



**Data Glacier**

Your Deep Learning Partner

# G2M Case Study

Virtual Internship

20-Jul-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:

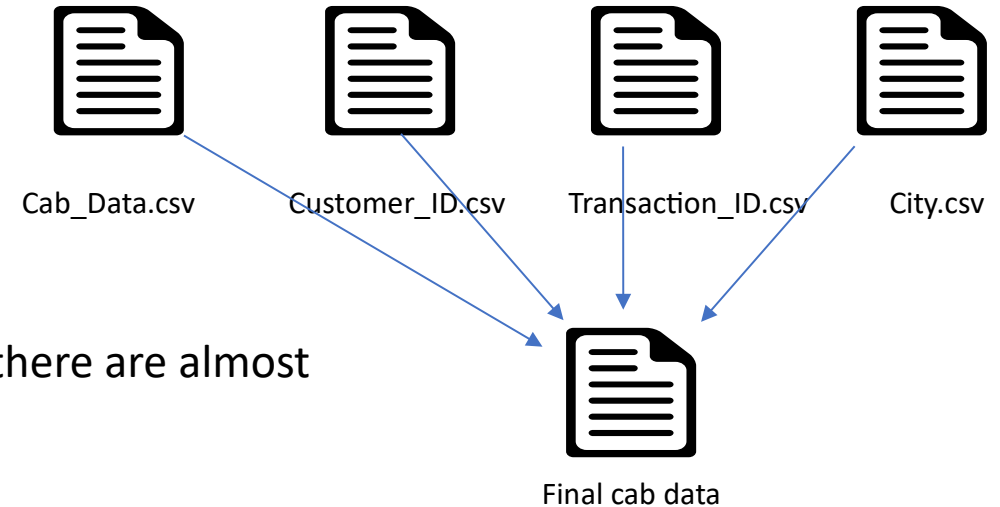
- Data Understanding
- Forecasting profit and number of rides for each cab type
- Finding the most profitable Cab company
- Recommendations for investment

# Data Exploration

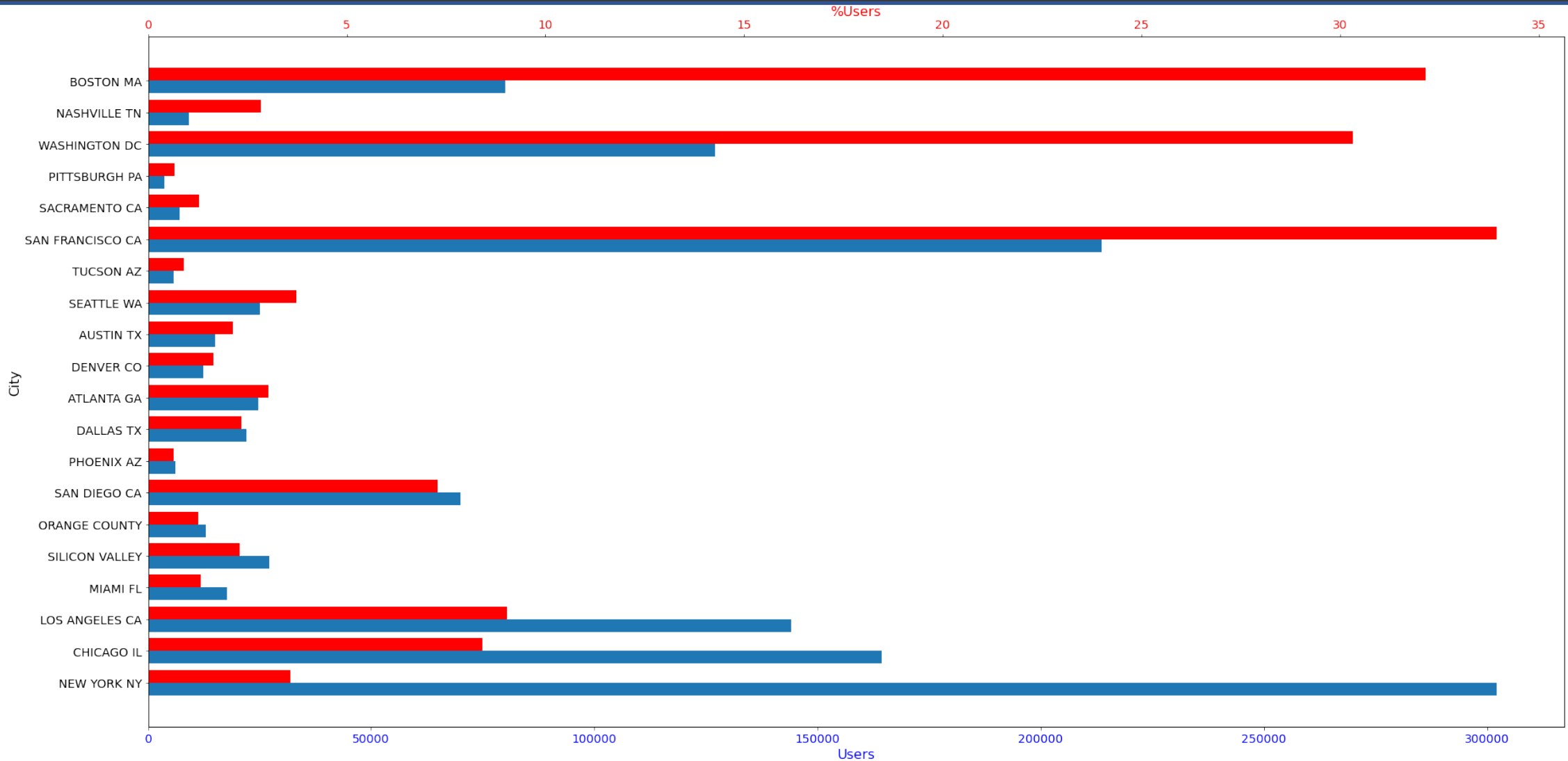
- 19 Features( including 4 derived features)
- Timeframe of the data: 2016-01-02 to 2018-12-31
- Total data points :359,392

## Assumptions:

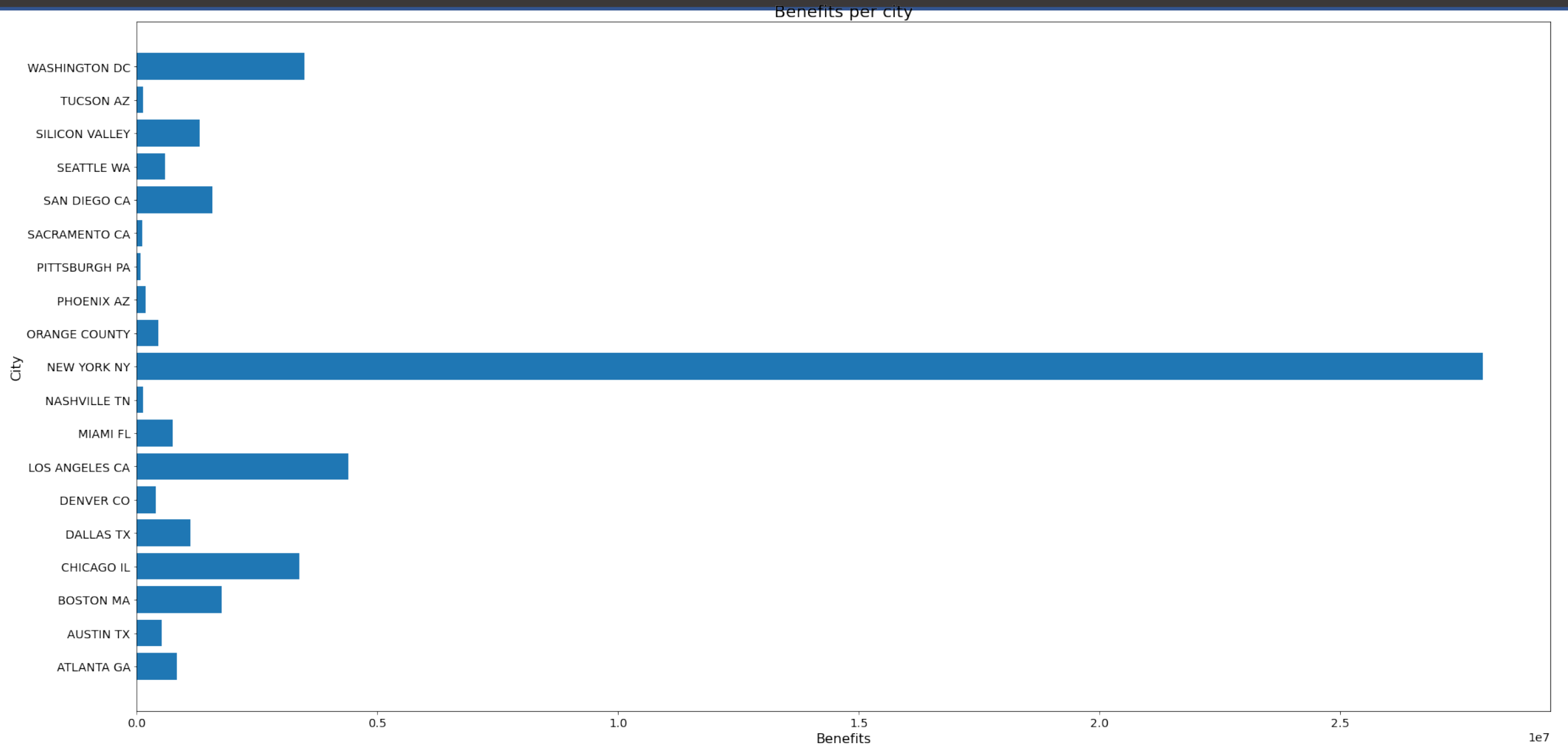
- There are more transactions in the Transaction\_ID but because there are almost no data we will ignore the rows that do not match.
- The benefits will be calculated by calculating the difference between the cost of the trip and the price charged for each trip.
- It is possible that one or two companies do not have any trip in one of the cities in City.csv.
- Not all ages and economical groups have the same behaviors in the matter of spending money in trips.
- Users feature of city dataset is treated as number of cab users in the city.  
We assume that is the sum of all users of cab companies. Therefore there are more than two companies in these cities.



# City users of Cab

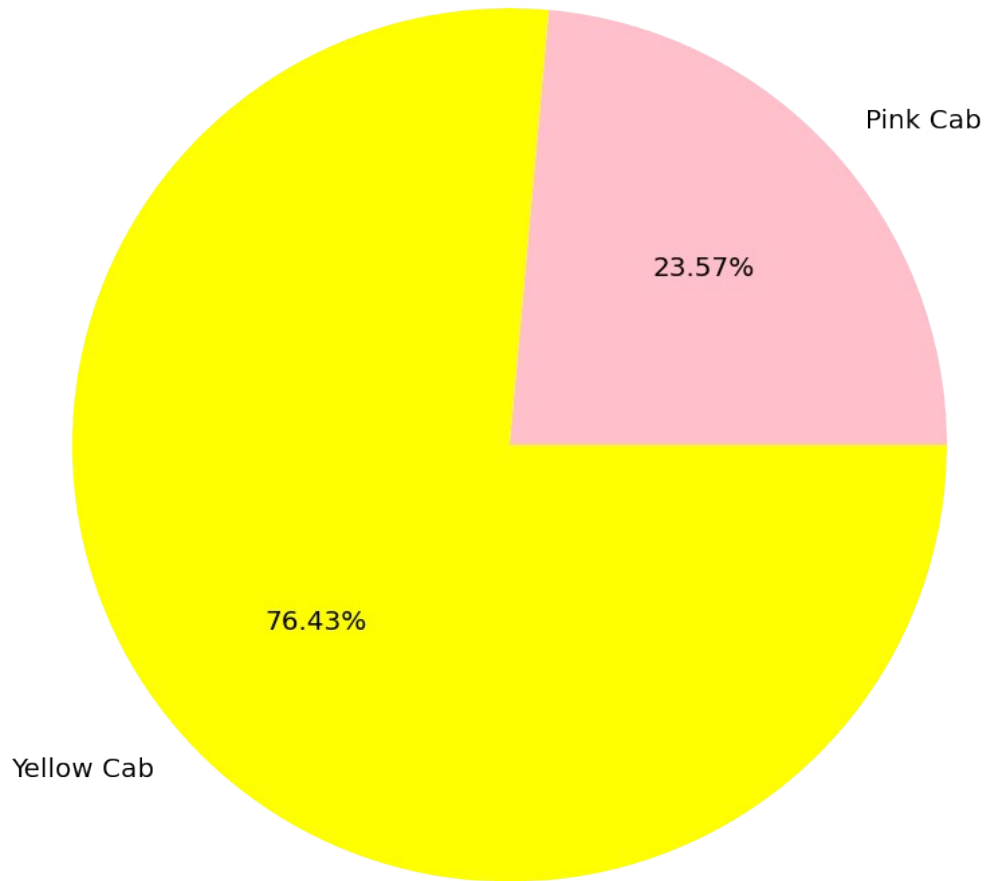


# Benefits seen in Cab Data

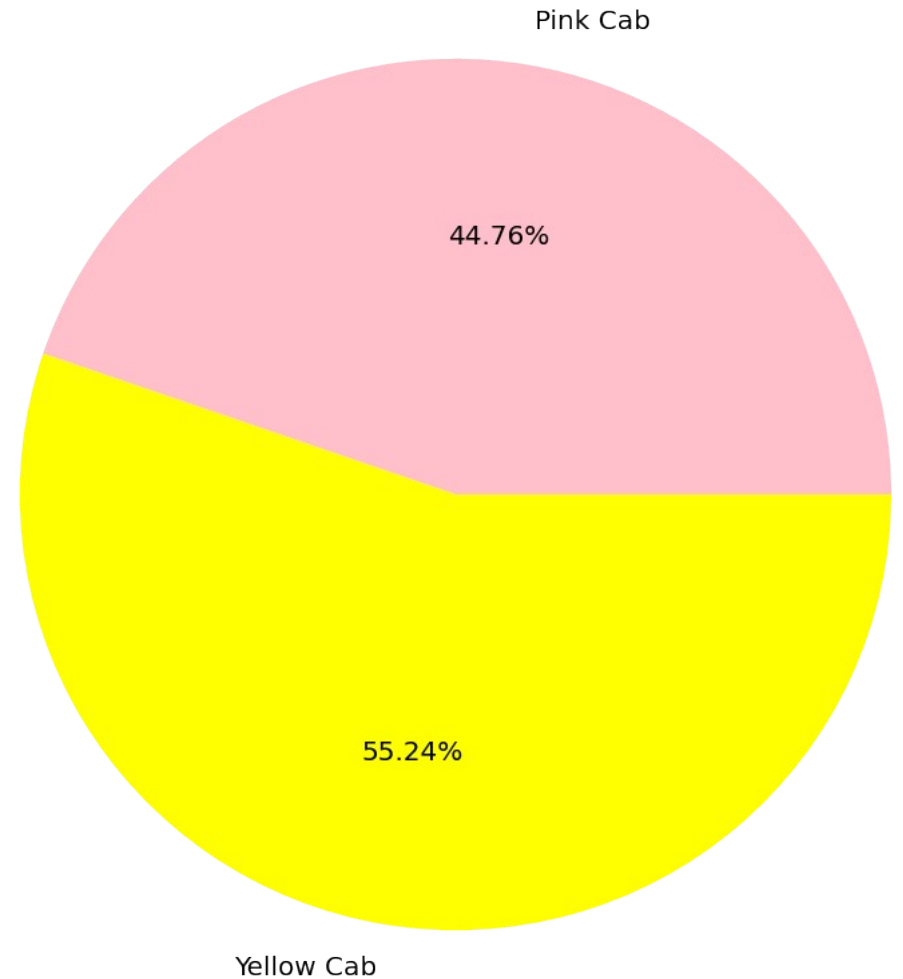


# Number of Trips per company and users

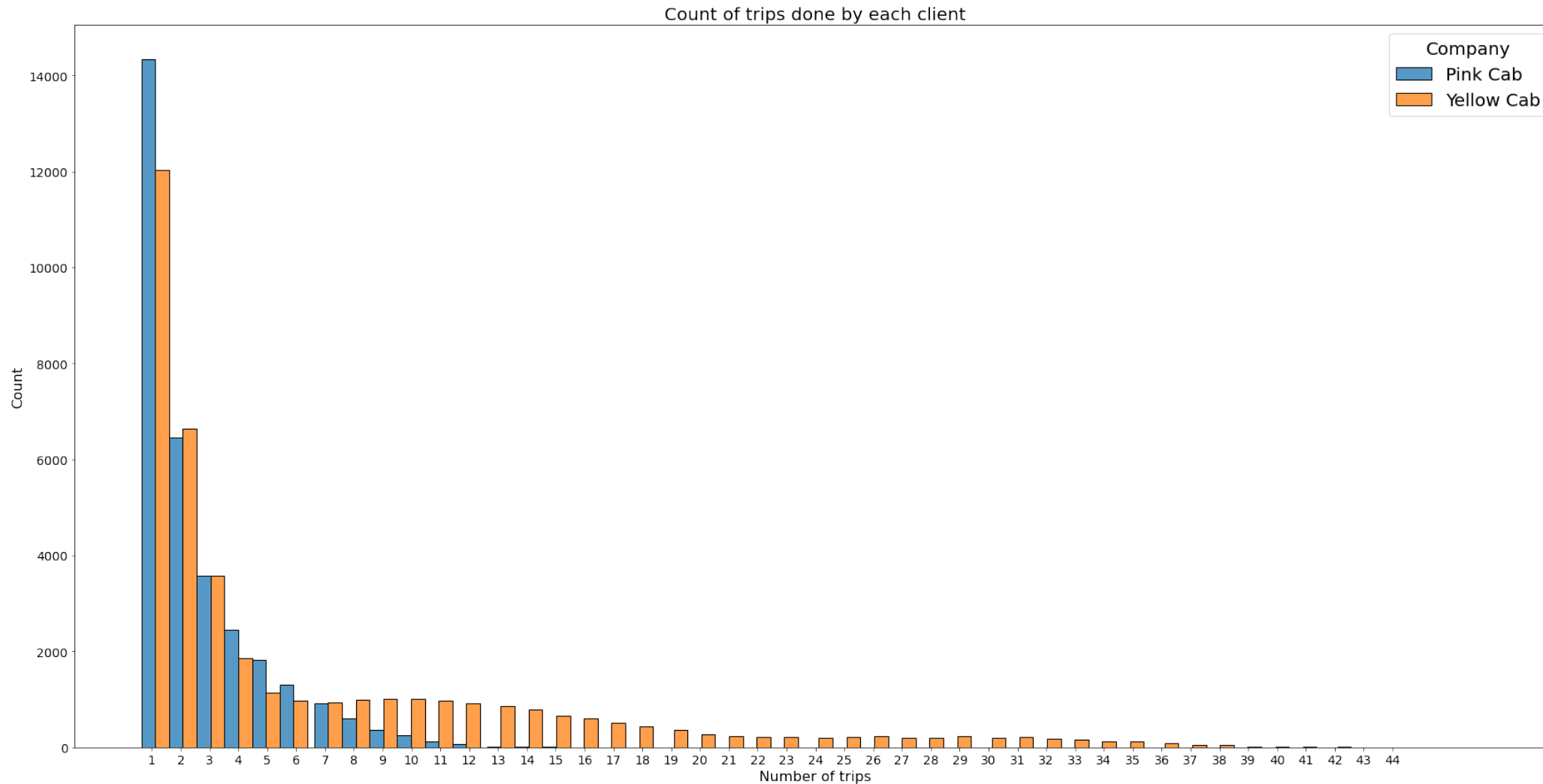
Trips per company



Users of each company

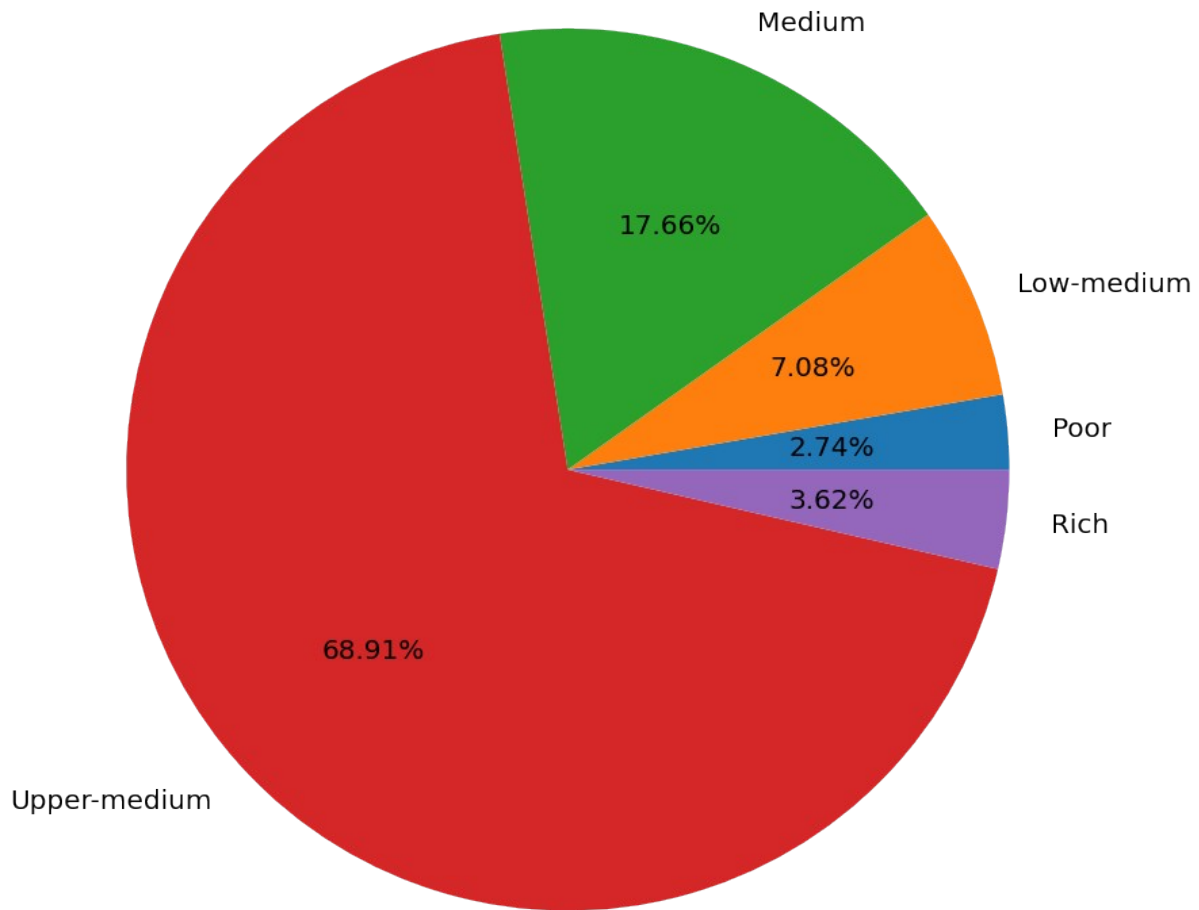


# How many people repeat trips

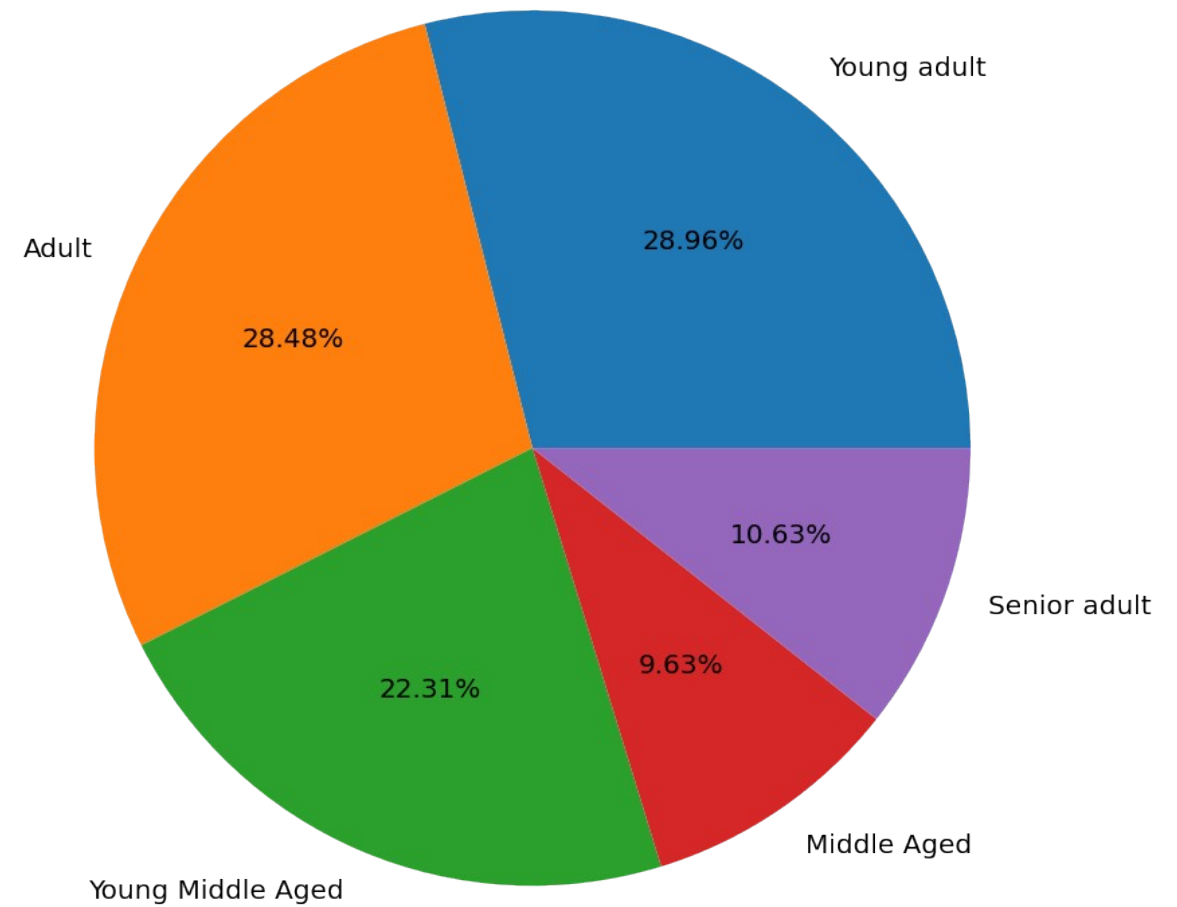


# Client Groups

Clients income group

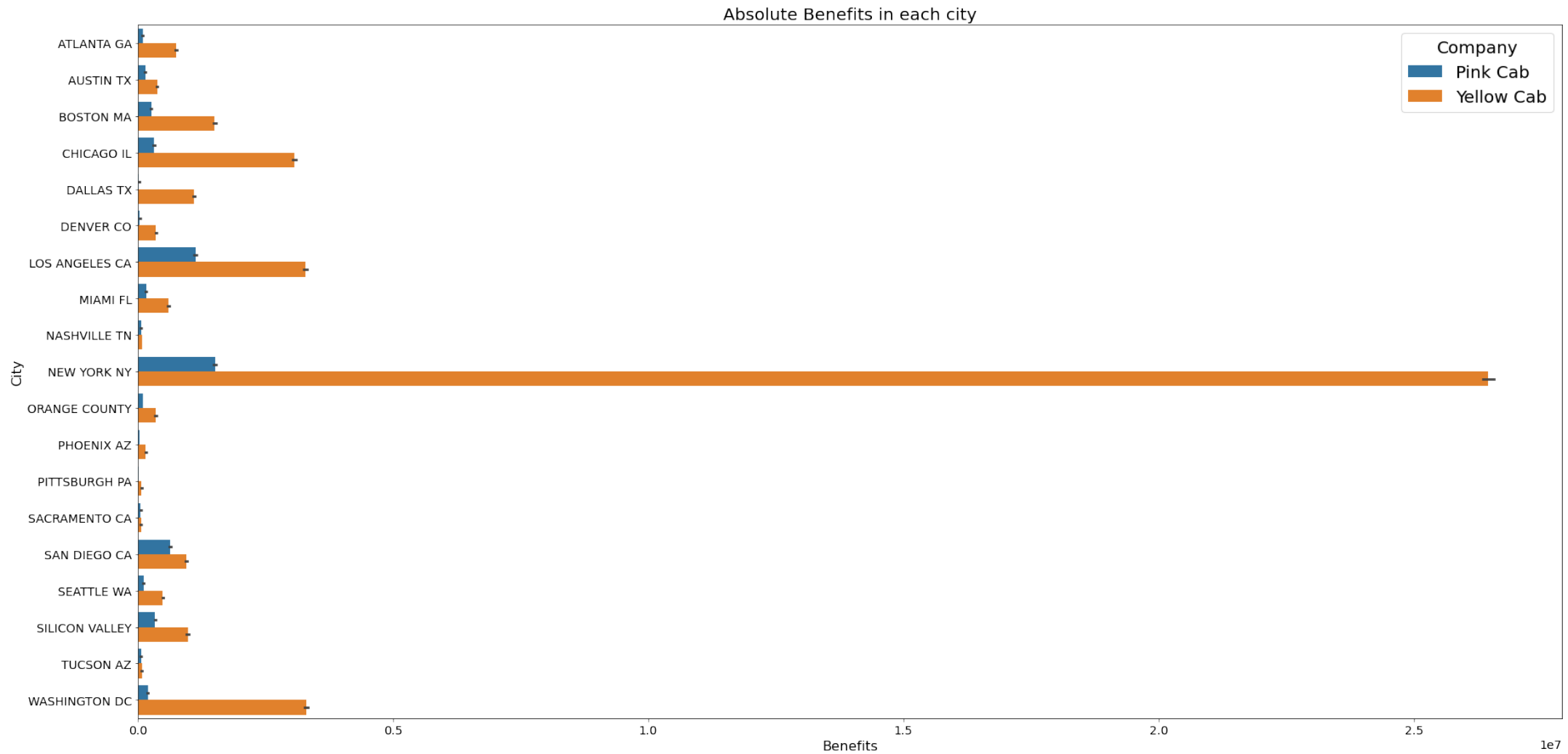


clients age groups

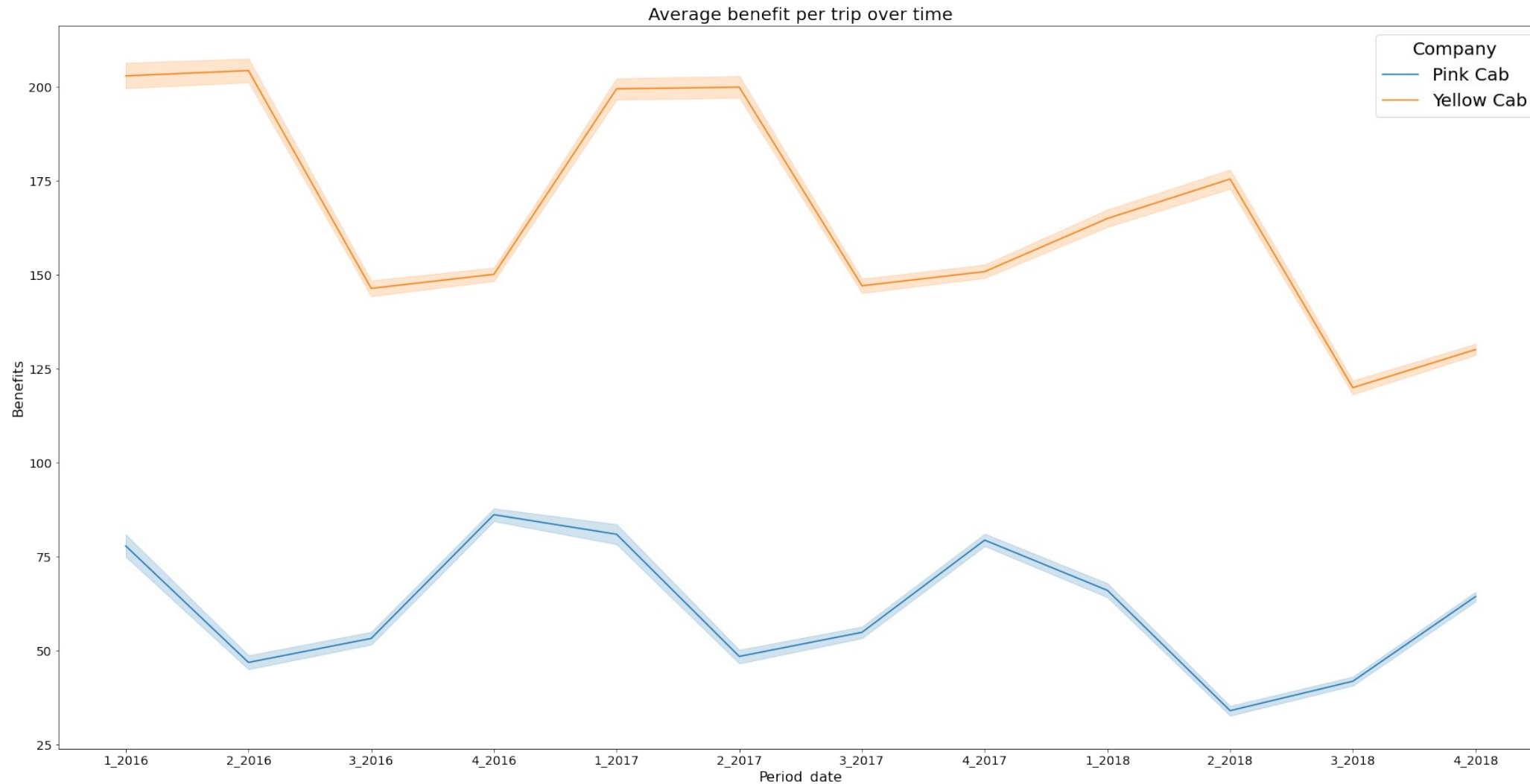




# Absolute benefits per company



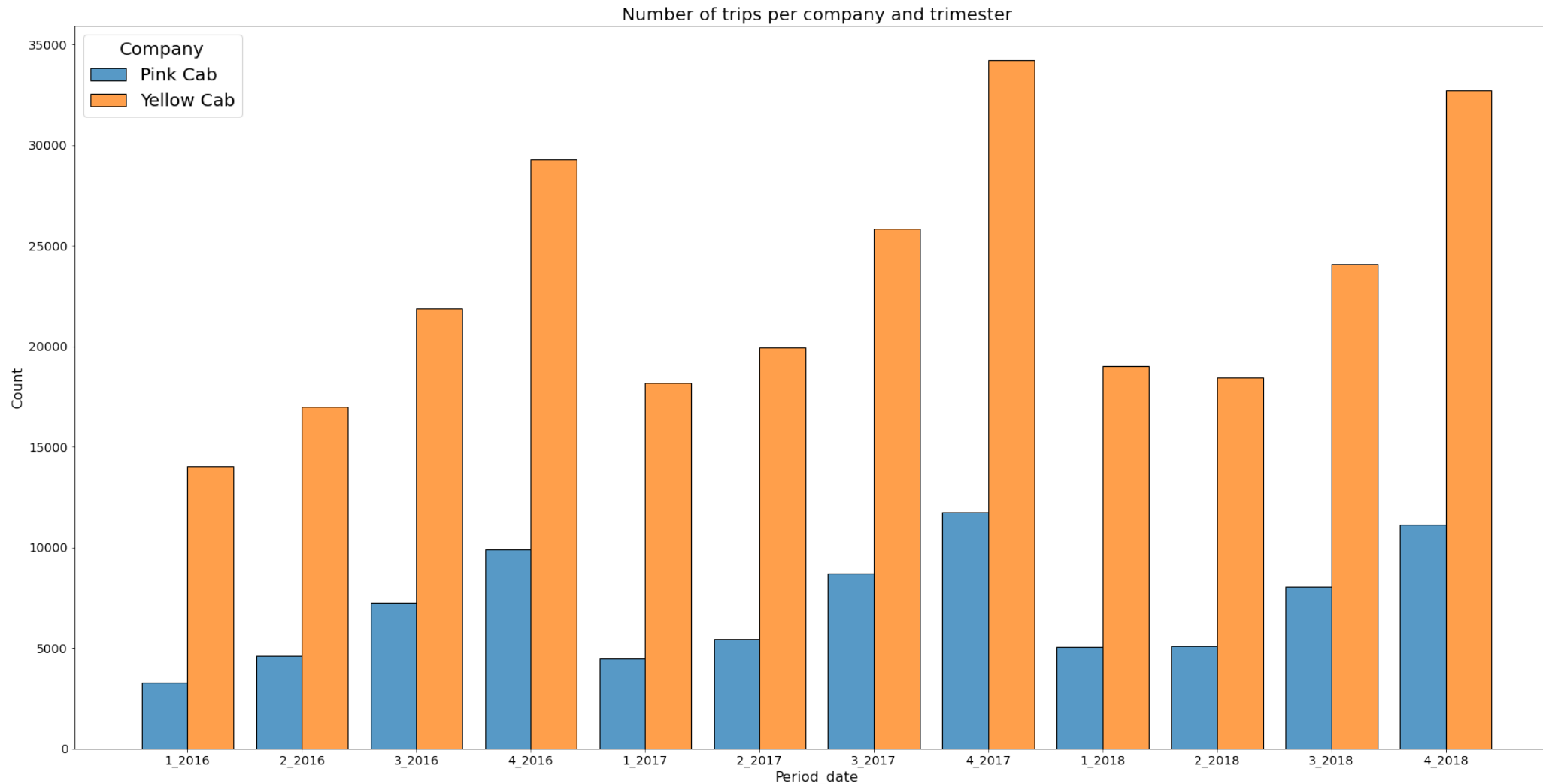
# Trimester profits



We see here that the benefits tend to decrease in the third and fourth trimester for Yellow Cab

For

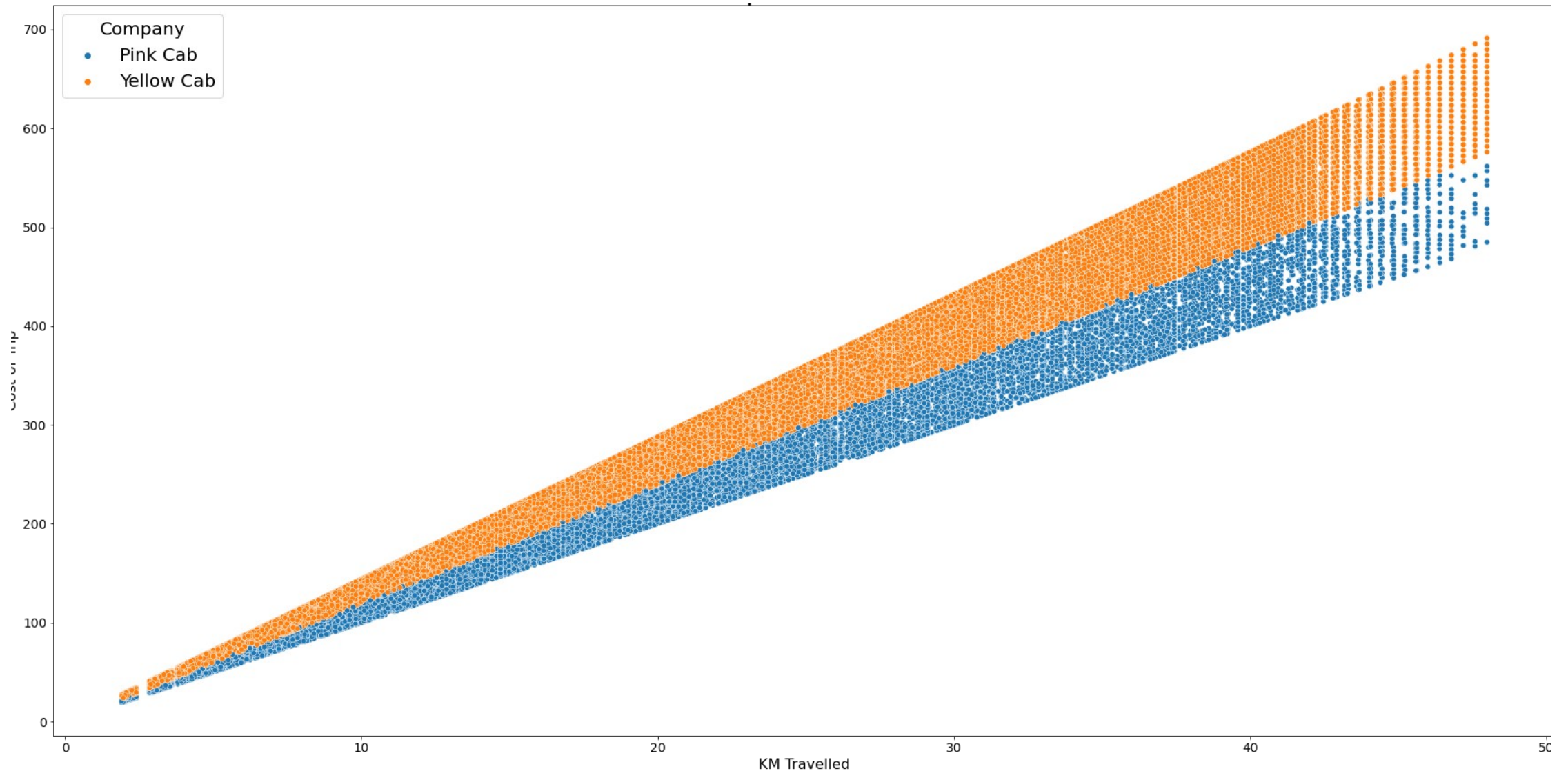
# Number of trips per trimester



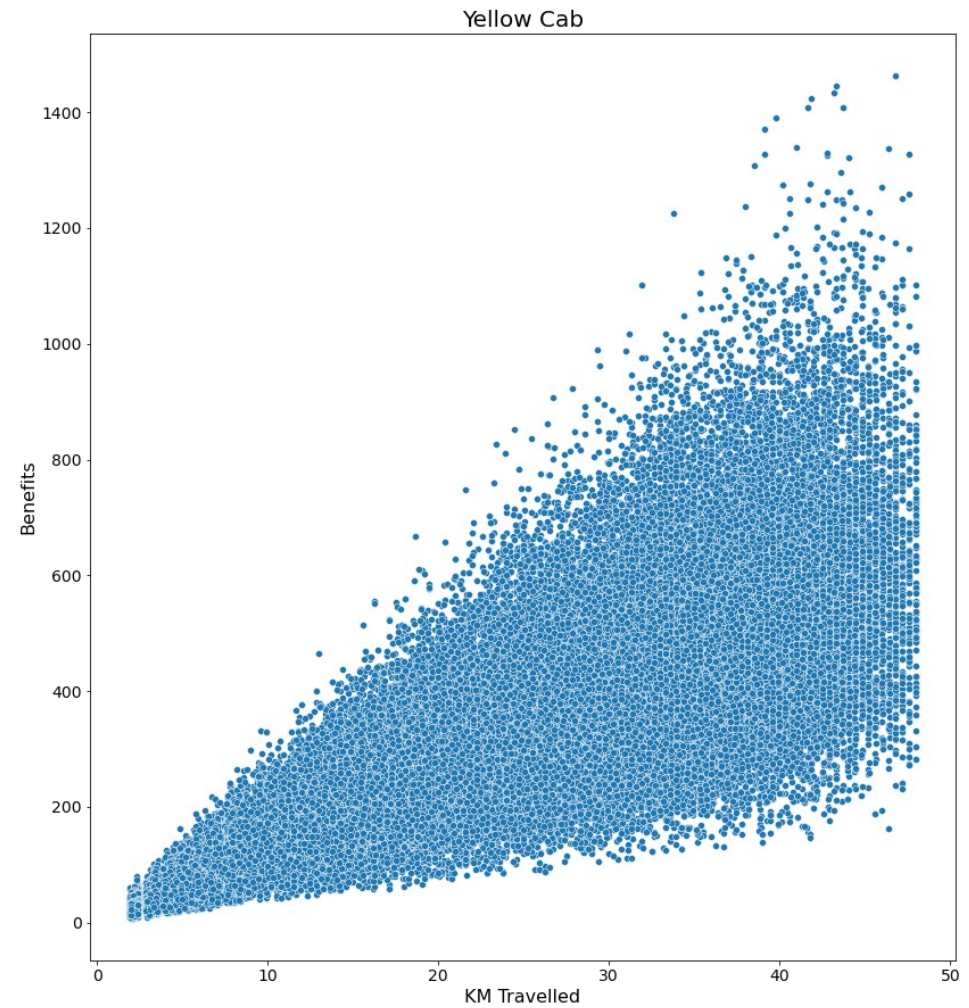
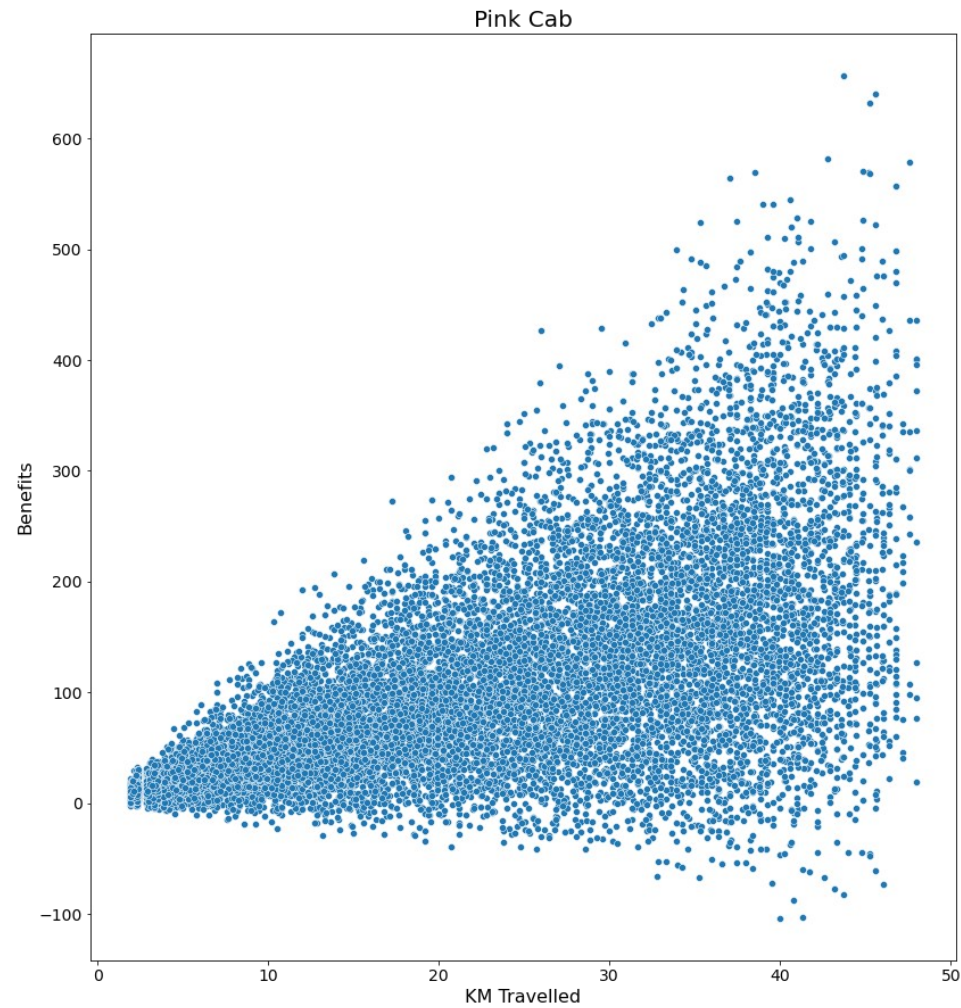
We see here that the number of trips are higher in the third and fourth trimester like I expected.

However it does not match with the benefits table.

# Example: Cost of trip per distance

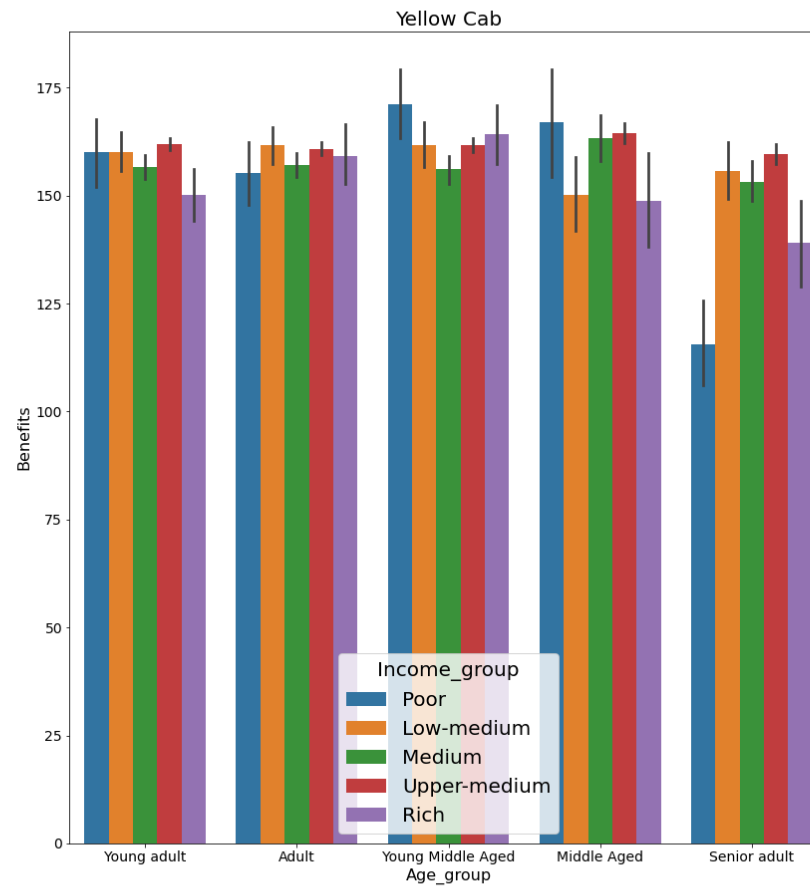
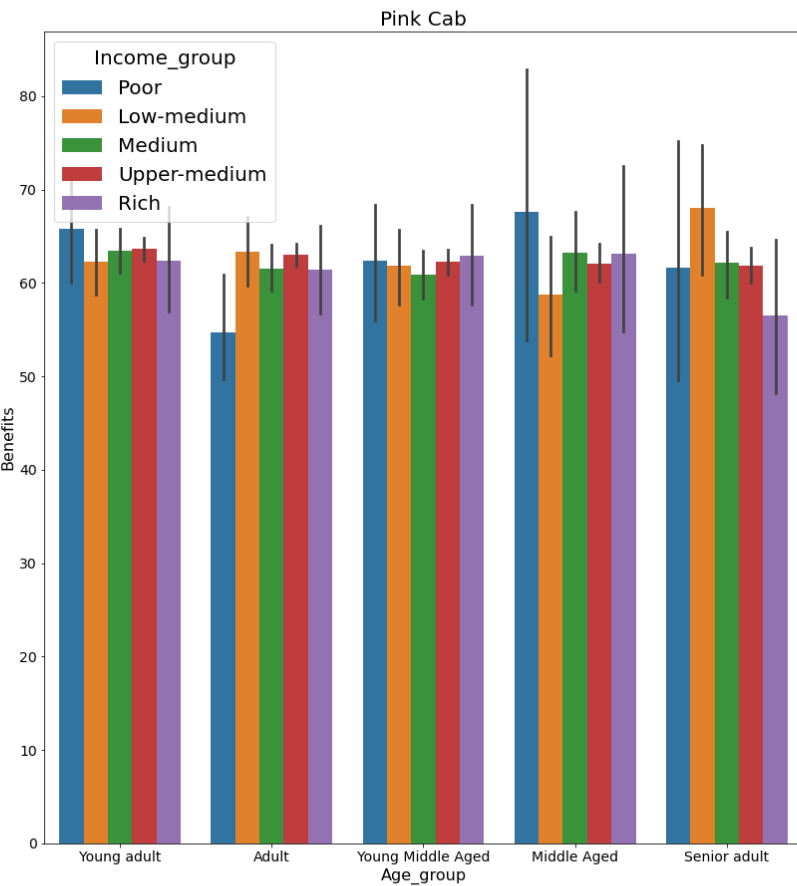


# Example: Benefits of trip per distance



# Benefits obtained by group

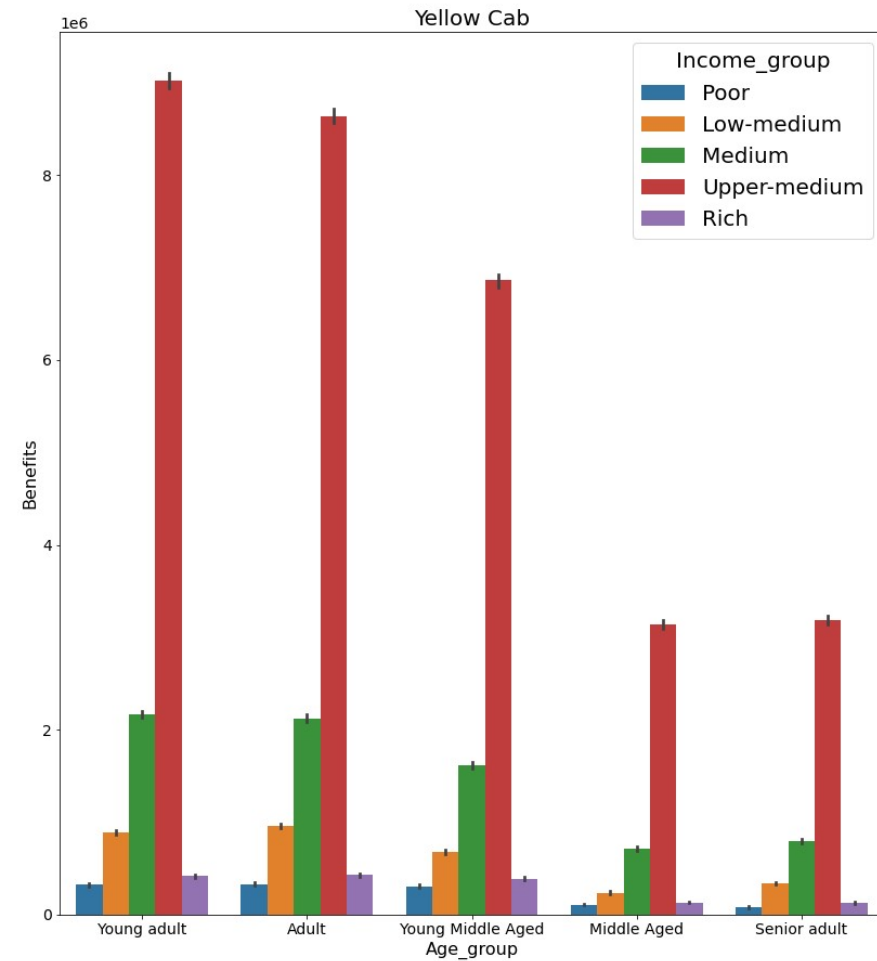
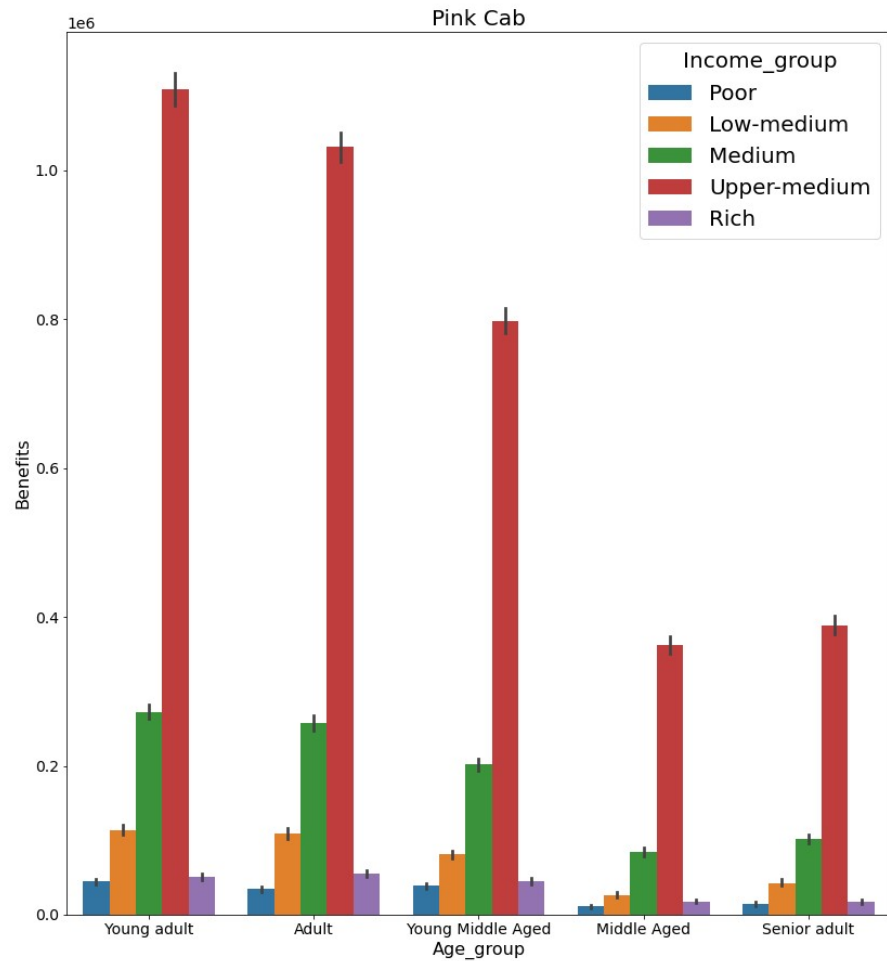
Money obtained per individual/age\_group



As we can see here. Yellow Cab obtains higher benefits from each group than the pink Cab.

# Benefits obtained by group total

Money obtained per individual/age\_group



# Recommendations

We have evaluated both the cab companies on following points and found Yellow cab better than Pink cab:

- **Customer Reach** : In general, we see that Yellow Cab has more clients than Pink Cab. This is specially visible because the cities where they have more clients tend to be the ones more populated
- **Customer Retention:** We have seen that most clients tend to return the yellow cab company while most of the users
- **Age wise Reach** : In both cases most of the clients are young people (between 16 and 35 years old) with a high economical income. This is logical as most of them will not be married neither will have saved enough money for
- **Average Profit per KM:** Yellow cab's average profit per KM is almost three times the average profit per KM of the Pink cab.
- **Income wise Reach** :Both the cabs are very popular in high and medium income class but here also Yellow cab is performing better than Pink cab in offering their services to all the three income class group (low, medium and high)

**On the basis of above point , we will recommend Yellow cab for investment.**



# Thank You



**Data Glacier**

Your Deep Learning Partner