



Data Glacier

Your Deep Learning Partner

G2M Cab Case Study

Virtual Internship

21-Jan-2023

Agenda

Problem Statement

Approach

Assumptions

EDA

EDA Summary

Recommendations

Problem Statement

The Client

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.

The Candidates

We are examining two candidate cab companies: Pink Cab and Yellow Cab.

Our Mission

Conduct exploratory data analysis using the given data to provide actionable insights to help XYZ make a decision about which cab company to invest in.

Approach

To evaluate which cab company is the better prospect, we will focus on the following 3 key metrics:

1. Profits - which company makes more money
2. Customer reach - this is split into two parts:
 - Which company sells more rides in the most cities
 - Which company has a larger customer base when considering the following:
 - Gender
 - Age groups
 - Income levels
3. Customer retention - which company has more returning customers

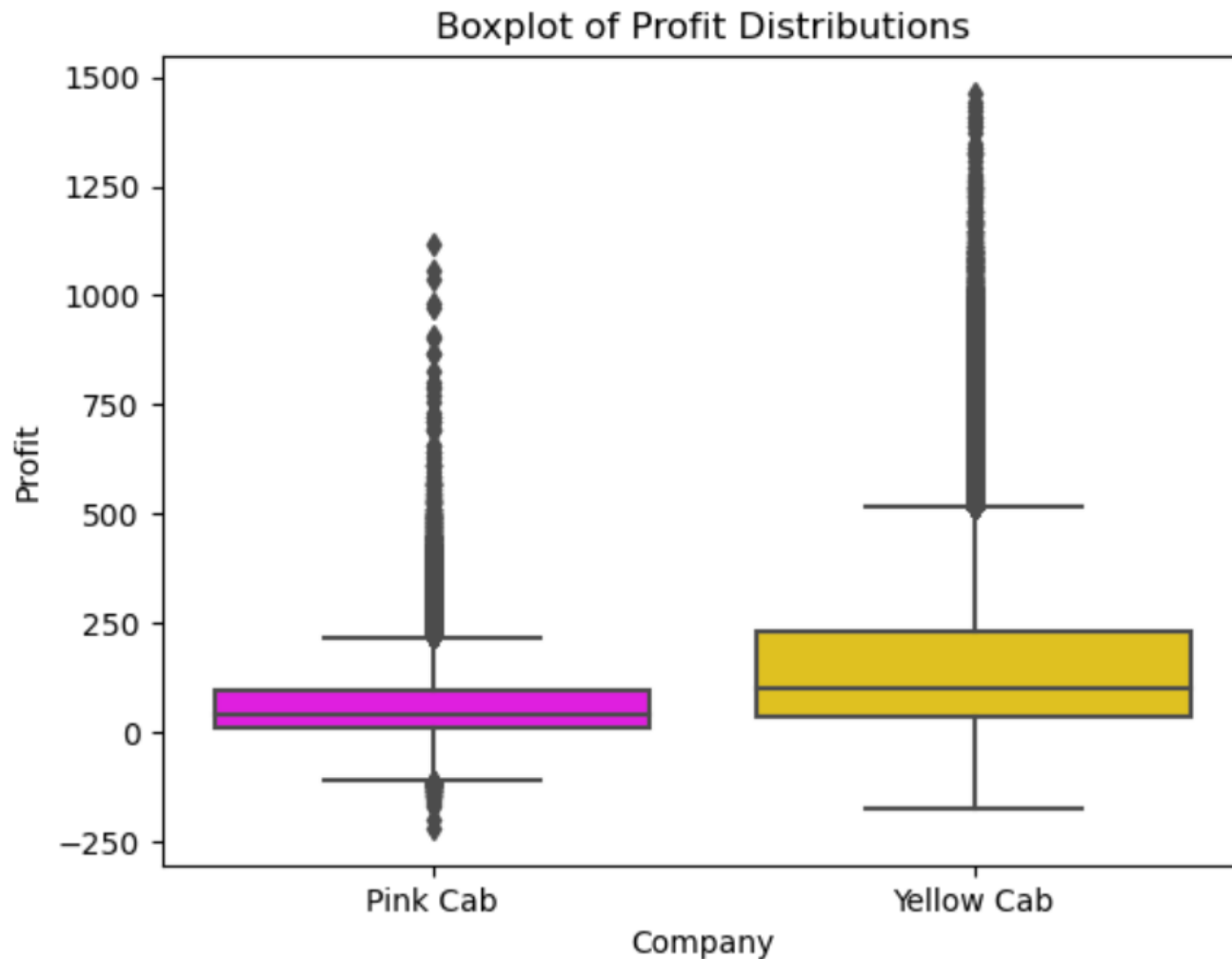
The cab company that performs better across these 3 metrics is the company that XYZ should invest in.

Assumptions

We make two key assumptions while conducting our analysis:

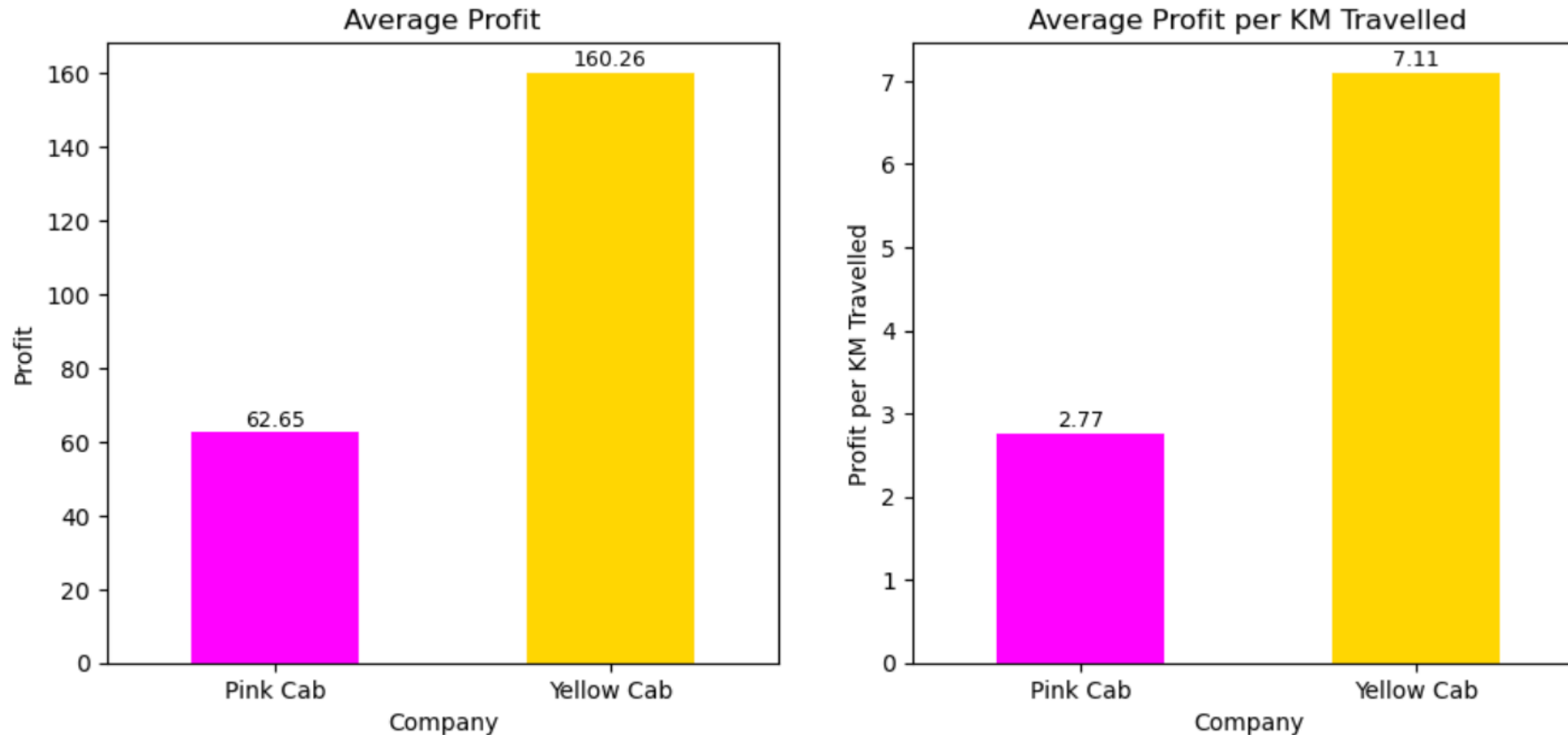
1. Profits can be calculated using only the cost of the trip and the price charged for the trip, no other factors are accounted for when calculating profit.
2. The “Price Charged” feature of the raw cab data (Cab_Data.csv) did appear to have some outliers. However, due to having insufficient evidence to prove that these outliers were mistakes, they were included in the analysis so that the provided data set is fully represented.

Profit Distributions



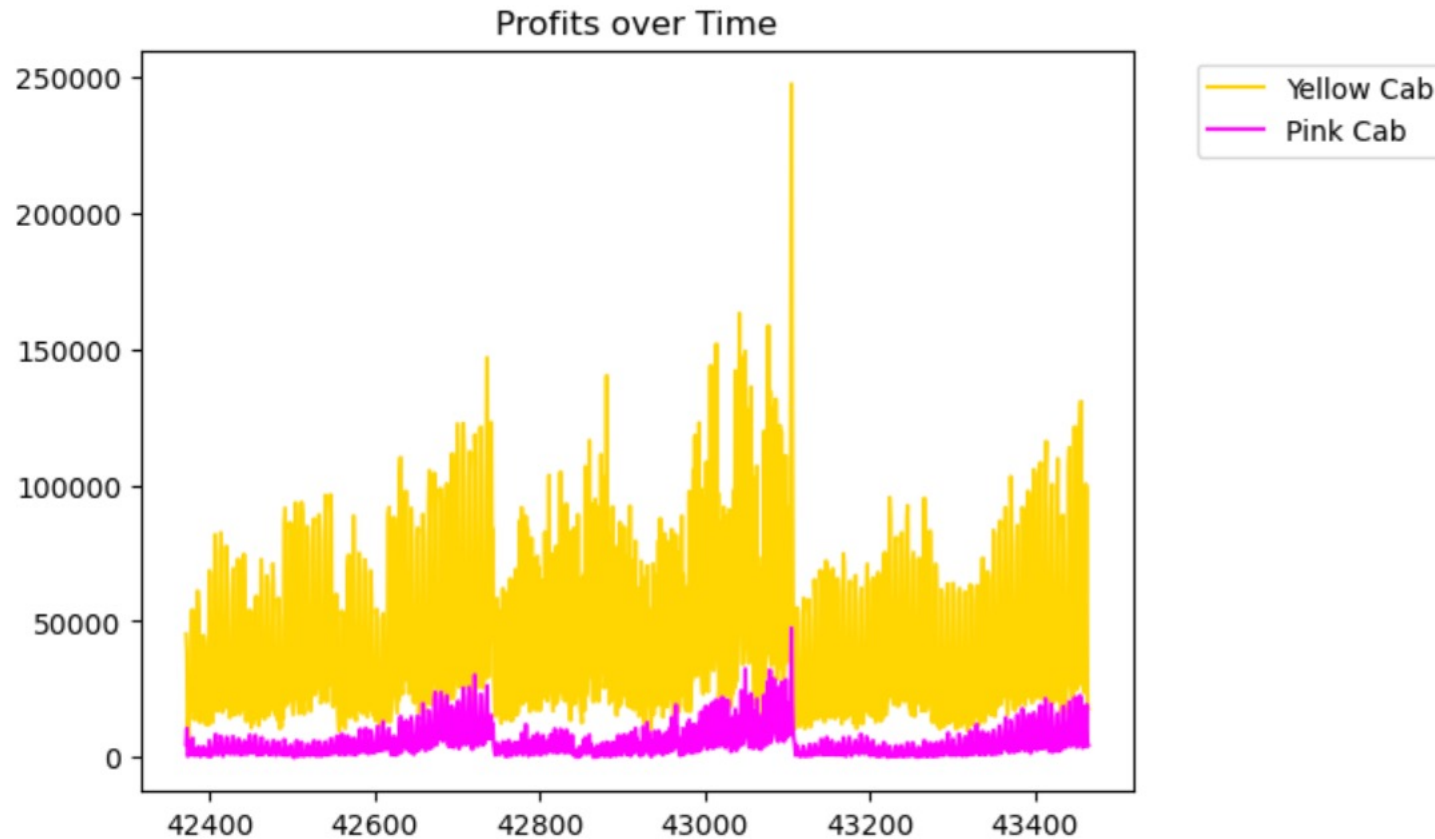
- Yellow Cab has a higher median profit than Pink Cab.
- Judging by the plot, the median profit of Yellow Cab is about the same as the upper quartile profit of Pink Cab.
- This indicates that Yellow Cab is significantly more profitable.

Average Profits



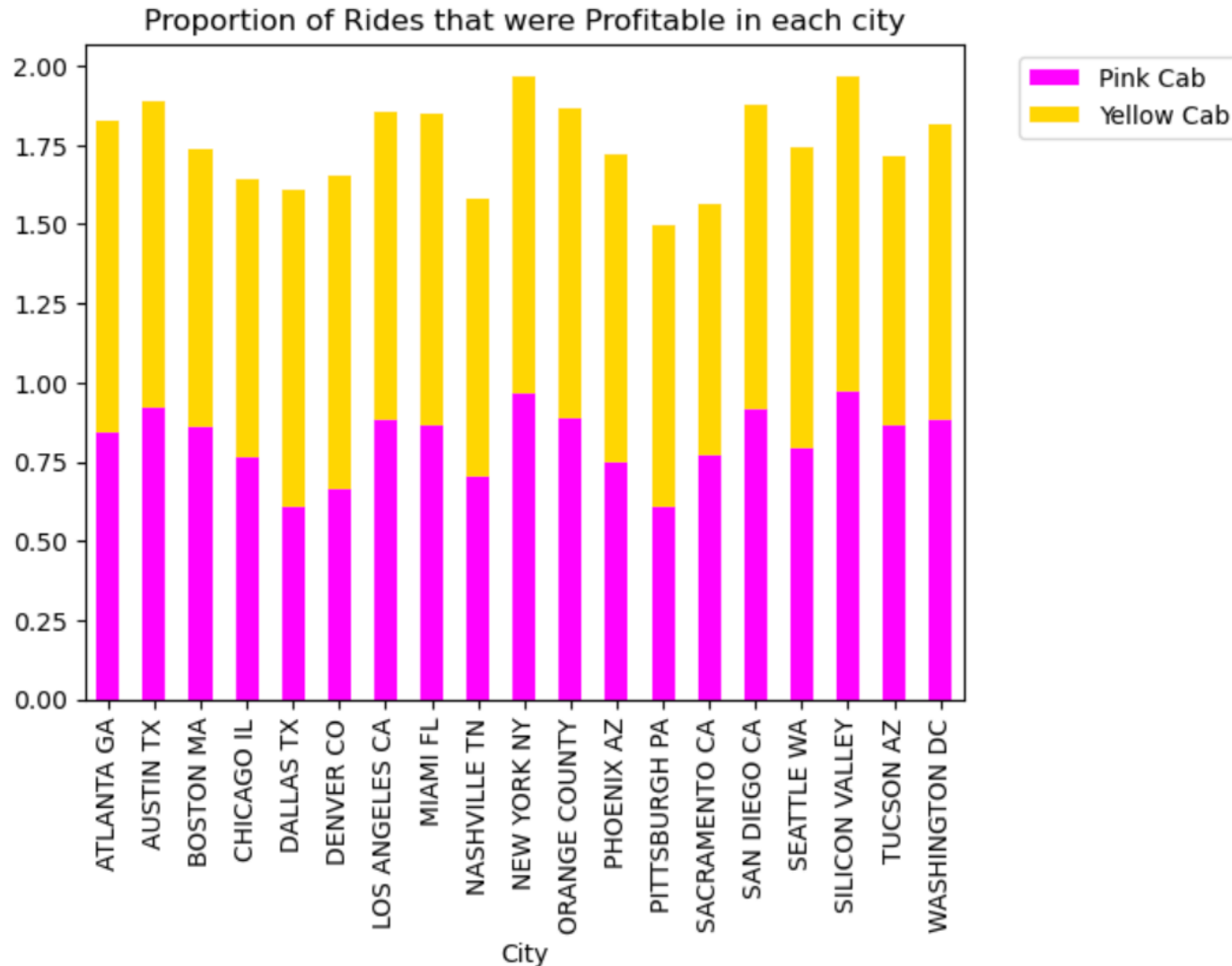
The bar plot shows that Yellow Cab reported about 2.6 times the overall average profit as well as about 2.6 times the average profit per KM travelled reported by Pink Cab.

Profits over Time



- Profits of both cab companies exhibit a similar pattern over time.
- Recently, it appears that profits have not been as high as in the past.
- The pattern is not smooth, which indicates that profits change frequently with time.
- The profits of Yellow Cab are significantly higher than those for Pink Cab for the entirety of the time represented in the data set.

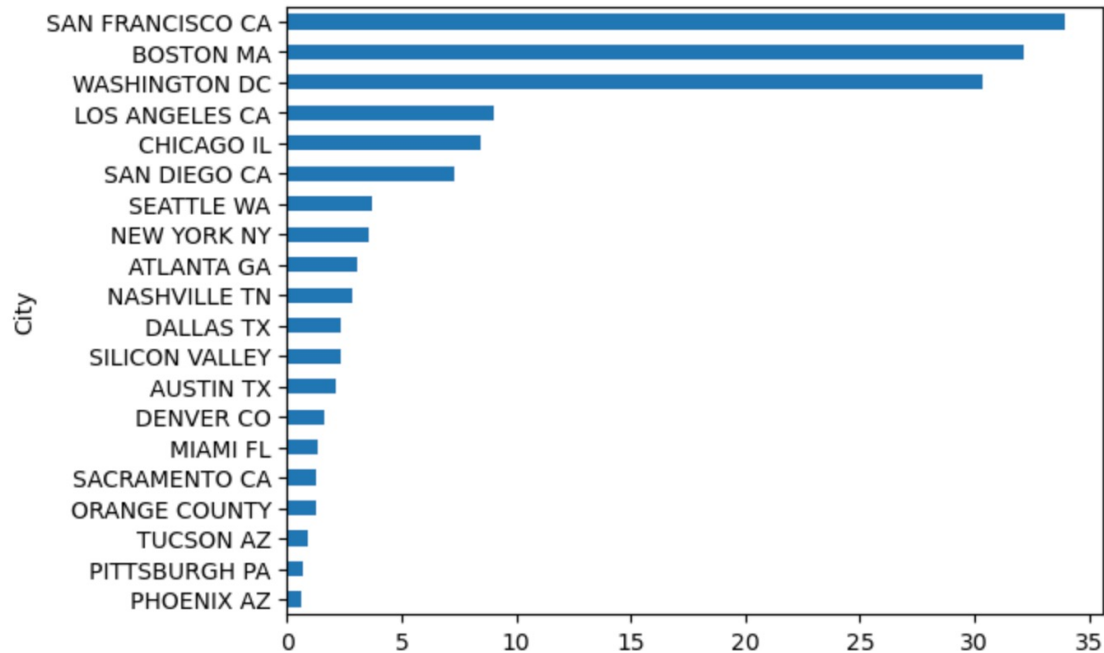
City-wise Profitability of Rides



- The plot shows the fraction of rides that made the companies a profit in each represented city.
- Overall, the pattern for both companies is quite similar, indicating that the city-wise percentage of profitable rides is usually about the same for both Pink Cab and Yellow Cab.

Company Presence in Different Cities

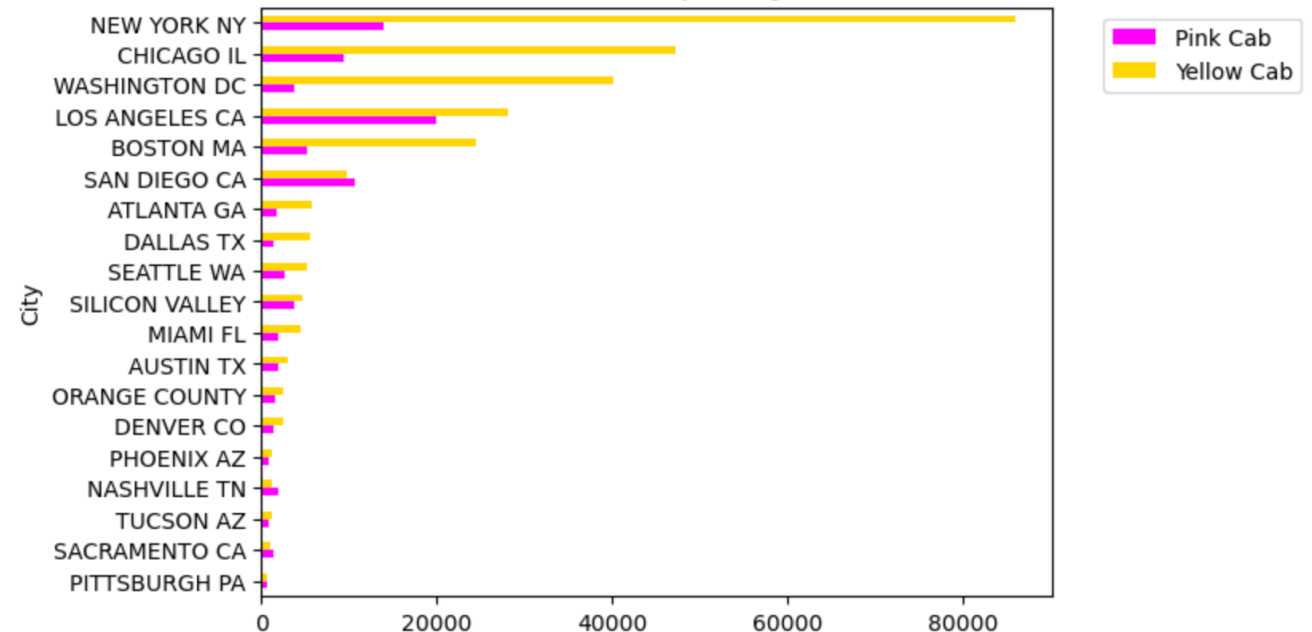
Percentage of City Population that uses Cabs



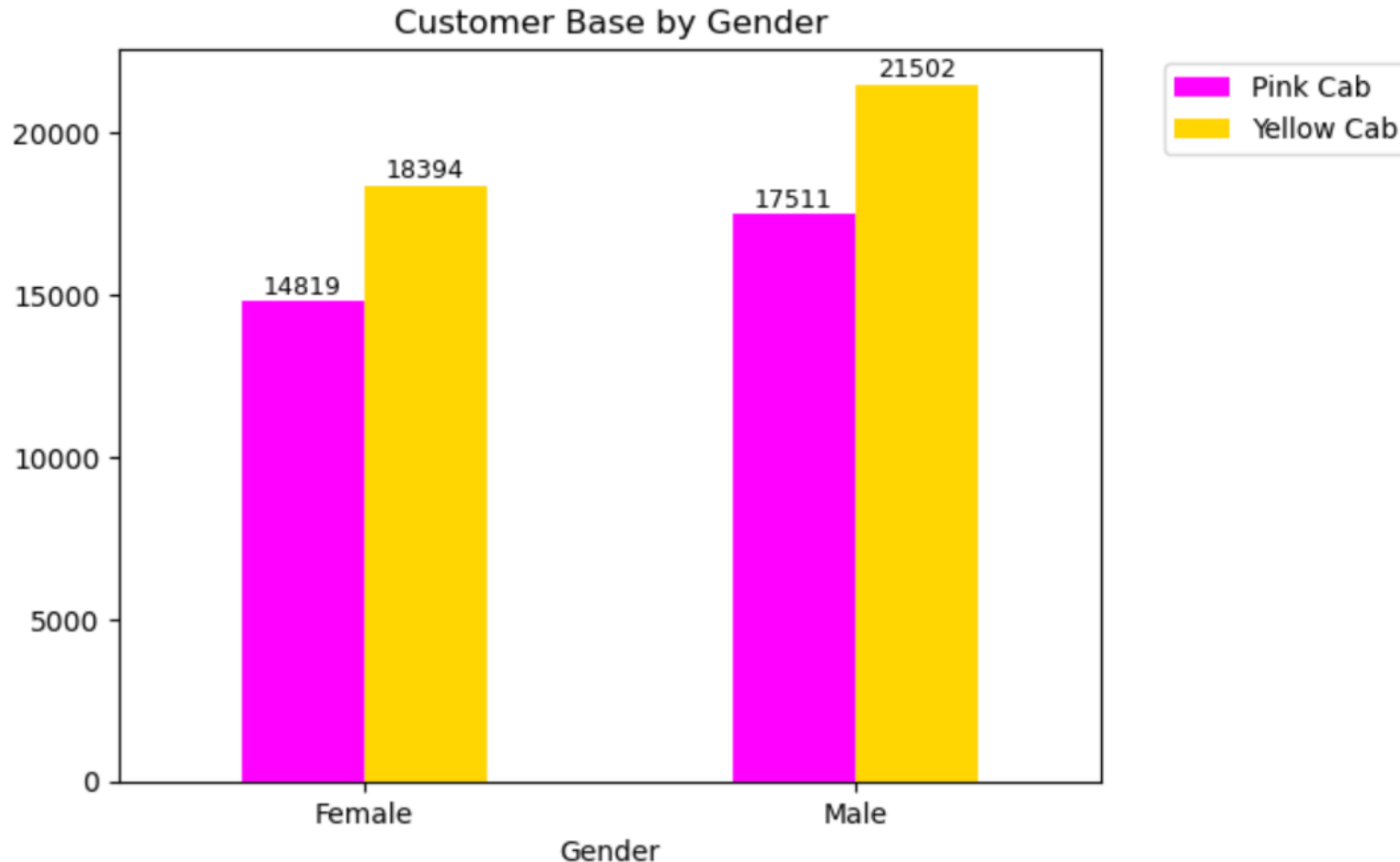
- Yellow Cab serviced more rides in 15 of the 19 represented cities.
- Yellow Cab has a higher presence in most cities.

- Cities in which a large portion of the population are cab users, such as Boston, Washington, LA, and Chicago, are dominated by Yellow Cab.
- In New York, the city with the most cab rides in total, Yellow Cab was the more popular choice by a large margin.

Total Rides per City

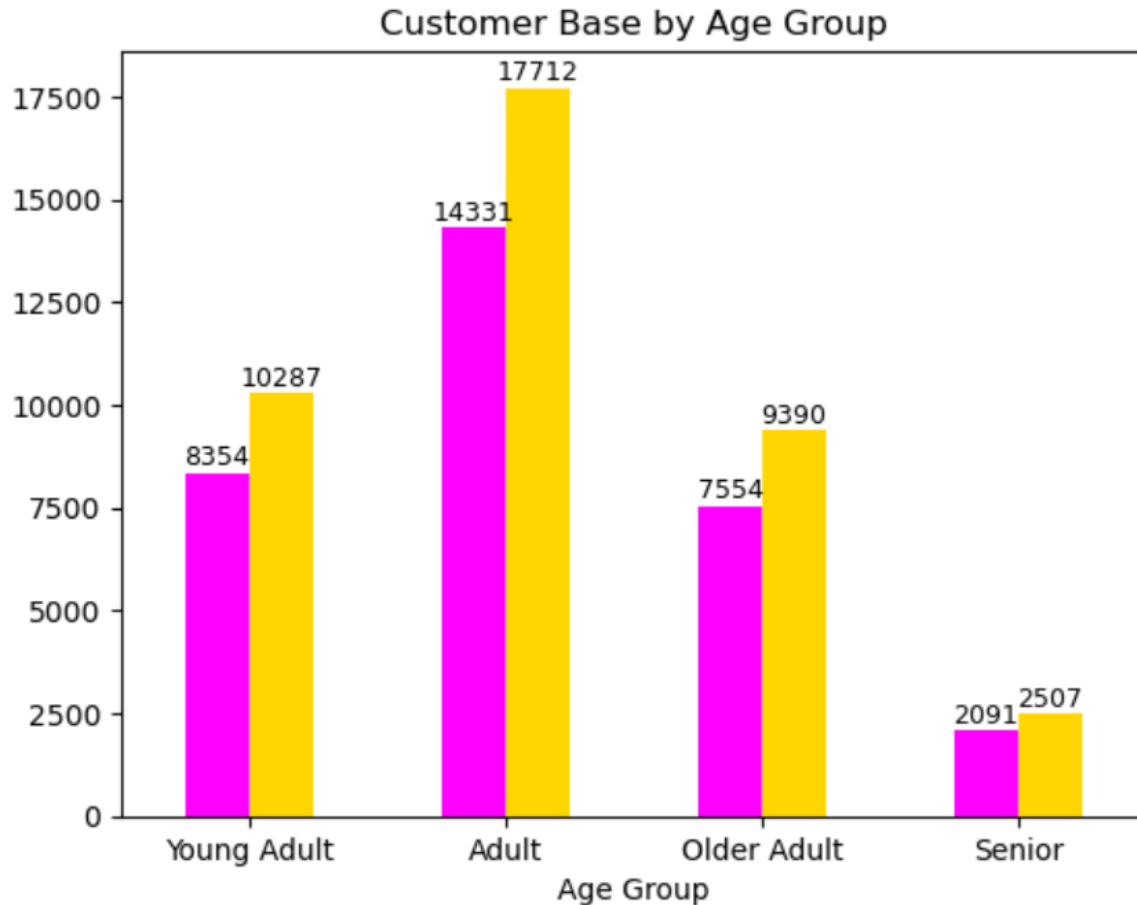


Customer Base by Gender



- Men make up the majority of the customer base for both cab companies.
- Yellow Cab has a larger customer base for both genders.
- Both more men and more women are Yellow Cab customers than Pink Cab customers.

Customer Base by Age Group

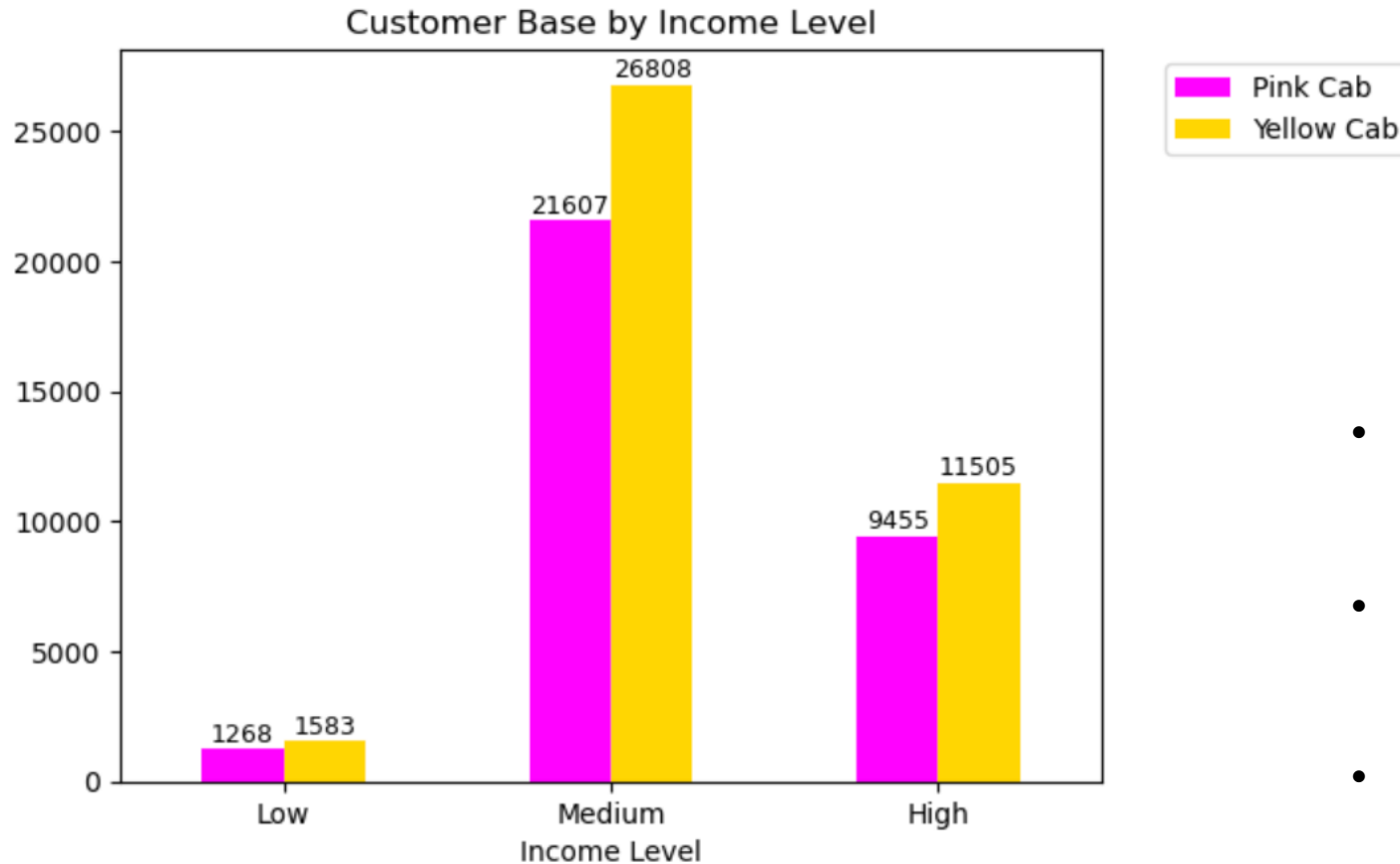


Four categories of age:

- 18 - 25 is "Young Adult"
- 26 - 39 is "Adult"
- 40 - 59 is "Older Adult"
- 60 and above is "Senior"

- The majority of customers for both cab companies are in the "Adult" age group.
- Yellow Cab has a larger customer base for all four age groups.
- Yellow Cab is more popular than Pink Cab with customers of all ages.

Customer Base by Income Level

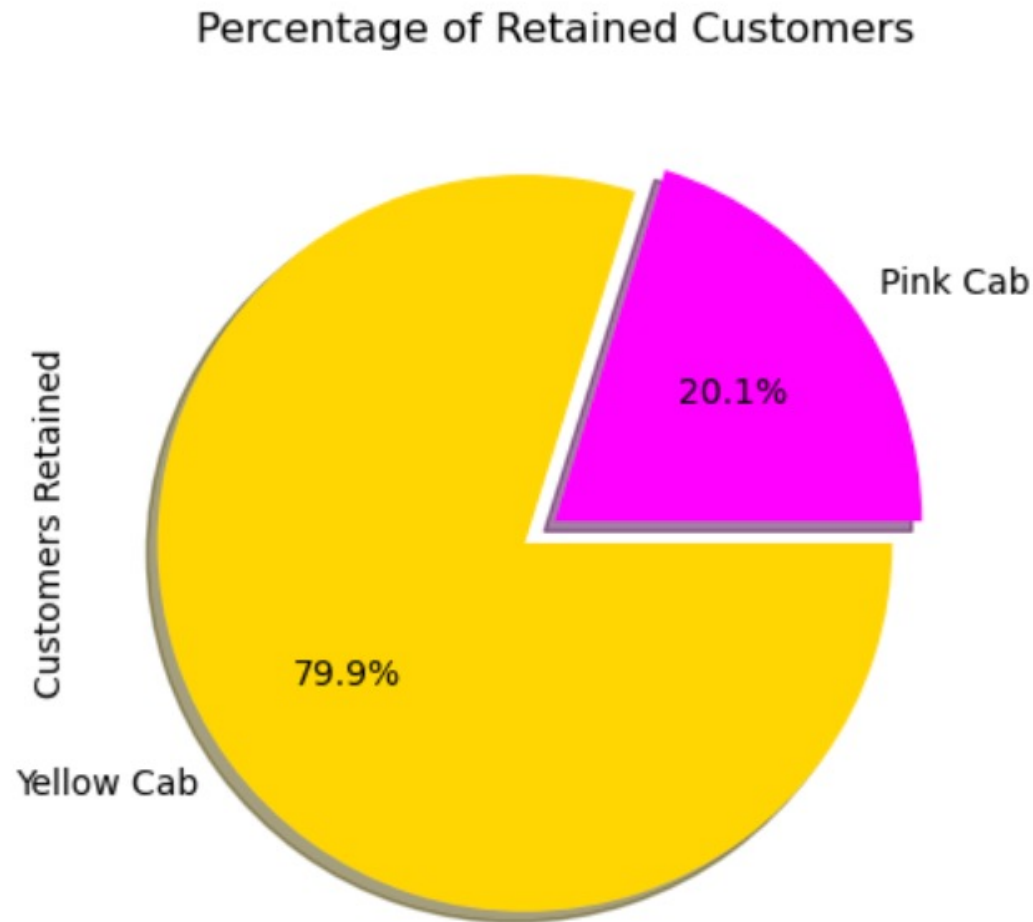


Three levels of income:

- < 3000 USD/Month is "Low"
- 3000 - 20000 USD/Month is "Medium"
- Above 20000 USD/Month is "High"

- The majority of customers for both cab companies are medium income earners.
- Yellow Cab has a larger customer base for all three income levels.
- Yellow Cab is more popular than Pink Cab with customers of all income categories.

Percentage of Retained Customers



Note: we consider a customer as being retained if they enlisted the services of the same cab company more than 5 times in the given time period.

- Yellow Cab exhibits higher customer retention levels than Pink Cab.
- When considering all customers who purchased more than 5 rides from the same cab company, 79.9% of these customers belong to Yellow Cab, whereas only 20.1% belong to Pink Cab.
- Thus, Yellow Cab has a customer retention rate that is approximately 4 times more than the rate for Pink Cab.

EDA Findings Summary

Profit:

- The city-wise percentage of profitable rides is about the same for both companies
- Yellow Cab has a higher overall average profit and average profit per KM travelled (in both cases, about 2.6 times more profitable)
- Yellow Cab was significantly more profitable for the entire time period represented in the data set
- Based on the above points, **Yellow Cab** wins in the profit category

Customer reach:

- Yellow Cab services more rides in most of the cities represented in the data set
- In New York, the city with the most cab rides overall, Yellow Cab was the more popular choice by a large margin
- Yellow Cab has a higher presence in cities where larger portions of the population are cab users
- Yellow Cab has a larger customer base for both genders
- Yellow Cab has a larger customer base for all age groups
- Yellow Cab has a larger customer base for all income levels
- Based on the above points, **Yellow Cab** wins in the customer reach category

Customer retention:

- Yellow Cab exhibits higher customer retention levels than Pink Cab (about 4 times more)
- Based on the above point, **Yellow Cab** wins in the customer retention category

Our Recommendation

Recommendation:

- XYZ should invest in Yellow Cab.

Reasoning:

- Our analysis has shown that Yellow Cab is the winning candidate for all 3 of our chosen evaluation metrics.
- Therefore, the evidence indicates that investing in Yellow Cab will result in a much better ROI than investing in Pink Cab.

Thank You



Data Glacier

Your Deep Learning Partner