



Data Glacier

Your Deep Learning Partner

G2M Case Study

Virtual Internship

2-8-2023

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Executive Summary

This presentation is for week 2 of Data Glacier Internship "G2M insight for Cab Investment firm" project, where I conduct an in-depth market analysis for XYZ, a private firm in the US, planning to invest in the Cab Industry. With significant growth in the Cab Industry and numerous key players in the market, XYZ aims to make an informed decision through actionable insights derived from multiple data sets.

Throughout the week, I deeply investigated the data, exploring its structure, conducting data cleaning and validation, and identifying patterns and relationships between different variables. This process involves the identification and handling of duplicates, handling missing data, and ensuring data quality and integrity.

Problem Statement

The Client:

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 and as per their Go-to-Market(G2M) strategy they want to understand the market before taking final decision.

Project delivery:

I have been provided with multiple data sets that contains information on 2 cab companies. Each file (data set) provided represents different aspects of the customer profile. XYZ is interested in using your actionable insights to help them identify the right company to make their investment.

The outcome of my delivery will be a **presentation to XYZ's Executive team**. This presentation will be judged based on the visuals provided, the quality of your analysis and the value of your recommendations and insights.

Approach

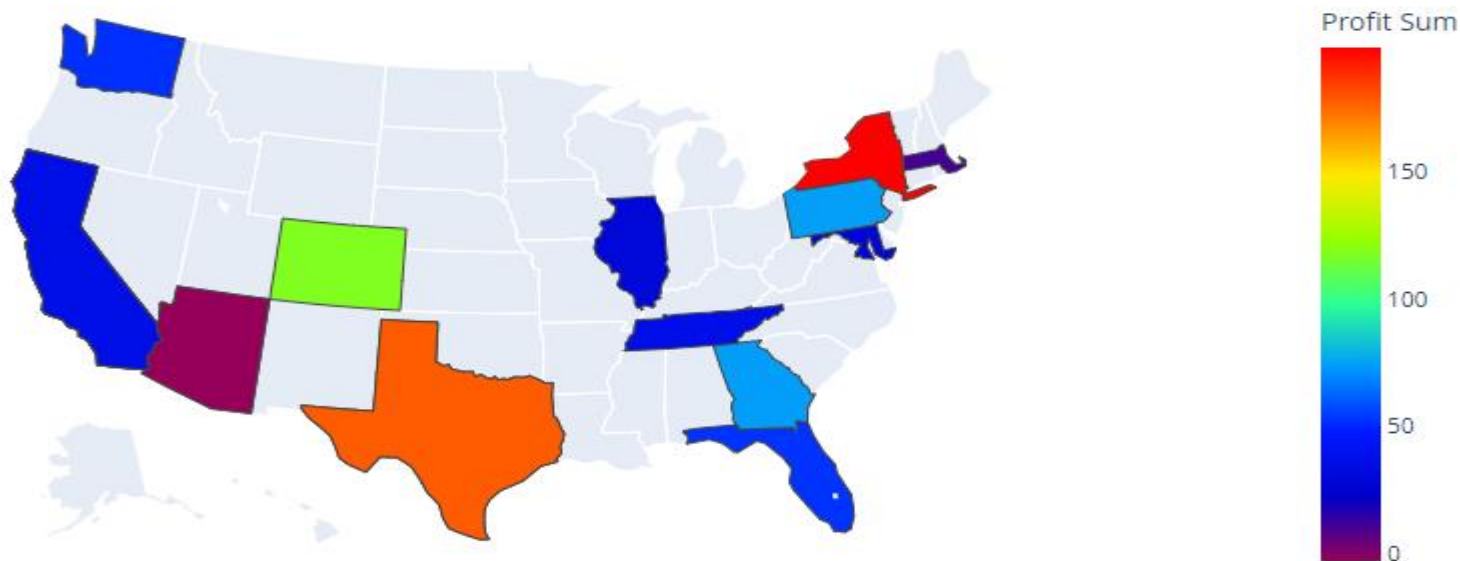
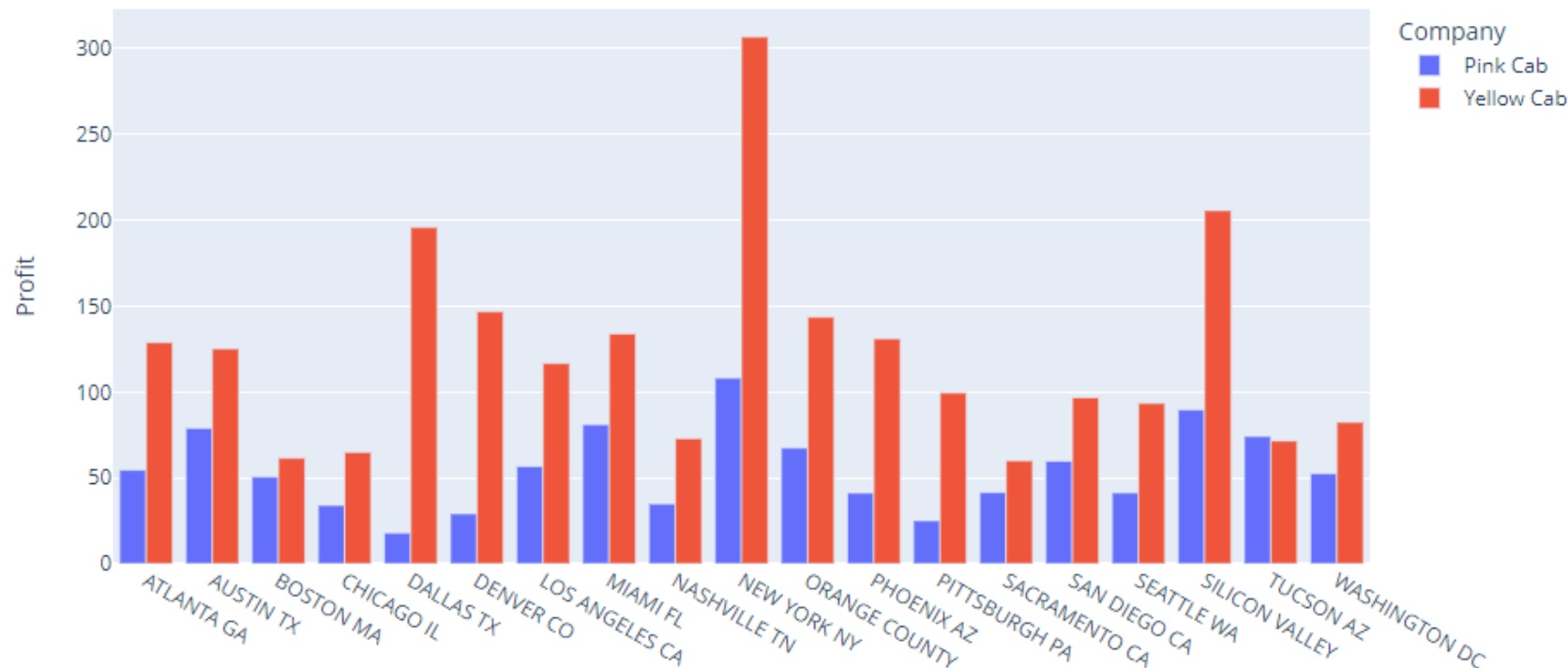
The approach (more detailed inside the notebook) to give insights both in marketing and company terms consist on answering:

1. Is there any crucial correlation between average/total profit with city?
2. Is the percentage impact of daily demand significant to the price per KM?
3. What's the correlation between age, gender and profit income for both companies?
4. What's the overall predictability of User's binary choice given trip and personal information?
5. What's the propensity of Users for features distance, income, day of the week and payment option as probability distribution?
6. What's the probability for positive profit at both companies?
7. How much scattered is cab demand given time periods?
8. What's the overall demand between each cab company?

Hypothesis 1

EDA

Given that no city has a positive weak correlation of 0.30 or within a 0.5 error margin with the profit sum, and supported by graphical information, **New York** has more than 56% of the profit while **Yellow Cab** has the majority of the overall cab industry profit.

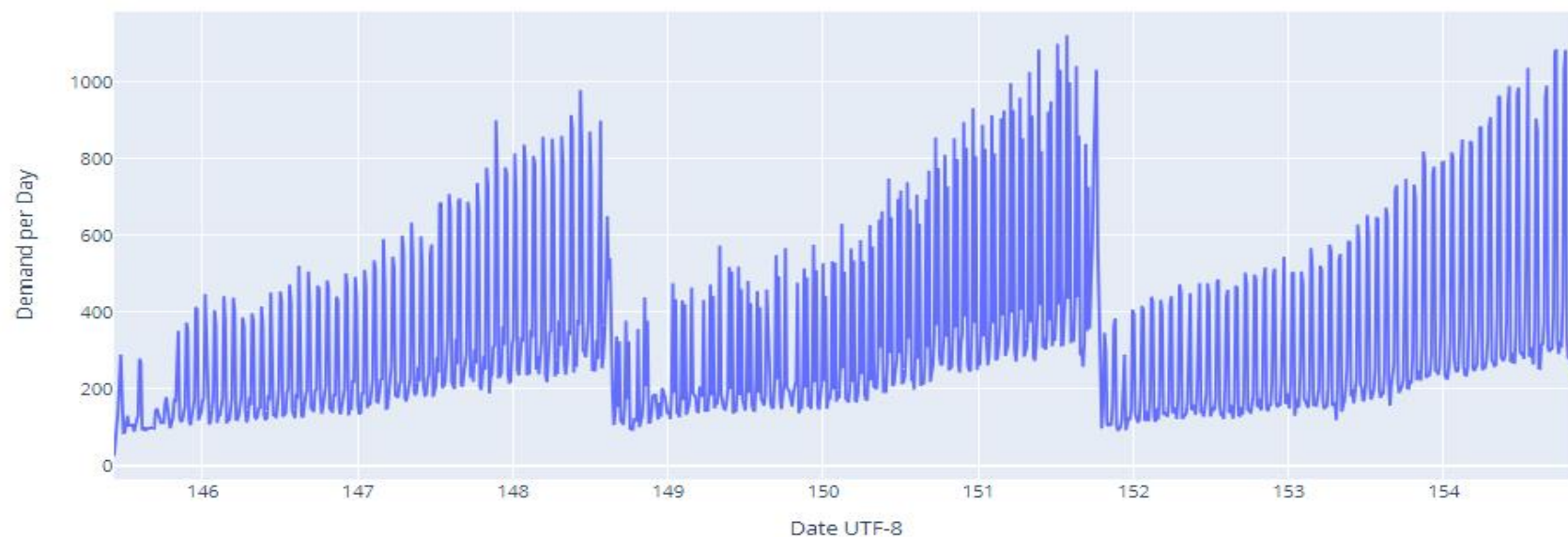
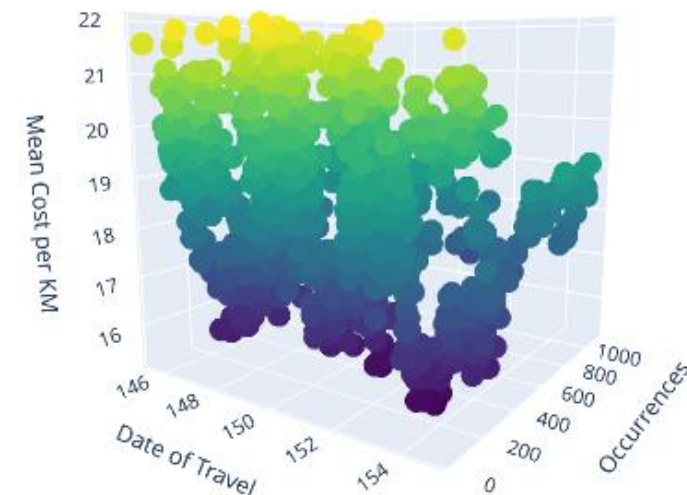
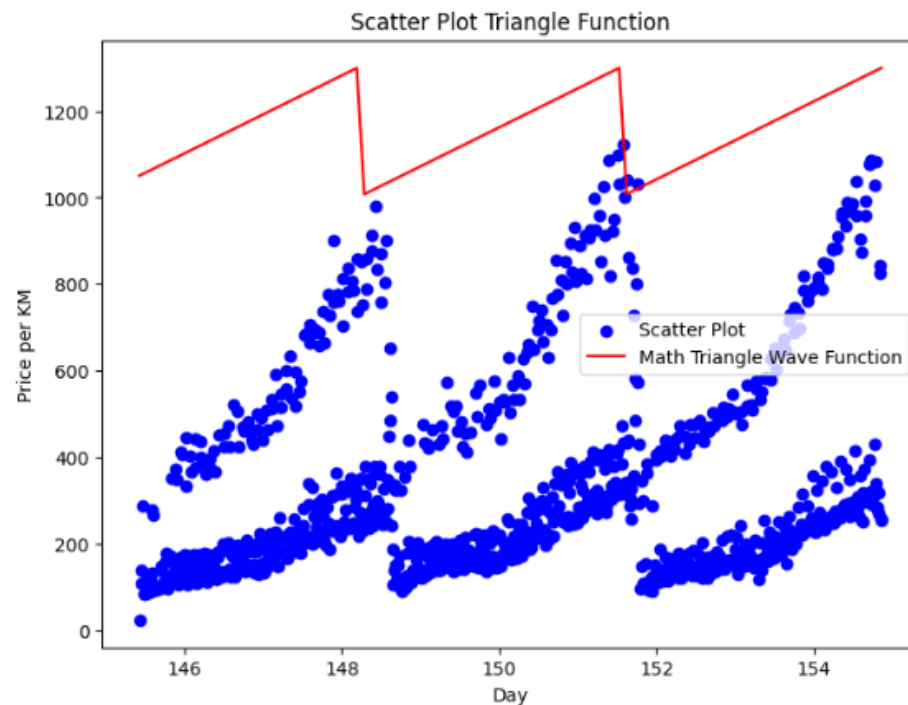


Hypothesis 2

EDA

During growing seasons, the daily demand every (Calendar days) fluctuates every other day at the same rate (11%).

There is not predictable price per km given only daily demand, however, for daily demand over 432 rides, minimum price per km is 17.

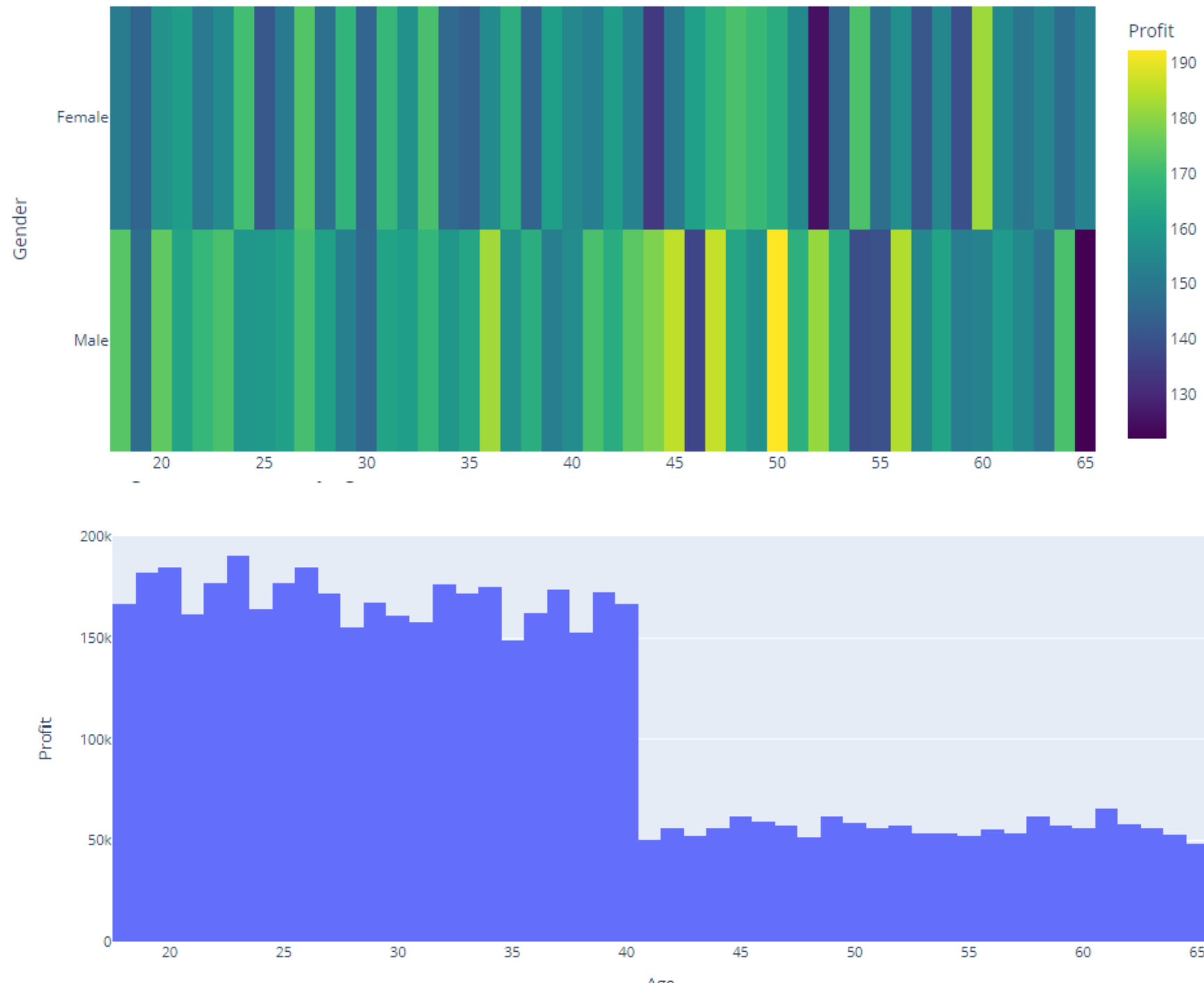


Hypothesis 3

EDA

Male users for both companies tend to give more profit average than females by 4%

Ages from 18 to 40 gives an average of 14M Profit for Yellow Cab and 170K for Pink Cab, for both there is an abrupt cut-off to 35%-40% when users turn 41 years old.

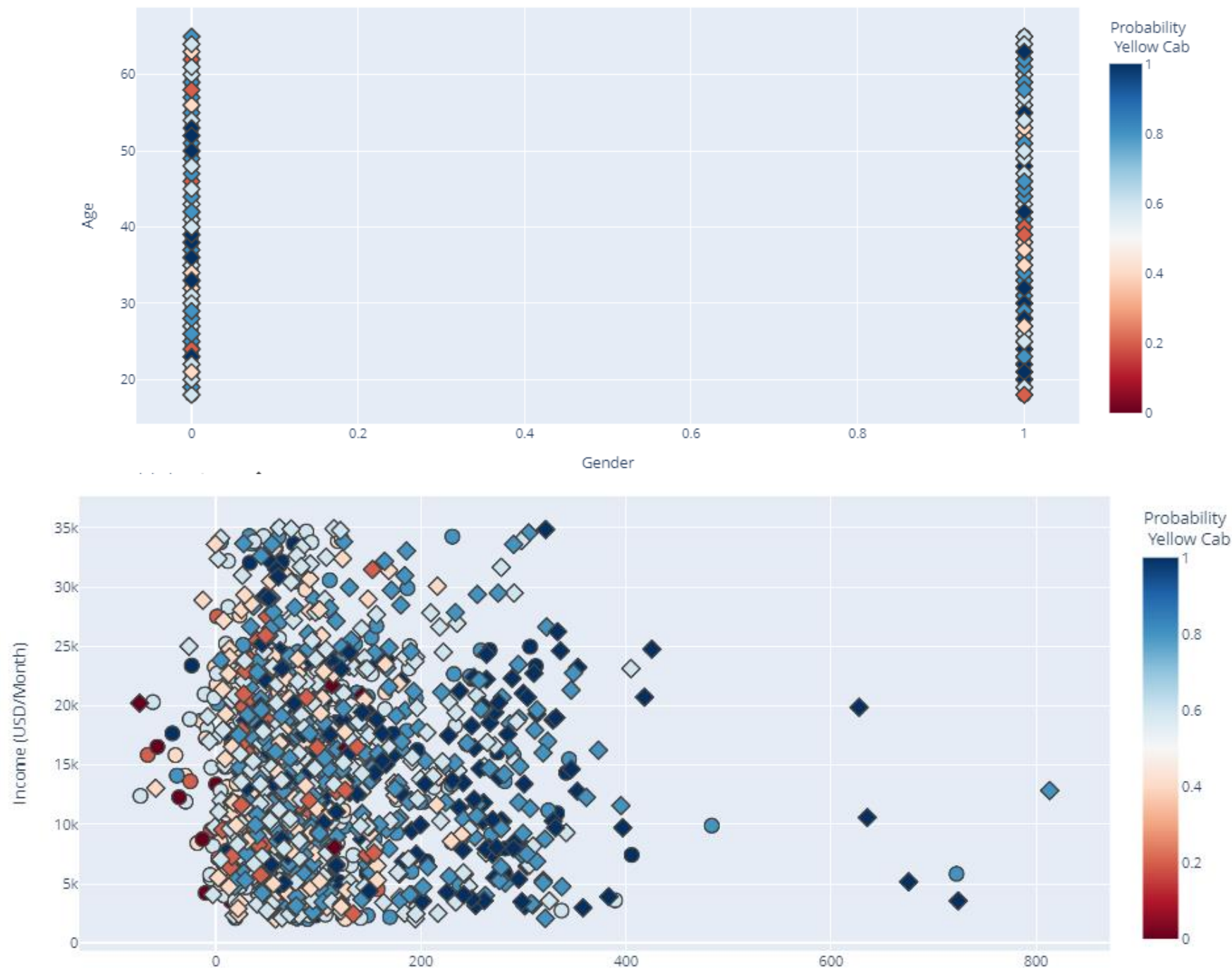


Hypothesis 4

EDA

As profit value increases, specially from 60, the probability of the User choosing Pink Cab over Yellow Cab decreases.

As for the income, there is no strong correlation between this feature and the binary choice (We can see that along the y axis probabilities remain the same)



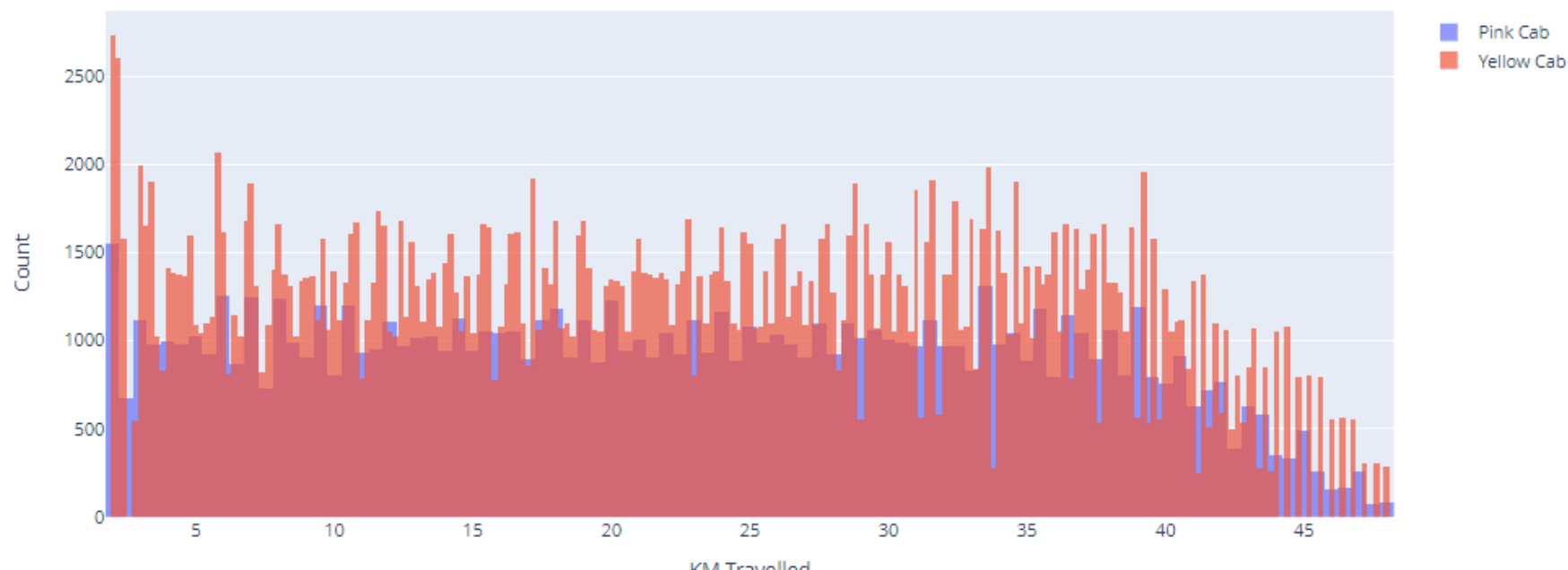
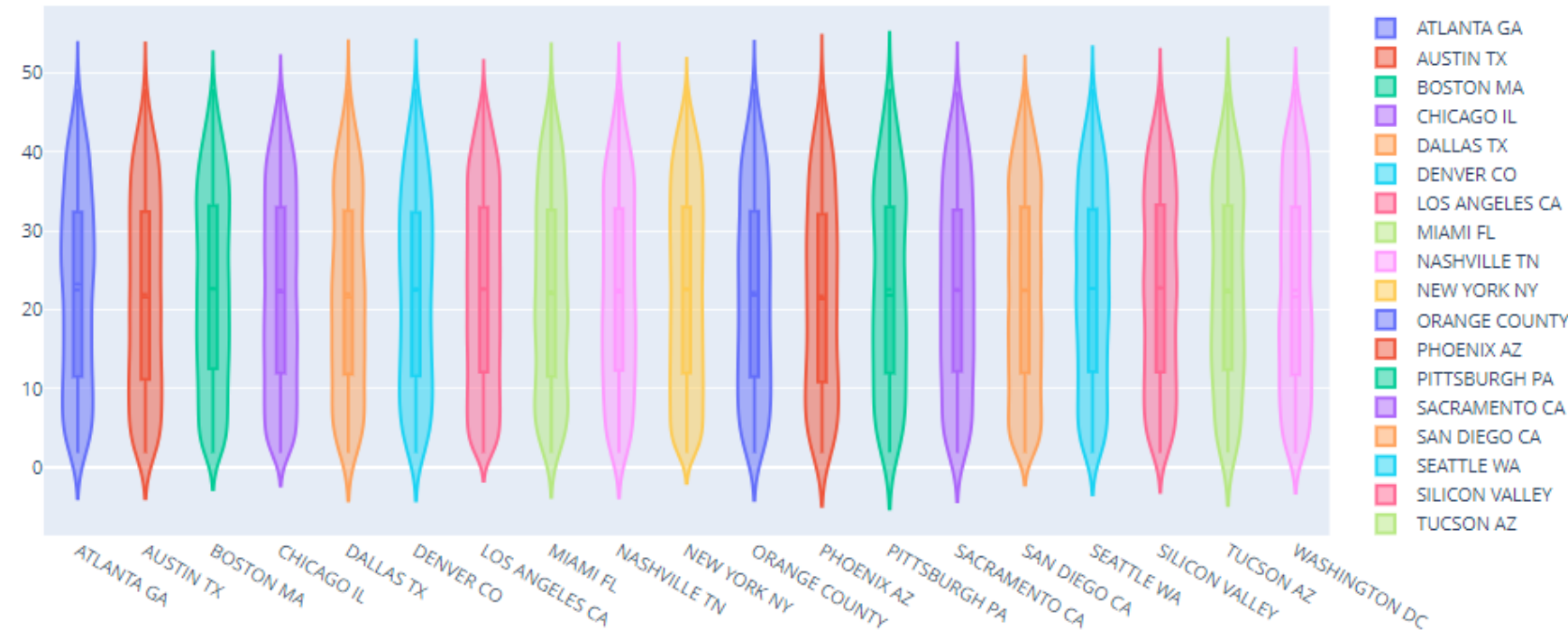
Hypothesis 5

EDA

Given the extremely high Standard Deviation for both cab companies in terms of distance, more relevant range would be as follows:

60.09 % of yellow cab users rather distances [10,30]

60.14 % of pink cab users rather distances [10,30]

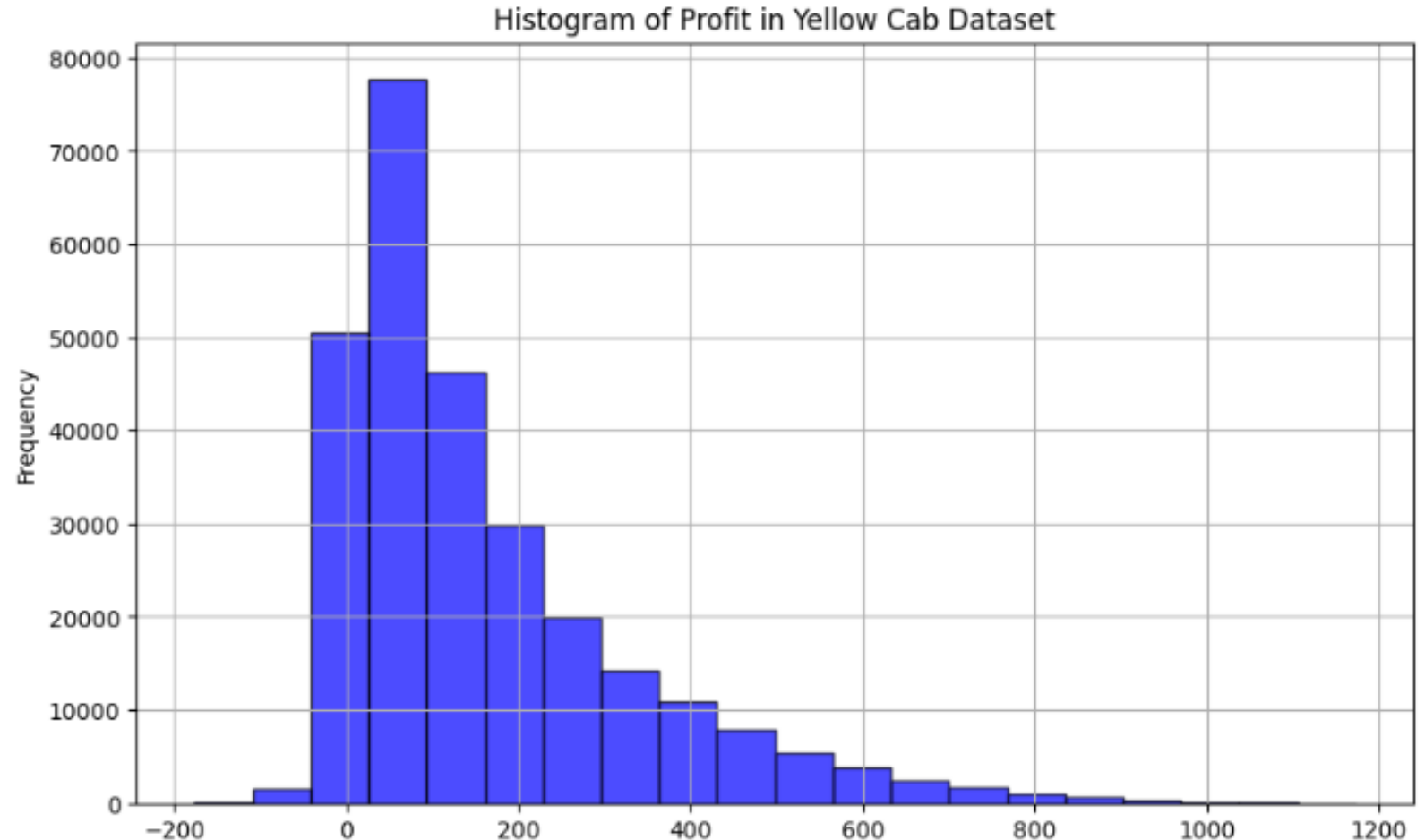


Hypothesis 6

EDA

There is a 95% chance for positive profit for Yellow Cab and 87% for Pink Cab

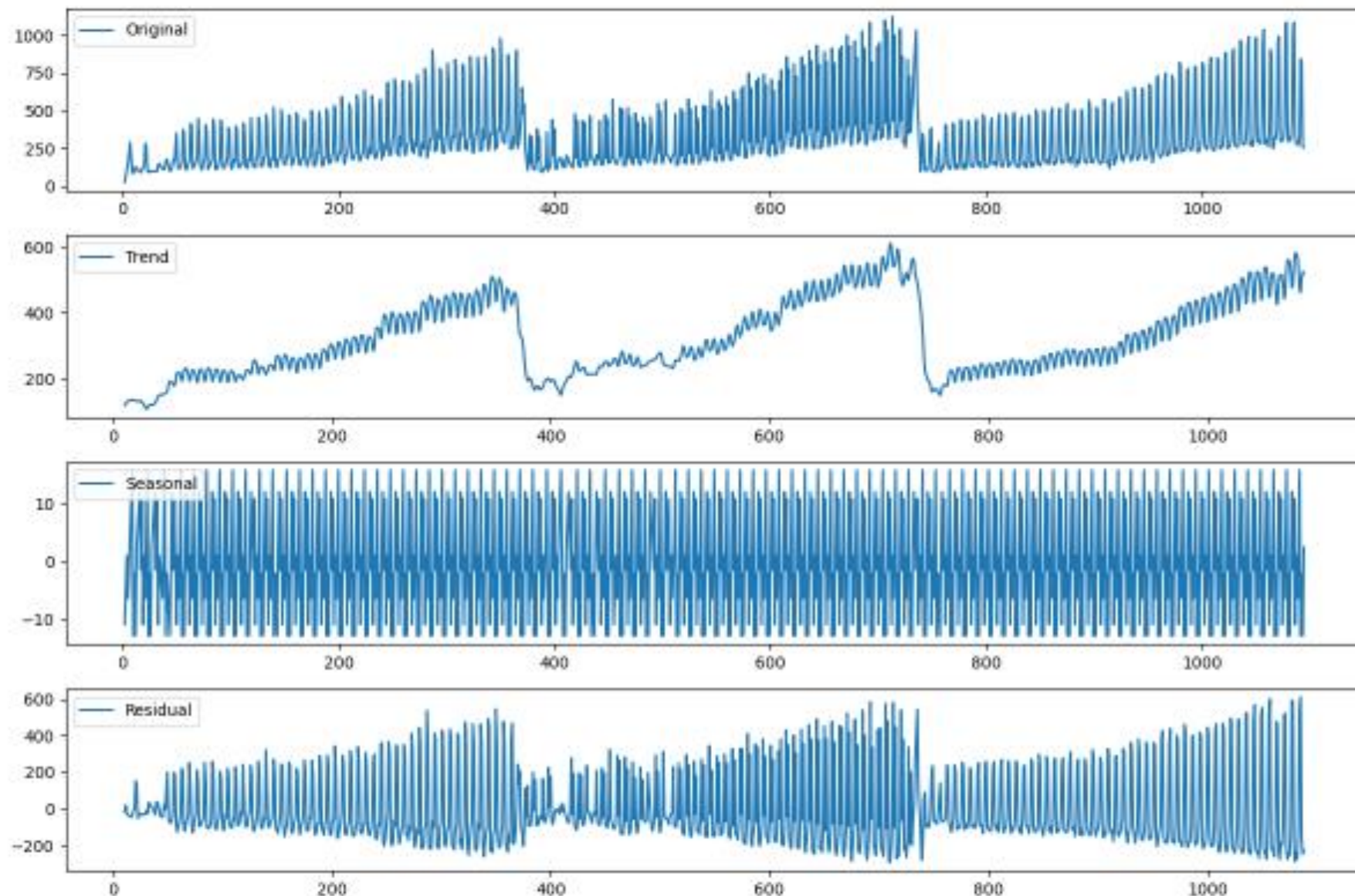
There is 95% chance any new profit for Yellow Cab will be [159, 160], while there is 95% chance any new profit for Pink Cab will be [62, 63]



Hypothesis 7

EDA

There is a 0.78% growth of daily demand per day during time periods of [360,385] days, then there is a sudden drop of 100% to start over again, and during this growing seasons it oscillates every other day by 90% of it's last growth.

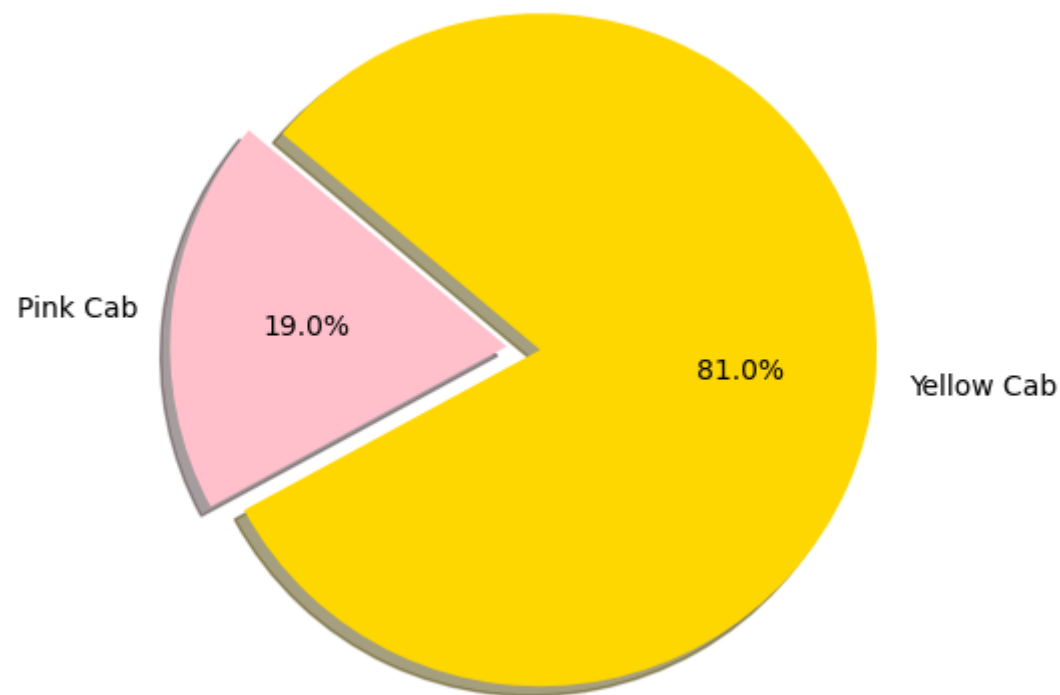
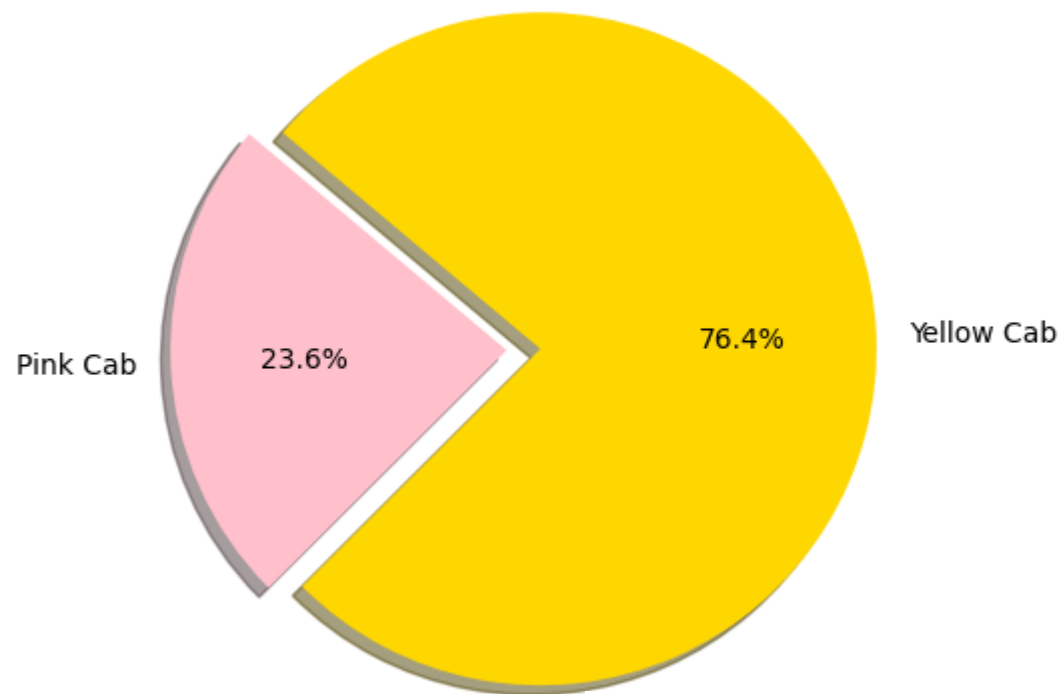


Hypothesis 8

EDA

Non loyal costumers rather Yellow Cab over Pink cab for a 76.4%

Loyal costumers rather Yellow Cab over Pink cab for a 81.0%



EDA Summary

After completing the Exploratory Data Analysis (EDA) process, the notebook aims to provide insights both in marketing and company terms by answering several key questions related to the Cab Industry dataset. Syntax approaches could be described as:

- City and Profit Correlation: Z-test
- Daily Demand Impact on Price: Z-test
- Age, Gender, and Profit Income Correlation: Heatp
- Predictability of User's Binary Choice: KKN
- Users' Propensity for Features: Confidence Inter
- Probability of Positive Profit: Confidence Inter
- Scattered Cab Demand over Time Periods: SDTS
- Overall Demand Comparison between Cab Companies: Percentage computation

Marketing Strategy Results:

- Since New York has a significant portion (more than 56%) of the overall profit, it would be wise to prioritize marketing efforts in this city to maximize profits.
- Male users tend to give more profit on average than females by 4%. The companies can create targeted campaigns to attract and retain male customers.
- Customers aged between 18 and 40 contribute significantly to the profit of both Yellow Cab and Pink Cab. The marketing strategies can be tailored to cater to this age group, offering discounts or incentives to attract them.
- Both companies have a 60-40 ratio for Card-Cash payment methods. Both can promote cashless payments and offer incentives to customers who choose this payment mode.
- During the periods of daily demand growth [360, 385] days, the cab companies can run seasonal promotions, offering discounted fares, loyalty rewards, or bonus points for frequent rides. This can attract more customers during the peak demand periods.
- To cater to customer preferences, offer special weekend promotions, discounts, or packages. Highlight the convenience and benefits of choosing the cab service during weekends, especially on Friday, Saturday, and Sunday.

Cab Company Results:

- Yellow Cab has the majority of the overall cab industry profit. This indicates that Yellow Cab is likely more established and profitable in the market.
- New York contributes more than 56% of the profit, and Yellow Cab has a significant share in this lucrative market.
- Yellow Cab users tend to choose the company more frequently for longer distances. This suggests that Yellow Cab may have a competitive advantage in providing services for longer trips, which can lead to higher revenue.
- The analysis shows a 95% chance of Yellow Cab having a positive profit, while the probability for Pink Cab is slightly lower at 87%. This indicates a higher level of confidence in Yellow Cab's profitability.
- Both loyal and non-loyal customers show a preference for Yellow Cab over Pink Cab. This suggests that Yellow Cab may have a stronger brand presence and better customer satisfaction, which can lead to repeat business and positive word-of-mouth.
- Yellow Cab can capitalize on the growing demand during specific seasons, as the analysis indicates that demand grows at an 11% rate during these periods. This provides opportunities for increased revenue generation.
- Both companies share a 60-40 ratio for Card-Cash payment methods, indicating a stable and established payment system for Yellow Cab.
- As the profit value increases, the probability of a user choosing Pink Cab over Yellow Cab decreases

Finally, given all of these advantages Yellow Cab has over Pink Cab based on the provided data, my final recommendation is to invest on **Yellow Cab** including also the recommended marketing strategies.

G2M Case Study

Thank You