Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M insight for Cab Investment firm

11/15/2023

Gao Mo

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Executive Summary

Datasets Analyzed:

1. Cab_Data.csv: Transaction details of two cab companies.
2. Customer_ID.csv: Customer demographics linked to unique identifiers.
3. Transaction_ID.csv: Mapping of transactions to customers and payment modes.
4. City.csv: Information on US cities, their population, and cab user statistics

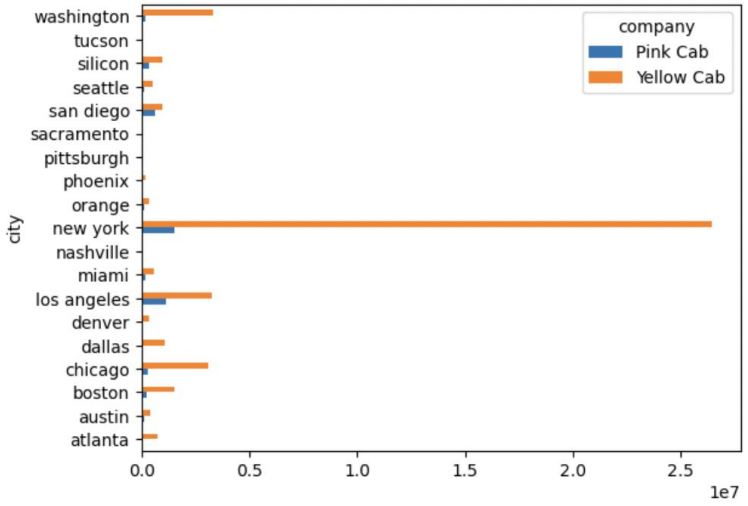
Methodology:

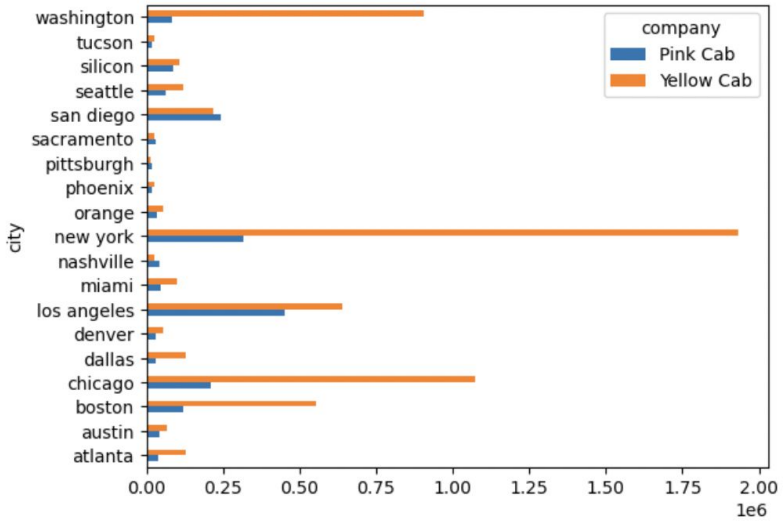
- Data Integration and Cleaning: Merged, transformed, and cleansed datasets to create a unified master dataset. Removed duplicates, handled missing values, and identified outliers.
- Hypothesis Testing: Generated and tested 5 hypotheses

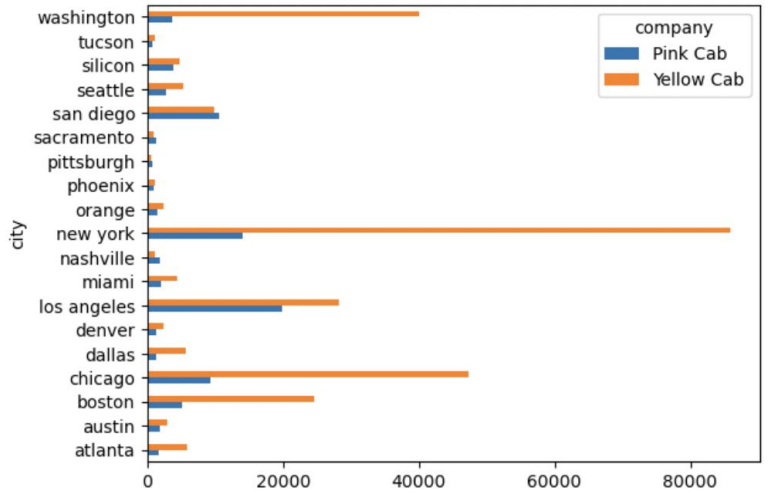
Problem Statement

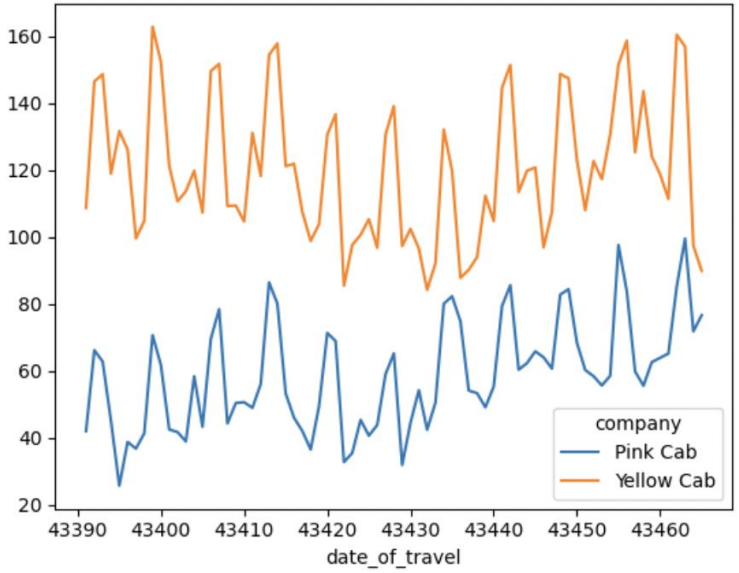
XYZ, a US business, is getting ready to put money into the fast-paced taxi industry. However, the task of finding the best company to invest in among a sea of competition is not easy. Grasping this, XYZ knows it needs a deep dive into detailed data from two leading taxi businesses to make a decision. This data review would have to cover transaction specifics, rider details, and how often cabs are used across diffe rent cities. It would also need to look at things beyond the taxi companies' control like holidays and what the weather's like. The aim? To understand trends in the market, how customers behave, and where the profits are. This way, XYZ can use the power of data to choose its investment wisely. The idea is to sync with its big-picture goals and make the most out of what this fluid sector has to offer.

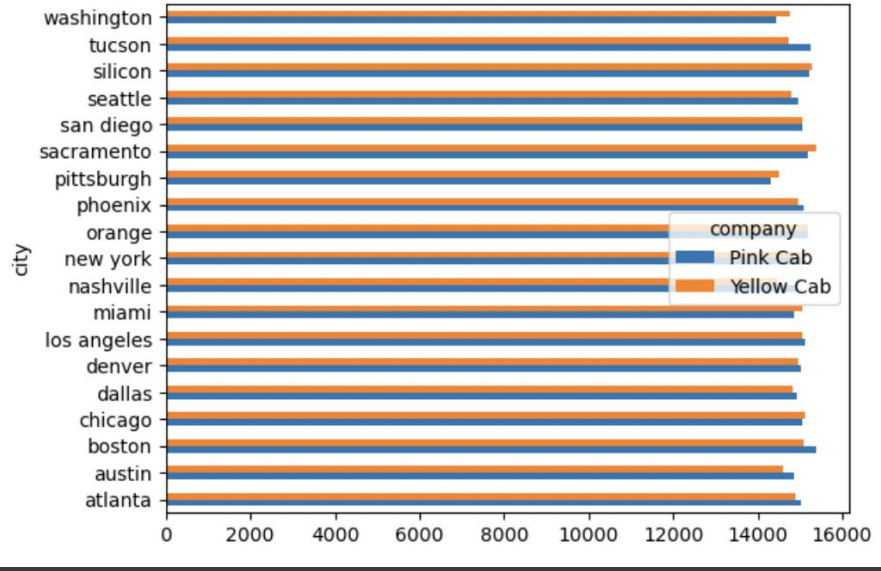
EDA











Data Glacier
Your Deep Learning Partner

EDA Summary

- Yellow Cab Company achieves a higher average profit margin than Pink Cab Company in all cities.
- Yellow Cab Company achieves a higher average travel distance margin than Pink Cab Company in all cities.
- Yellow Cab Company achieves a higher amount of transactions margin than Pink Cab Company in all cities.
- In the last 150 days, Pink Cab shows a general growing trend in terms of profit gain while Yellow Cab shows a little to no change in the trend of profit gain. The overall profit gain of Yellow Cab is still higher than that of the Pink Cab

  Both Companies users have a similar average of income

Recommendations

Final Recommendation

It appears that the Yellow Cab company has a higher profit gain than Pink Cab across all cities during the period from January 31, 2016, to December 31, 2018. Concurrently, Yellow Cab has more transactions and a higher amount of distance traveled in each city, indicating that Yellow Cab is more active than Pink Cab. Moreover, Yellow Cab does not show a trend of diminishing values in the last 150 days in terms of profit gain. Finally, the quality of the users is similar between Yellow Cab and Pink Cab. Therefore, it is more advisable to invest in Yellow Cab.

Thank You