Exploratory Data Analysis

Comparative Analysis of Competing Cab Companies Using Transaction, Customer, and City Data

March 20th,2023

Agenda

Case Study Summary

Problem Statement

Approach

EDA

Summary & Recommendations

Case Study Summary

Comparative Analysis was performed using transaction and customer data from two competing cab companies across 19 American cities. The purpose of the analysis is to determine which company is a better investment for XYZ Investment Firm.

The Key Data features explored are Profit Earnings & Customer Engagement.

Each of these features are broken down further to see their dependencies on other factors:

- Profit per Mile Driven
- Profit Earned per Customer
- Profit Earned per City
- Profit Earned per Customer per City
- Seasonality in Profit Earnings by City
- Market Engagement per City
- Unique Riders per Month by City
- Customer Segmentation based on Gender, Age, and Income

What was found is that the single biggest factor in Profit Earnings is the market engagement by city. The larger the customer base in specific US cities, the greater predicted Profit Earnings for that company.

Problem Approach

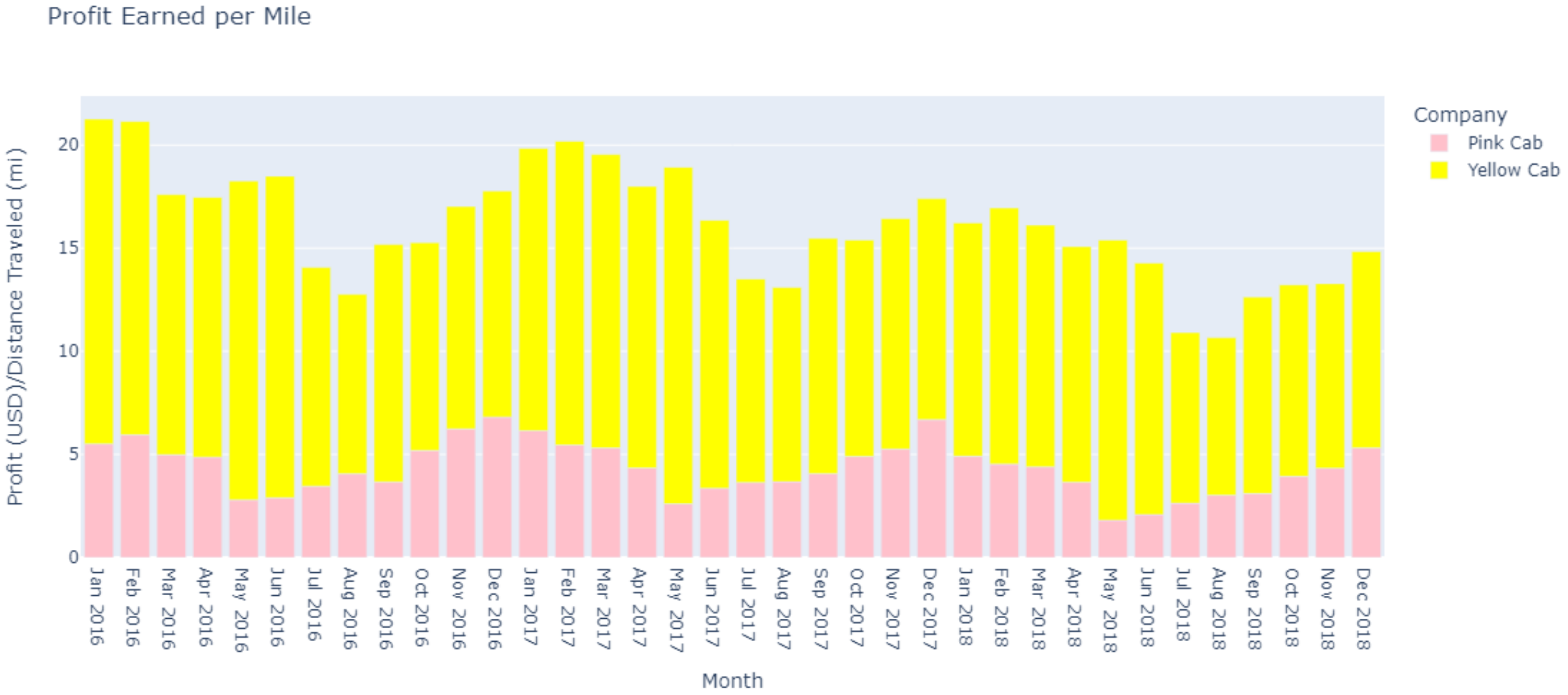
Comparative Analysis was performed using transaction and customer data from two competing cab companies across 19 American cities. The purpose of the analysis is to determine which company is a better investment for XYZ Investment Firm.

The Key Data features explored are Profit Earnings & Customer Engagement.

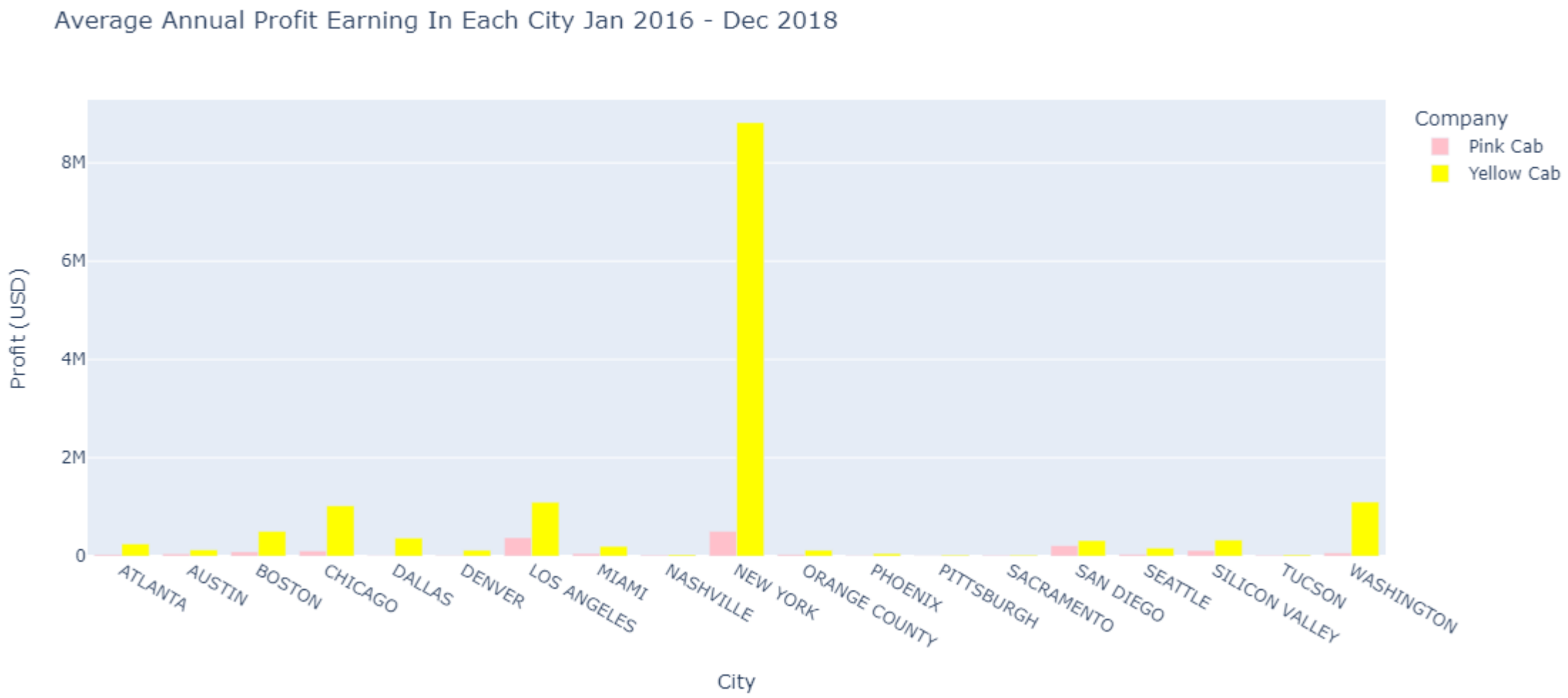
Each of these features are broken down further to see their dependencies on other factors:

- Profit per Mile Driven
- Profit Earned per Customer
- Profit Earned per City
- Profit Earned per Customer per City
- Seasonality in Profit Earnings by City
- Market Engagement per City
- Unique Riders per Month by City
- Customer Segmentation based on Gender, Age, and Income

What was found is that the single biggest factor in Profit Earnings is the market engagement by city. The larger the customer base in specific US cities, the greater predicted Profit Earnings for that company.

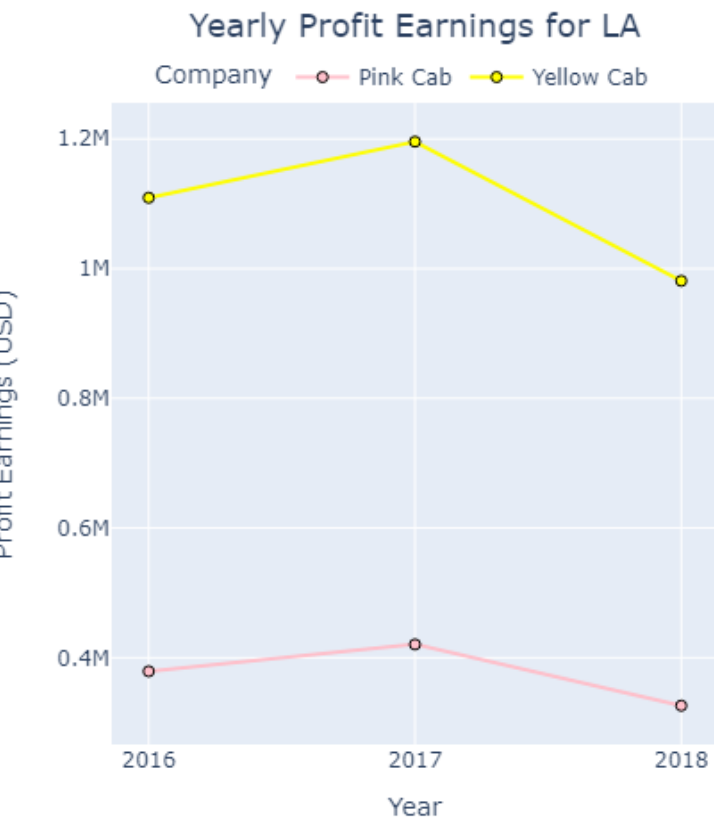


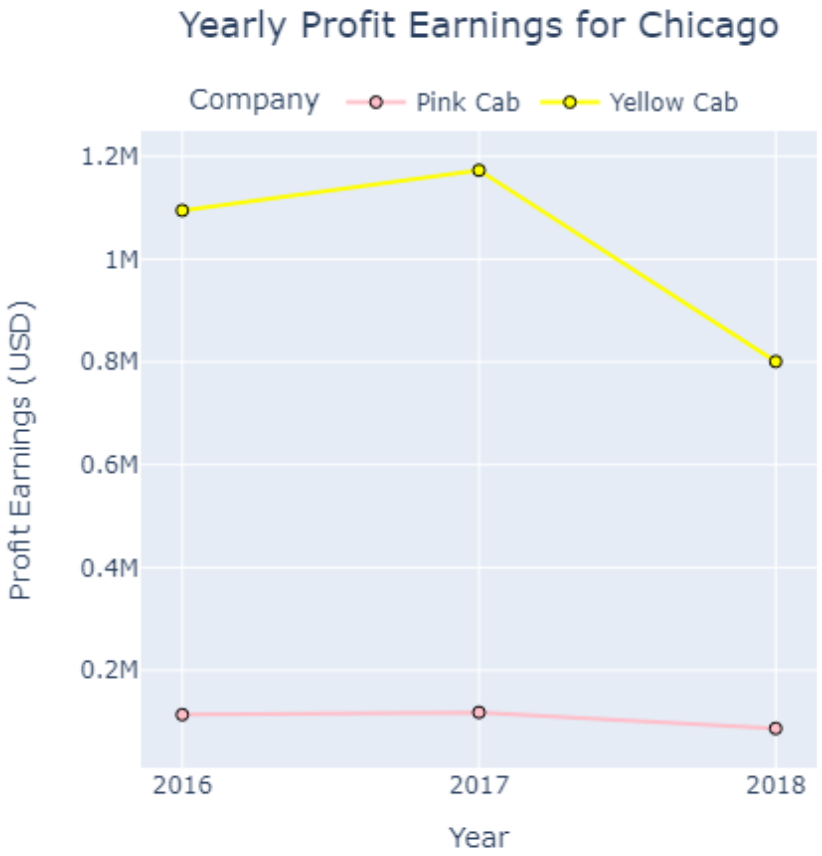
The is a significant maintenance costs associated with keeping a fleet of functioning taxicabs; the more miles driven, the higher the operational costs. Maximizing the profit earnings per miles helps offset this inherent operational cost. The data below suggests that Yellow Cab is significantly and consistently out-performing the Pink Cab Company in this metric. Since the data is controlled for location, and by extension fuel prices, the assumption can be made the extra cost is due to the vehicles themselves. Pink Cab Company may have a higher proportion of older or more fuel inefficient vehicles in their fleet.

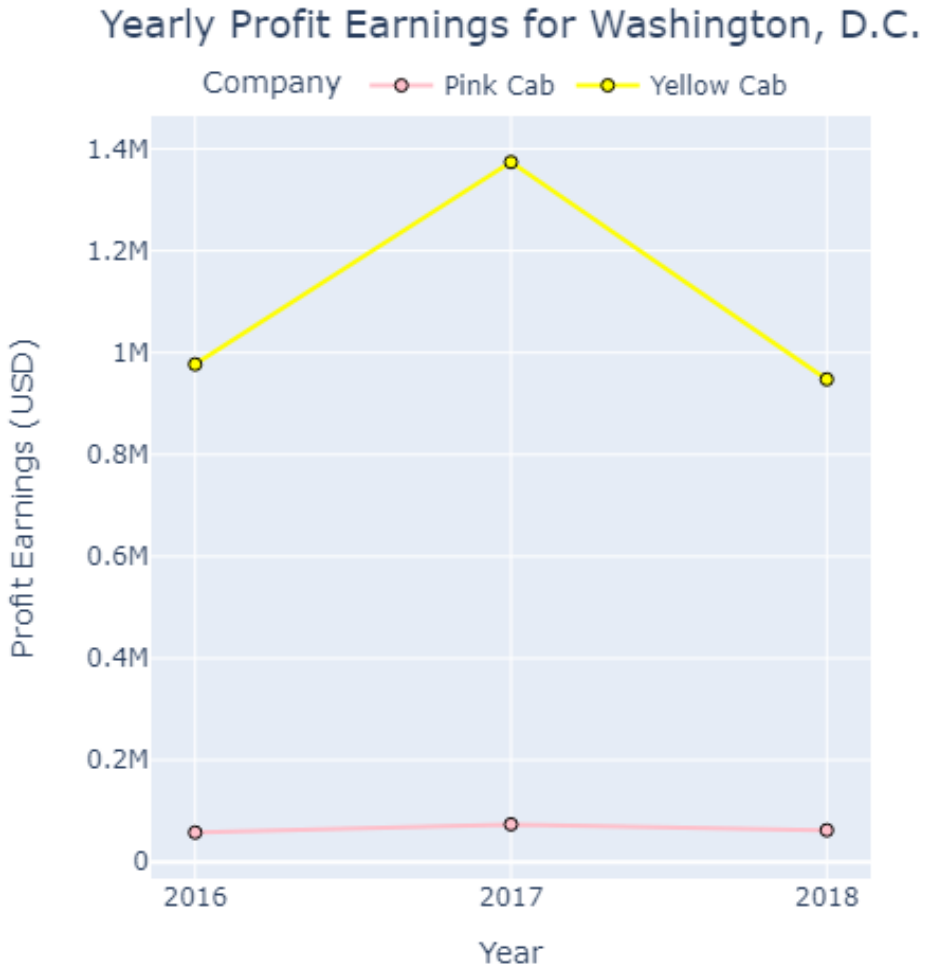


The profit earning per city illustrates very clearly the drastic difference in the cab market across the US. New York City is such a marked outlier compared to all other US cities that it has generated more profit than all the other markets combined. There is a larger profit margin per ride within NYC, but more importantly, the potential and actual customer base is significantly larger there than in any other US city.

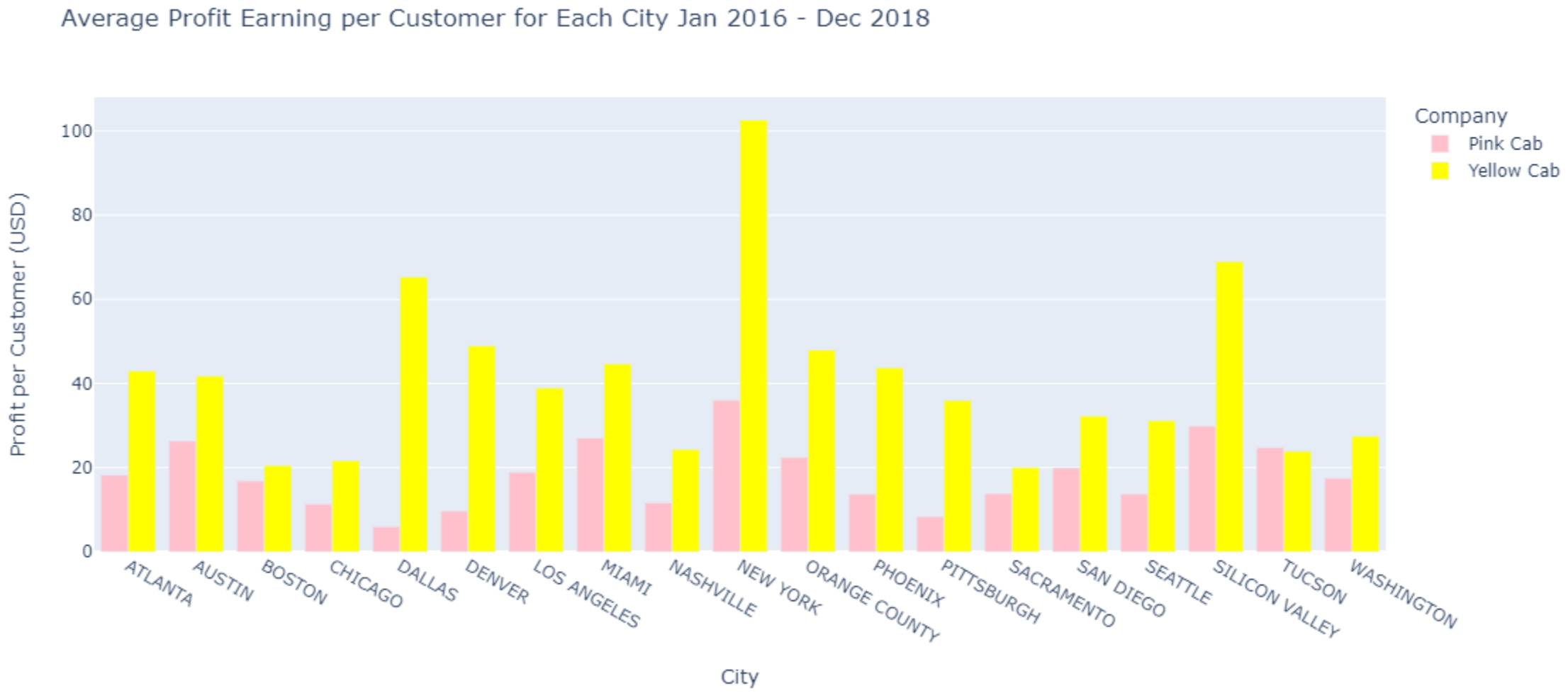
Given the importance of the New York cab market, it is significant to note that the Yellow Cab company is out-performing the Pink Cab Company by a large margin.



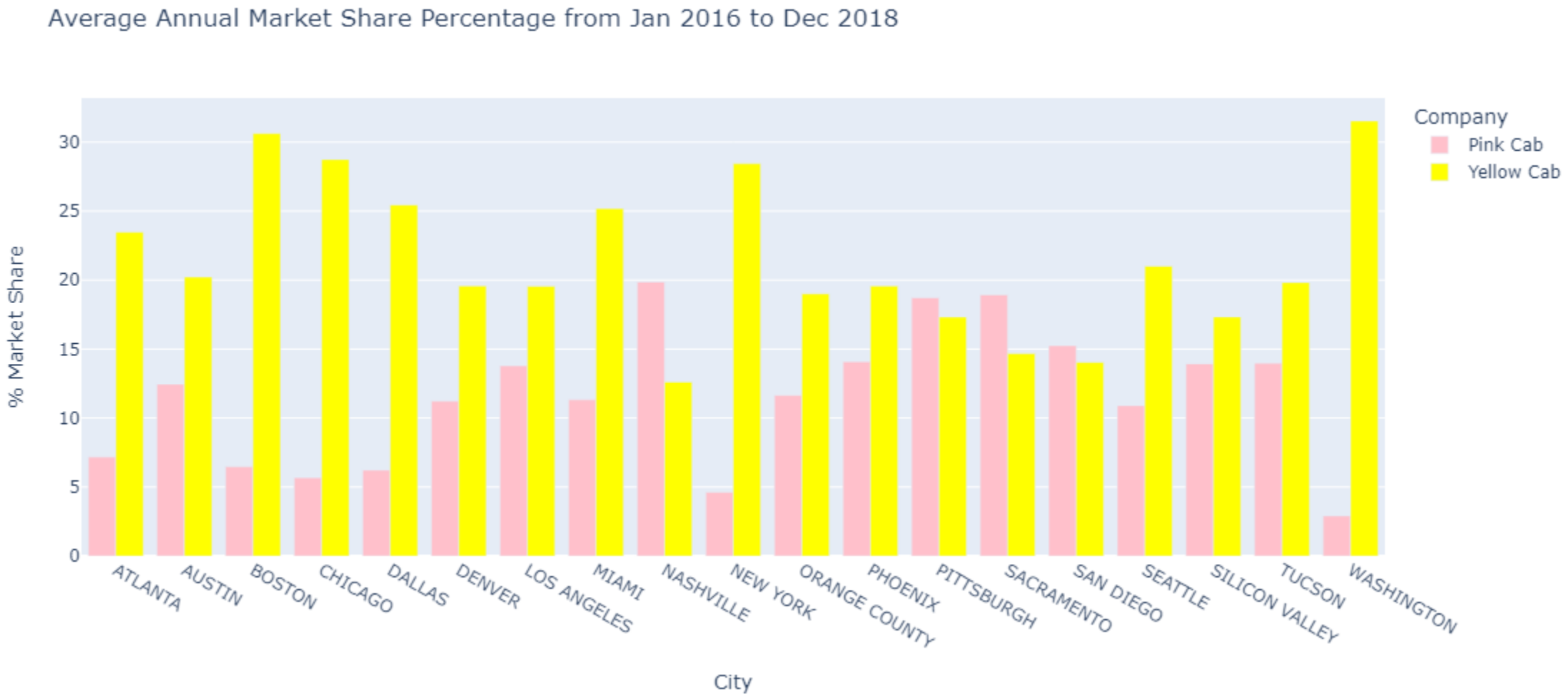




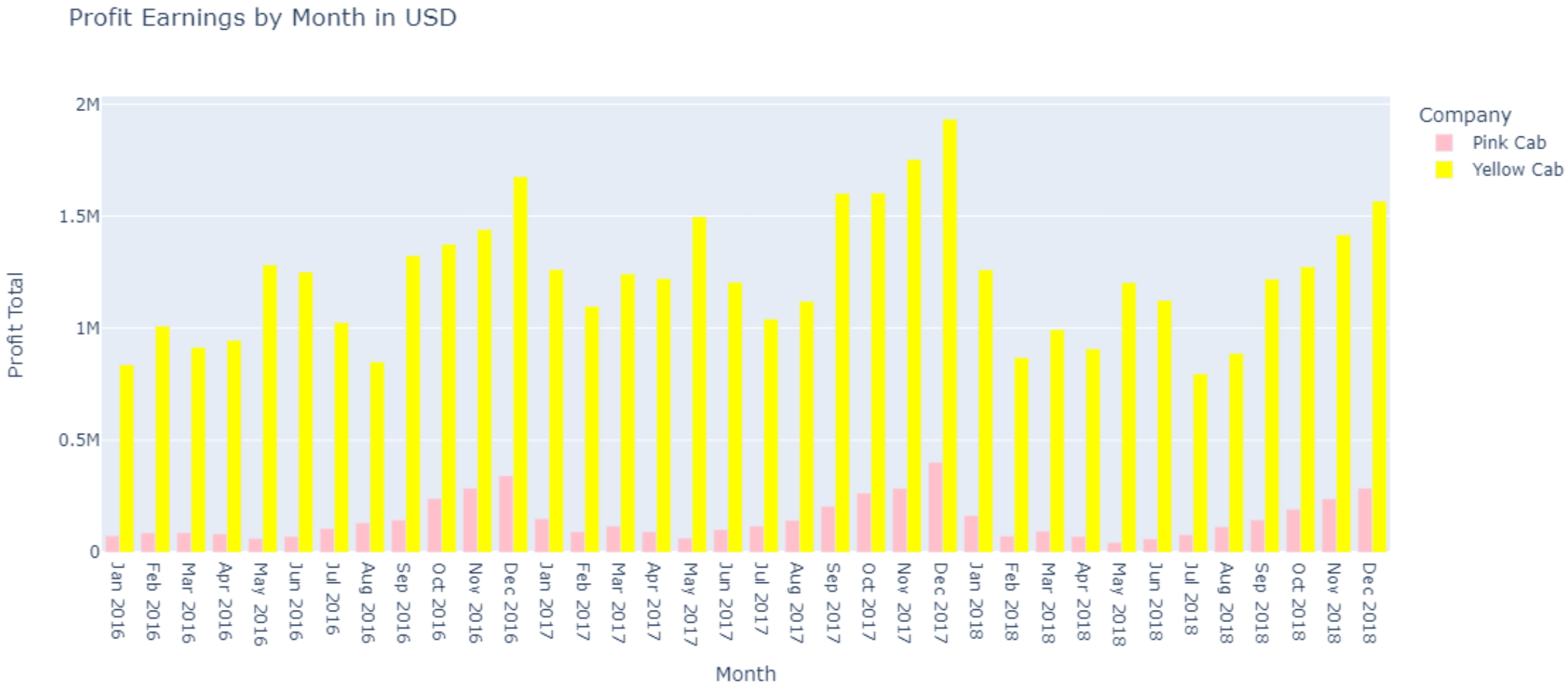
And in the smaller, but still significant markets, The Yellow Cab Company is pulling in much higher earnings.



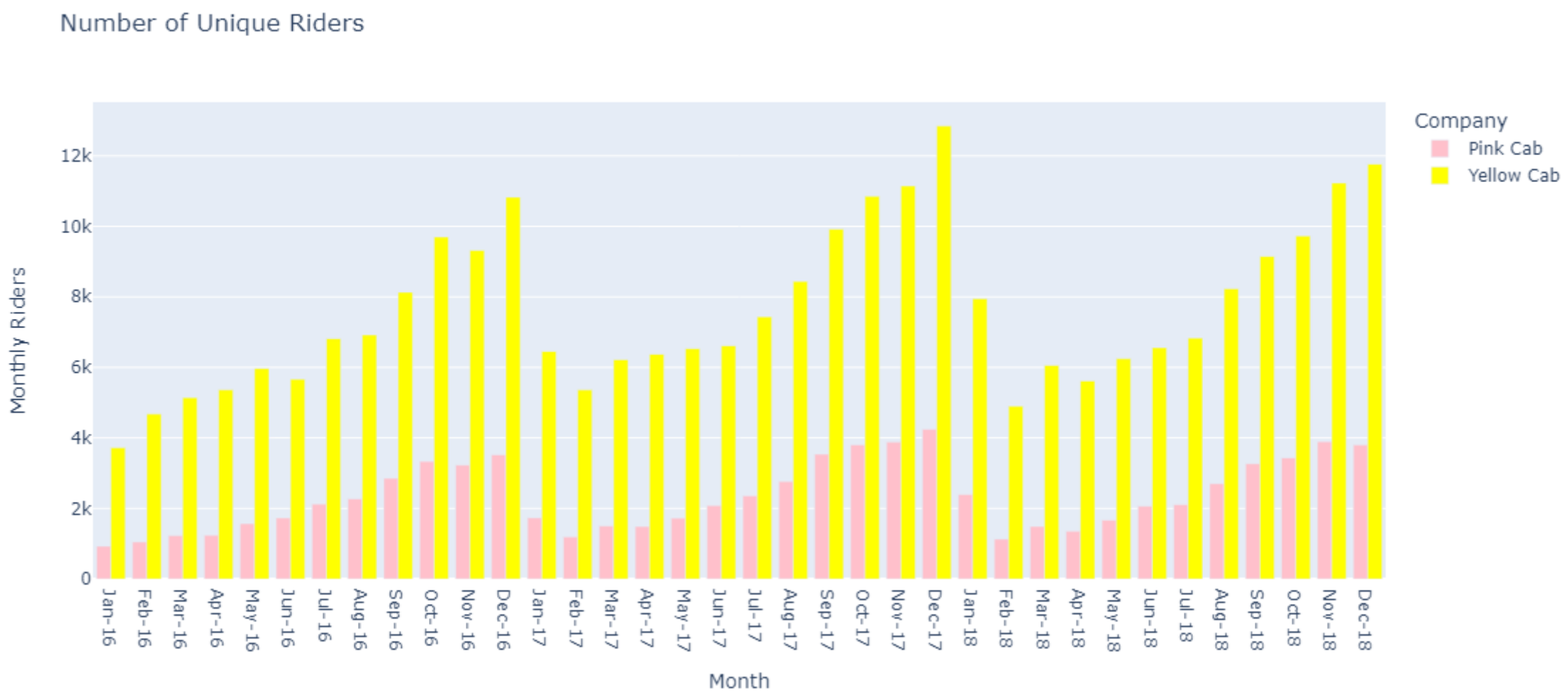
When we adjust the earnings by customer, we can see that New York still stands as an outlier. There are a few other US cities where the profit earnings per customer are significant, however the customer base is quite small. For those cities that are experiencing population and job growth, we can expect to see the customer base grow over time. These markets may be good potential investments in the coming years (such as in Dallas, TX). The Yellow Cab company has greater earnings per customer across all cities measured.



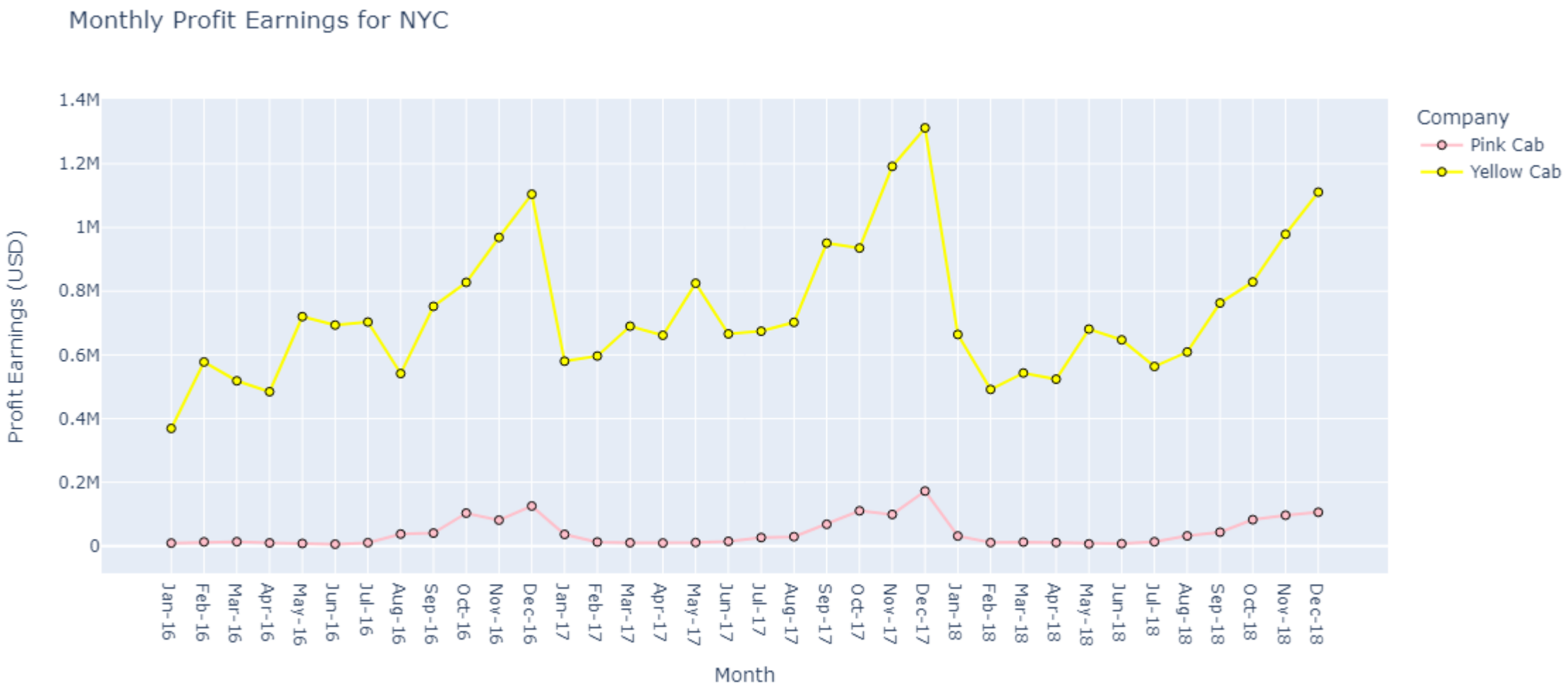
In the 4 most profitable markets (NYC, Chicago, LA, and Washington, D.C.), the Yellow Cab Company has deeper market penetration. The Pink Cab Company only has a larger market share in 4 of the 19 cities measured.



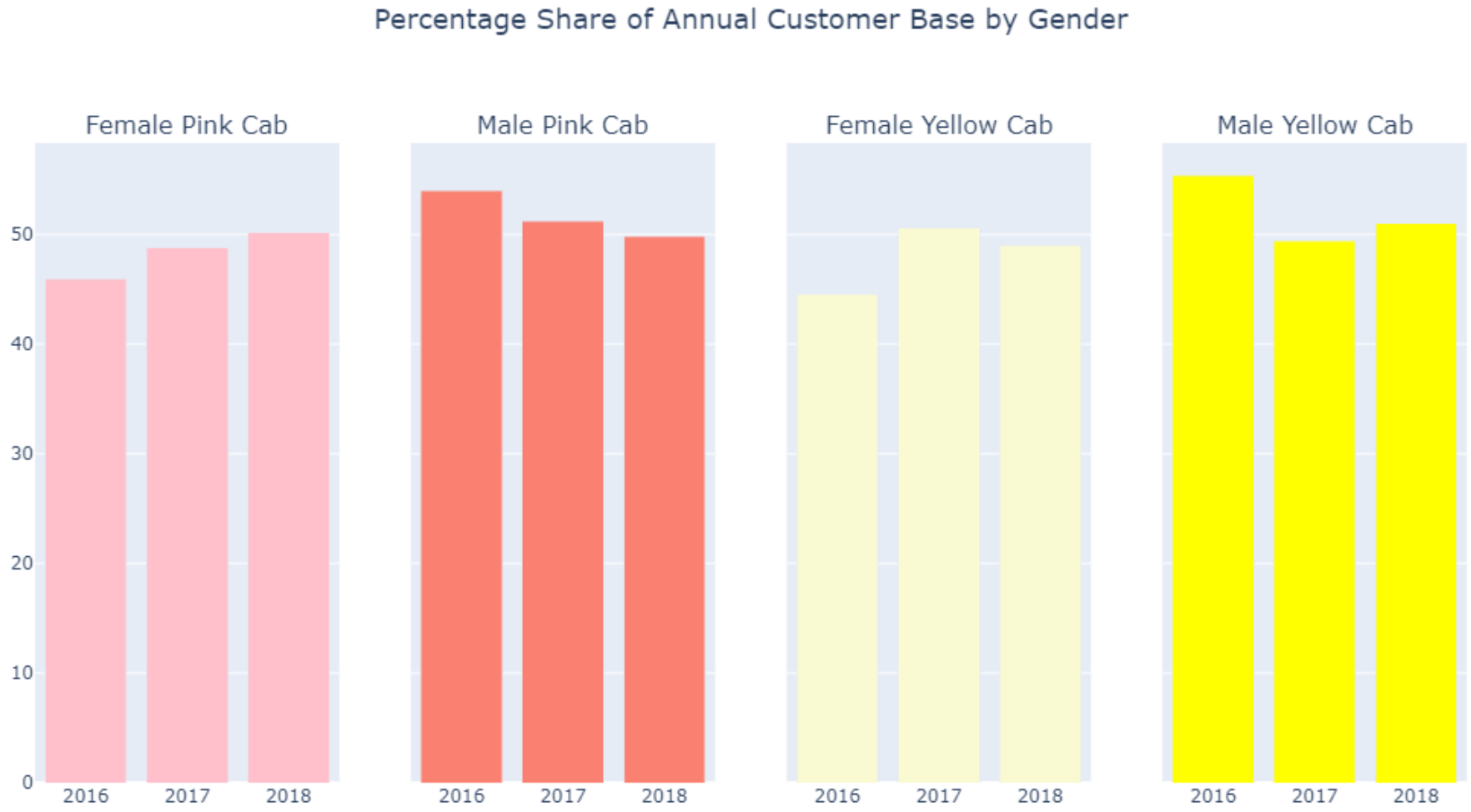
When we look at earnings across months, we can see that there is a seasonal earning cycle in line with the US travel seasons. The highest peak occurs during the US "Holiday Season" of Nov – Dec and smaller, but still significant peaks occur in line with the Memorial Day Weekend (Late May/ Early June) and Labor Day Weekends (Late Sep/ Early Oct). These are expected increases for any service provider in the travel industry within the US and travel volume increases during these times annually.



More interestingly than overall increase in volume is the increase in unique riders per month during these seasonal cycles. This data reflects engagement with entirely new customers, which implies a higher chance of retaining those customers long term.



In the New York Market in particular, we see a huge influx of riders during the holiday season.



Customer Segmentation across gender shows there is a general trend towards parity and a greater engagement with female customers over the three years measured.

Summary and Recommendations

It has been demonstrated clearly that the Yellow Cab company is better positioned across larger and smaller markets within the US. They have deeper market penetration and greater profit margins per mile across the board. It is our recommendation that investment should be focused in the Yellow Cab company.

In terms of where potential investments should be focused, the most profitable market, by far, is New York City, where small increases in market share are compounded both by the over all larger market available, and wider profit-per-customer margin.

Thank You