



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

Cab Data Analysis

16th February 2022

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Problem Statement

Our executives have lately been interested in the cab market in the US. However, since the market is divided between two cab companies, our executives have been restricted from investing. We therefore have performed EDA (Exploratory Data Analysis) which should provide our stakeholders a more clear perspective.

Approach and Background

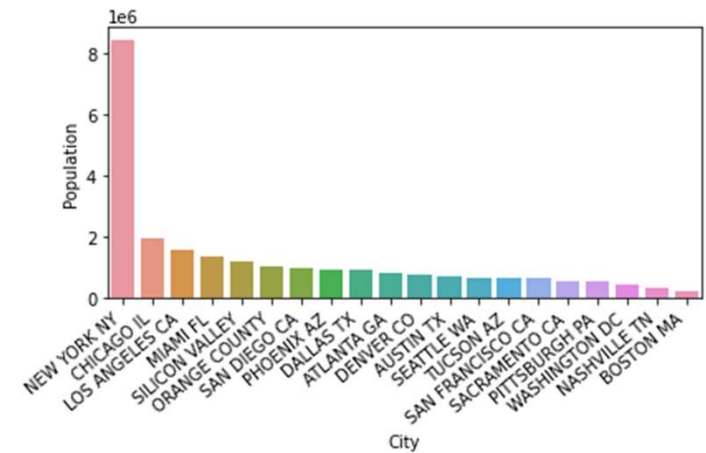
- The Cab industry has remarkably grown in the past 3 years which is getting our executives in XYZ to invest in the market
- Yellow and Pink Cab have been old players in the cab market, wherein each have operated in different states for the past 10 years.
- These companies however blew in the past 3 years with the online cab hailing services.
- Our cab market currently seems fresh and young with lots of growth potential
- We have divided our analysis in two parts
- The first part involves us of performing analysis on where its best for our company to operate.
- We analyzed this by comparing Population, Ages, and Income of our target Audience
- The second part focuses more on company specific analysis. Where we look at which company generated more profit, travelled more often etc.

Exploratory Data Analysis

1. Population Analysis
2. Profit per City
3. Card or Cash
4. Age of Customers paying by Cash and Card
5. Cost of Travel
6. Profit Comparison between two Companies
7. Total KM Travelled
8. EDA Summary and Conclusion

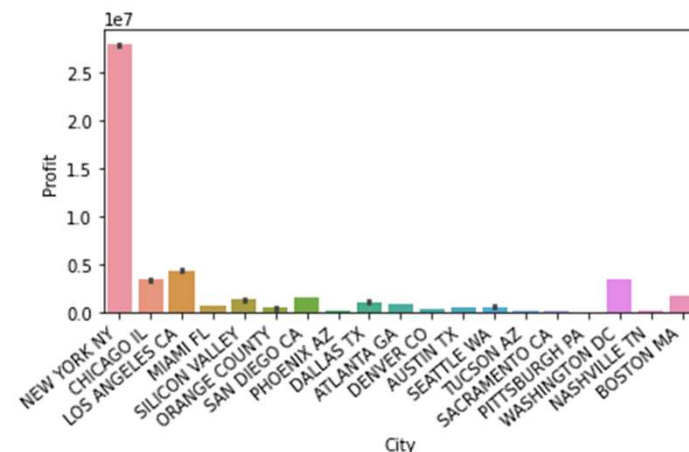
Population Analysis

- We began by performing an analysis of population for the cities our investors were interested in
- By the graph to our right we can see cities like New York, Chicago, Los Angeles, Miami, Silicon Valley, Orange County and others should prove to be viable investments for our investors.
- This is mainly because of the large population they carry hence, the probability for obtaining viable customers from these cities is higher.



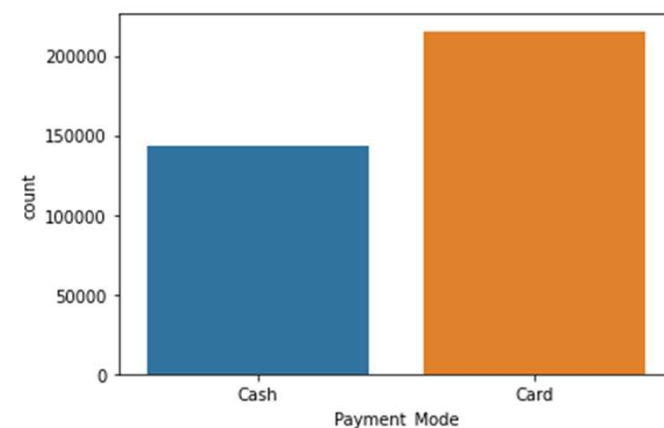
Profit Per City

- If we were to follow the analysis done on the previous slide wherein, we discovered that cities like New York, Chicago, Los Angeles, Miami, Silicon Valley and Orange County were the ones with most populous, we would've recommended these cities.
- However, with the graph of Profit per City this wouldn't seem like a valid recommendation.
- Cities like Boston MA which had the least population of all cities ended up generating profits greater than cities like Austin which had a higher population.
- The top cities according to profit would be New York, Chicago, Los Angeles, Washington DC, Silicon Valley, San Diego and Boston MA.



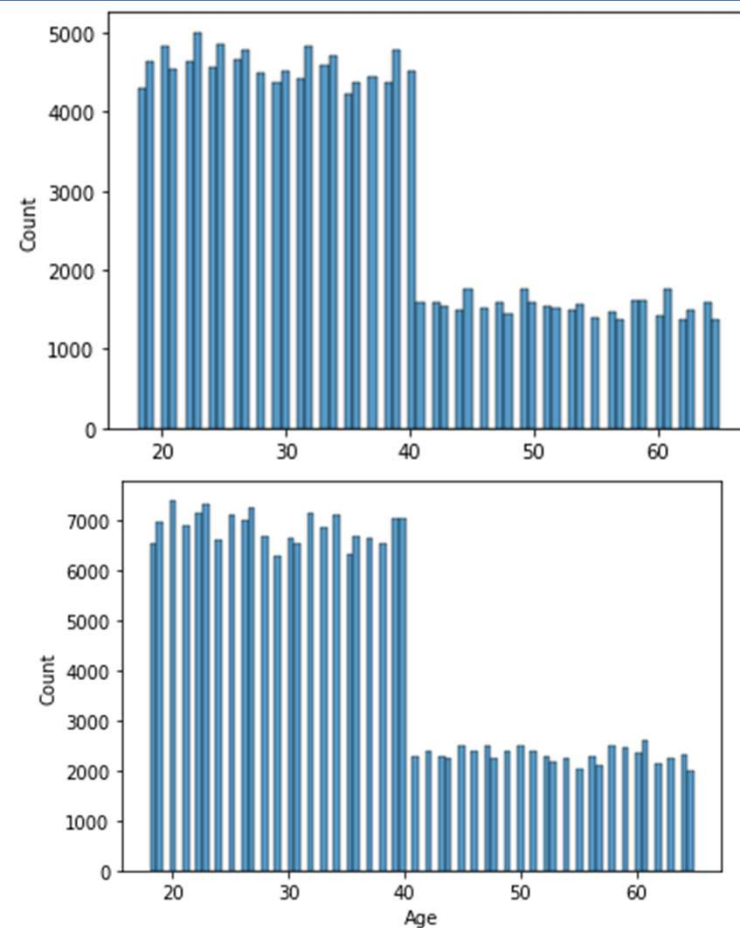
Card or Cash

- Customers don't always carry cash do they?
- Well the practice seems to be more common than we think
- According to the data provided customers used Card 1.5x more often than cash.
- This makes acceptances of all bank Card payments a valid concern for our investors.
- We'll want to make sure that our app ensures acceptance of all banking services available in the US.



Ages of Customers Paying by Cash and Card

- The plots to the right let us get an average idea of the age of our base users
- Our customers mainly lie in the 12-40 age bracket indicating they're more likely to be using our cab services for work purposes which would require us to be punctual and on time.
- Being punctual though is supposed to be something pre assumed by any company, however this analysis informs us on how improving algorithms for calculating distance can benefit us.



Continuation

- The analysis and questions answered in the previous slides are more geared toward the general things our investors should focus on after buying a company.
- However, we yet haven't provided analysis on which company our investors should invest in.
- The following slides will include plots in which we try to guide our investors on which company they should aim at bidding for.

Cost of Travel



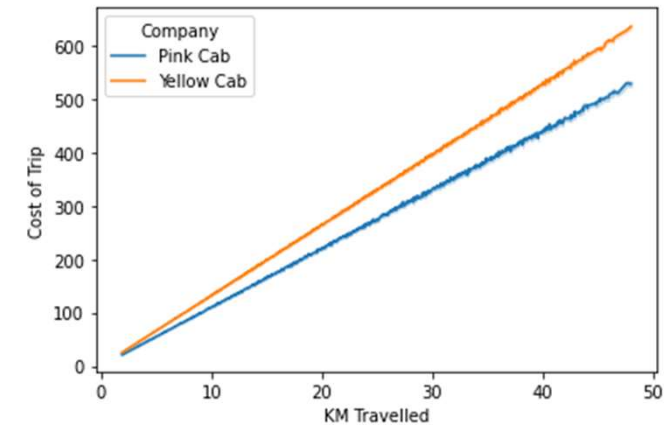
The plot to the right gives us an idea of the usage of resources by the companies under question



Pink Cab seems to utilize resources more efficiently



This plot indicates that Pink Cab would be a valid investment. However, the Profit plot in the following slide would love to differ.



Profit Comparison between the two Companies



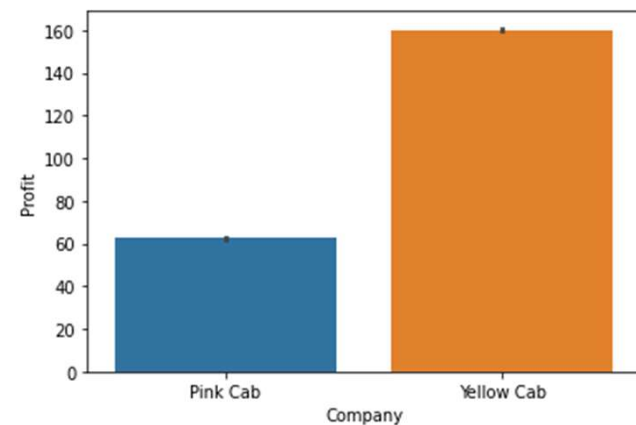
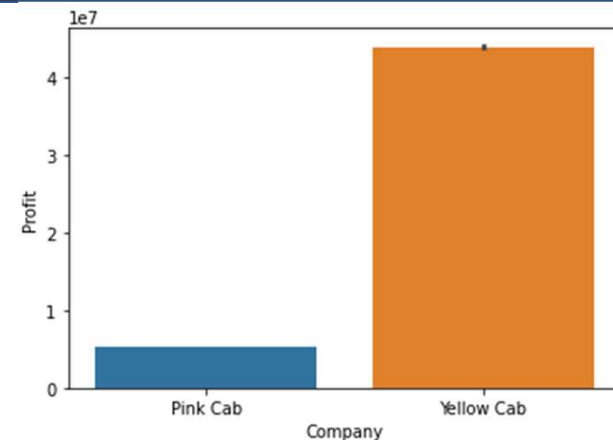
The plots to our right indicate the profit distribution between the two companies.



We can spot that Yellow Cab is far more profitable compared to the Pink Cab. This analysis though does not go sound with the Cost of Travel plot, we get a very good idea of which company can be a valid investment.



The plot to the top right indicates Total Profit for the two companies while the one below indicates the average profit for each company per trip according to the data provided.



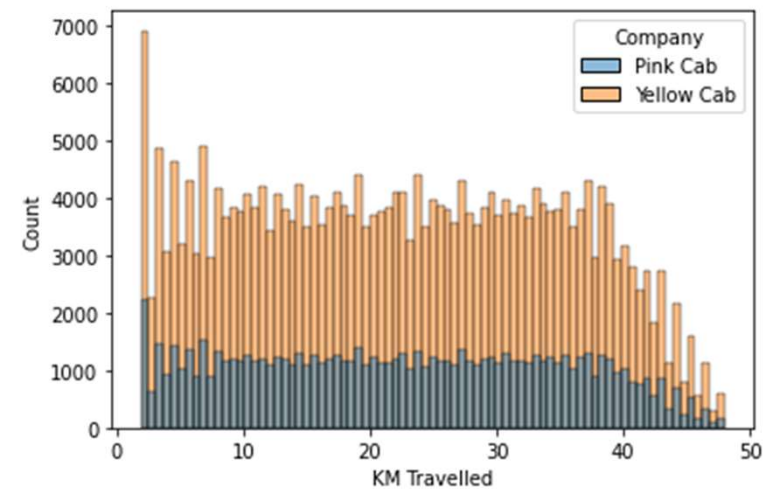
KM Travelled



Analysis performed in this plot tells us on how much each company's cab travelled



It can be easily spotted that Yellow Cab travelled more compared to Pink Cab. Wherein Yellow Cab surpassed Pink Cab in number of trips for almost every KM bracket.



Profit Comparison between the two Companies

Which Company to choose?

- After exploring all the hypothesis shown above, I would recommend investors to invest their money in Yellow Cab despite the high cost of travel
- They carry massive profit margin compared to their competitor
- By the plots in the previous slide, we can also infer that Yellow cab is used more than Pink cab. Which is why investors will end up generating more return on investment
- We also inferred that more people pay by card, however this does not negate the usage of cash by users. Massive population of users use cash which is why it will also have to be catered

Profit Comparison between the two Companies

Where to focus post purchasing?

- We also saw which cities were possibly the most populated from which we inferred New York was the most populated while others were somewhat close
- However, when we saw total profits for each city, we got to know that cities like Boston and Silicon Valley despite lower population ended up yielding profits greater than cities with high population
- Therefore, we would recommend investors to put their money in Yellow cab and focus on cities like New York, Chicago, Los Angeles, Washington DC, Silicon Valley, San Diego and Boston MA

The End

Thank You