



**Data Glacier**

Your Deep Learning Partner

# G2M CASE STUDY

XYZ company analysis

Erwin Linares

07/15/2023

# Agenda

Problem Statement

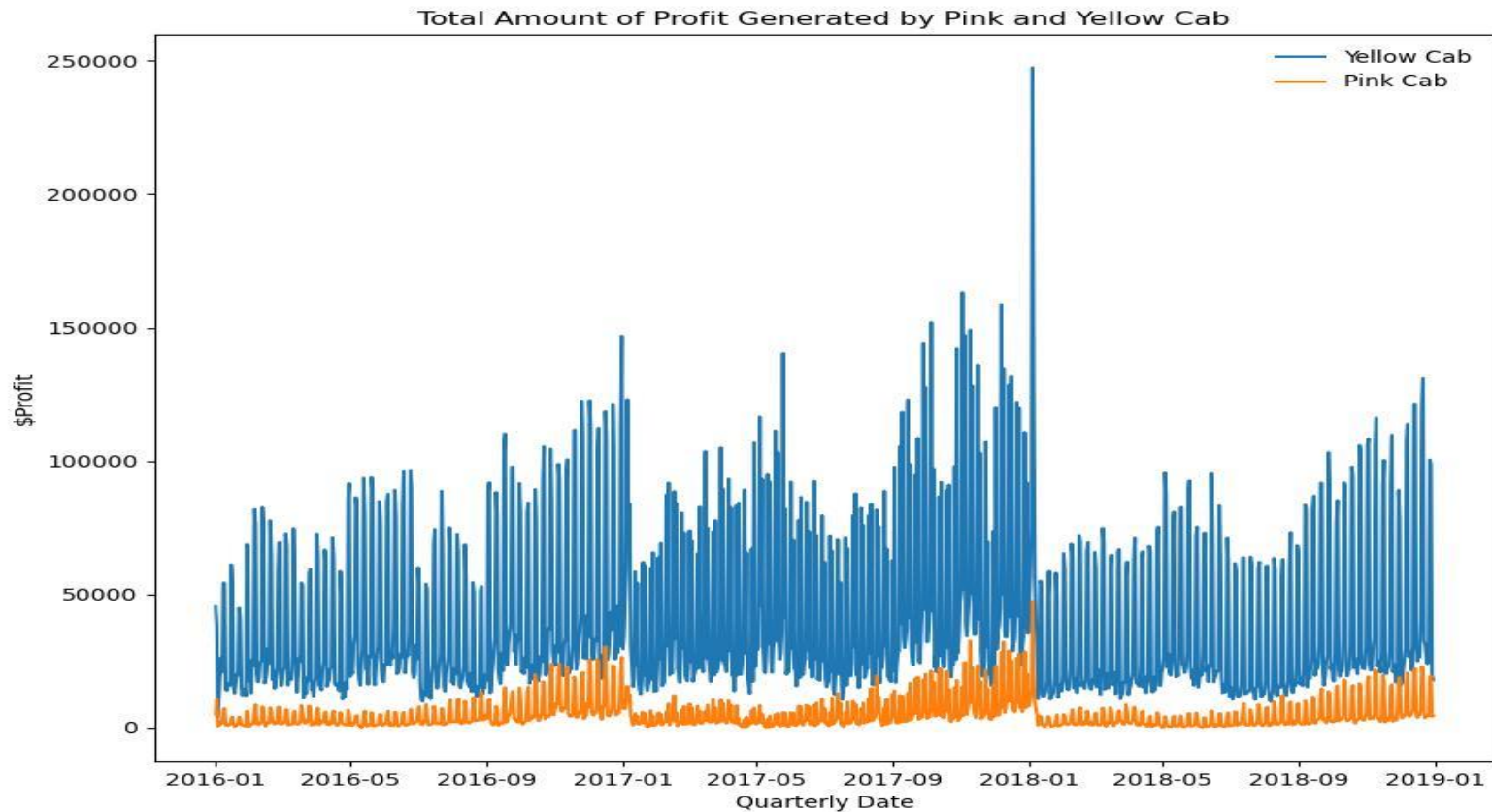
EDA Summary

Recommendations

# Problem Statement

- Company XYZ is a private firm in the US , due to a growth rate of the taxi cab market they are considering investing in the market . Their Go-to-Market(G2M) strategy means they want us to carry out a data analysis and help them come up with a final understanding.

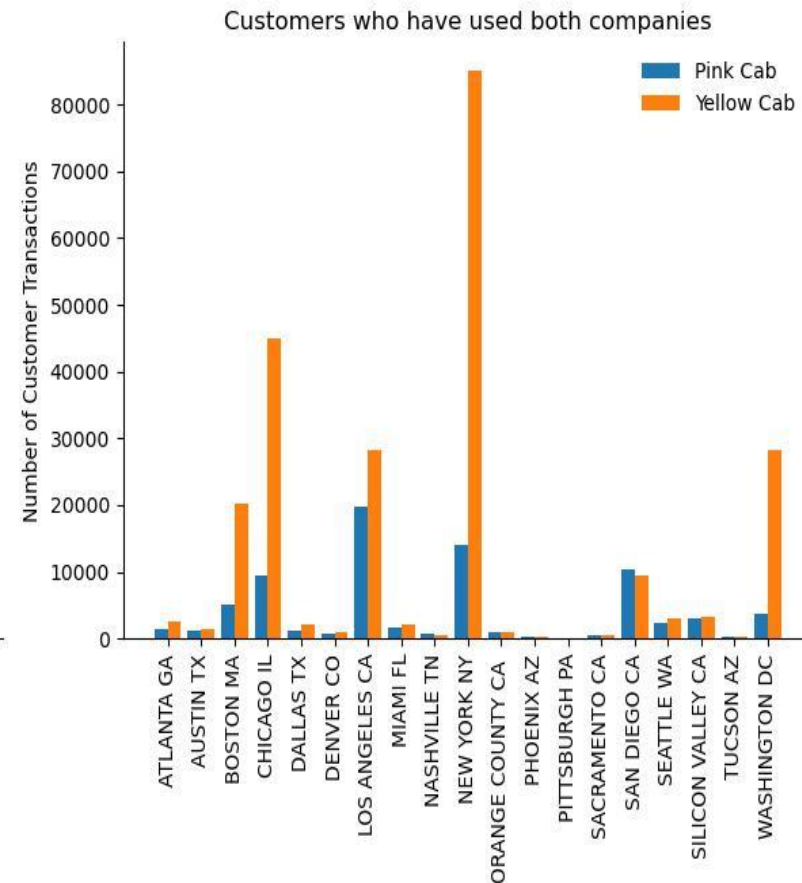
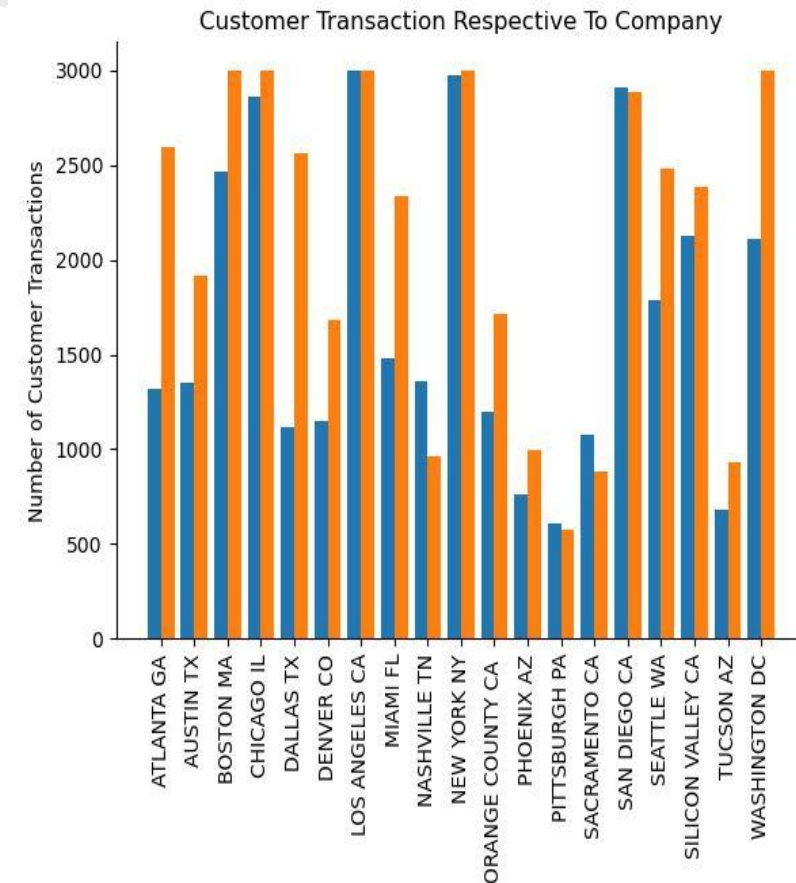
# Profit Analysis



Within the Graph that depicts the total amount of profit generated by Pink and Yellow Cab we can see that:

1. There is a clear dominance by Yellow Cab
2. There are signs of seasonality shown here where towards the end of the year both companies see an increase in sales.

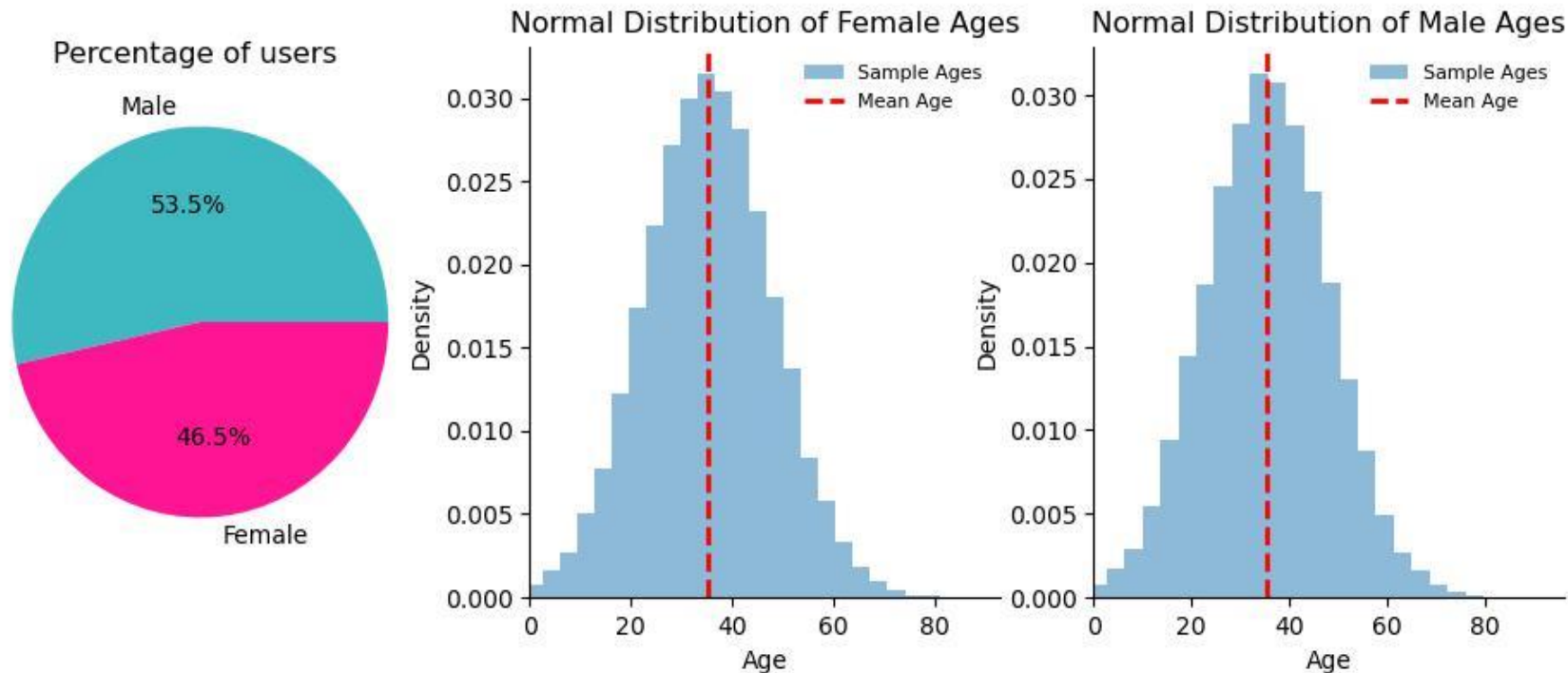
# Customer Retention Analysis



The left side of the visualization shows the customer transactions for Pink Cab and Yellow Cab in a city. However, this doesn't reveal the true customer preference as it only considers the total transactions for each company without accounting for factors like pricing and brand loyalty.

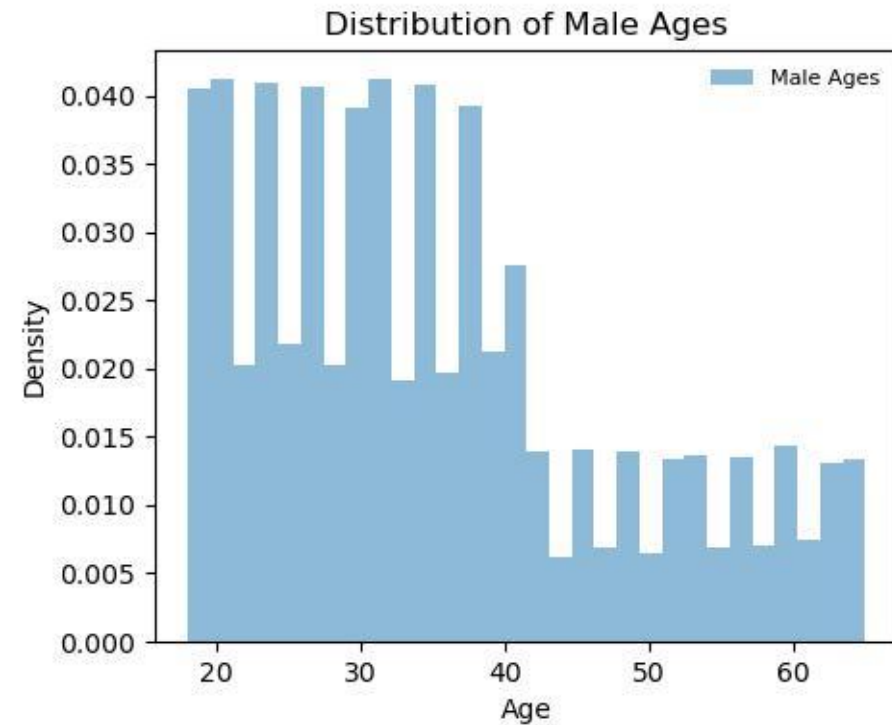
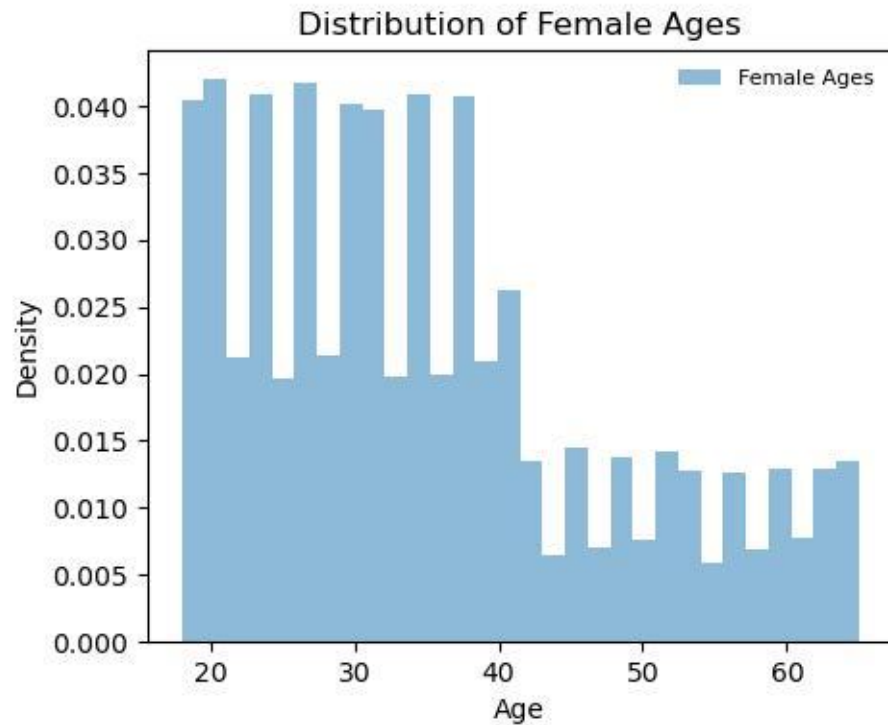
The right side provides a more insightful view by showing customers who have used both companies and which one dominates their transactions. Yellow Cab appears to have a significant advantage in about 5 cities, but the reasons behind this dominance are not clear due to the limited data. Possible factors could be larger demand in those cities, higher mean income, or better hospitality and luxury.

# Customer Demographics

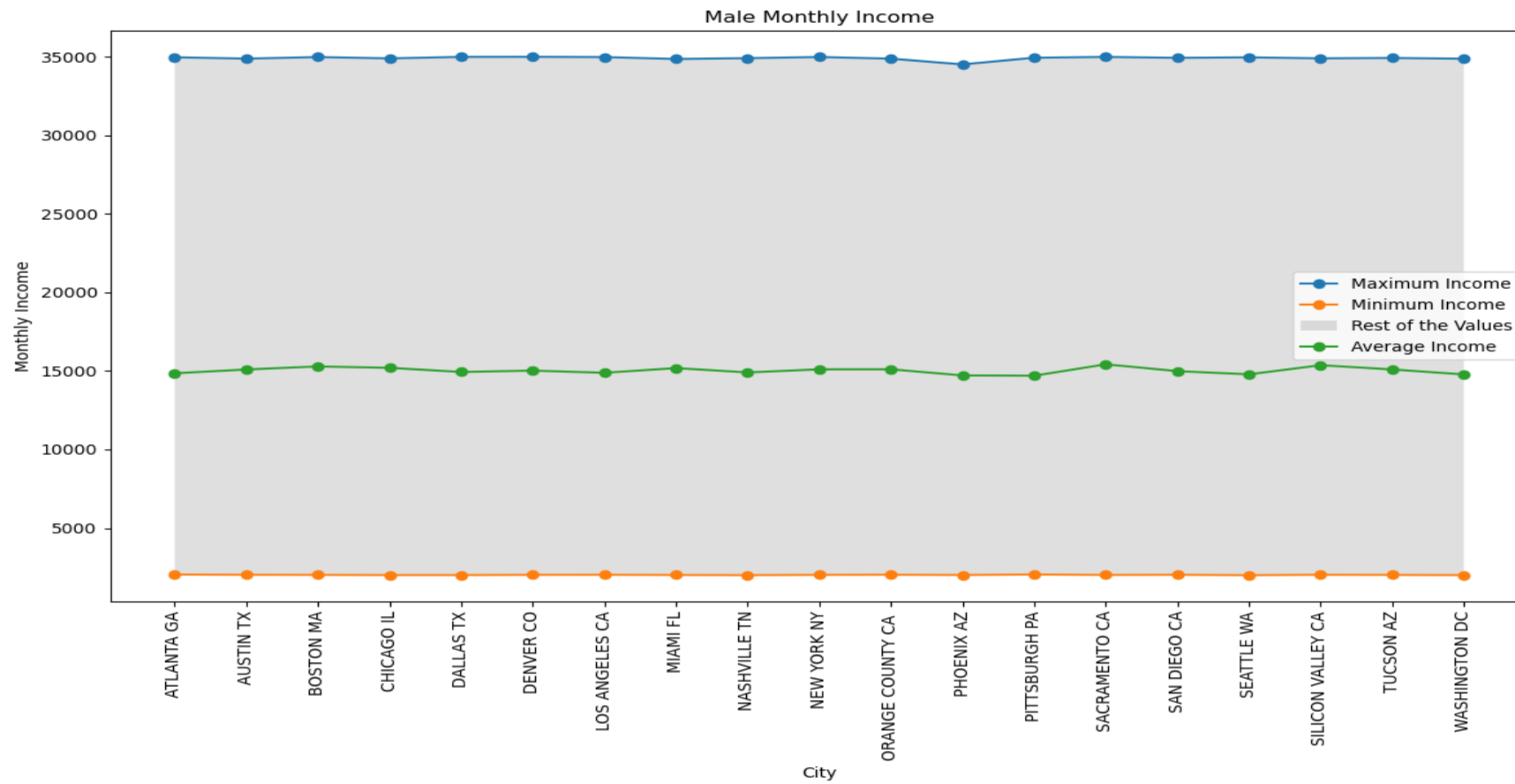


In the above graphs the graph on the left shows the percentage of users that are male and female, the graphs in the middle and right show the normal distribution of both male and female user to get an idea of what the target market is like

# Actual Distribution of User Ages

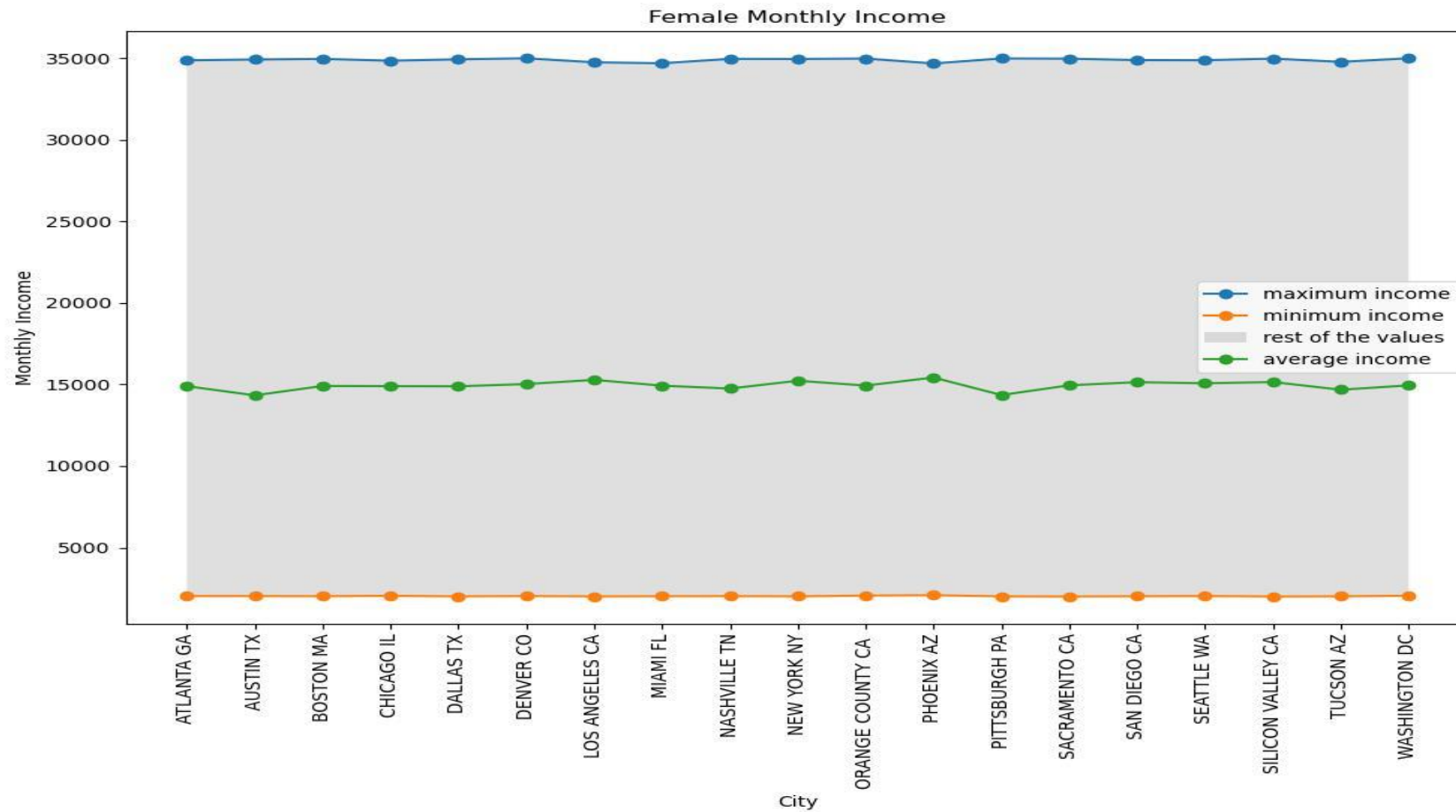


# Male Monthly Income By City

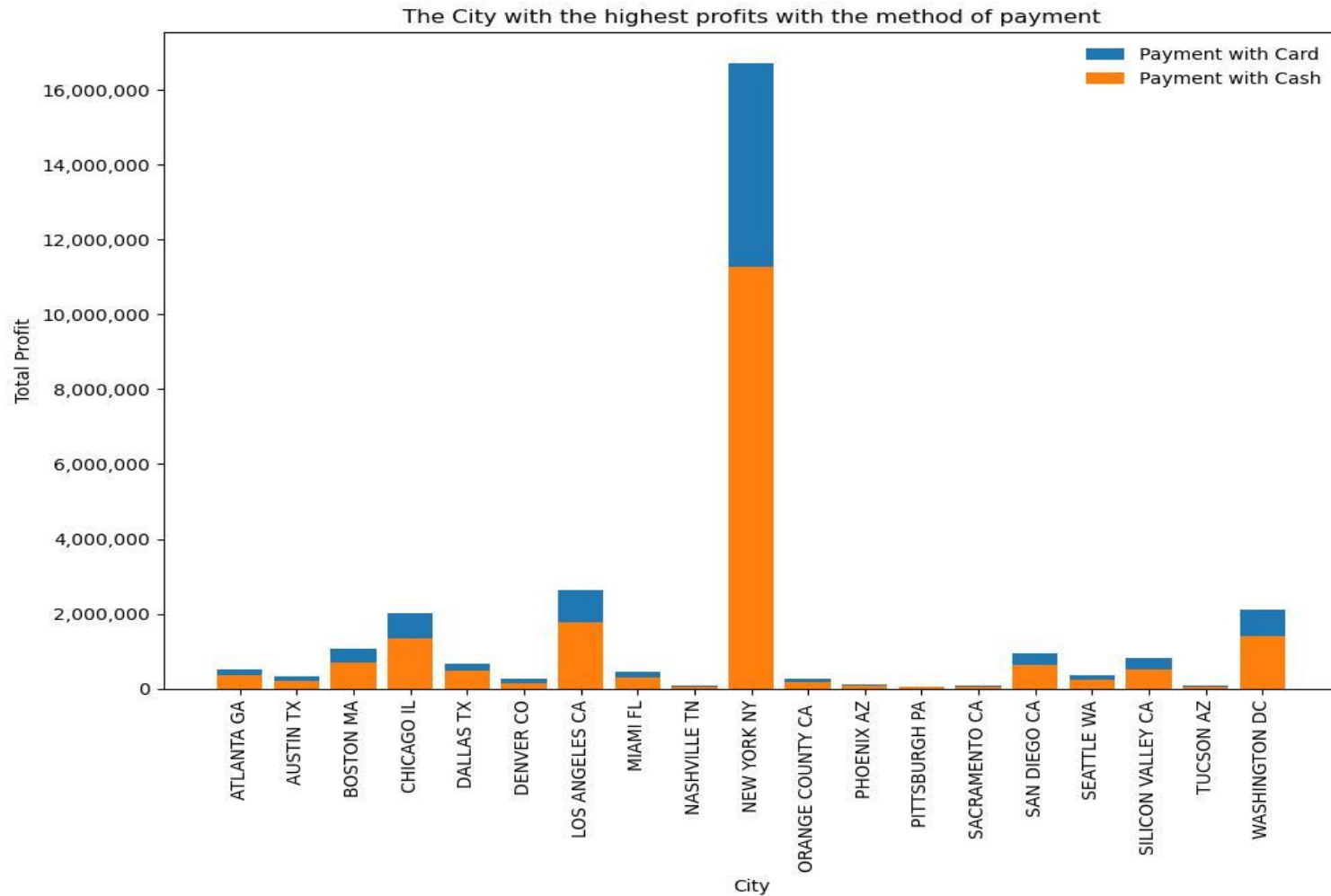




# Female Monthly Income By City



# Preferred Method of Payment



Here we see:

1. Highest total profit by city
2. Also most notably the preferred method of payment being with card.

# summary of the data and what to do moving forward

In conclusion :

- after analyzing the data set it is apparent that the target demographic should be male and female users between the ages of 20 and 50. The Taxi cab fair service could also slide into the middle class of the taxi cab companies because people don't mind paying high amounts for yellow cab and in fact most prefer them over pink cab based on the fact that they have more users. Offering a middle ground price charged could give the new taxi cab service an incredible amount of market share and help reach breakeven faster.
- My suggestion is to make it so that users with an average income of around 1300 monthly would suffice.
- The method of payment should definitely accommodate those that pay with card so having technology that could accept such payments because card is such a prominent method of payment and preferred over cash even in most cities.
- As for the city market to enter, if the taxi cab business model pursues a lower price cost than yellow cab and is able to offer what they offer and still differentiate the taxi cab company identity/brand then I suggest entering the Boston, Chicago, Los Angeles, New York, and Washington cities and trying to take away customers from Yellow Cab. However no matter what market we enter it is still risky because there are still unanswered questions about customer behavior. So the preferred state would be to obtain more data on this.
- It is a good idea to also look into the San Francisco market as well considering there are a considerable amount of users it could be a potential new market that has been left rather untouched.

# Summary of Analysis

How I went about my analysis was first getting to know the data set by looking at all data sets for this analysis, seeing what relationships there were and mapping them out. Then I matched the context of the problem with the facts of the data for example the date wasn't saved correctly so that had to be changed or the distance was in kilometers but our setting was the United States so that had to be changed to miles. I also checked for null values and anything else that need to be dealt with during the data cleaning phase. Then I wrote down some predicted hypothesis or things that just from a glance I could see a trend going on. After that I visualized each one. This is a short summary of how the analysis happened and what were the steps I took.

# Thank You