



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M insight for Cab Investment firm

25th October 2022

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

EXECUTIVE SUMMARY

- XYZ is a private equity firm in US which plans to invest in the Cab industry and they want to have a good understanding of the market
- The data available is of two cab companies: Pink Cab and Yellow Cab

PROBLEM STATEMENT

- There are two Cab companies considered over the period of two years and their performance recorded.
- Objective: Analysis of the data and deriving useful recommendations for investment

APPROACH

Approach to the research is given as follows:

- Data Exploration
- Find the most patronized companies
- Find out reasons for consumer behaviors
- Recommendation for investment

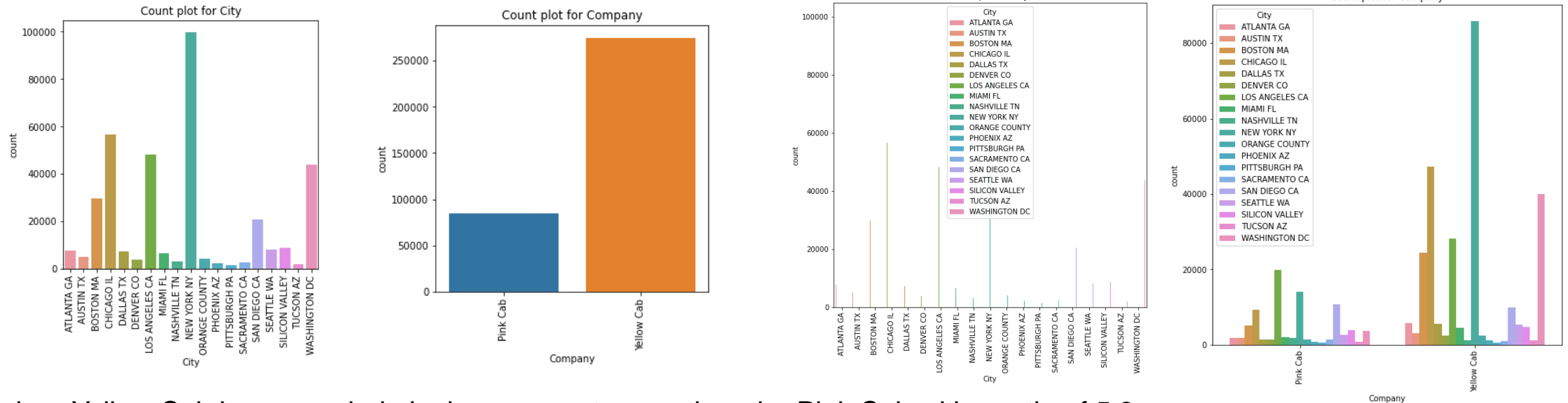
EDA

Data preparation:

- We are given four datasets and they are merged as appropriate
- These data include Cab Data, City data, Customer data and transaction data
- These datasets are checked individually for null values before moving on to exploratory analysis on each of the data.

EDA

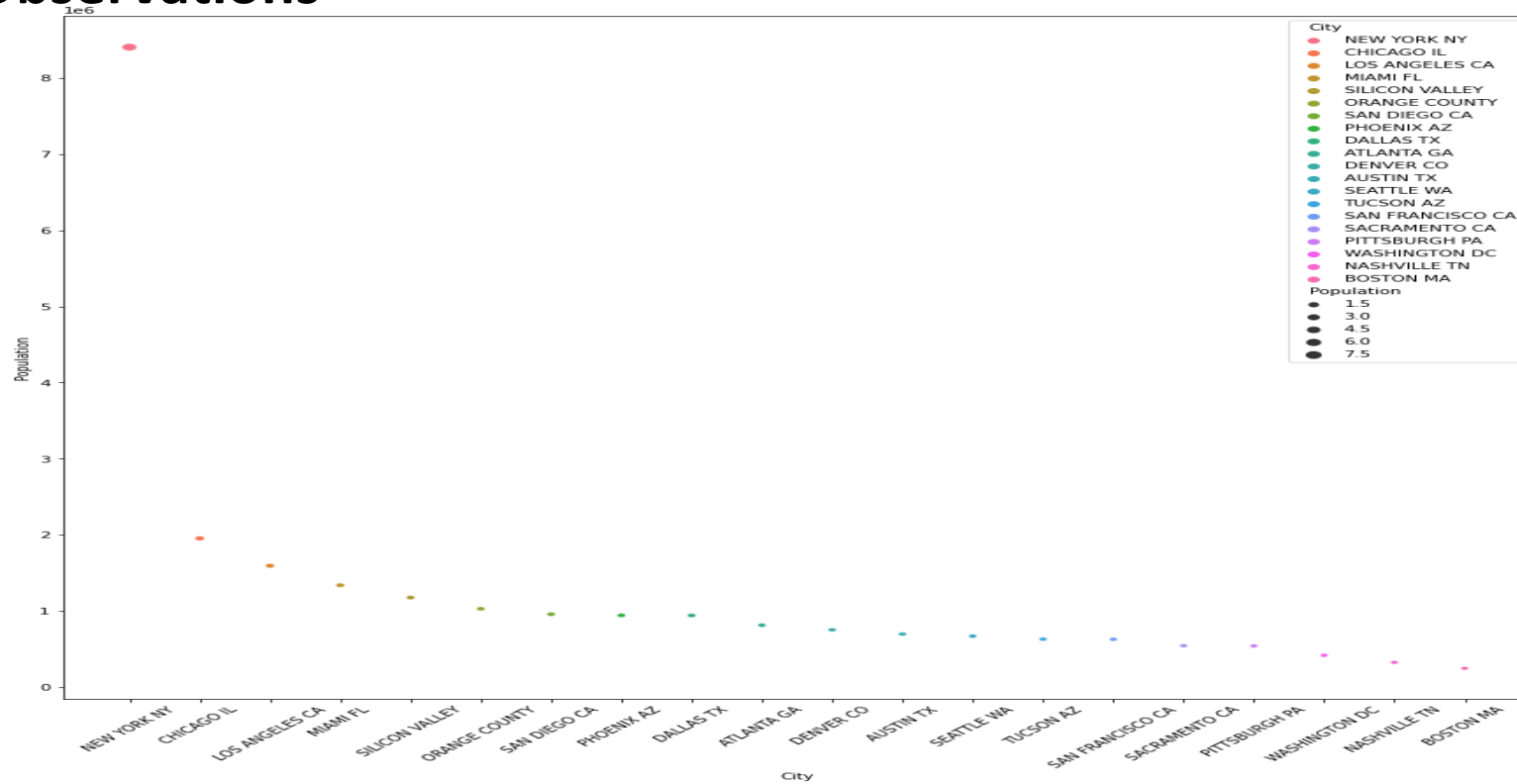
Cab data observations



- Yellow Cab has overwhelmingly more customers than the Pink Cab with a ratio of 5:3
- New York has the total number of customers when the Yellow and Pink Cabs are considered, but broken down by City, the Yellow Cab has its highest number of customers in New York while Pink cab has its highest patronage in Los Angeles

EDA

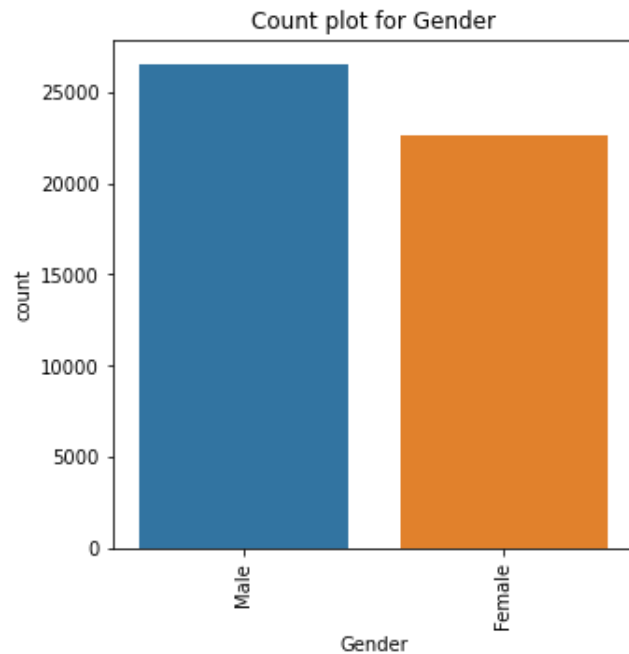
City Data Observations



New York represents the highest number of users as well as the largest population

EDA

Customer Data Observations

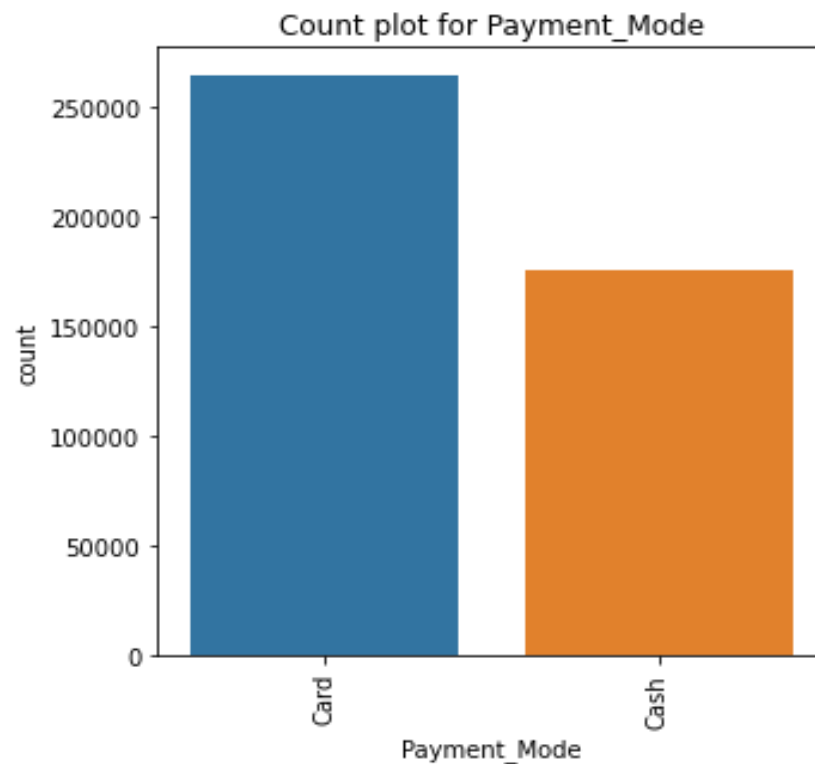


	Customer ID	Age	Income (USD/Month)
count	49171.000000	49171.000000	49171.000000
mean	28398.252283	35.363121	15015.631856
std	17714.137333	12.599066	8002.208253
min	1.000000	18.000000	2000.000000
25%	12654.500000	25.000000	8289.500000
50%	27631.000000	33.000000	14656.000000
75%	43284.500000	42.000000	21035.000000
max	60000.000000	65.000000	35000.000000

- 1.The average age of the customers is 35 years (Millenials)
- 2.The average income of the customers is approximately 15,000 dollars
- 3.More men than women use the cab services with a ratio of about 5:4

EDA

Transaction data Observations



1. People overwhelmingly preferred to use cards for payments rather than cash

EDA SUMMARY

- Yellow cabs are more popular due to their use in the more populous cities
- Customers preferred using cards as opposed to cash for payment due to its relative convenience

RECOMMENDATIONS

- It is better to invest in the Yellow Cabs at the moment as opposed to the pink cab due to its popularity and its use in the Most populous city, New York.
- Encourage card payments for the convenience of both the drivers and the customer. This can be done by applying some sort of discount after a number of payments with cards
- The marketing for the cab should be aimed at millenials and the male population since they are more likely to use the cab services

Thank You