



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M insight for Cab Investment Firm

15.07.2025

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Executive Summary

- 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.

Problem Statement

- Objective: To assist XYZ firm in determining the most profitable and optimal investment solution through clear visualizations and actionable insights.

Approach

- Data Understanding
- Data Cleaning and manipulation
- Analysing and visualizing data to find best cab company
- Finding the most profitable and used Cab company
- Recommendations for investment

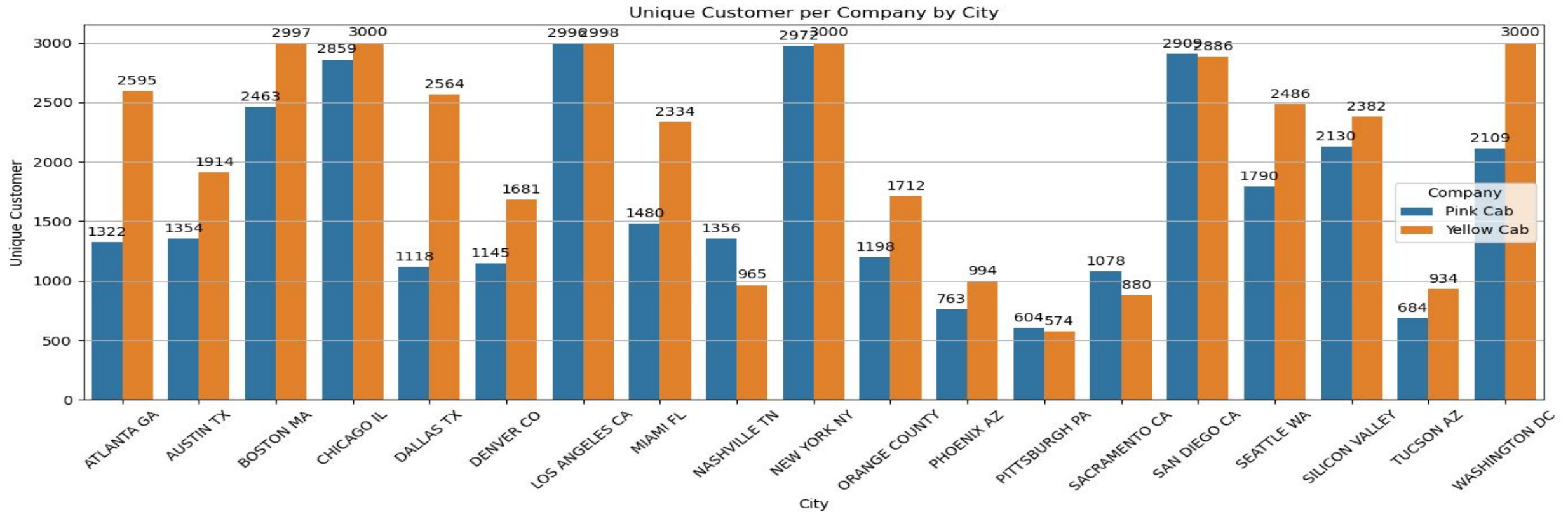
Data Exploration

- 23 features in total including 9 derived features
- Time Frame: 2016-01-31 to 2018-12-31
- Total Data: 359392
- Merged every dataset into one master data

Note:

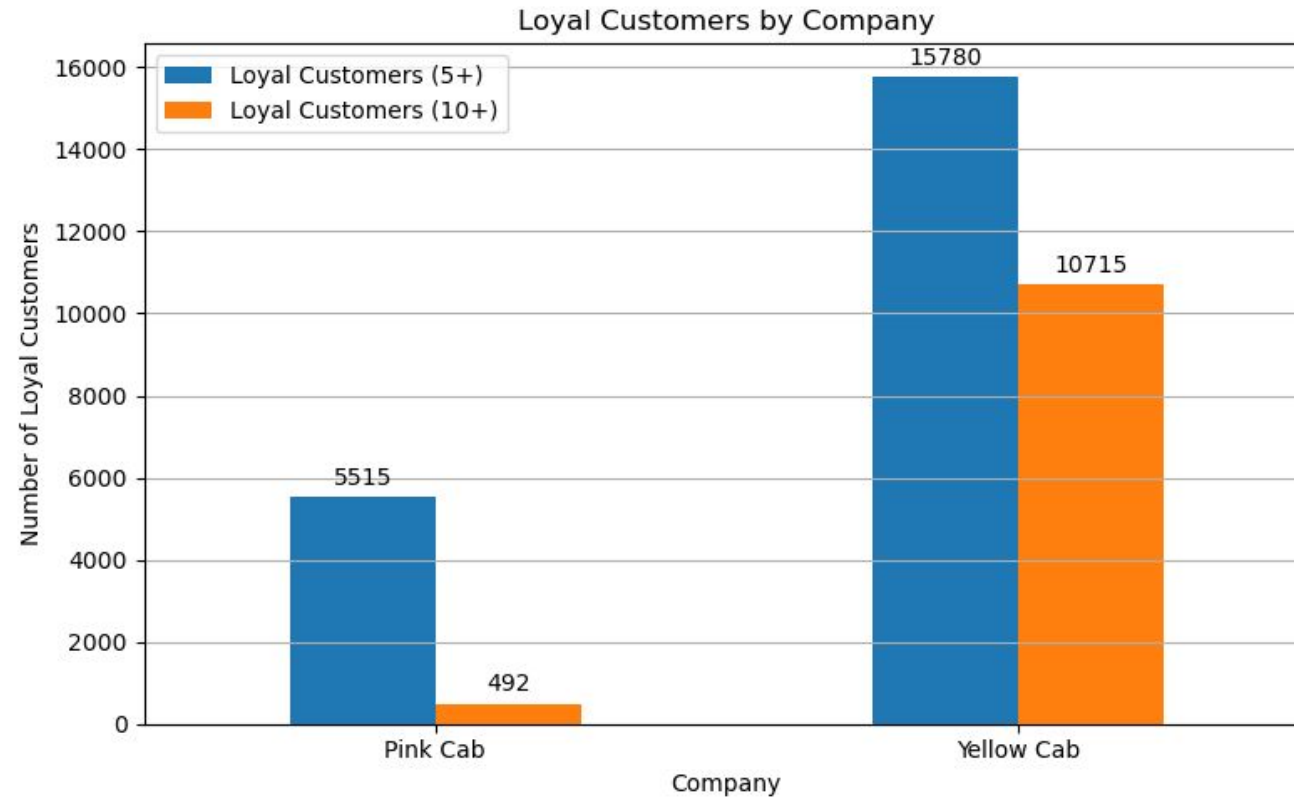
- We observed a high number of outliers in the Price Charged feature. However, instead of removing them, we've chosen to retain this feature for analysis because the price can be influenced by multiple external factors not captured in the dataset.
- The two core engineered features are: Profit, Profit per KM.
- The other 7 derived fields were created solely to assist with segmentation and deeper understanding.
- Profit is calculated on the basis of difference between Price Charged and Total cost and Profit per KM is calculated on the basis of Profit and Km Traveled.
- Users feature considered as Total number of users for cab company.

Yellow Cab has higher unique customer reach across U.S. cities



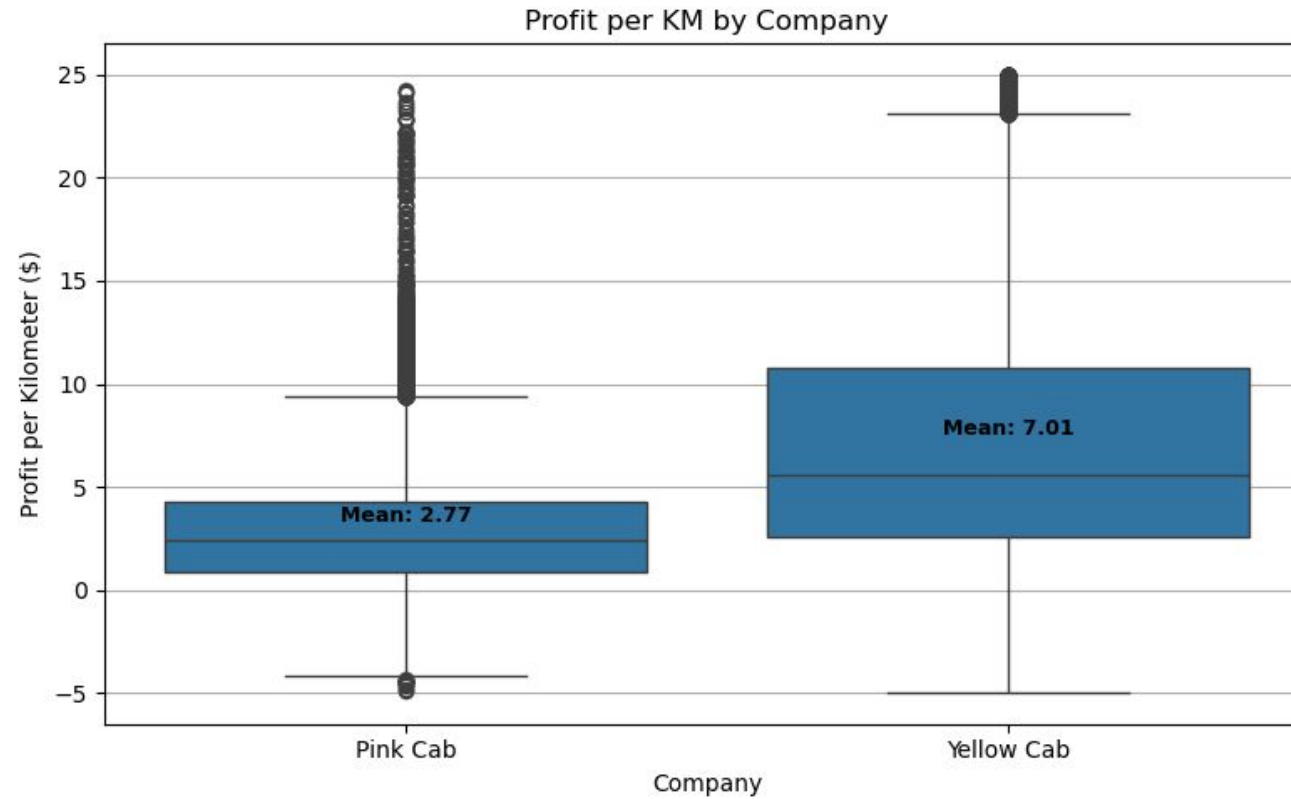
- Except for Los Angeles, New York, and Sacramento, Yellow Cab has a higher Unique customer reach across most US cities. This suggests stronger brand presence and market penetration, making it more attractive from a growth and user perspective.

Yellow cab retains more loyal customers (5+ or 10+ rides) than Pink cab



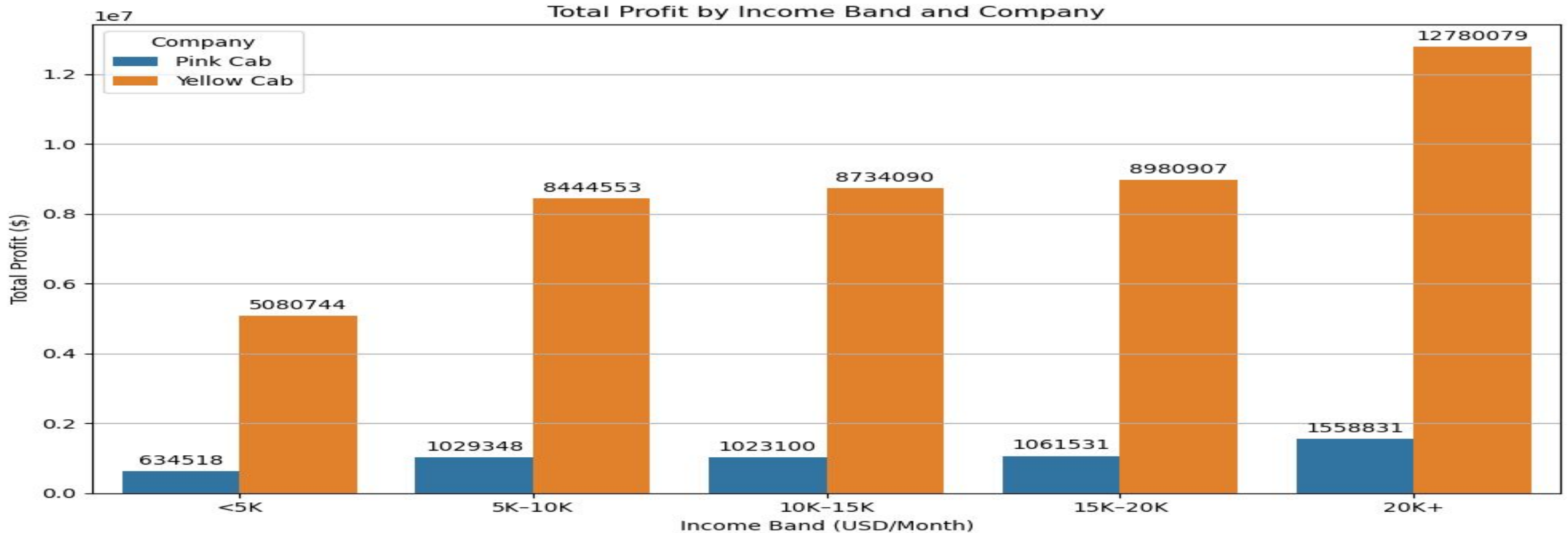
- Yellow Cab retains significantly more loyal customers than Pink Cab at both 5+ and 10+ ride thresholds. This reflects better customer satisfaction, engagement, and long-term value, which is a strong indicator of business health and investment potential.

Yellow cab has higher average profit per kilometer than Pink cab



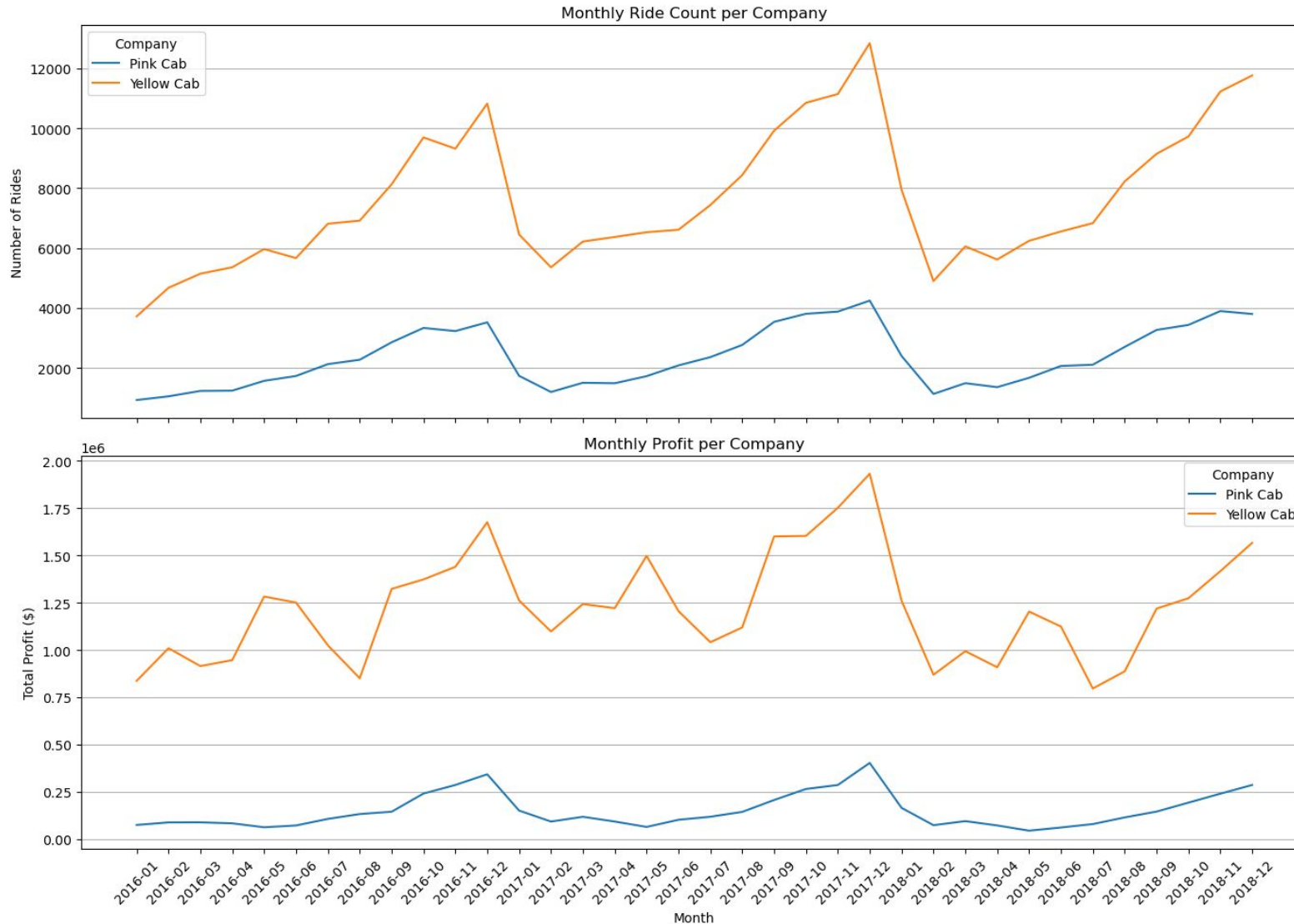
- We have removed rows where Profit per KM > 25 to reduce skewness.
- Yellow Cab demonstrates significantly better profitability per kilometer, making it a more cost-efficient and revenue-optimized operation compared to Pink Cab.

Higher income groups contribute more to overall profit, especially for Yellow cab



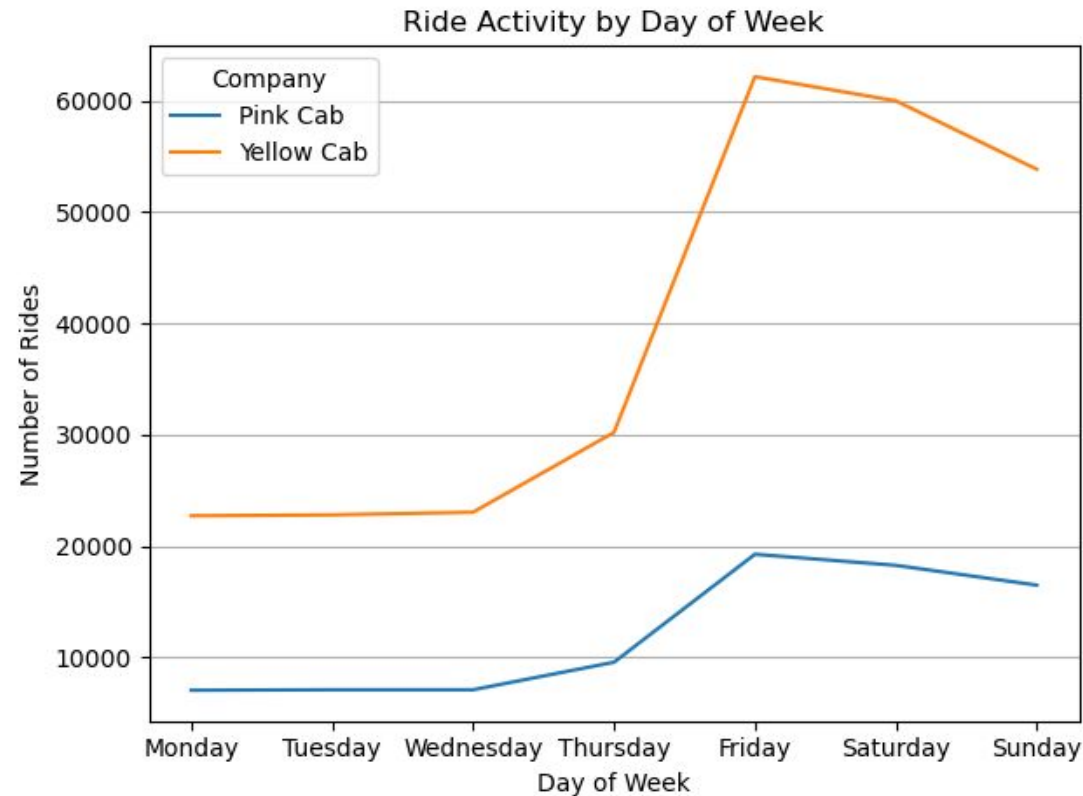
- Higher income customers contribute more to overall profit.
- Yellow Cab consistently earns more profit than Pink Cab across all income groups, with the largest difference seen in the 20K+ band.

There is strong seasonality in demand and profit for both companies



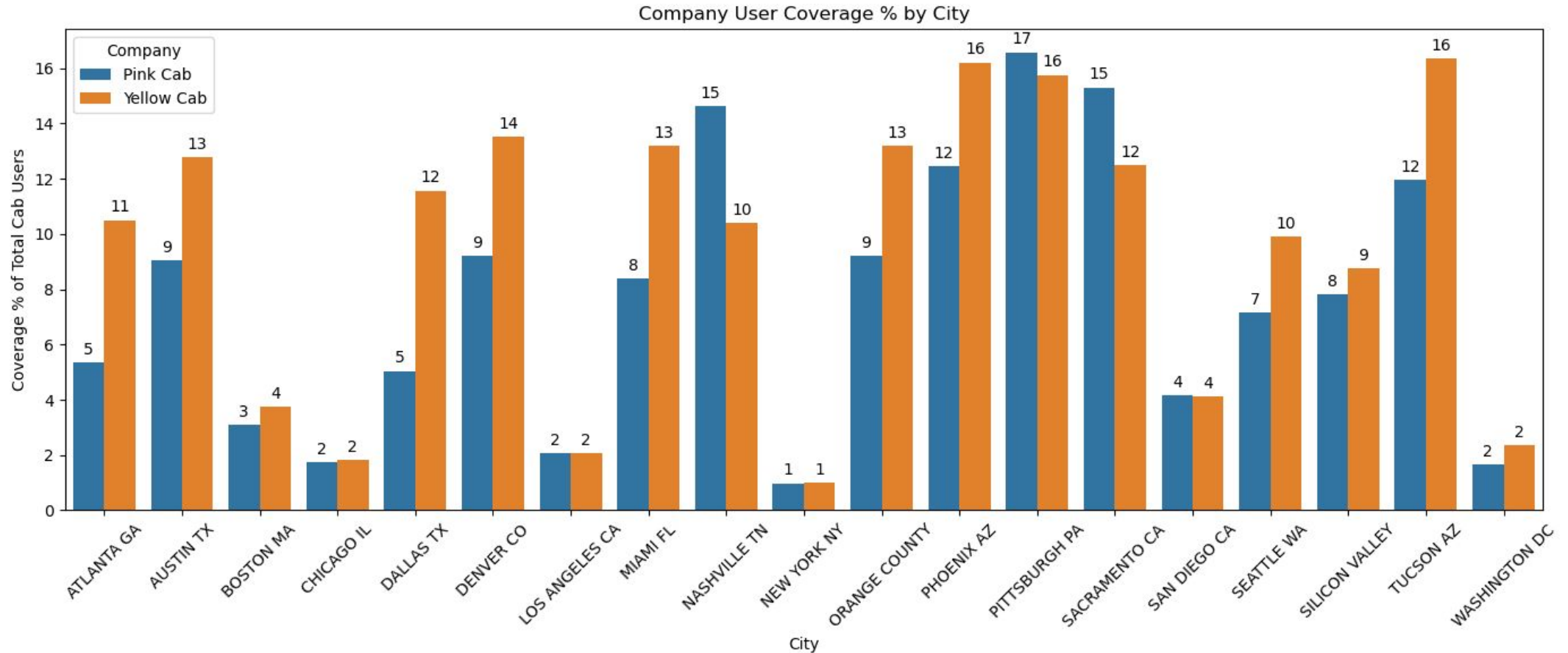
- Both Yellow and Pink cabs experience seasonal changes in ride demand and profit.
- Peaks are visible in months August-December, indicating higher travel during holidays and vacations.
- The cab industry shows clear seasonality in both demand and profitability.

Yellow cab sees more ride activity on weekends and holidays than Pink cab



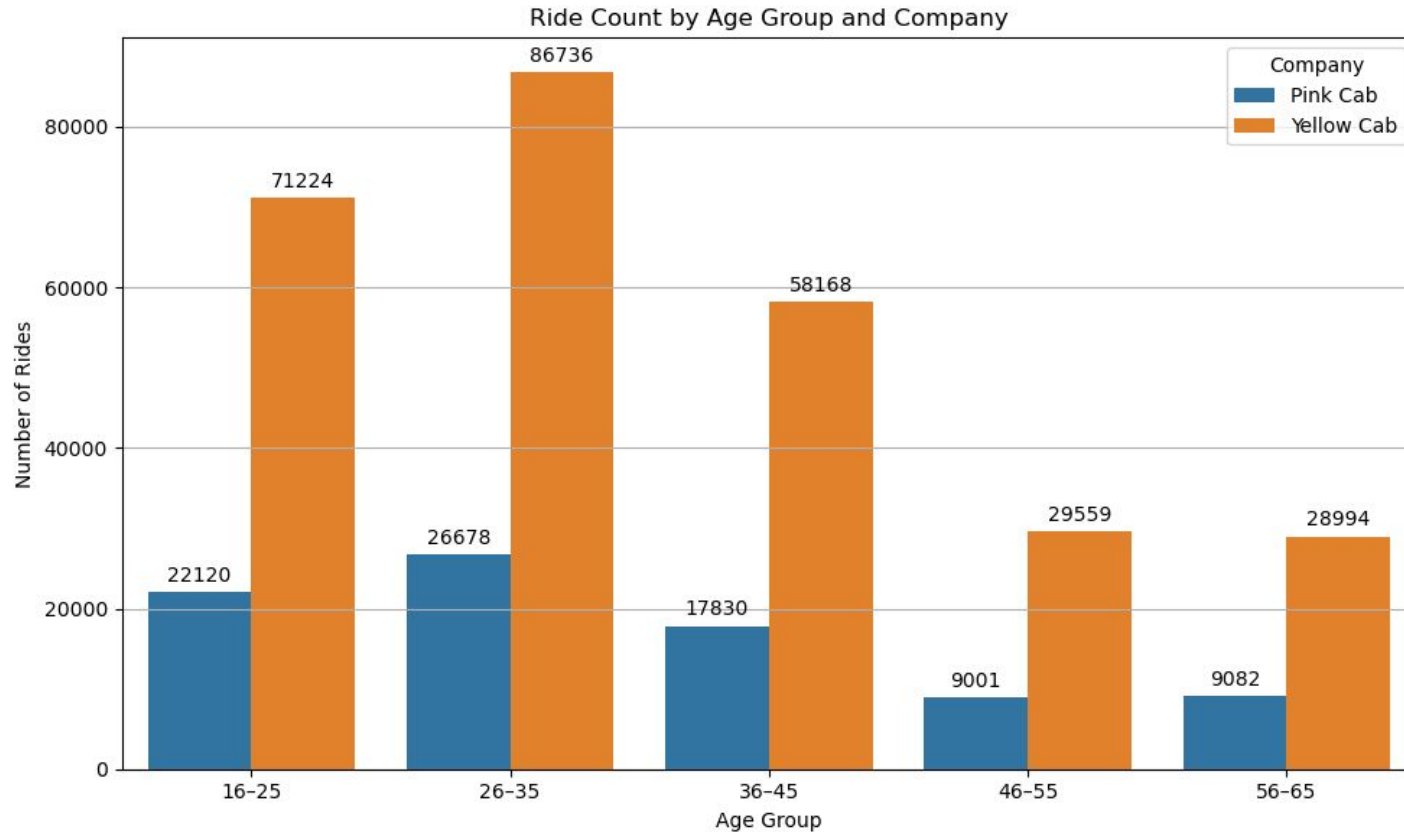
- Both companies show increased activity on weekends, but Yellow Cab sees higher ride volumes, indicating stronger weekend preference, which may indicate better service availability or popularity for personal travel use.

Yellow cab covers a higher percentage of total cab users in most cities.



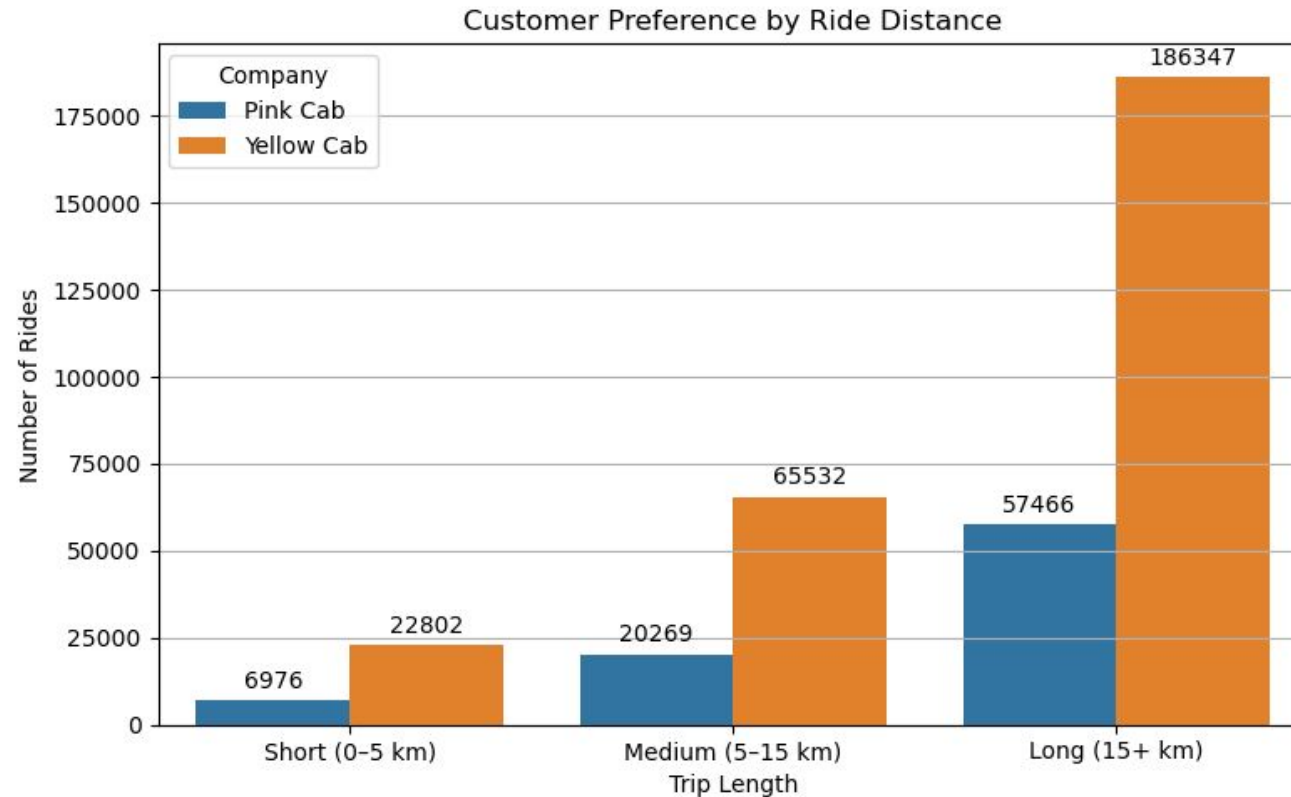
- However Pink Cab leads in only a few cities, Yellow Cab serves a larger portion of the available cab market in more cities than Pink Cab, making it a stronger candidate in terms of geographical market share.

Yellow cab has more balanced reach across all age groups, including senior citizens



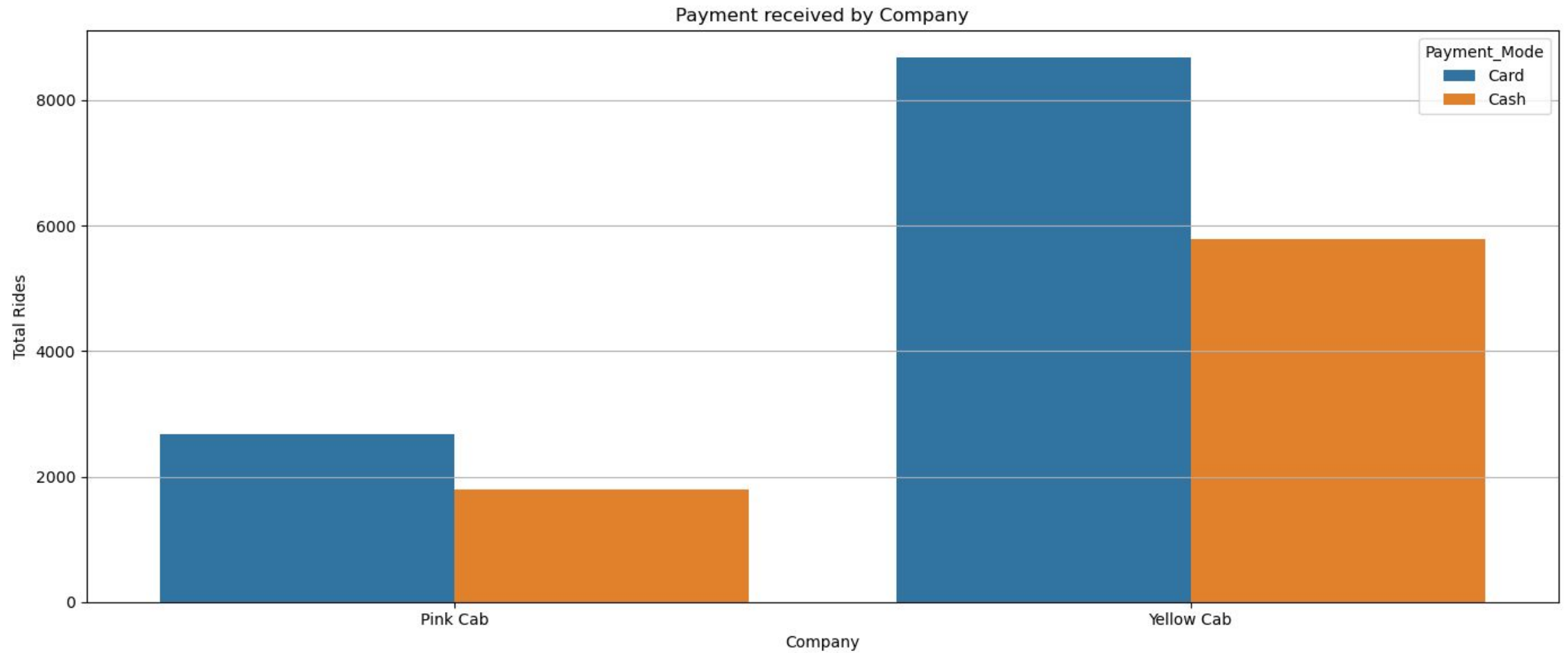
- Yellow Cab has a well-balanced distribution across all age groups, especially seniors, which broadens its long-term customer base and makes it a more inclusive investment option.
- In contrast, Pink Cab usage drops off in older age bands, suggesting it's less favored by senior customers.

Ride distance affects customer preference differently across companies

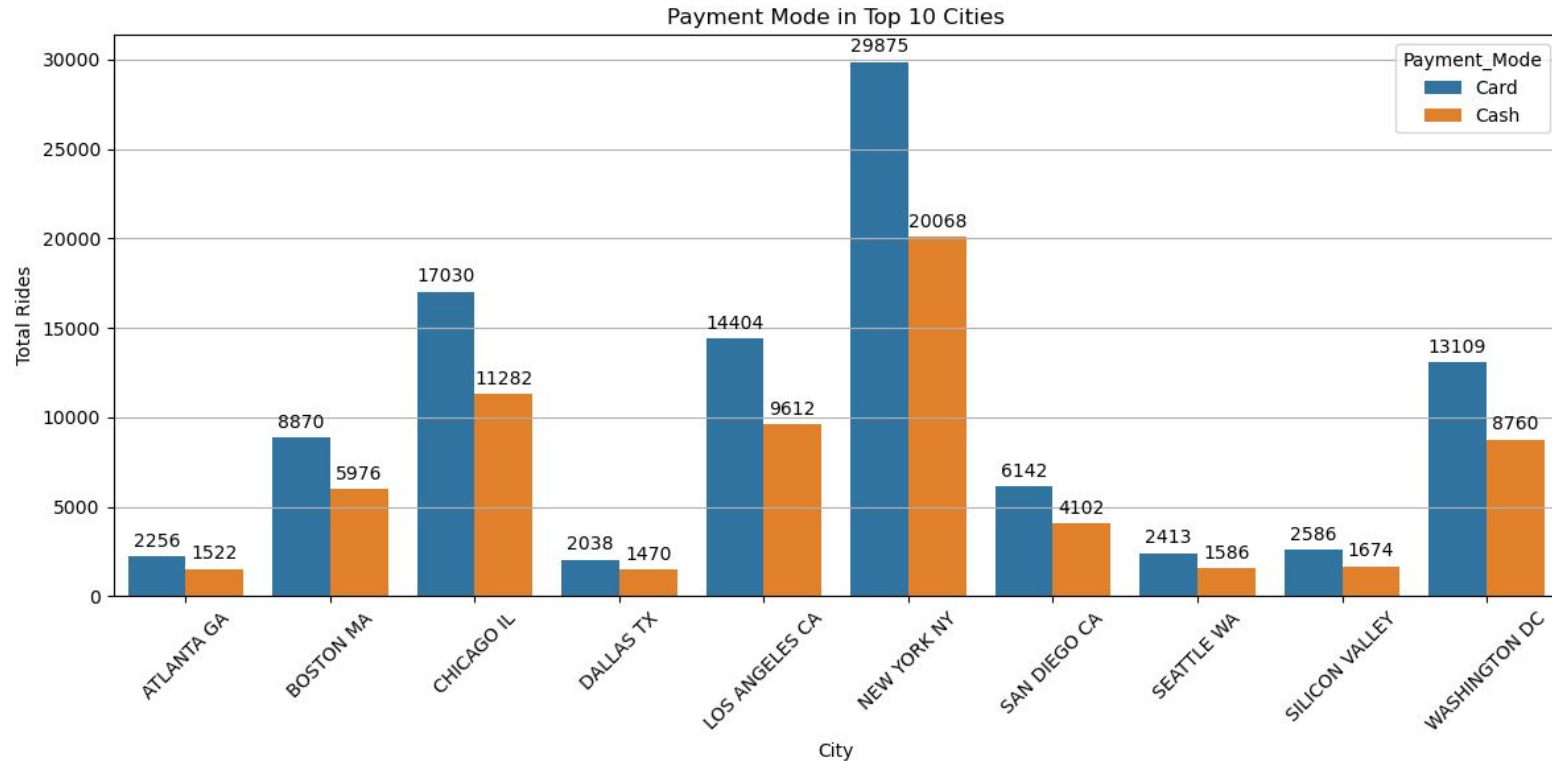


- Yellow Cab has higher ride volume across all trip distances, including short, medium, and long trips.
- This indicates that Yellow Cab is consistently preferred for all types of travel, whether quick local trips or longer commutes.

Payment mode preference differs by city and company



Payment mode preference differs by city and company



- Both cab companies tend to be more card-oriented overall, possibly due to their digital presence and operation in urban areas.
- However, Pink Cab shows a more balanced ratio of cash and card users in certain cities, suggesting a mix of digitally active and traditional customer bases.

Recommendation and Analysis

Profit Analysis:

- Profit Per Cab Company Overall: Yellow Cab has almost 9 times more profit than pink cab.
- Profit per City : Yellow Cab has more profit in every city than Pink Cab.
- Profit Per Kilometer : Yellow Cab has more profit per kilometer in all cities than Pink Cab. Also, Overall Yellow cab has almost 3 times more profit per KM than Pink Cab.
- Profit Per Month and Year : Yellow Cab has almost 10 times more profit per all months than pink cab. Moreover, Yellow Cab has almost 9-10 times more profit per Quarter than pink cab.

Customer Reach:

- Customers Reach in Cities: Pink Cab has more Customer reach than Yellow cab in only 4 cities.
- Customers Reach Time wise : Yellow cab has certainly higher customer reach than pink cab in every month of all three years observed.
- Customer Reach Age wise: Yellow Cab has better customer reach for every age group than Pink Cab

Recommendation:

- Based on above analysis I will recommend you to use Yellow cabs than Pink Cabs as Yellow Cab has dominated in all the areas whether it is Profit or Customer Reach.

Other Analysis and Recommendation:

- Boston, Los Angeles, Washington DC, Chicago, and San Diego have the highest percentage of their population using cab services, indicating strong local adoption. Meanwhile, New York stands out with the largest absolute number of users, driven by its large population base. Therefore, expanding cab operations in these cities would likely yield the most benefit.
- Additionally, approximately two-thirds of users prefer paying by card rather than cash, highlighting the importance of supporting digital payment methods for customer convenience and market alignment.
- Lastly, both cab services experience a surge in usage during the second half of the year. Hence, it is advisable to operate at maximum capacity during Q3 and Q4 to capitalize on seasonal demand.

Thank You