



**Data Glacier**

Your Deep Learning Partner

# Exploratory Data Analysis

G2M insight for Cab Investment firm

20-Sep-2022

# Background –G2M(cab industry) case study

XYZ is a private equity 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 :** Provide actionable insights to help XYZ firm in identifying the right company for making investment.

The analysis has been divided into four parts:

- Data Understanding
- Finding the most profitable Cab company
- Recommendations for investment

# Data Exploration

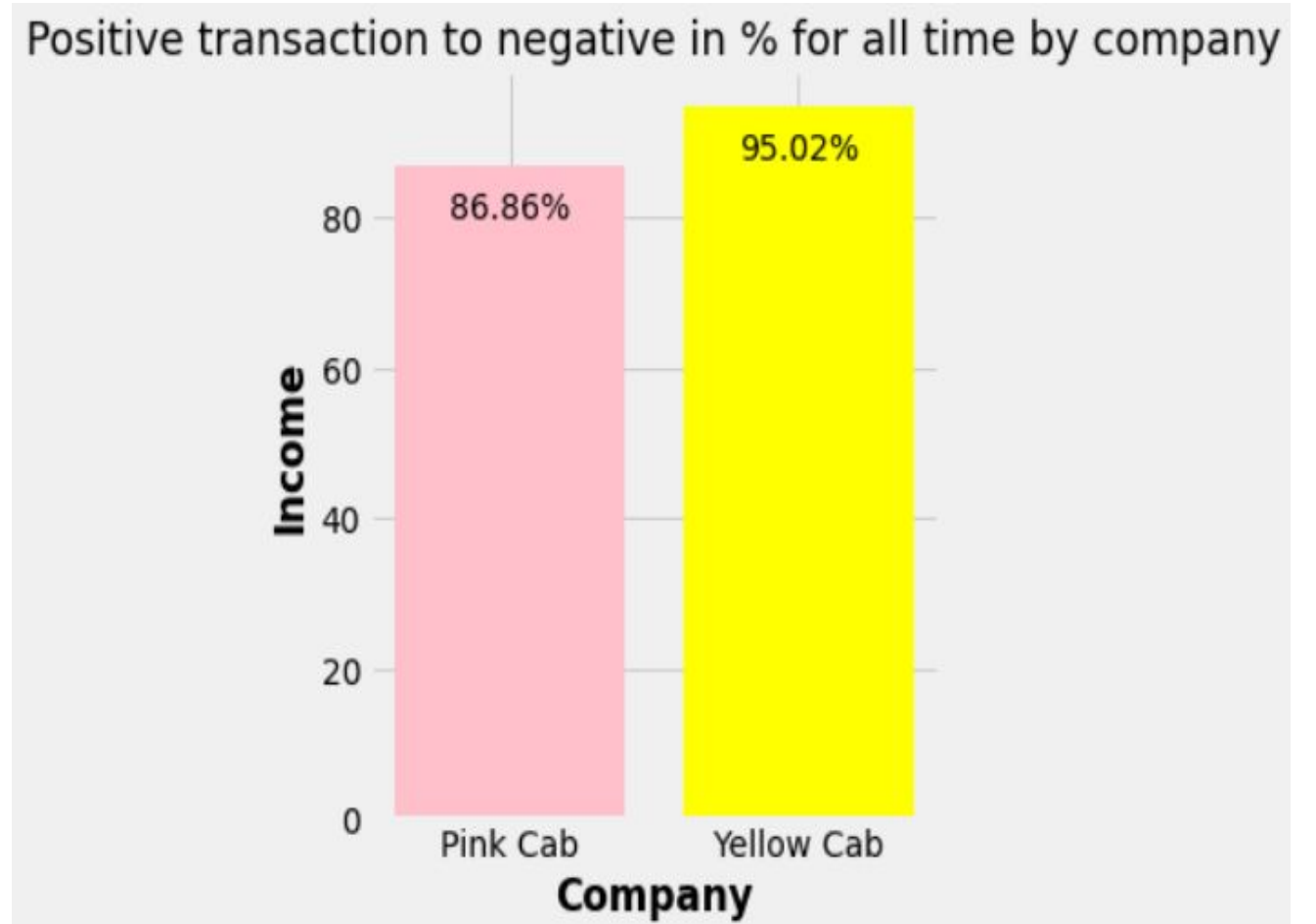
- 4 databases with 14 unique categories:

cab\_data (359392 x 7), city (20 x 3), customer\_id (49171 x 4), transaction\_id (440098 x 3)

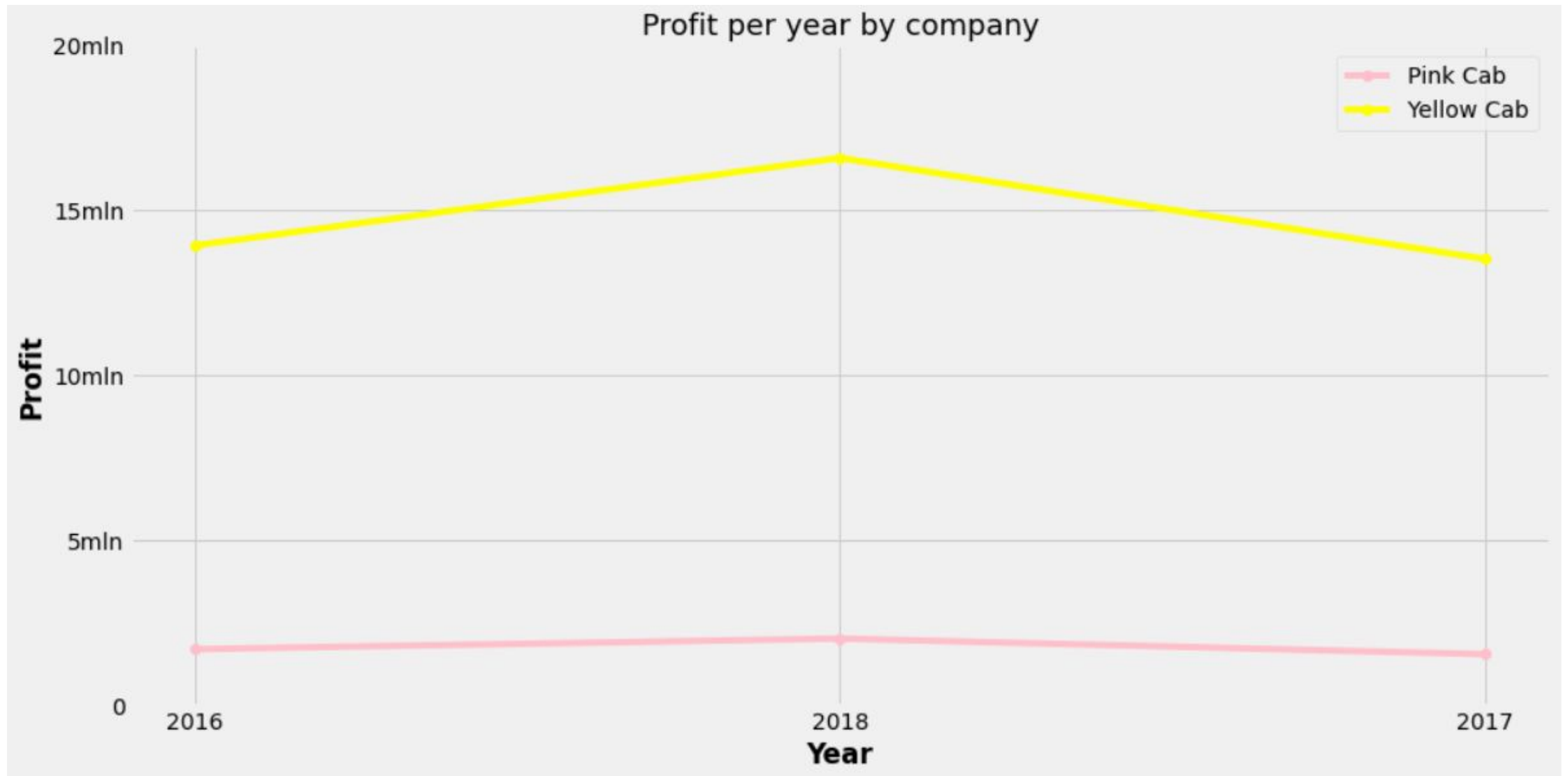
- Timeframe of the data: 2016-01-31 to 2018-12-31

1. All dates are stored as integers representing the number of days since January 1, 1900, which need format to normal type.
2. Price\_Charged and Cost\_of\_Trip functions used to calculate profit.
3. Users feature of city dataset is treated as number of cab users in the city. We have assumed that this can be other cab users as well(including Yellow and Pink cab)
4. "Customer ID" is repeted feutures, but "Transaction ID" - no.

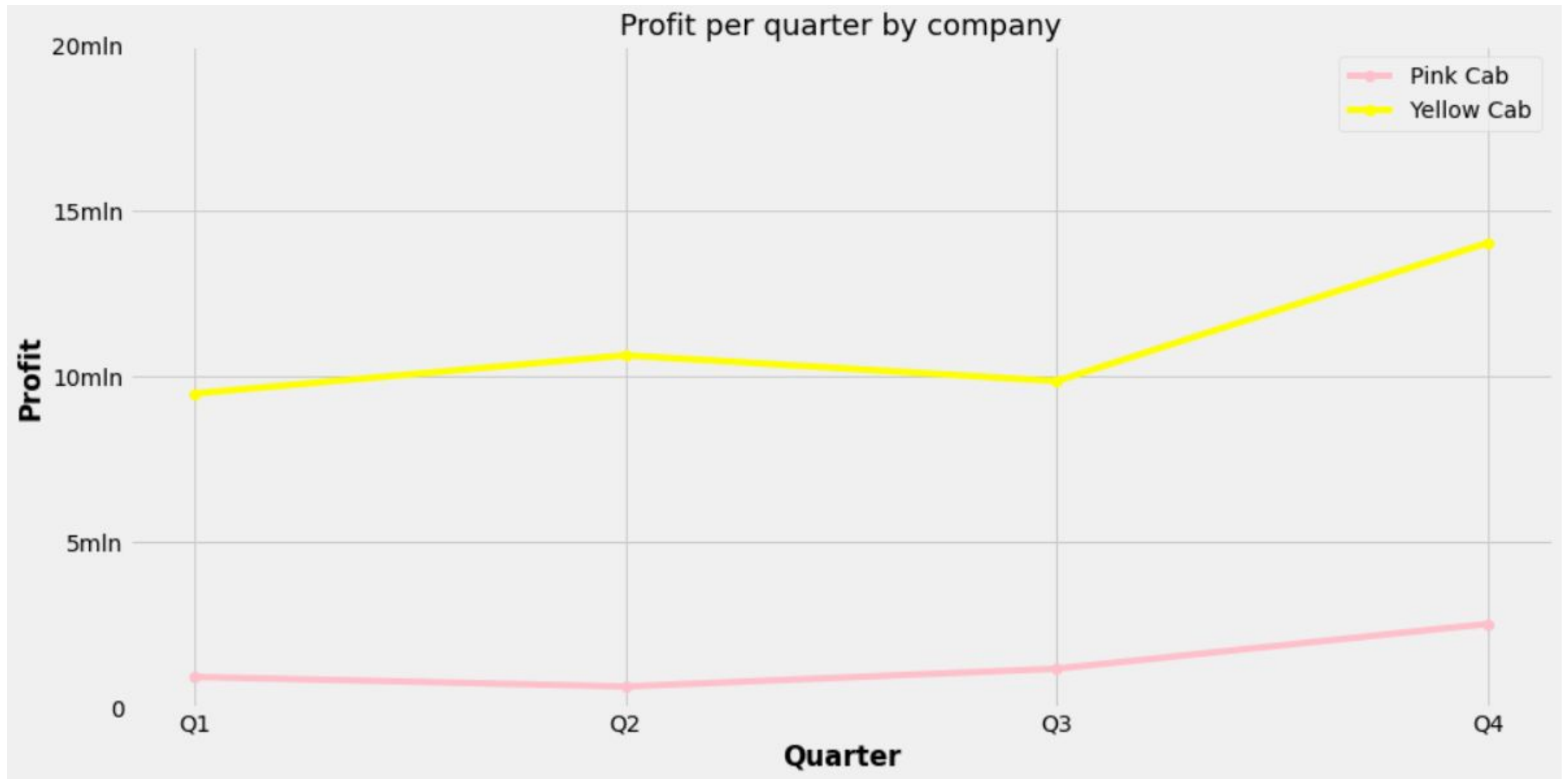
# Profit Analysis



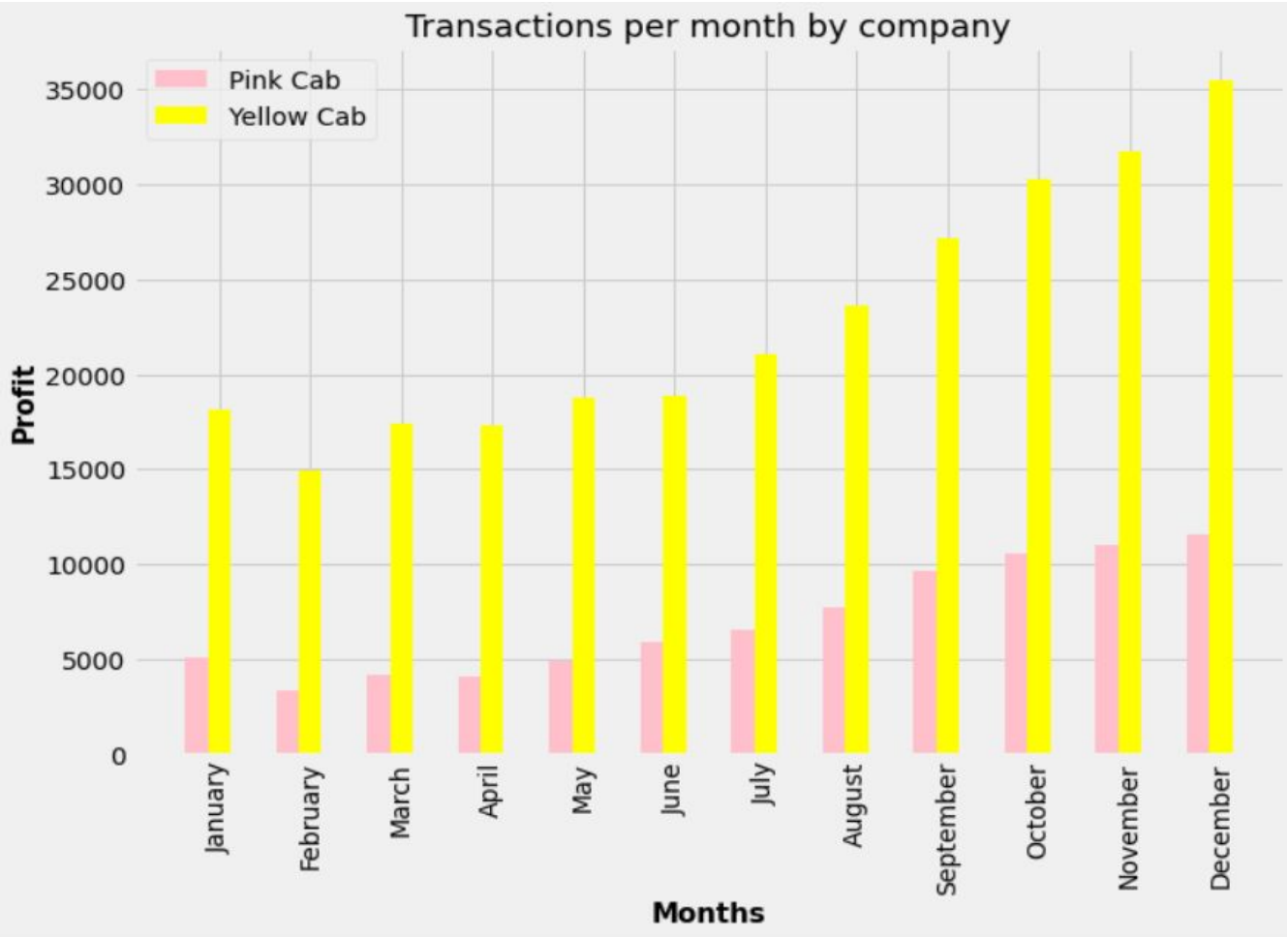
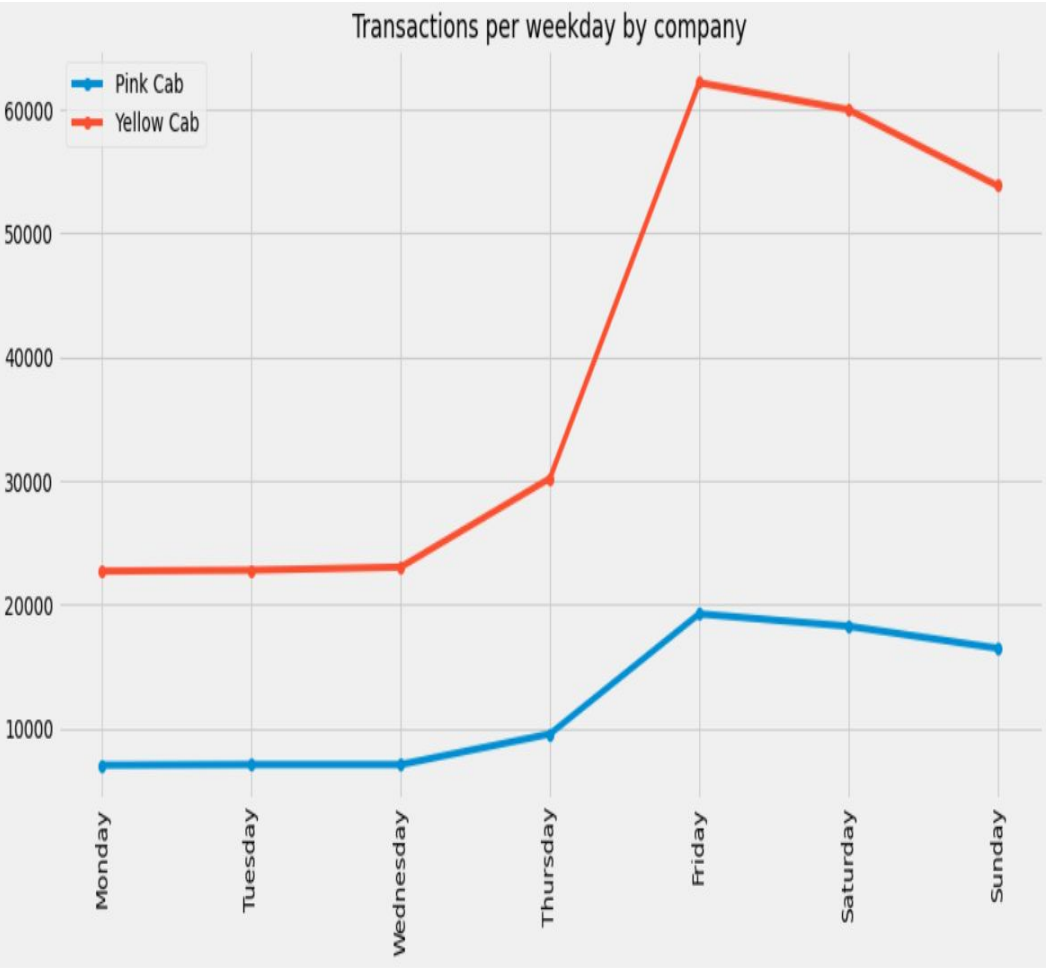
# Yearly Profit Analysis



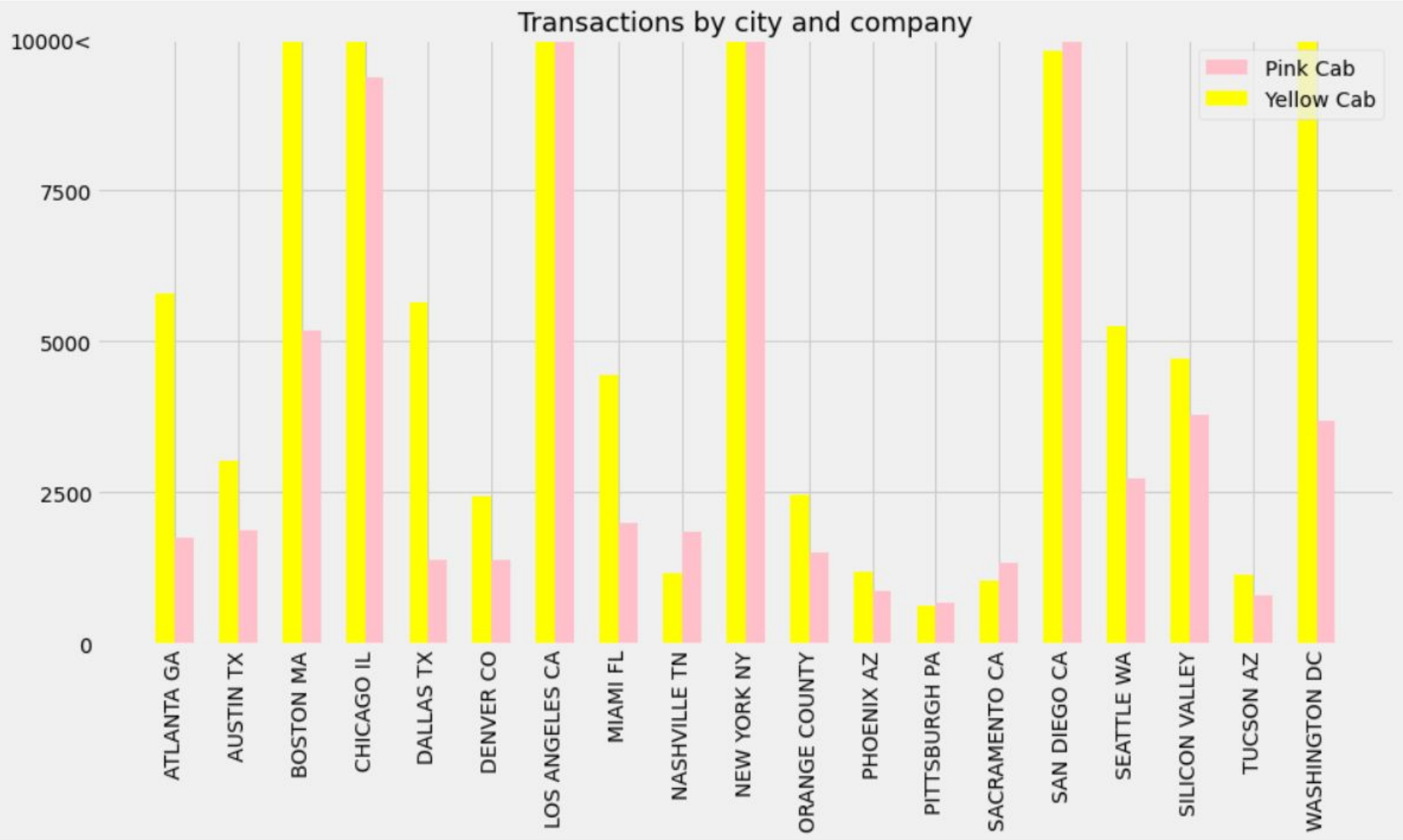
# Quarterly Profit Analysis



# Number of transactions per week/ per month



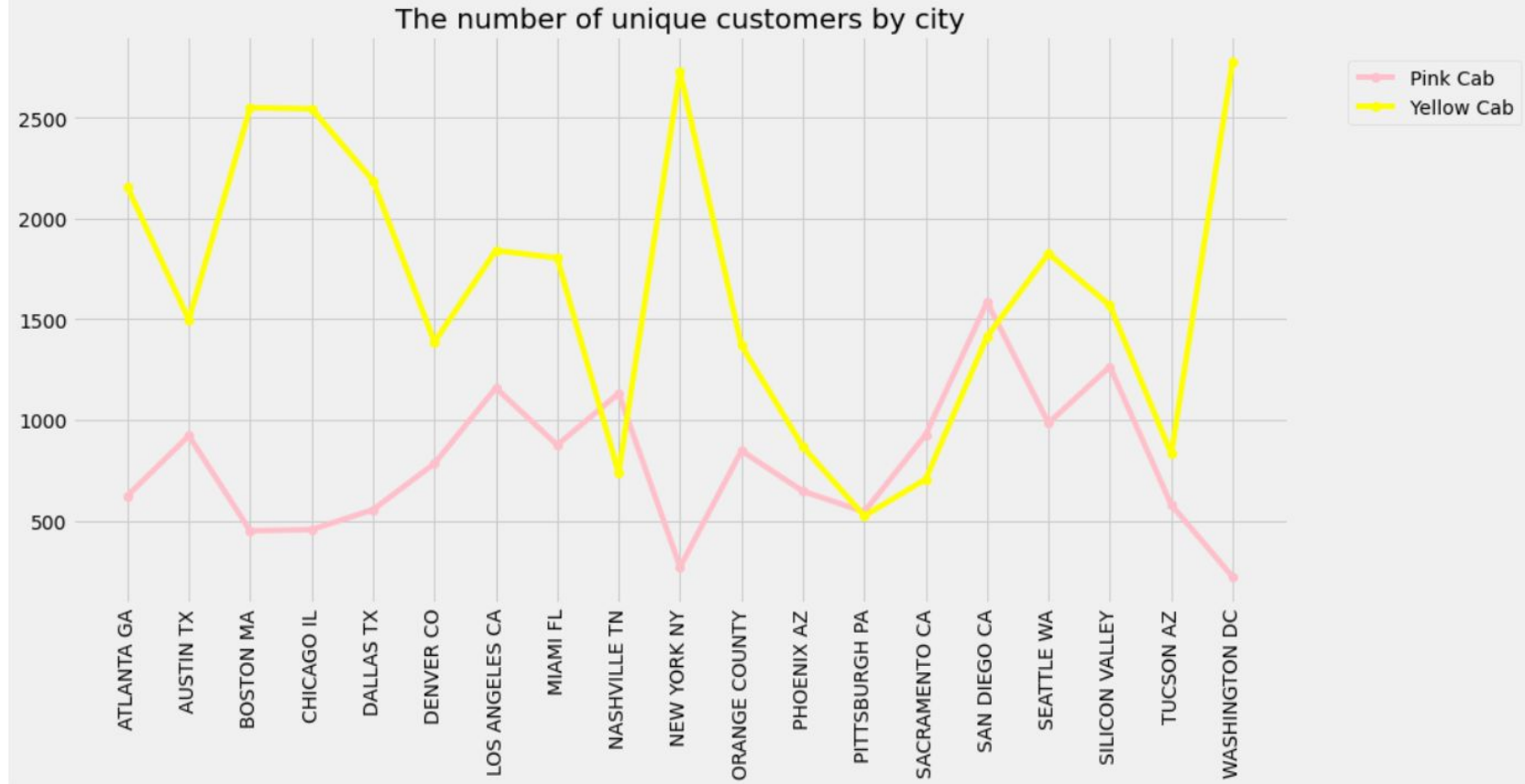
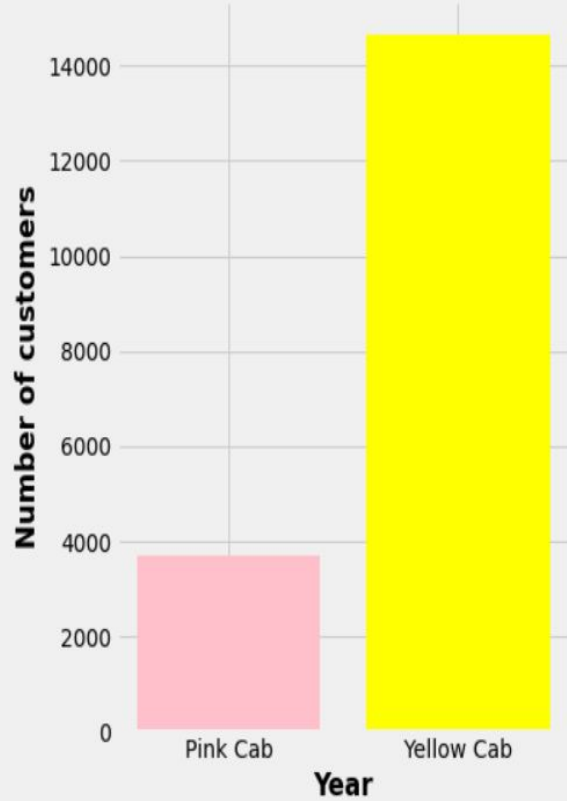
# Popularity of taxi companies by city



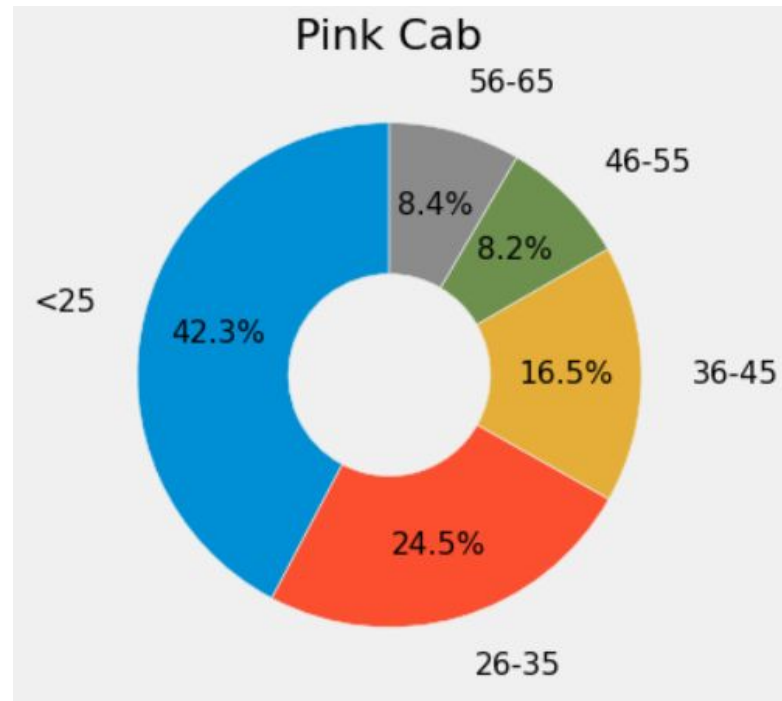
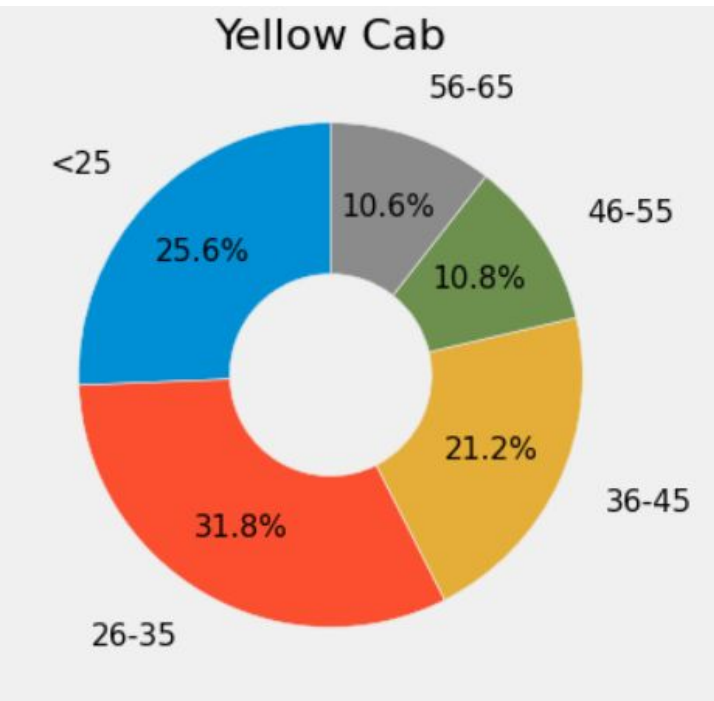


# Number of unique customers

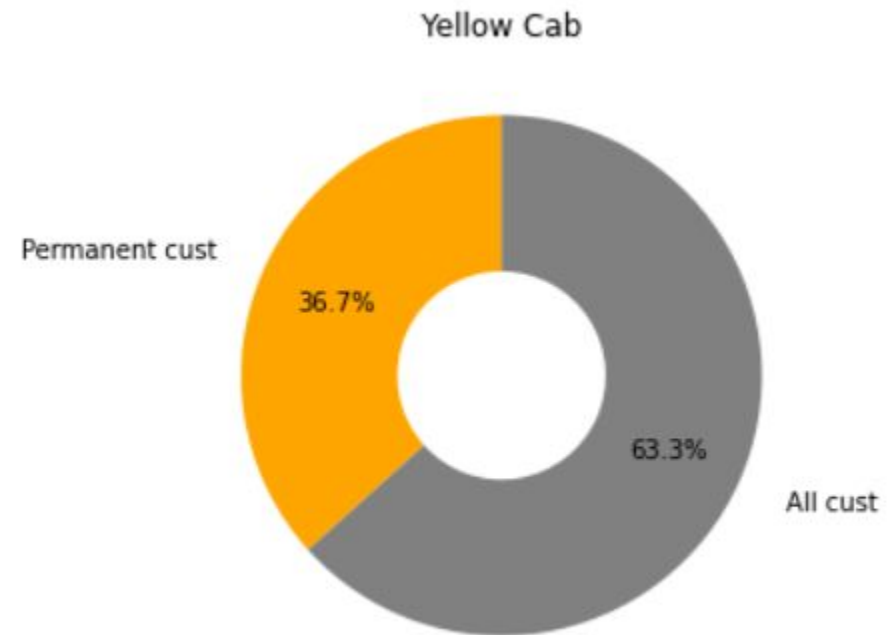
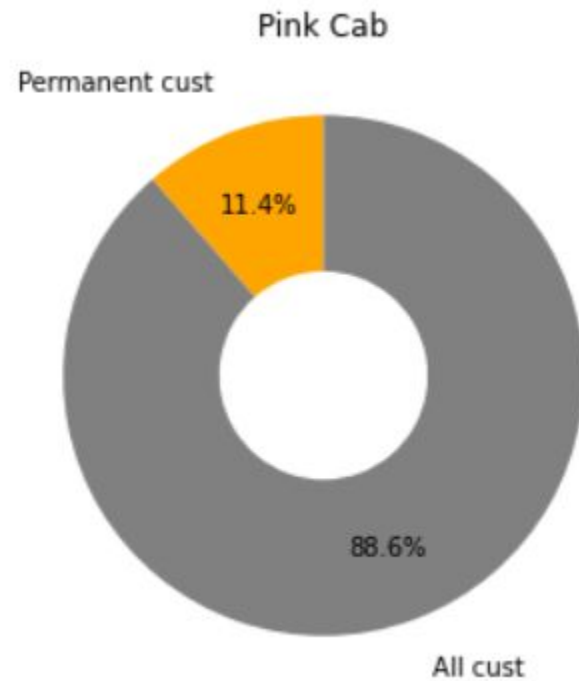
The number of permanent customers >5 by company



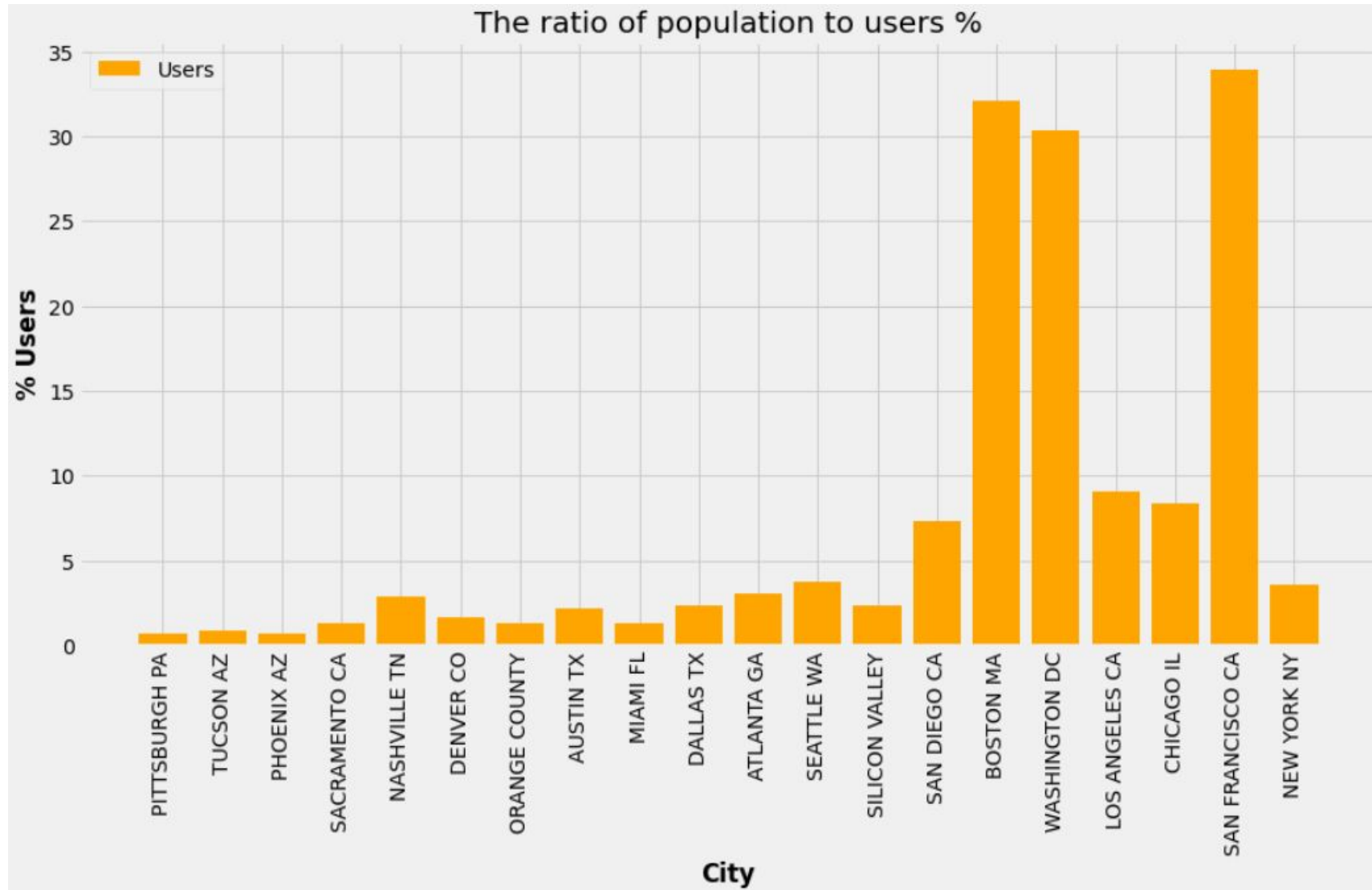
# Number of customer base Analysis Age & Gender



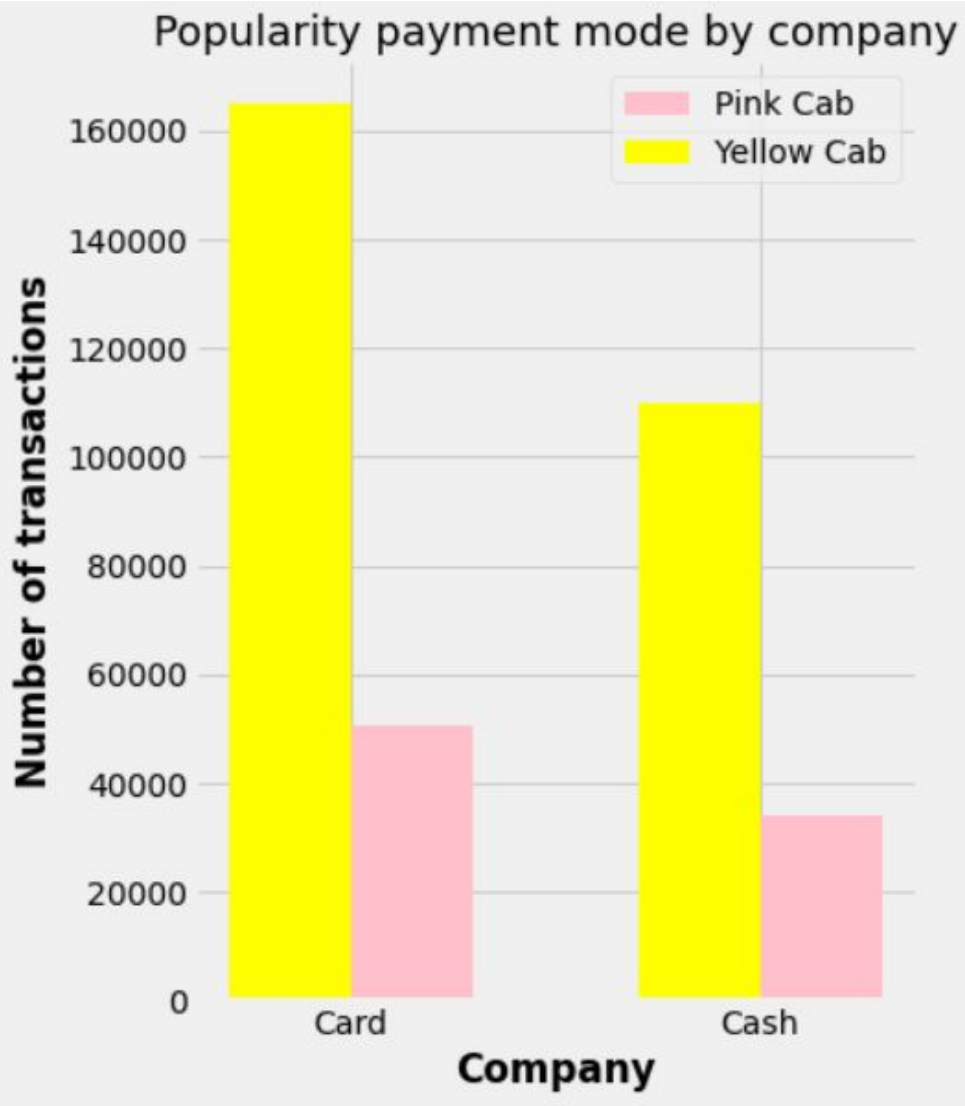
# The number of unique permanent customers % by company



# The ratio of population to users % by city



# Popularity payment mode by company



# Summary

**Based on the visualization results, we can draw the following conclusions:**

1. The maximum revenues for the period 31.01.2016-31.12.2018 belong to the yellow cabin company.
2. Profitable transactions make up 95% of all transactions, as at that time the pink cab has only 86 percent.
3. 2018 can be considered a more profitable year, although the general trend does not have sharp changes, which indicates stability.
4. The most productive quarter for both companies is 4 (October, November, and December)
5. The number of transactions by month and day of the week is approximately two times greater in the yellow cab company than in the yellow cab, although the trends are as similar as possible.
6. Regarding the popularity of using these taxi companies in cities: in most of them it is equal, sometimes one or the other is better.
7. Women use taxi services more actively in both companies.
8. Pink cabs are more popular among young people under 25, and yellow cabs are more popular among customers aged 26-35.
9. The number of unique customers per company depends on the city. New York, Washington - yellow cab, San Diego, Nashville - pink cab.
10. The number of customers who used a taxi is more than 5 times greater in the yellow cab company.
11. The percentage of regular customers to the total number is also higher in the yellow cab company, which indicates that people are interested in this company. 11.4% pink, 36.7% yellow cab.
12. The best ratio of the city's population to the number of taxi users is in: San Francisco, Boston, Washington.
13. A popular method of payment is the card method, however, for the Pink Taxi company, there is no significant difference between the card and cash payment methods.

**The recommended company for investment will be Zhovtyi kab, as it has the most positive conclusions.**

# Thank You