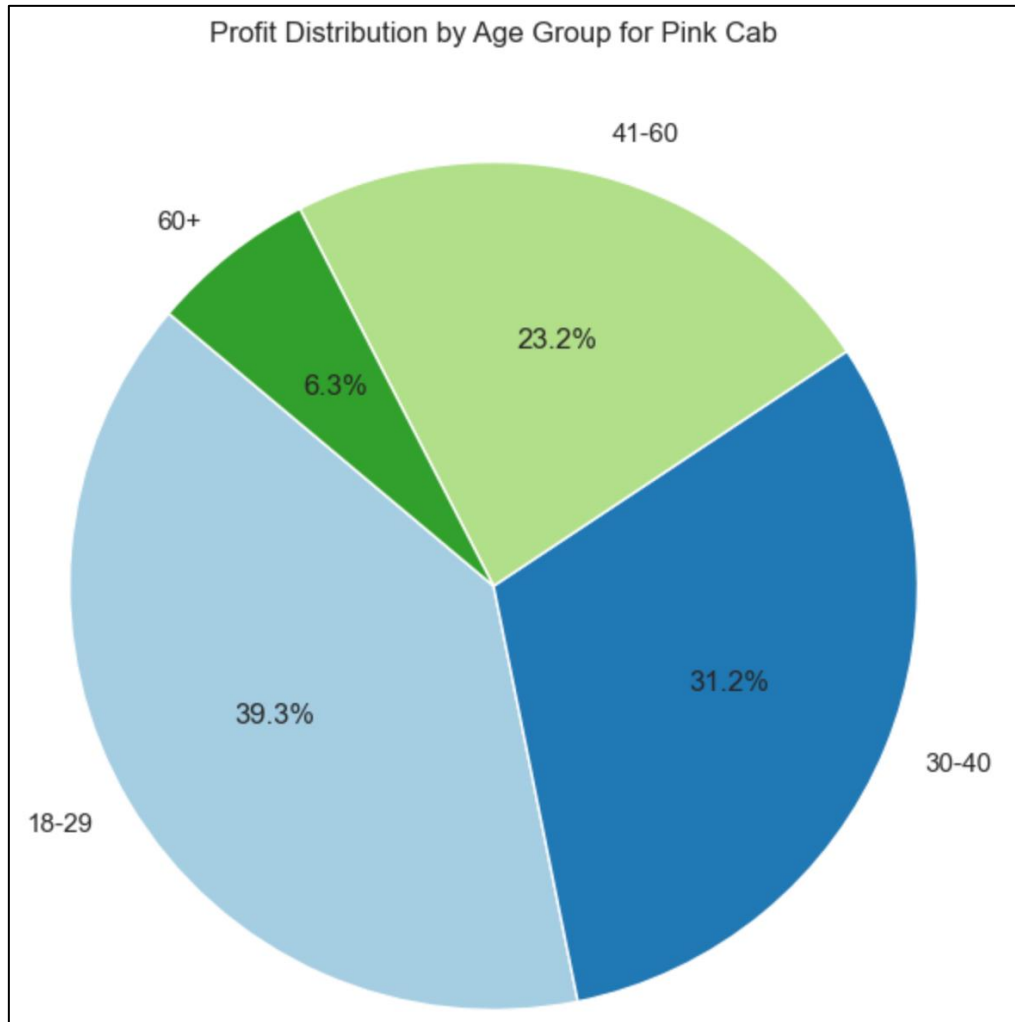
Both cab companies have highest profit contribution from age group of 18-29. Other than that, both companies has almost equal % profit share from other age groups.

Profit and Customer base Analysis Gender wise

Profit Distribution by Gender for Pink Cab

Male

56.1%

43.9%

Female
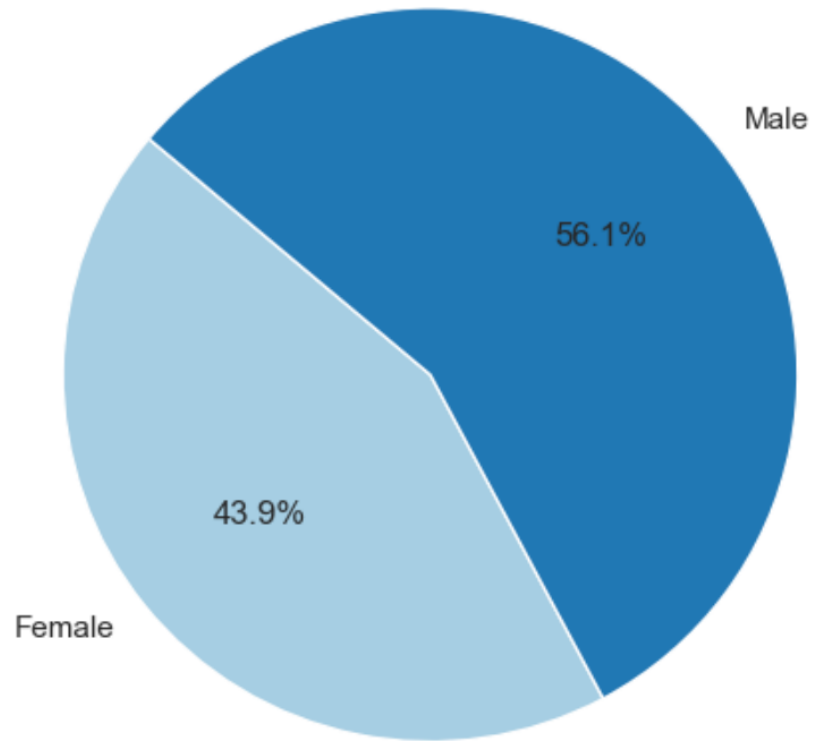
Profit Distribution by Gender for Yellow Cab

Male

58.8%

41.2%

Female
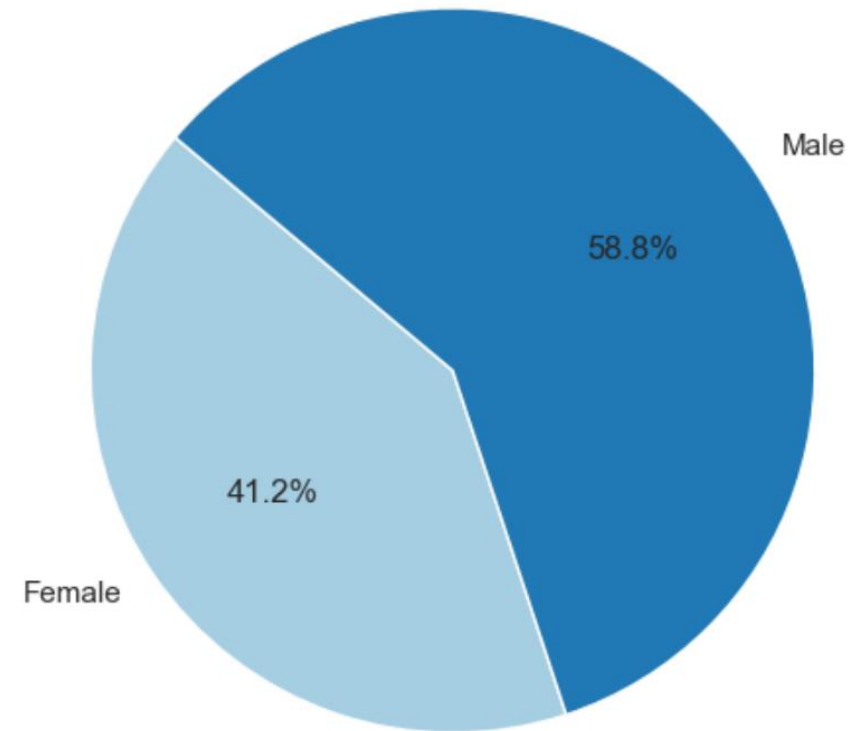
The distribution of gender in both the profit and customer base appears to be relatively balanced for both Pink Cab and Yellow Cab. There is no significant gender bias, as the percentages of profit contribution and the customer base are fairly evenly distributed between male and female customers for both cab companies.
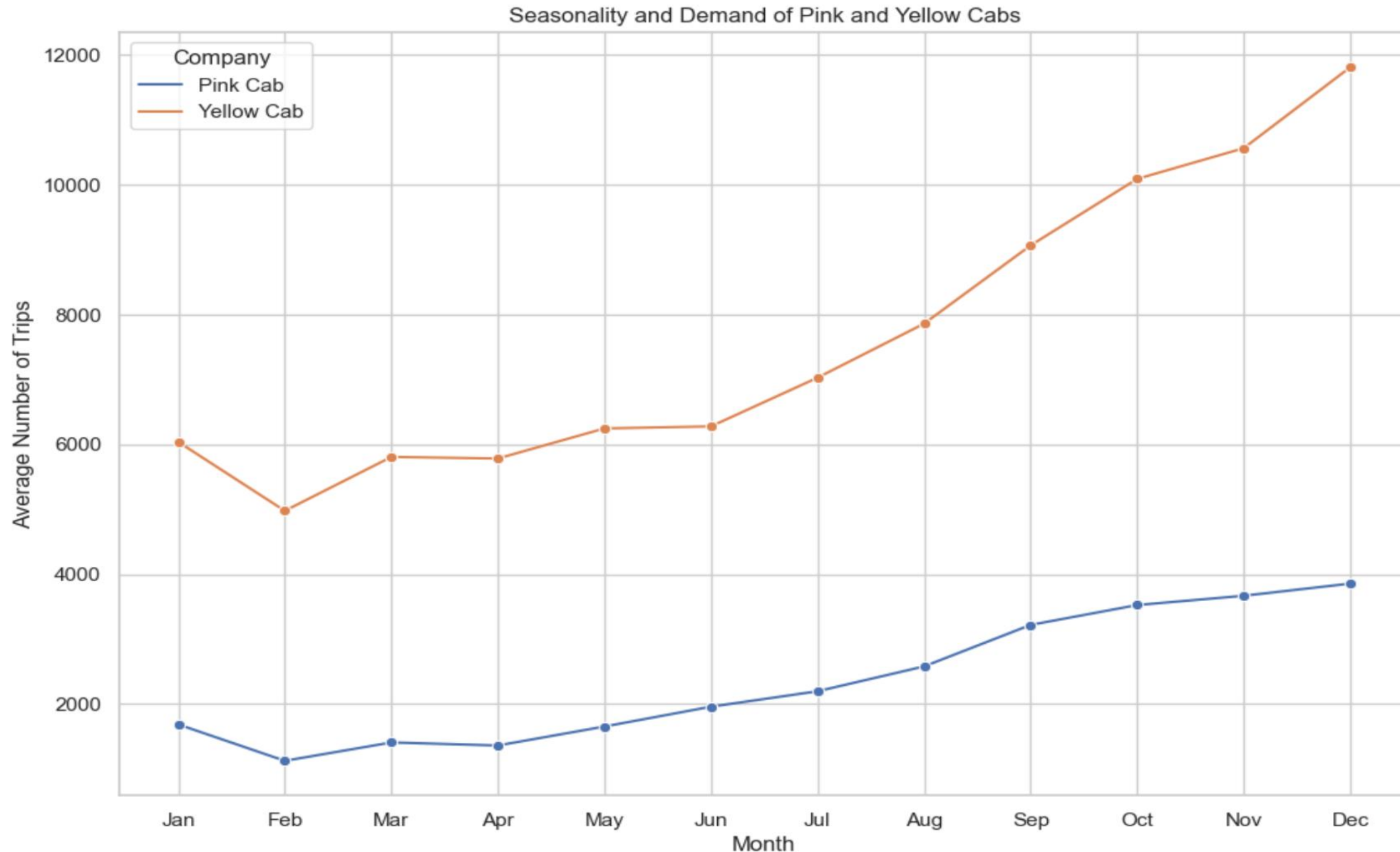
User Presence of cab companies' city wise

Customer Presence of Yellow and Pink Cab City-wise
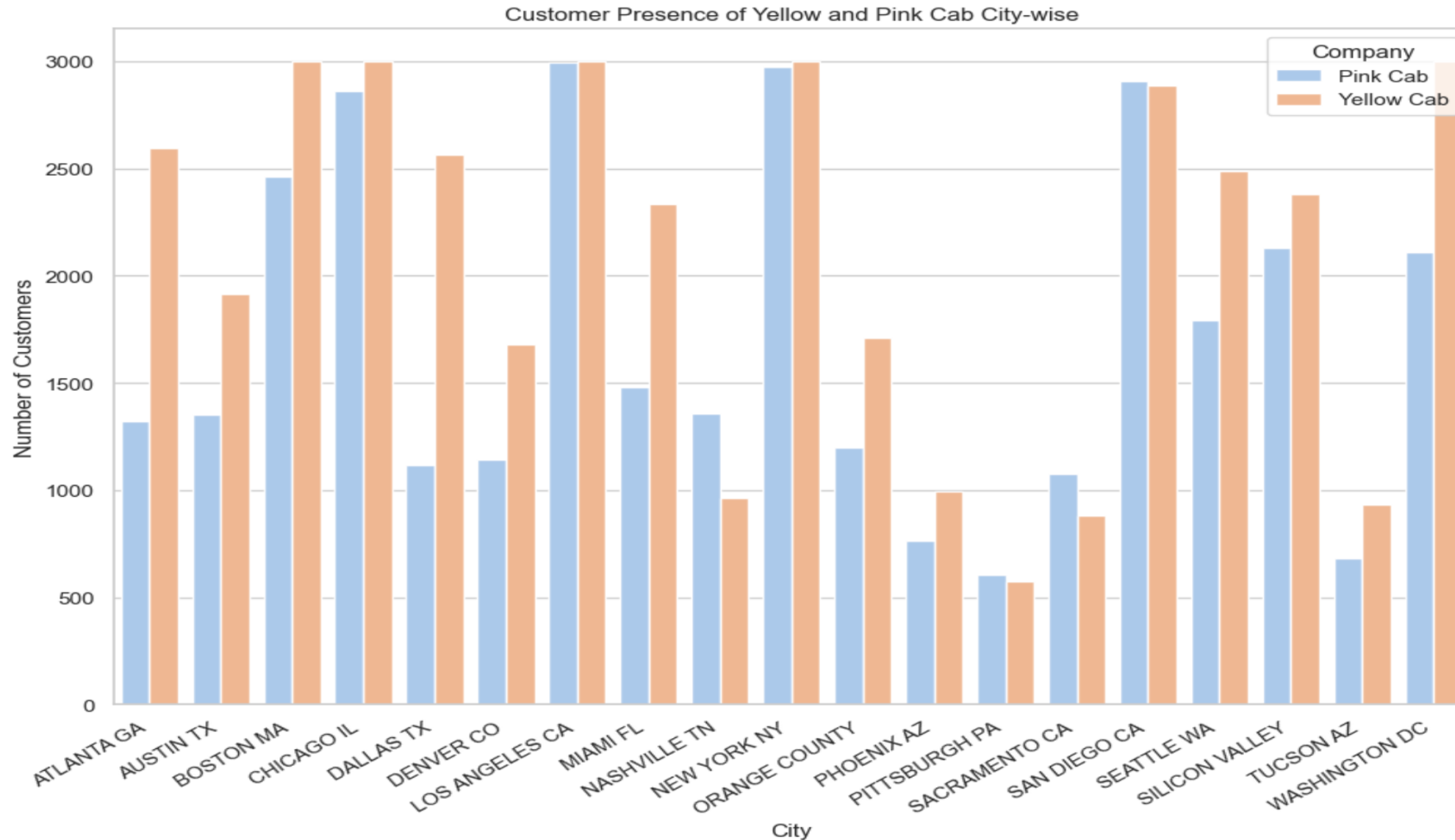
Yellow cab has 15 cities with more users than Pink Cab.
Pink cab has a greater customer reach in following 4 cities:
1. San Diego
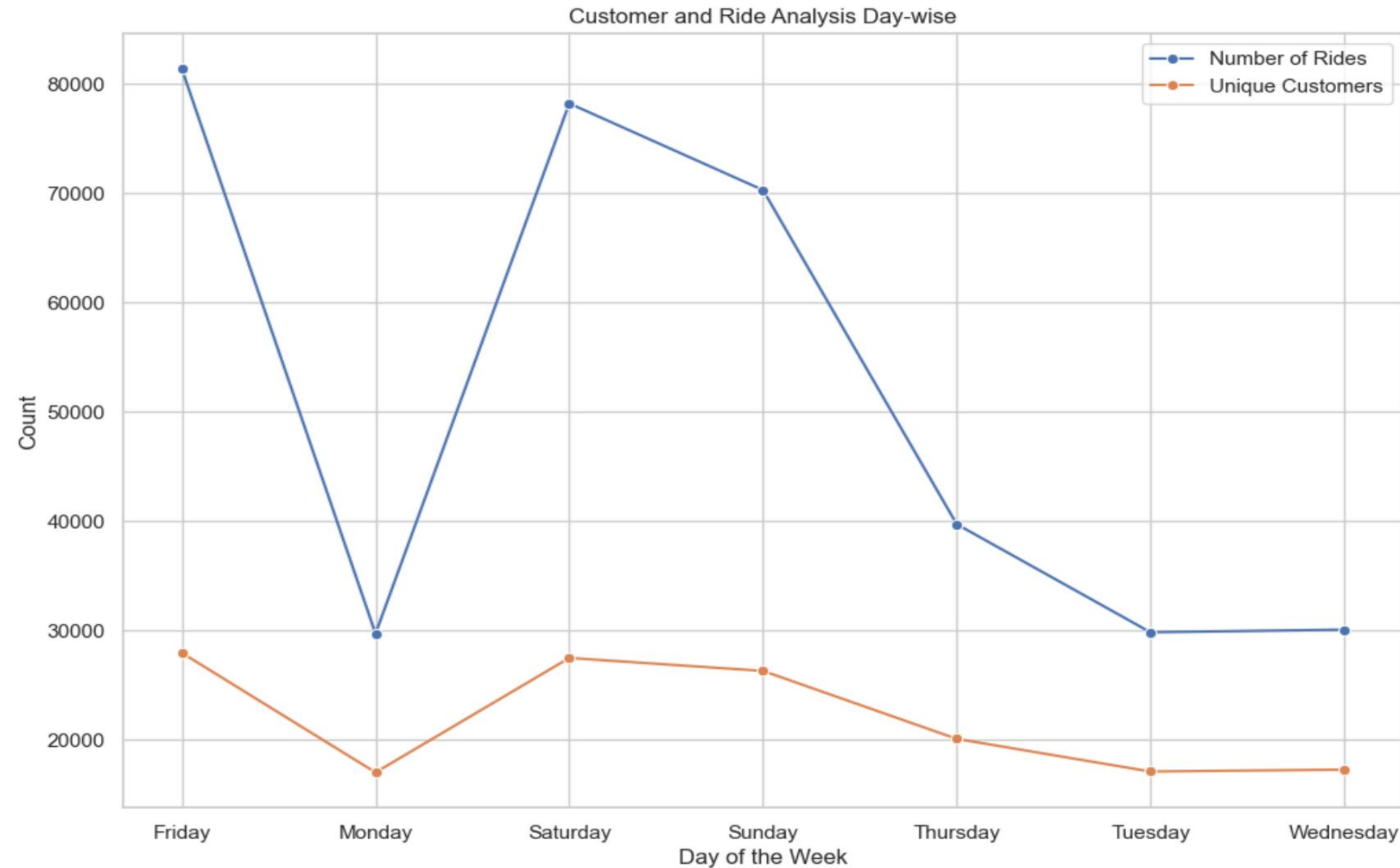2. Nashville
3. Sacramento
4. Pittsburgh

Demand of Pink and Yellow cabs Monthly

Seasonality and Demand of Pink and Yellow Cabs

Average number of trips for both companies increases from July to December. December being the highest, while the lowest trips are in the month of February.

Demand of Pink and Yellow cabs Day-Wise

Customer and Ride Analysis Day-wise

Both companies show higher demand during weekends. While the rides are least on Monday.

Ride And Customer Analysis Day Wise

Customer and Ride Analysis Day-wise for Pink and Yellow Cab
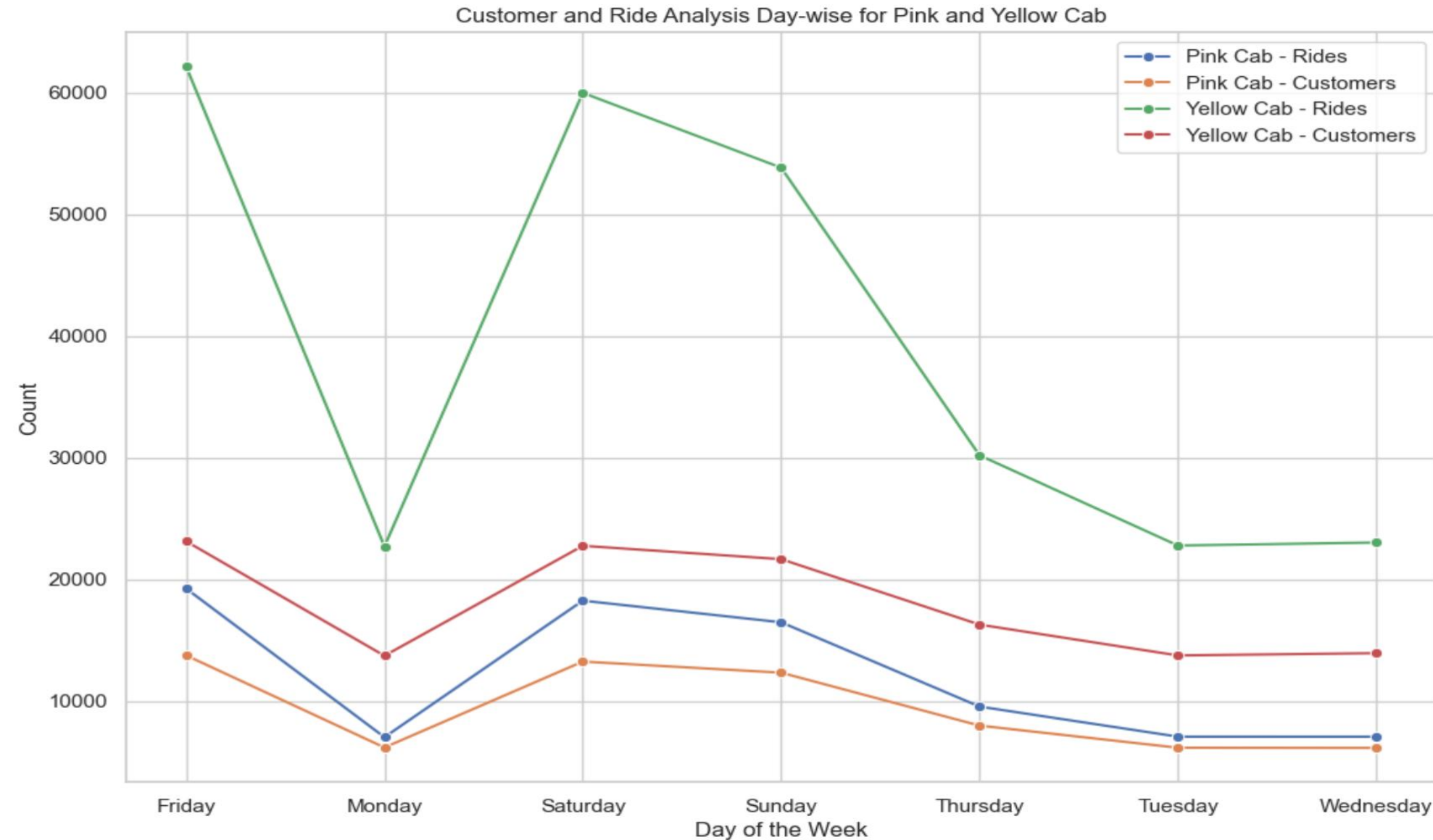
Both companies show higher demand during weekends. While the rides and customers are least on Monday.

Customer Analysis on KM Travelled

Customer Analysis Based on Ride Distance
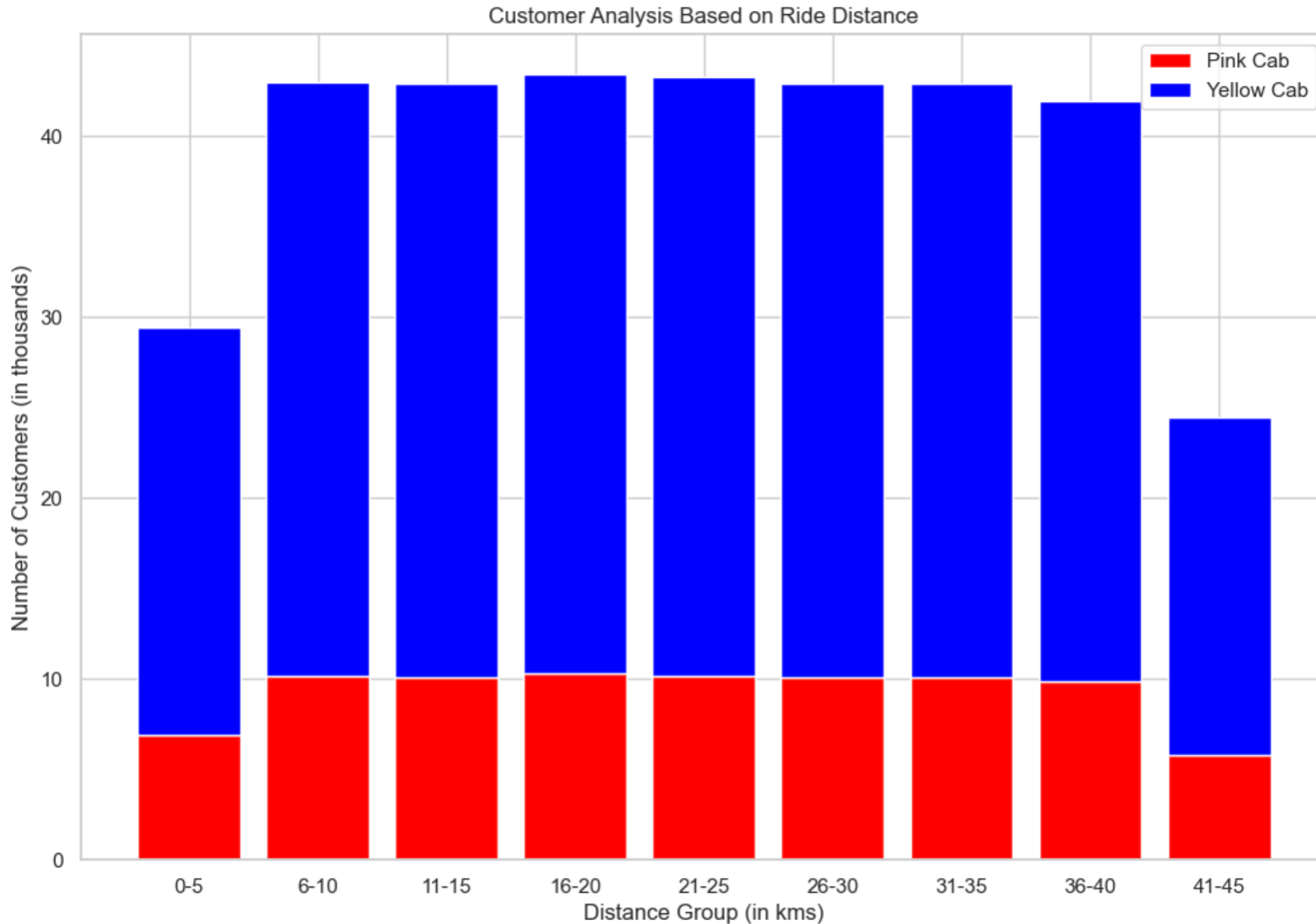
- Yellow Cab demonstrates a consistent customer distribution across short, medium, and long trips, indicating an effective customer plan for various trip lengths.
- Yellow Cab is way better than Pink Cab when it comes to attracting customers for both short and long trips.
- Yellow Cab has a strong customer base across various trip lengths, showing that it has a smarter plan compared to Pink Cab. This highlights Yellow Cab's superior strategy for attracting customers, especially in shorter and longer trips.

Recommendations

Profit Analysis:

1.Yellow Cab consistently outperforms Pink Cab in terms of profit over the three years.

2.Yellow Cab's profit per kilometre is more than twice that of Pink Cab, indicating a higher efficiency in generating revenue.

Customer Reach:

1.Pink Cab has a stronger presence in only four cities, whereas Yellow Cab dominates the rest.

2.Yellow Cab exhibits superior customer retention with a higher number of repeat customers.

Age-Wise Analysis:

1.Both companies share a nearly equal distribution of customers across different age groups, with the 18-29 age group contributing the most to profits.

Gender-Wise Analysis:

1.Males contribute more to profits than females for both Yellow and Pink Cabs.

Month-Wise and Day-Wise Demand:

1.Both companies experience an increase in the average number of trips from July to December, with December being the peak month. February sees the lowest demand.

2.Weekends witness higher demand, while Monday shows the least activity in terms of rides and customers.

Customer Analysis based on Kilometers Traveled:

1.Yellow Cab attracts a diverse customer base across various trip lengths, indicating a well-planned strategy compared to Pink Cab.

Conclusion: Based on the comprehensive analysis, I recommend investing in Yellow Cab due to its consistently higher profits, superior customer reach, and effective strategies for attracting and retaining customers across different factors.

Thank You