



**Data Glacier**

Your Deep Learning Partner

# Cab Data Analysis

MUHAMMED SEFA SÖZER

21-June-2022

# Background –G2M(cab industry) case study

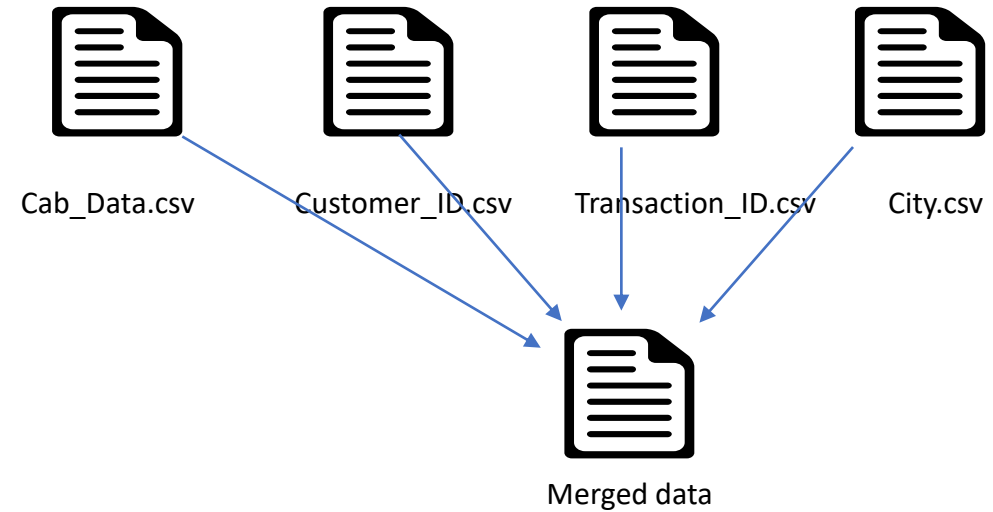
- XYZ is a private equity firm in US. Due to remarkable growth in the Cab Industry in last few years and multiple key players in the market, it is planning for an investment in Cab industry.
- Objective : Provide actionable insights to help XYZ firm in identifying the right company for making investment.

The analysis has been divided into four parts:

- Data Understanding
- Forecasting profit and number of rides for each cab type
- Finding the most profitable Cab company
- Recommendations for investment

# Data Exploration

- 24 Features( including 9 derived features)
- Timeframe of the data: 2016-01-31 to 2018-12-31
- Total data points :355,032



**Cab Data:** This file contains the transaction details of 2 taxi companies.

**City Data:** this file contains list of US cities, their population and number of cab users

**Customer Data:** This data is a data that matches the customer number of the cabin data.

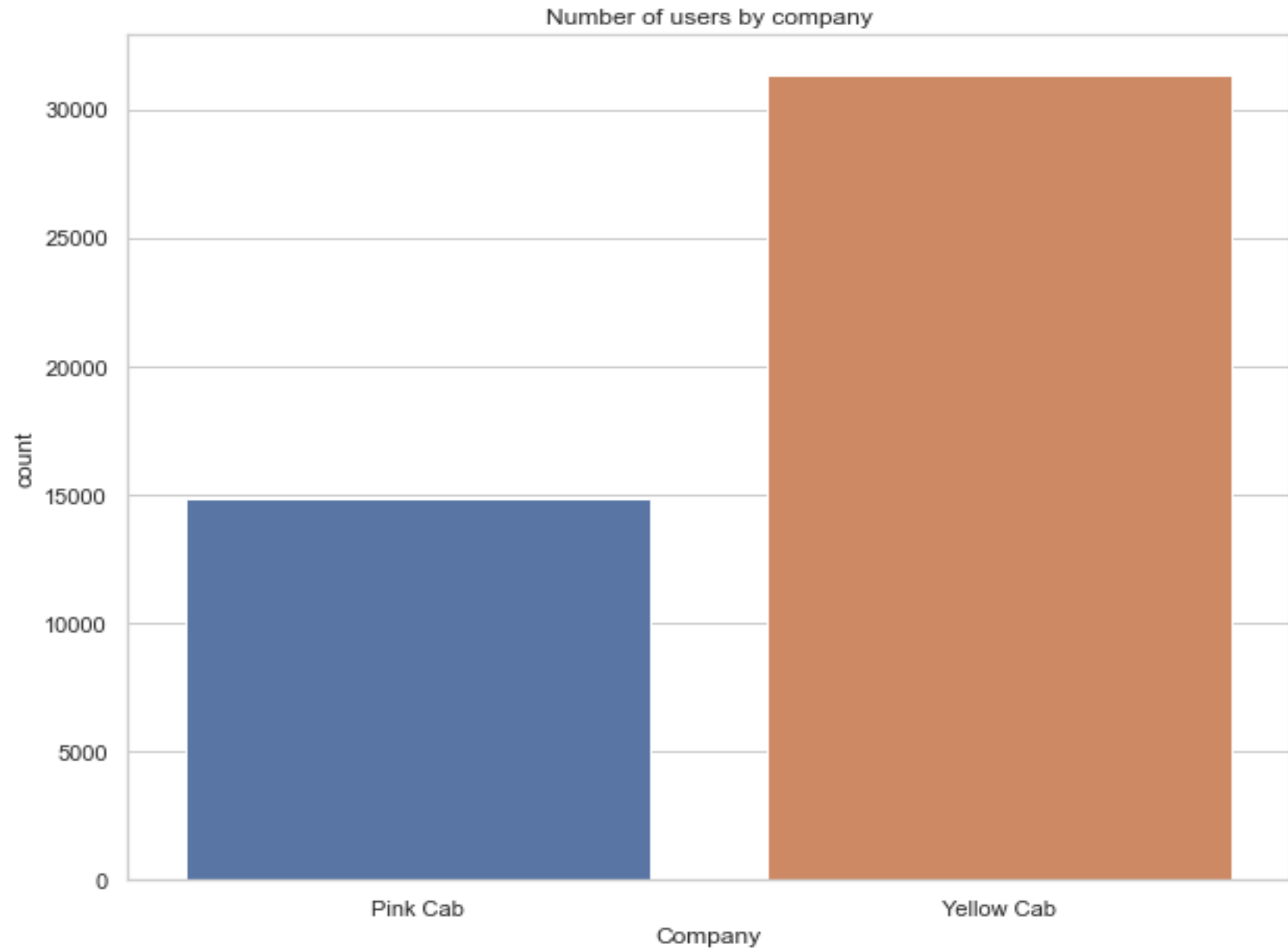
It is the data set that includes the gender, age and monthly expenditure amount of the customer.

**Transaction Data:** this is a mapping table that contains transaction to customer mapping and payment mode

**Merged Data:** Cab Data, City Data, Customer Data, Customer Data and Transaction Data combined.

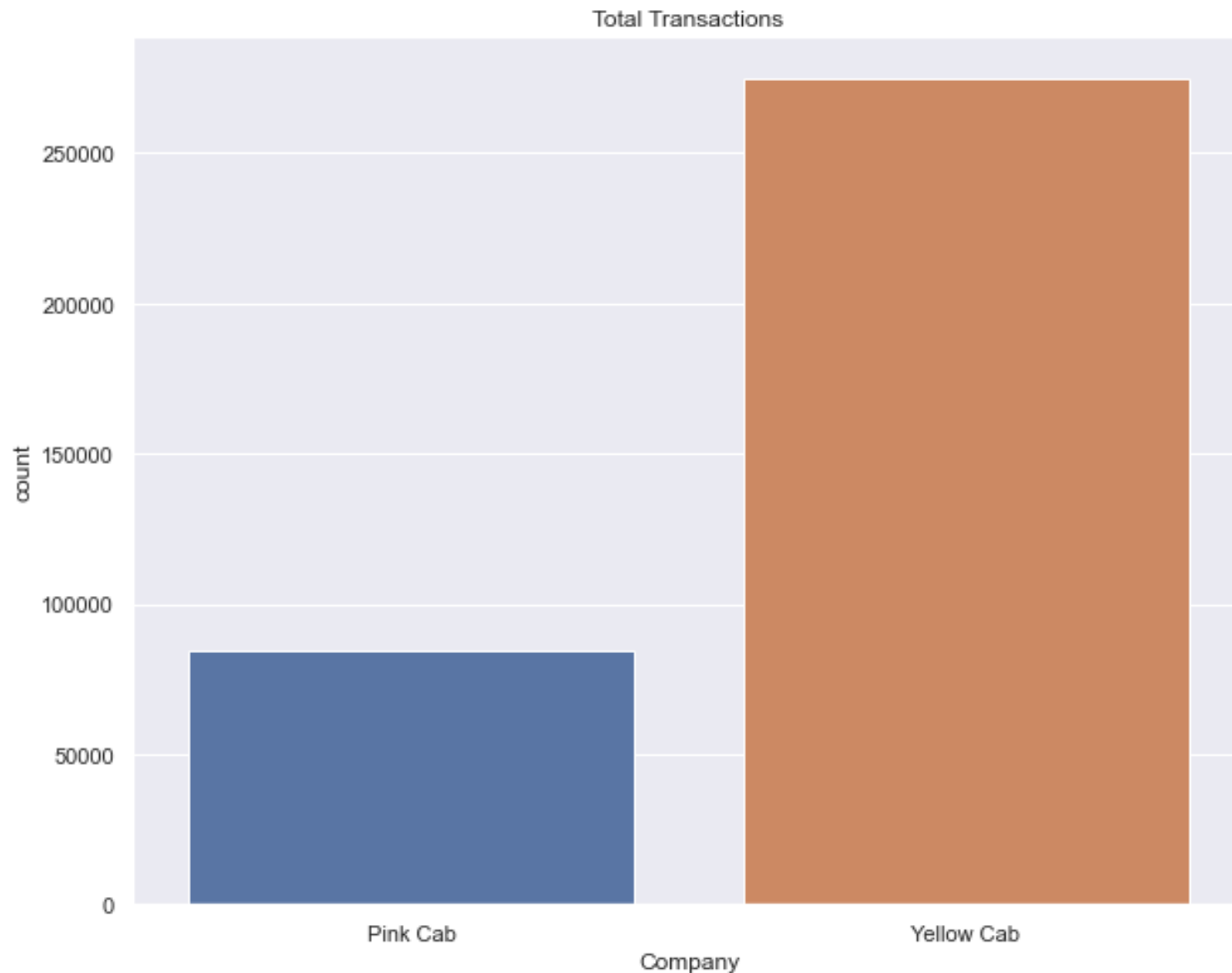
# Number of users by company

- The number of customers of the yellow taxi company is clearly ahead of the pink taxi company, as can be seen from the graph.



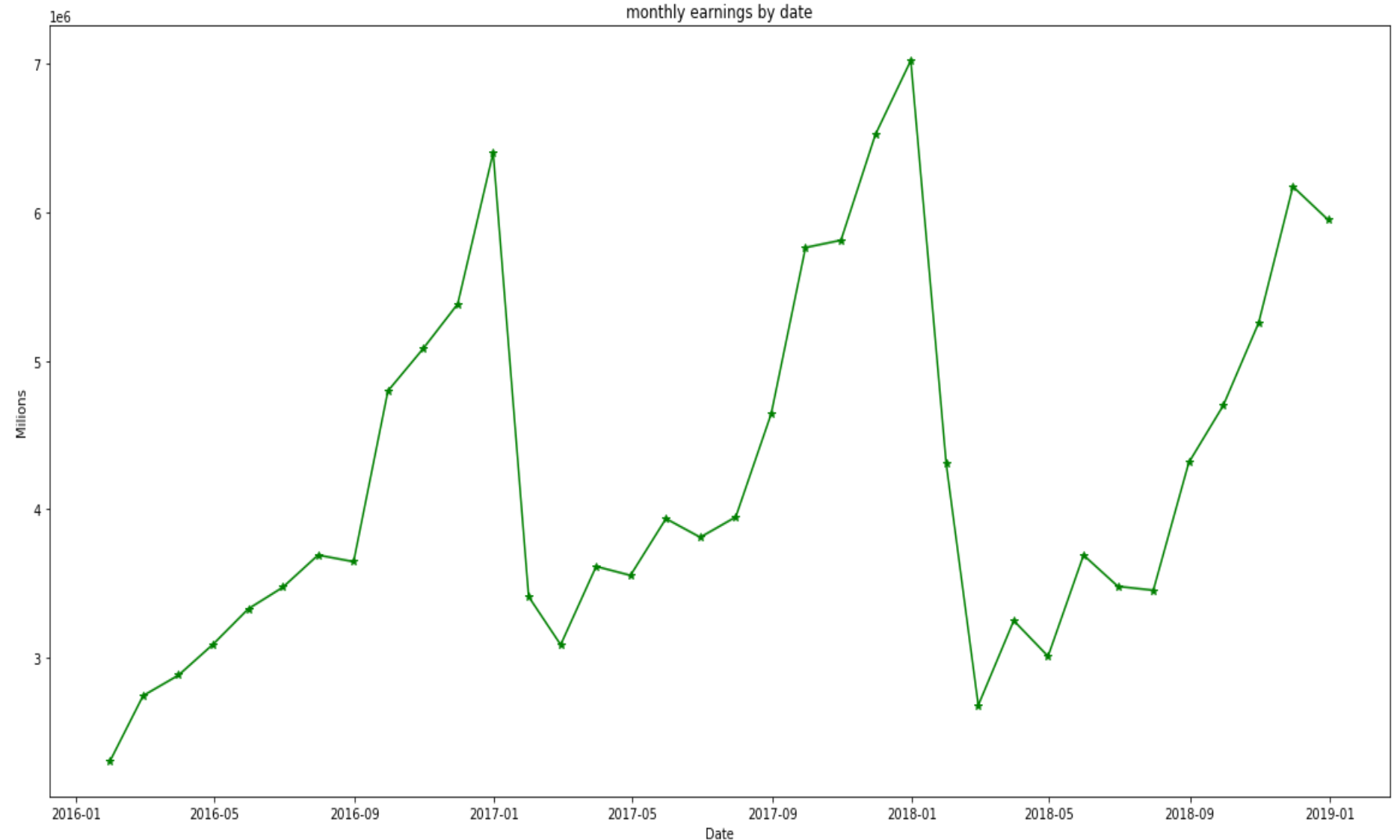
# Number of transactions made in companies

- The use of the yellow taxi company is clearly ahead of the pink taxi company, as can be seen from the graph.



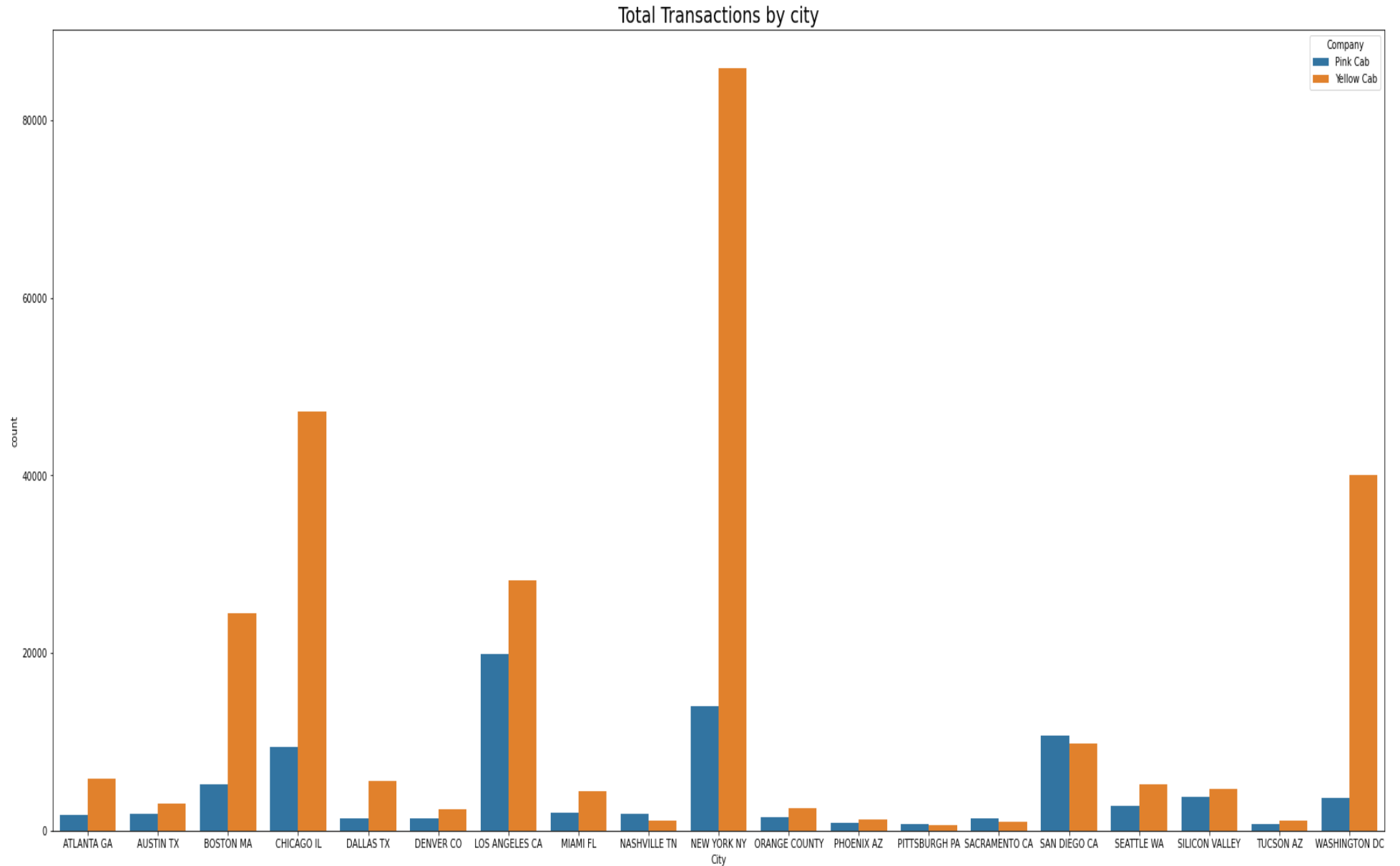
# Monthly earnings by date

- In this chart, we observe that the highest gain is in January and the lowest gain is in February.
- Overall, we observe an increase in earnings from February to the following January.



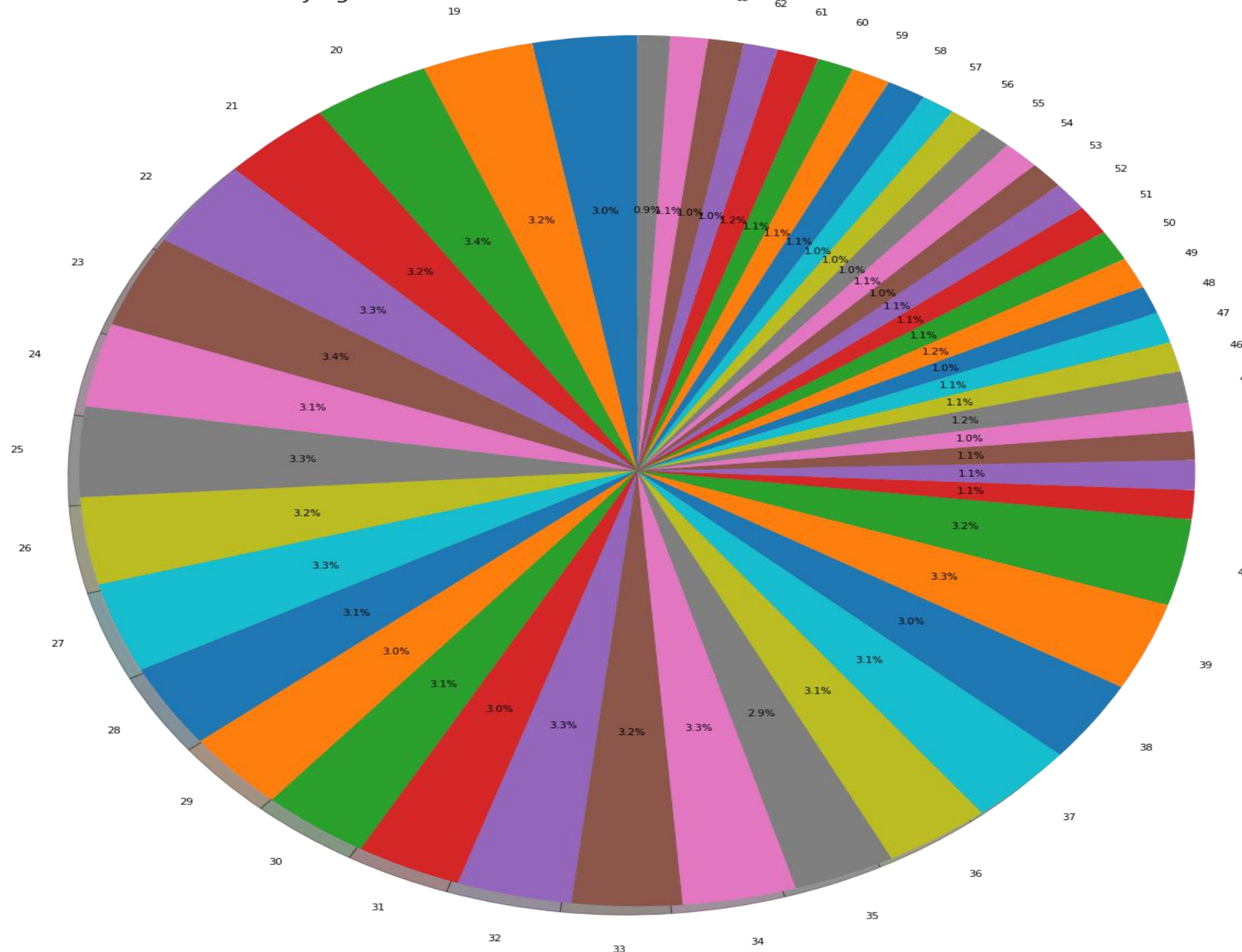
# Number of transactions by taxi companies by city

- The yellow taxi company is most used in New York and the pink company is most used in Los Angeles.



# Number of transactions by age

Number of transactions by age



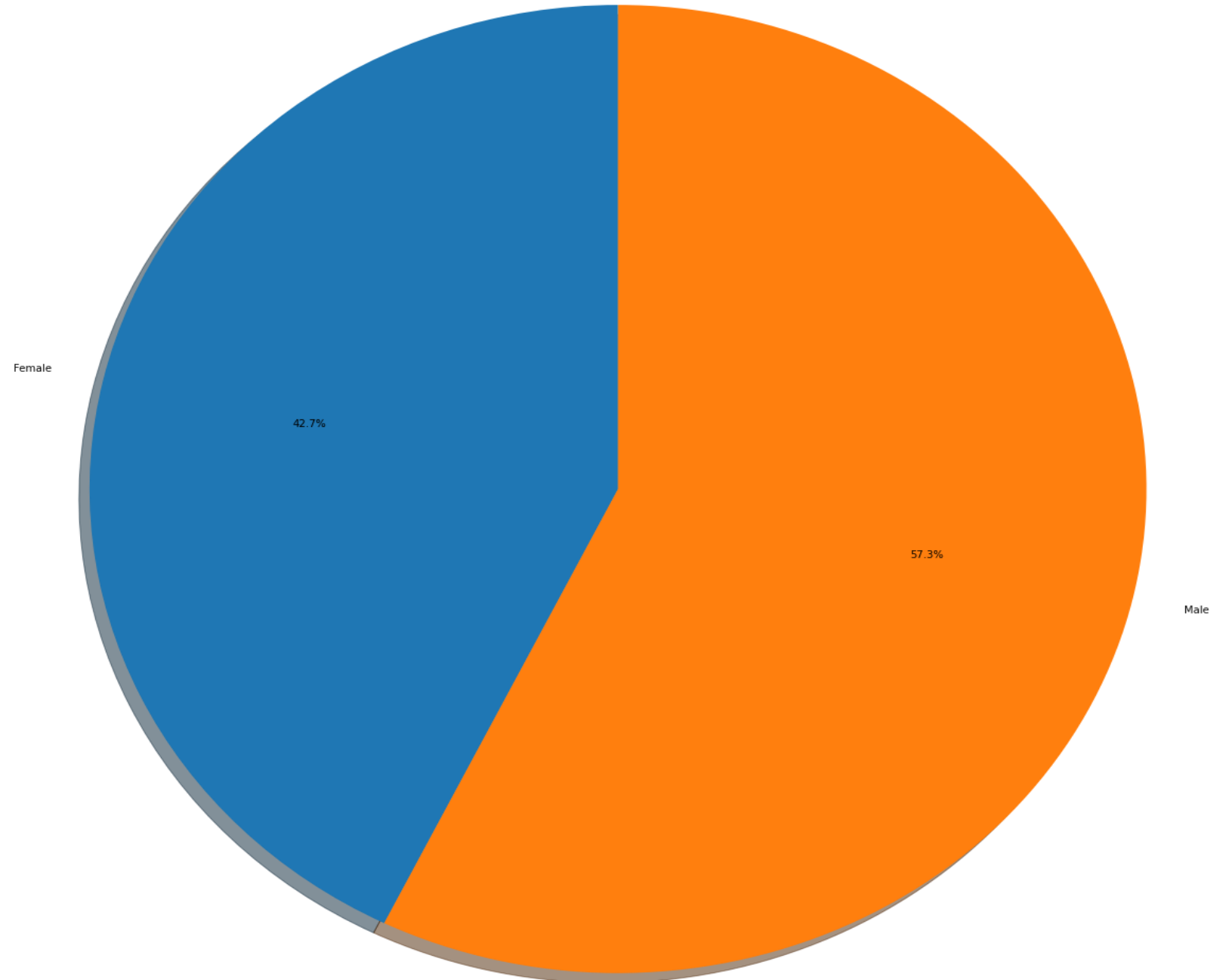
- In general, the rate of use is higher between the ages of 18-40, and the rate of use is low for those over the age of 40.



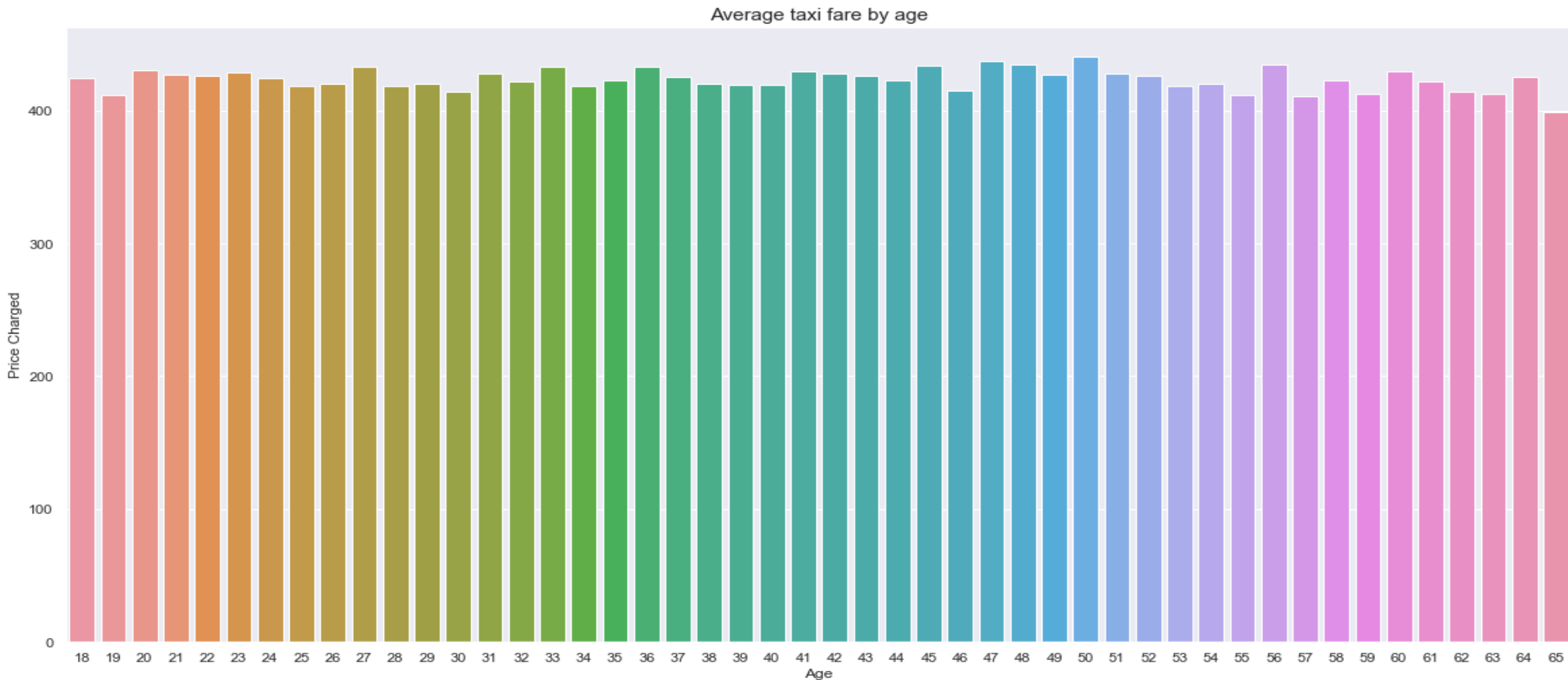
# Number of transactions by gender

- From this graph, we can see that men use it 7% more often than women.

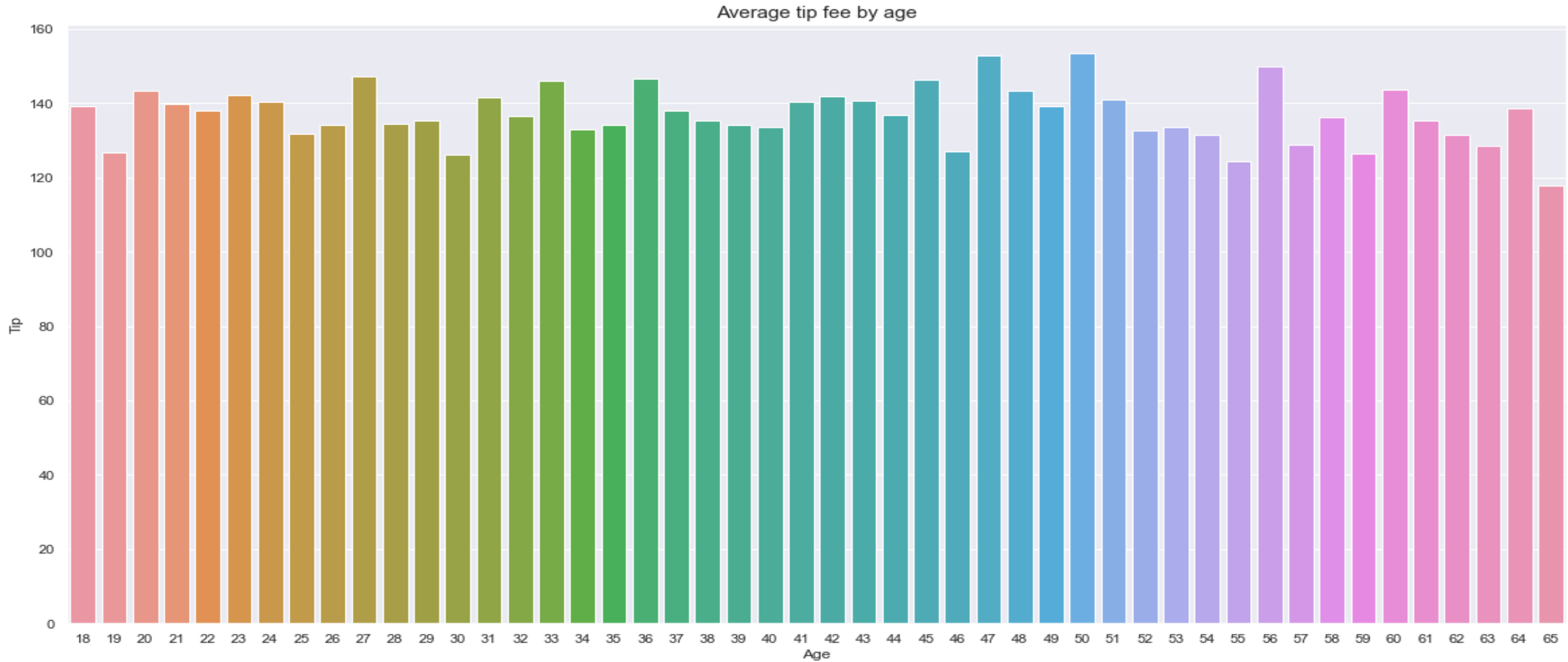
Number of transactions by gender



# Average amount paid by age

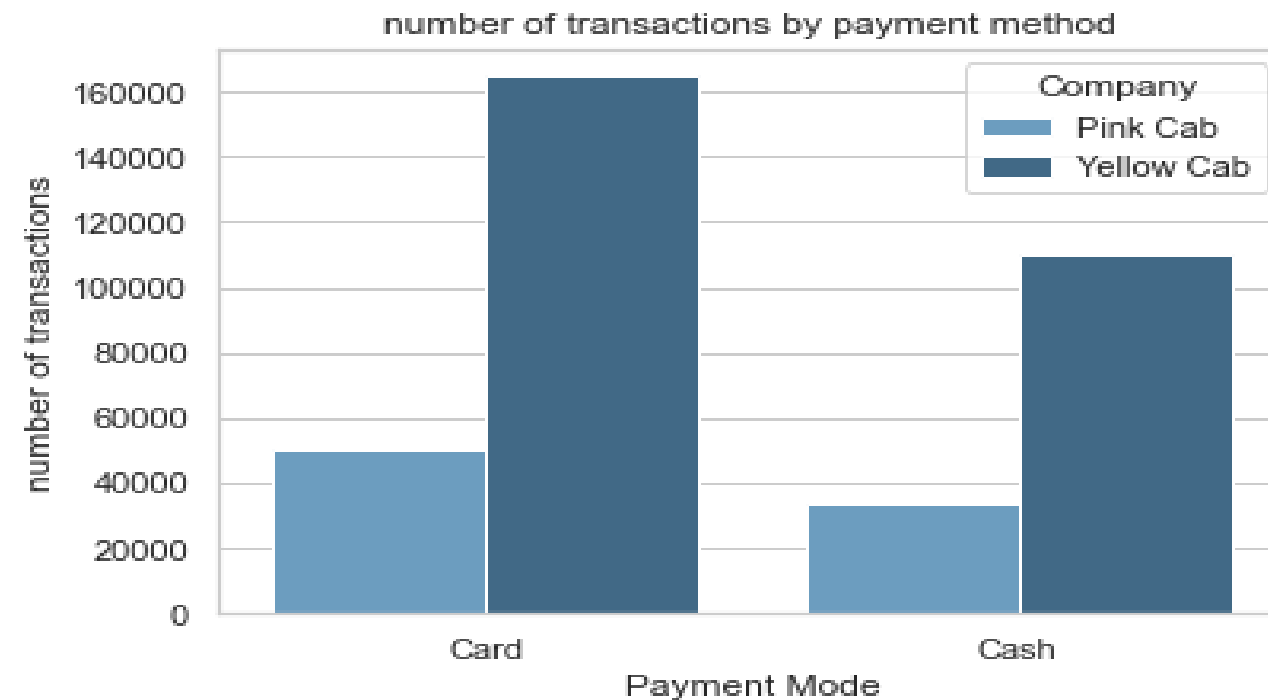
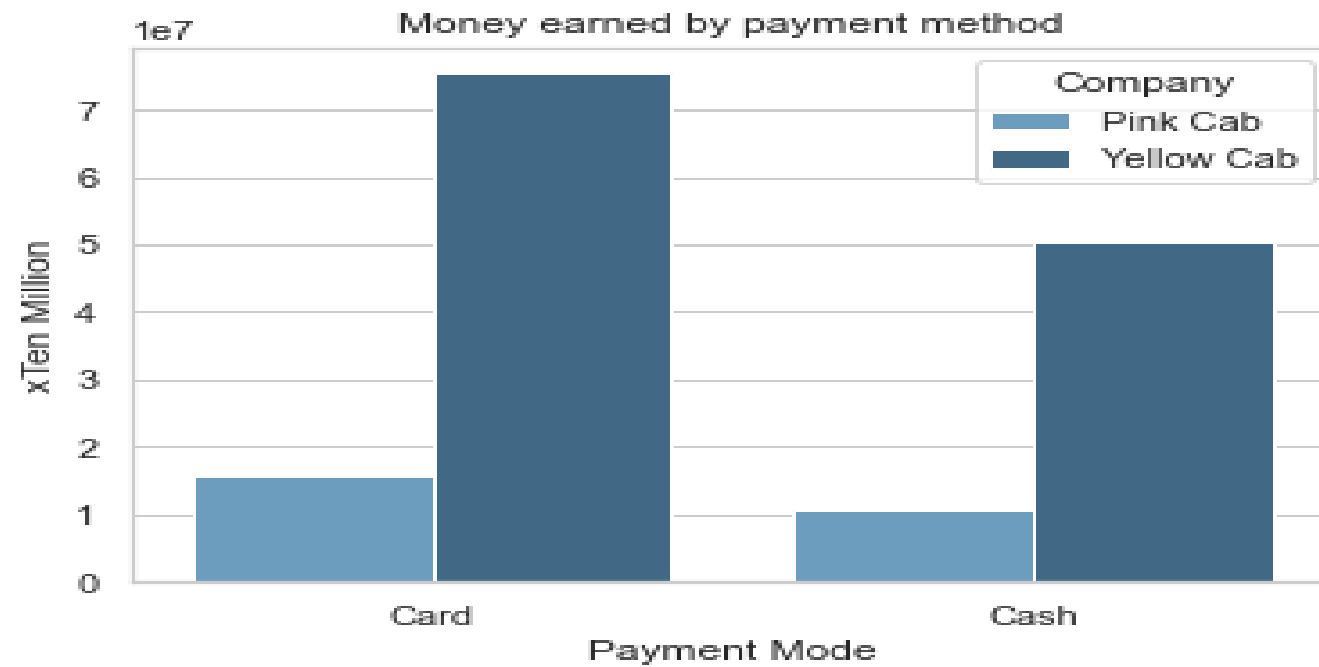


# Average tip fee by age



# Number of transactions by payment method

- As seen in the chart, both companies use the card payment method more frequently.



# recommendation

- When we examine the graphics, the yellow taxi company is far ahead of the pink company. If the pink taxi company is smaller than the yellow taxi company, it needs to grow itself, if it is the same as the company size, it needs to make campaigns and gain customers.
- Taxi usage rate over the age of 40 is low. Taxi opening fee may not be charged in order to gain customers over the age of 40.
- Usually, payment is made by card. Installing POS devices without POS devices in taxis will increase the company's turnover.
- When we examined the graphs, we observed that the location of the companies affected the earnings a lot. For this reason, companies should invest according to locations.

# Thank You