



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M – Cab Investment Firm

13/11/2022

Agenda

Problem Statement

Approach

EDA

Recommendations

Problem Statement and Approach

- 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.
- Objective: XYZ is interested in using our actionable insights to help them identify the right company to make their investment.

The analysis has been divided into 3 parts:

- Data understanding
- A comparison between the Pink and Yellow companies; profit, customer and user wise.
- Conclusions and Recommendations.

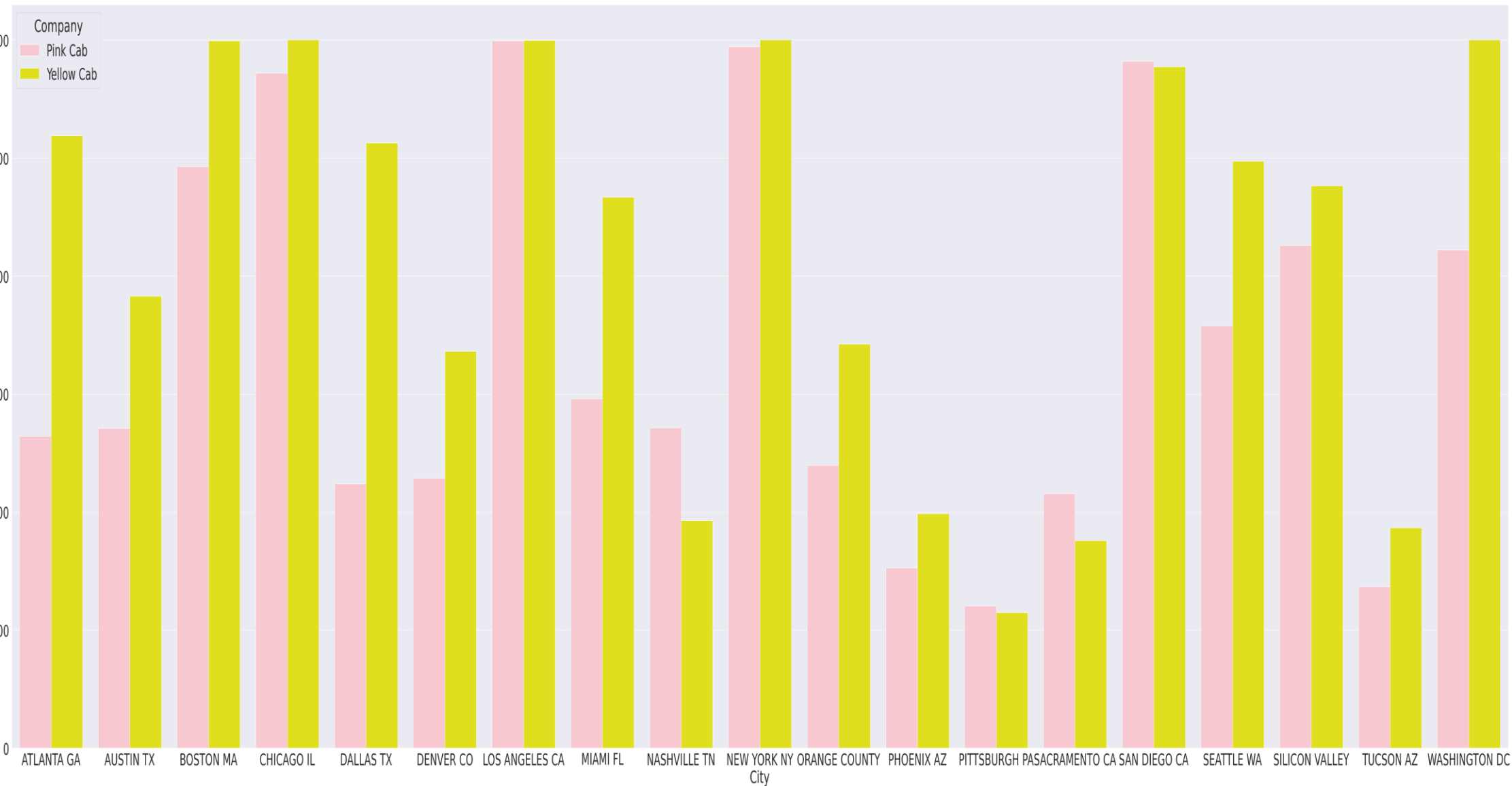
Data Exploration

- 19 Features (including 7 derived features)
- Timeframe of the data: 2016-01-02 to 2018-12-31
- Total data points: 359392

Assumptions:

- Profit is calculated by subtracting the Cost of Trip from the Price Charged.
- Outliers in the Price Charged feature are not removed, since we don't have any proof of being incorrect data.
- The Population and Users features remain the same for a period of 3 years. This is unrealistic. Also, the number of unique customers does not match the number of users based on the Users column. Hence, we drop these two columns.

Number of Users in each City

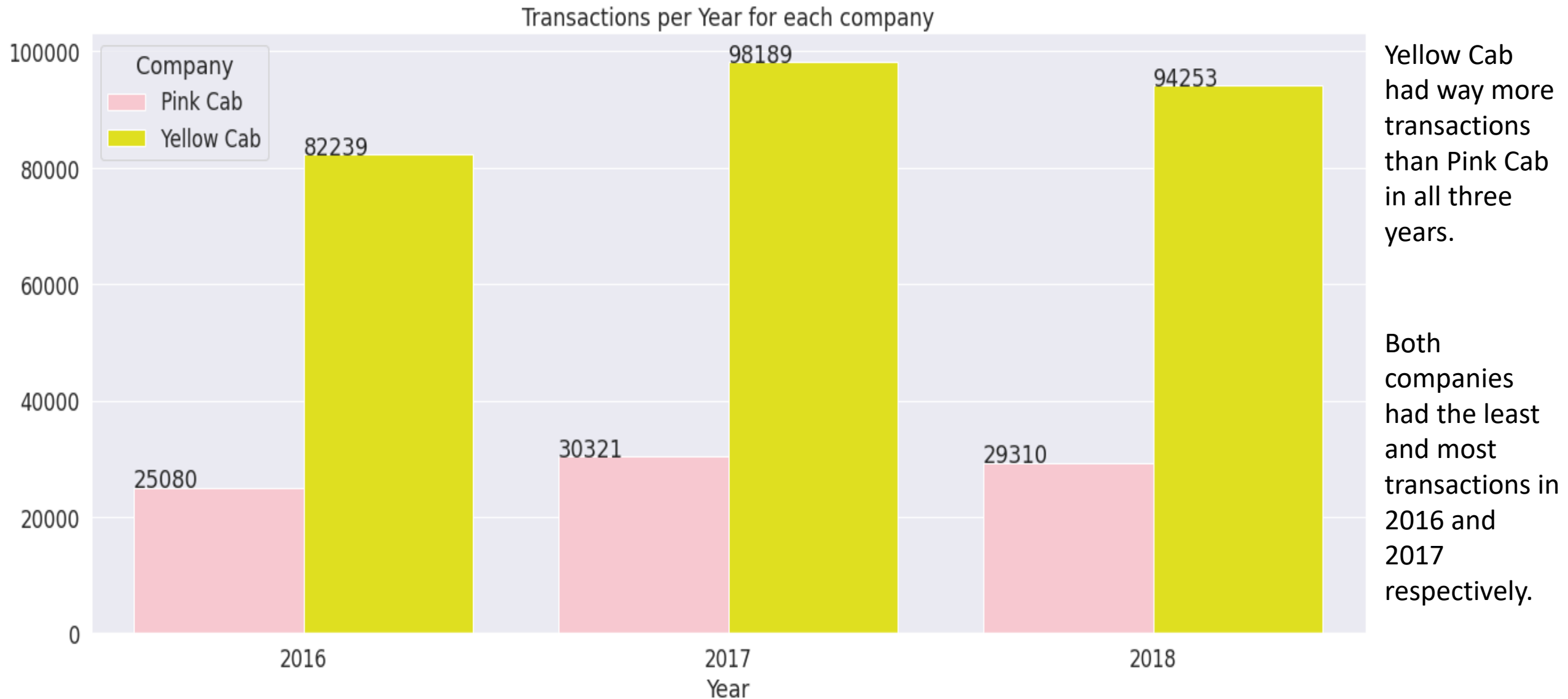


In most cities the number of users of the Yellow Cab company is more than that of the Pink Cab company.

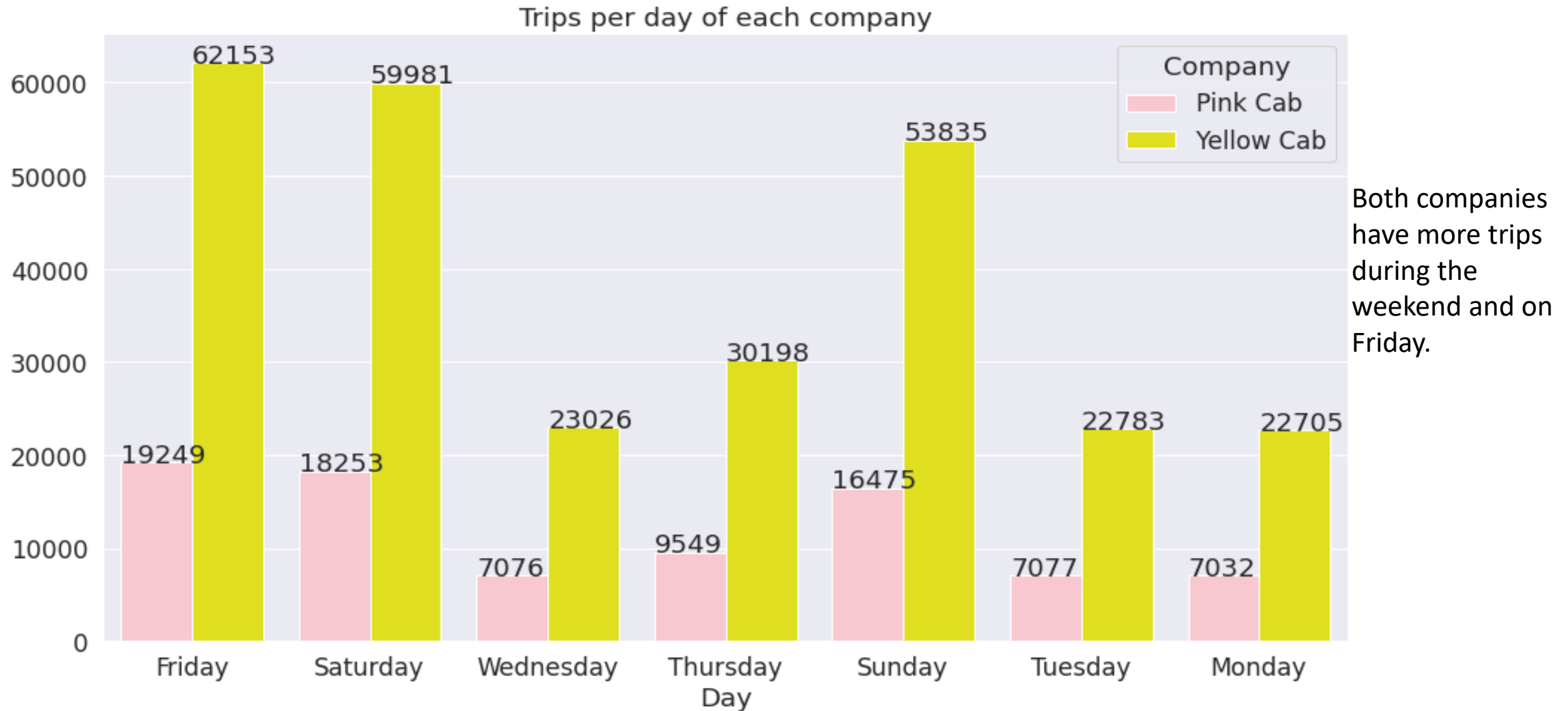
Only in NASHVILLE TN, SACRAMENTO CA, and SAN DIEGO CA the number of users of the Pink Cab company is more than that of the Yellow Cab company.

The companies have almost the same number of users in LOS ANGELES CA and NEW YORK NY.

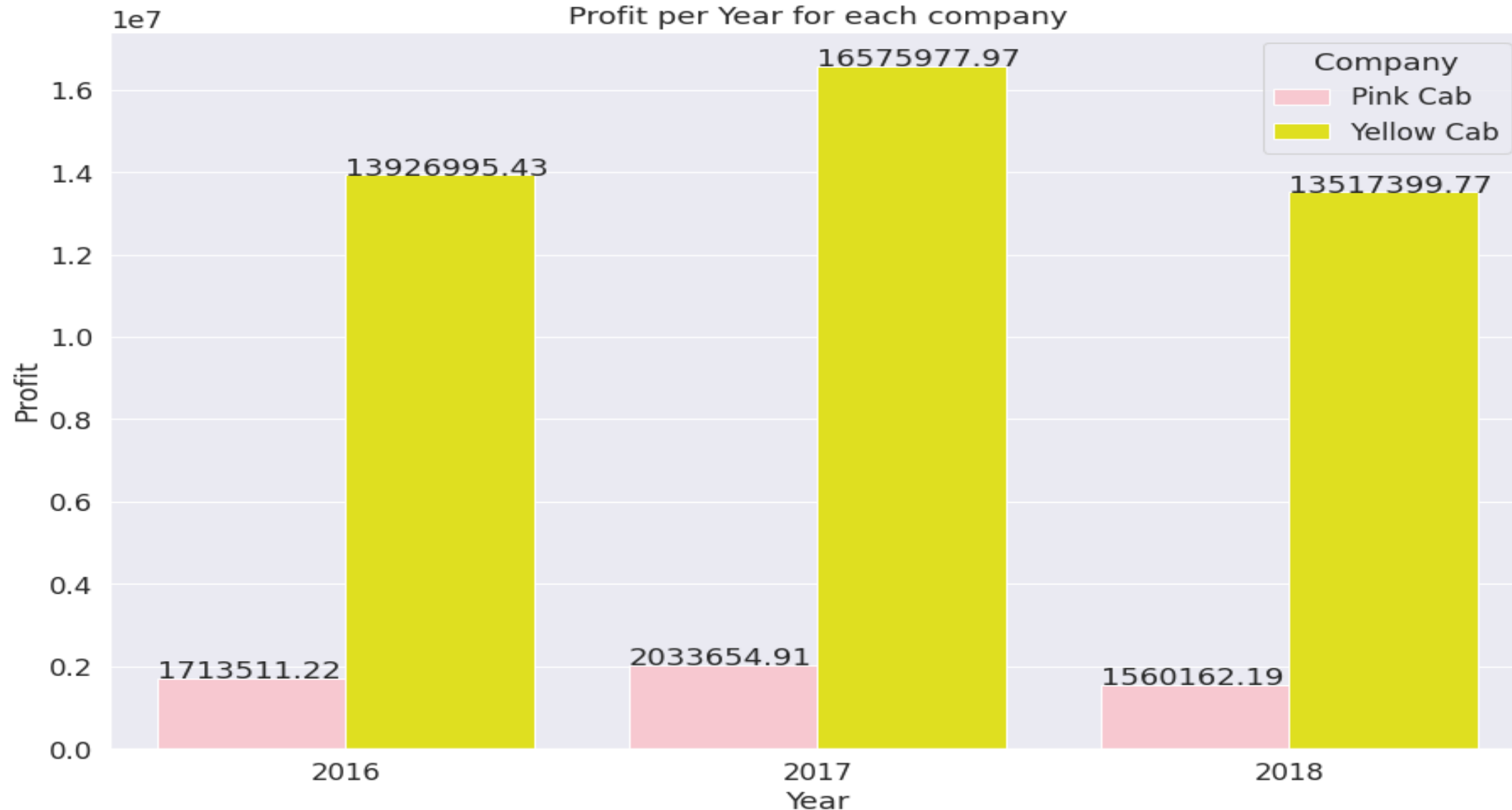
Yearly Transactions



Trips per day



Yearly Profit

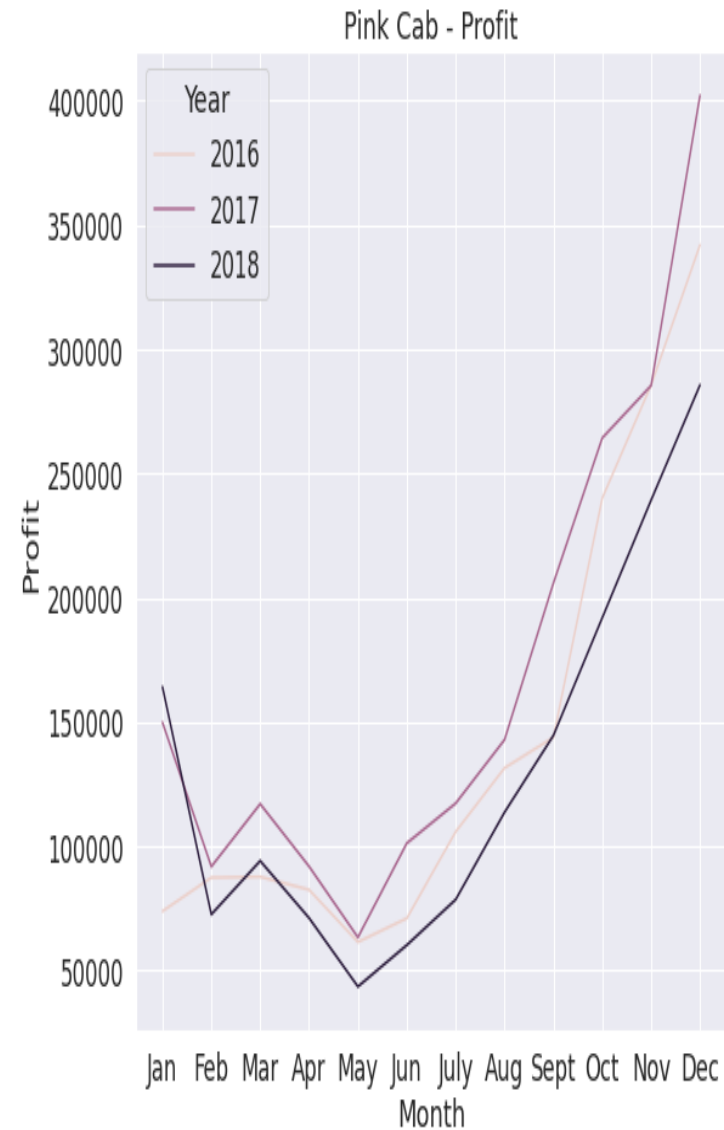
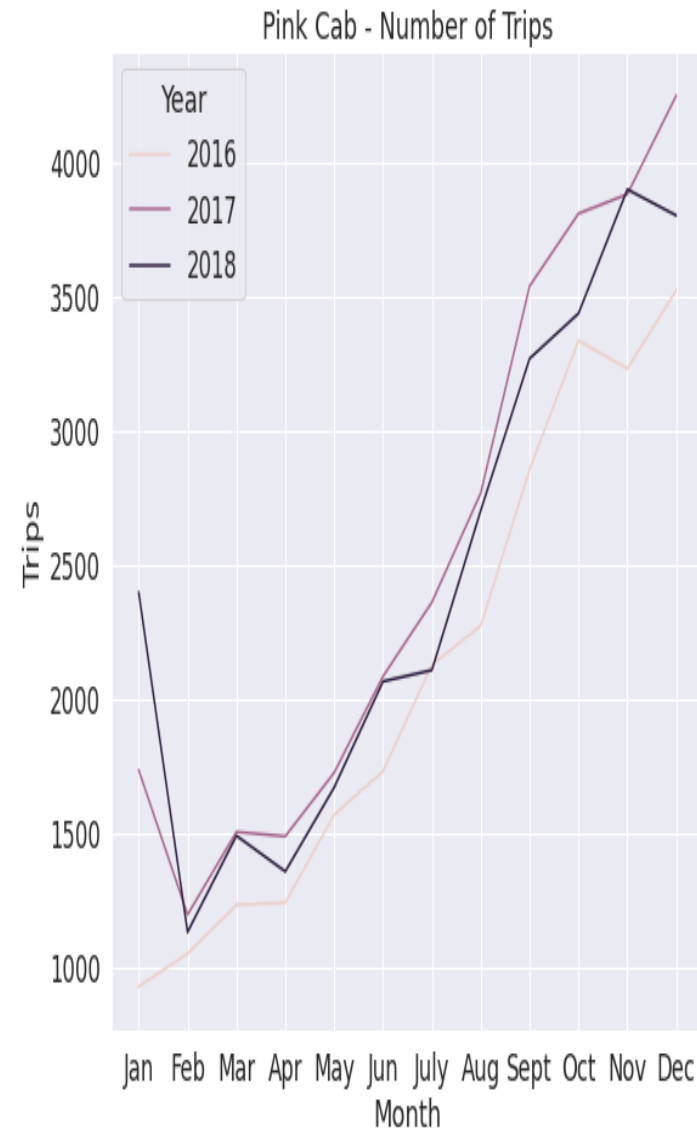
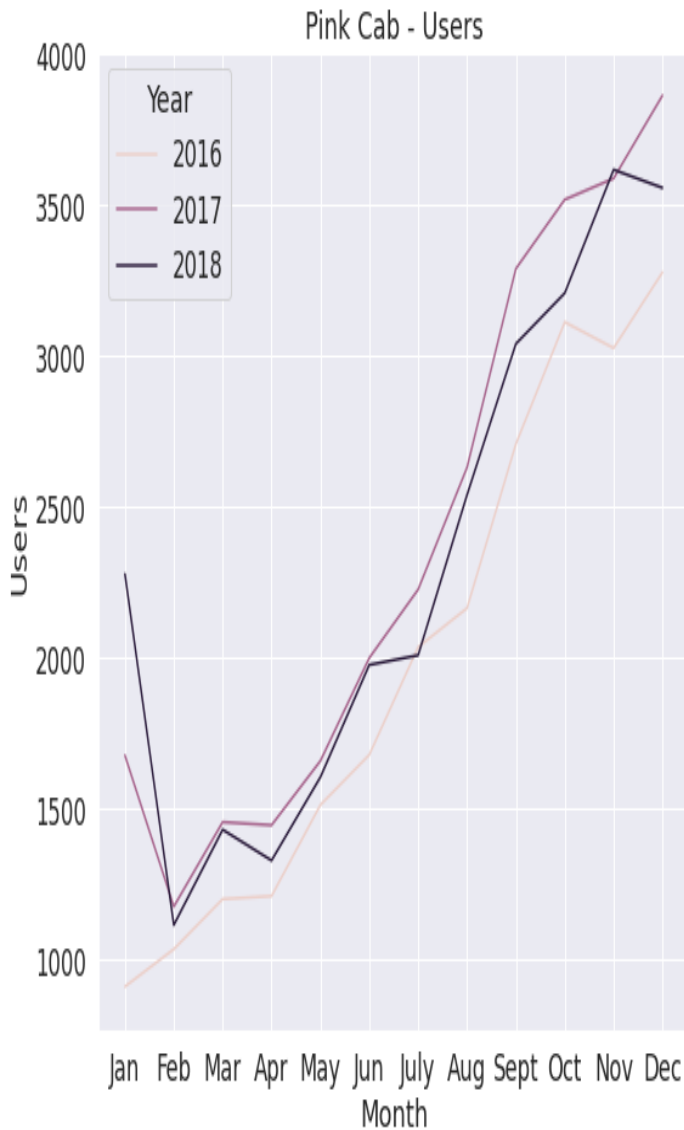


There is a huge difference in Profit between the two companies. Yellow Cab has a significantly larger Profit than Pink Cab.

2017 was the most profitable year for both companies.

2016 was more profitable than 2018 for both companies, despite having fewer transactions.

Users vs Trips vs Profit – Pink Cab



In Feb there was a drop in Profit, in Mar there was an increase in Profit, a big drop during Apr-May, and then the Profit gradually increased until the end of the year.

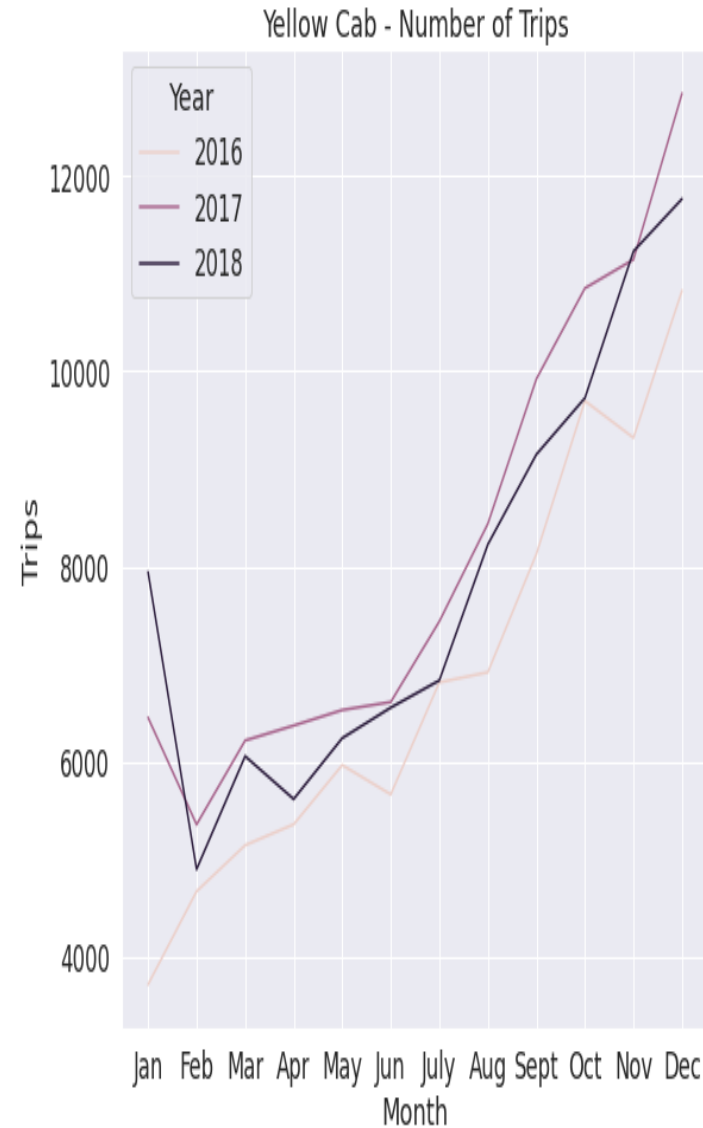
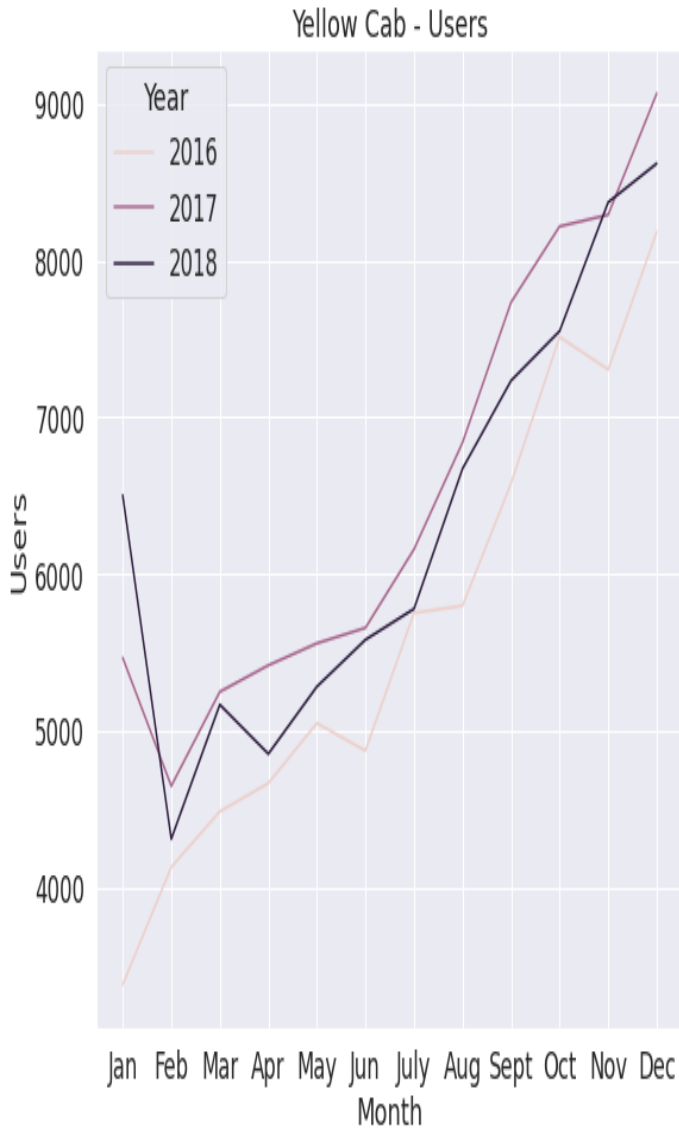
There is a gradually increase in numbers of trips and users in all three years.

Overall:

Profit and Number of trips seem to be highly positively correlated.

Despite the increase in number of trips in Mar-May, the profit decreased significantly.

Users vs Trips vs Profit – Yellow Cab



2017 and 2018 follow the same trend. A bid drop in Profit in Feb, Jun, and July and a big increase in Profit in May and Sept onwards.

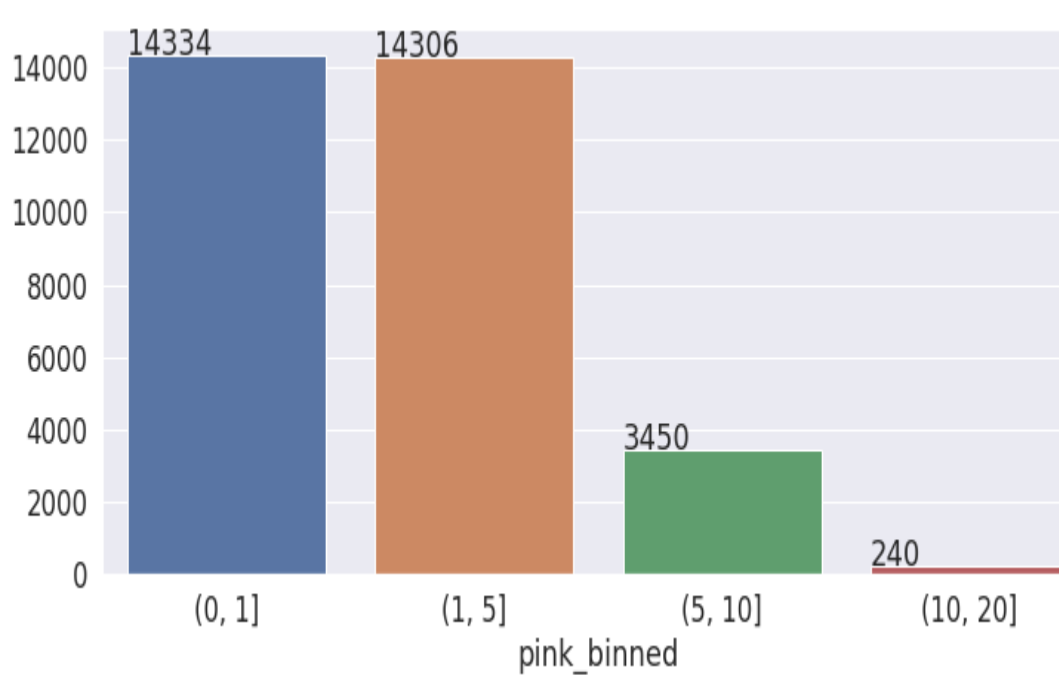
In 2016 there was an increase in Profit in Feb, May, and Sept onwards. There was a huge drop in Profit in Aug.

Overall:

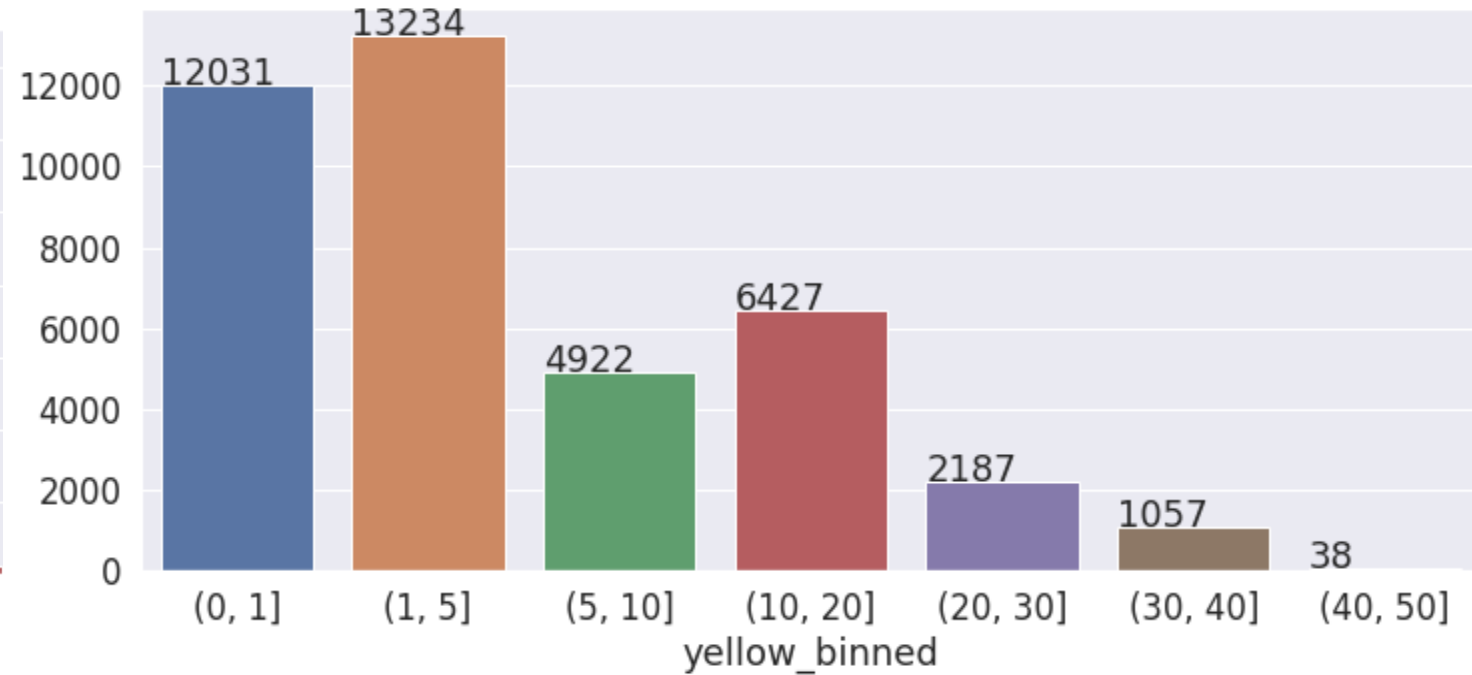
Profit and Number of trips seem to be positively correlated.

Despite the steady increase in number or trips during the summer months, the profit decreased dramatically.

Customer retainability



Most customers use the Pink Cab company once or up to 5 times.

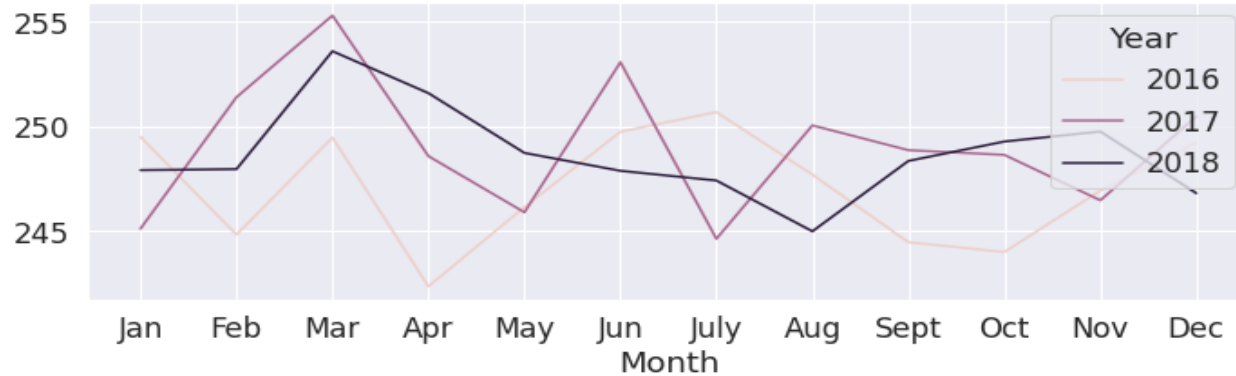


The Yellow Cab company does a better job at retaining its customers than the Pink Cab company.

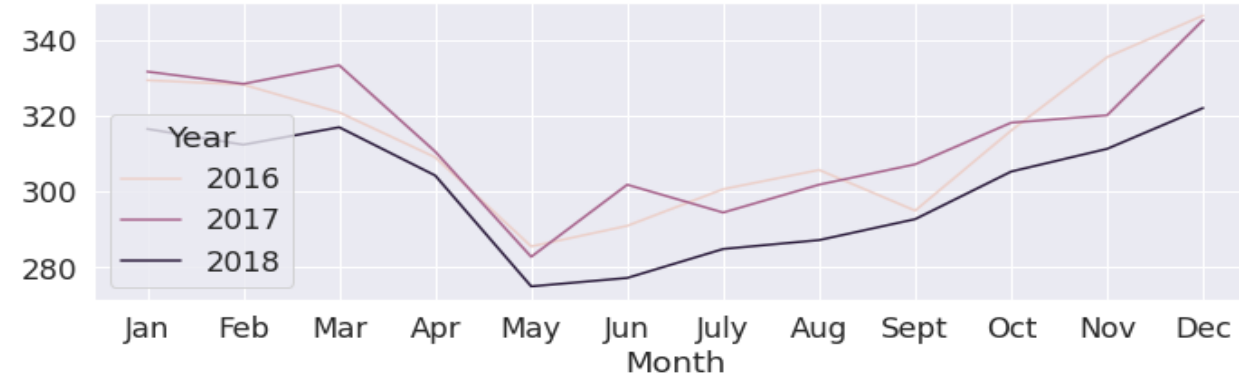
Most customers use the Yellow Cab company more than one time. About one third of its customers use its services more than 5 times.

Cost vs Price of Trips

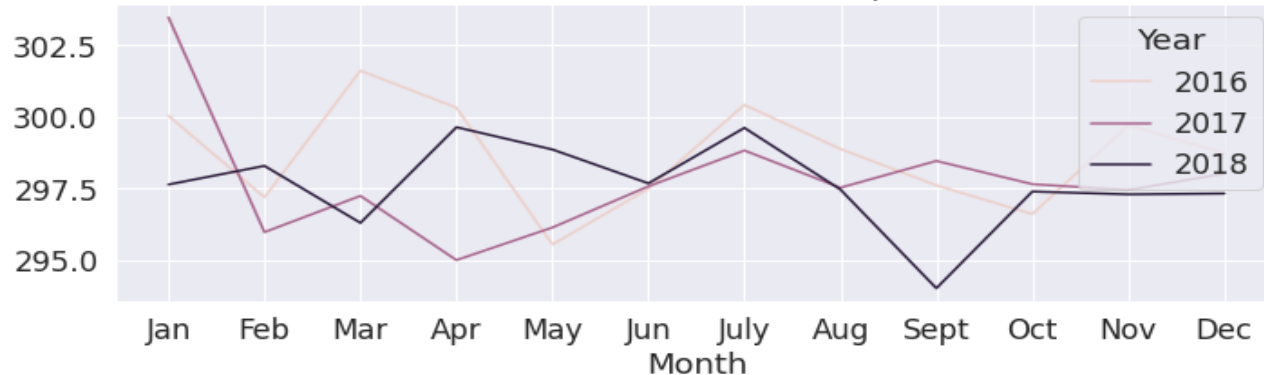
Pink Cab - Cost of Trip



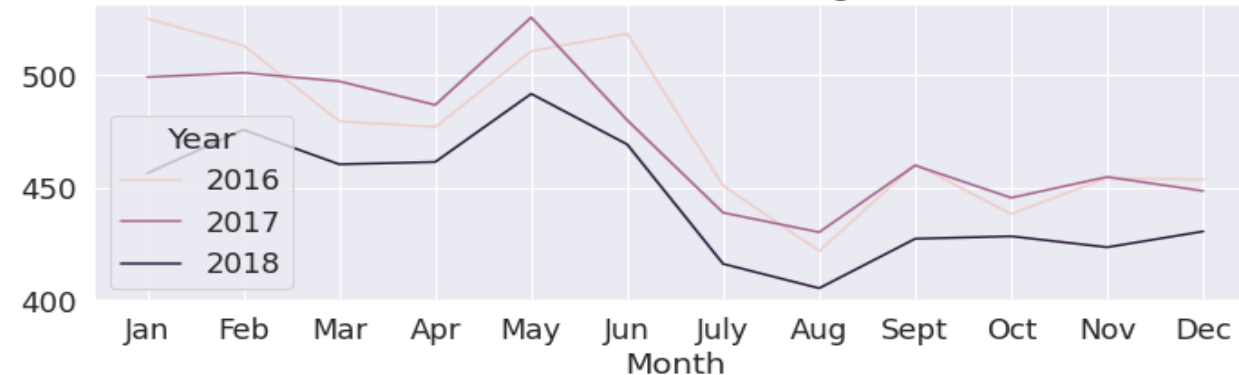
Pink Cab - Price Charged



Yellow Cab - Cost of Trip



Yellow Cab - Price Charged

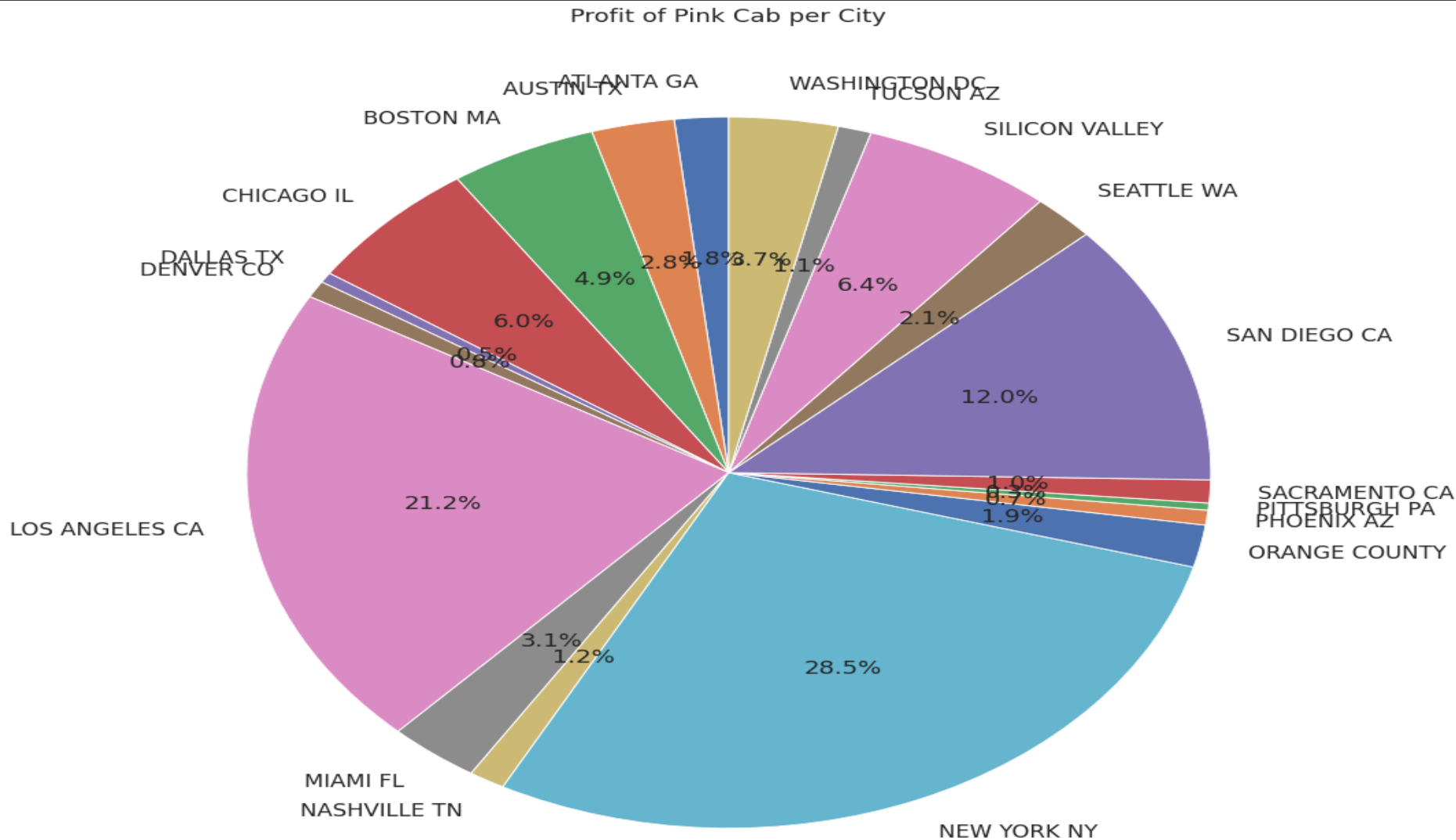


It is clear that during the months Mar-May the Cost of Trip decreased slightly, whereas the Price Charged dropped significantly. This caused the big drop in profit during these months for the Pink Cab Company.

During the summer months the Cost of Trip slightly increased, whereas the Price Charged decreased dramatically. This caused the big drop in profit for the Yellow Cab company.

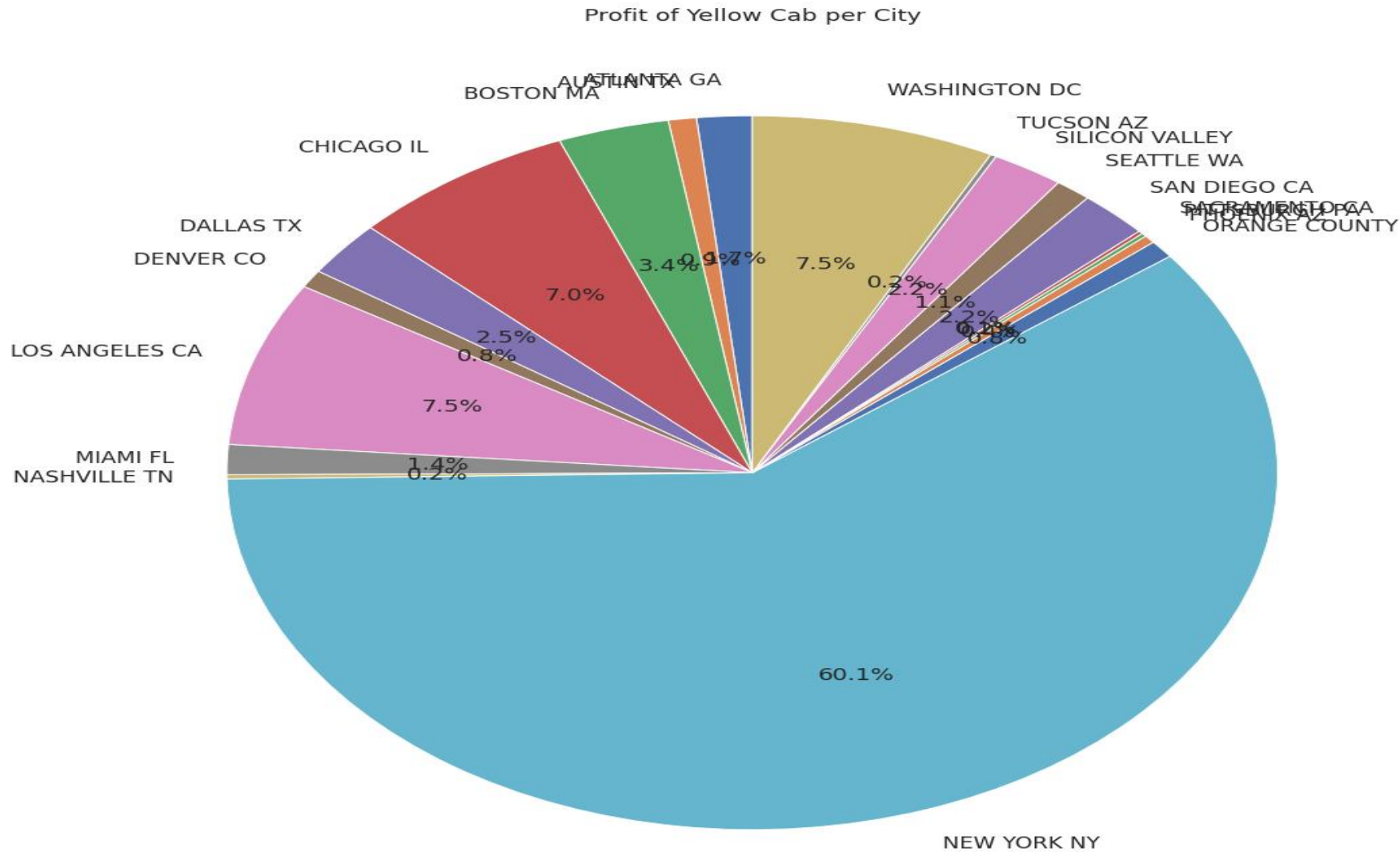
There is a distinct difference in Price Charged in May. Yellow Cab increased their prices, whereas Pink Cab dropped them significantly.
NB: The prices of Yellow Cab are much higher than Pink's Cab year round.

Profit per City – Pink Cab



The Pink Cab company makes the most profit in NEW YORK NY, LOS ANGELES CA, and SAN DIEGO CA.

Profit per City – Yellow Cab



The Yellow Cab company makes 60% of its profit in NEW YORK NY!

Conclusions and Recommendations

- In most cities the number of users of the Yellow Cab company is more than that of the Pink Cab company. Only in NASHVILLE TN, SACRAMENTO CA, and SAN DIEGO CA the number of users of the Pink Cab company is more than that of the Yellow Cab company. Yellow Cab has twice as many users as Pink Cab. Yellow Cab company has about three times more number of trips than the Pink Cab company.
- Yellow Cab had way more transactions than Pink Cab in all three years.
- There is a huge difference in Profit between the two companies. Yellow Cab has a significantly larger Profit than Pink Cab.
- During the months of Mar-May the Cost of Trip decreased slightly, whereas the Price Charged dropped significantly. This caused the big drop in profit during these months for the Pink Cab company.
- During the summer months the Cost of Trip slightly increased, whereas the Price Charged decreased dramatically. This caused the big drop in profit for the Yellow Cab company.
- The Pink Cab company makes the most profit in NEW YORK NY, LOS ANGELES CA, and SAN DIEGO CA.
- The Yellow Cab company makes 60% of its profit in NEW YORK NY!
- The Yellow Cab company does a better job at retaining its customers than the Pink Cab company.

Based on our analysis, we recommend investing in the Yellow Cab company.

Thank You