



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

<Data Science Virtual Internship>

<19 Nov 2022>

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Introduction

Problem statement: Company XYZ is looking to invest in the cab industry. XYZ is going to choose one winner from two cab companies to fund.

Objective: Analyze the datasets from 2 Cab companies to determine which one has the more potential to fund.

The outline of this study is as follows:

- Profit Analysis
- Customer Analysis
- Recommendation

The Overview

The total profits over the course of 3 years is shown by the trend plot in Fig. (a). It can be seen from a joint scatter plot in Fig. (b) that for the same amount of distanced travelled, Yellow Cab company charges more than the Pink cab company.

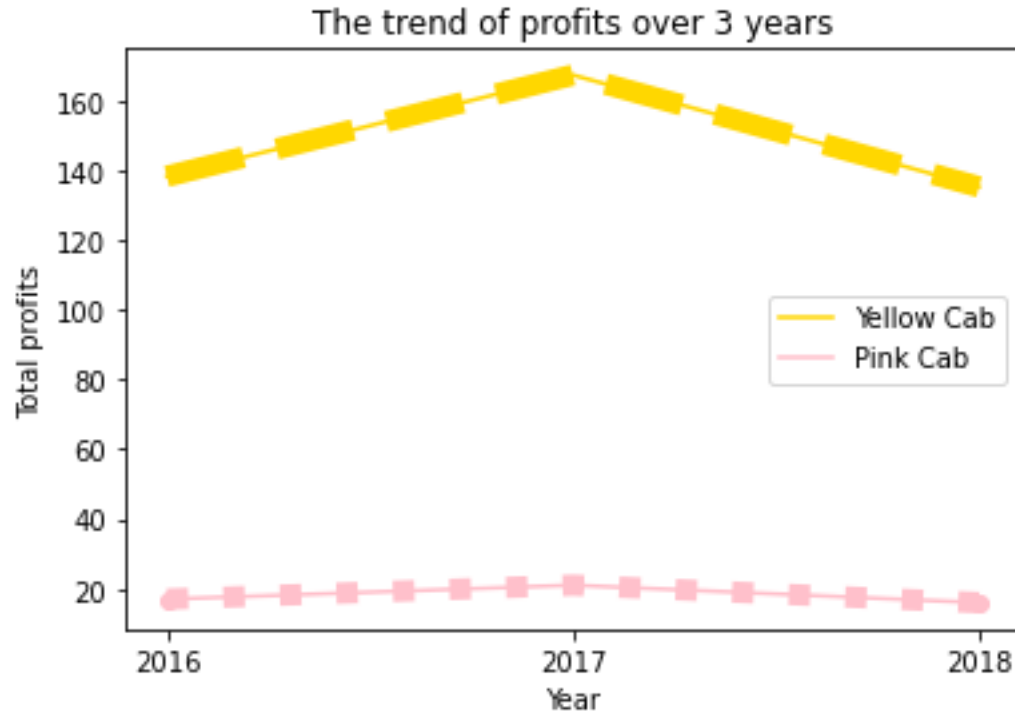


Figure.1

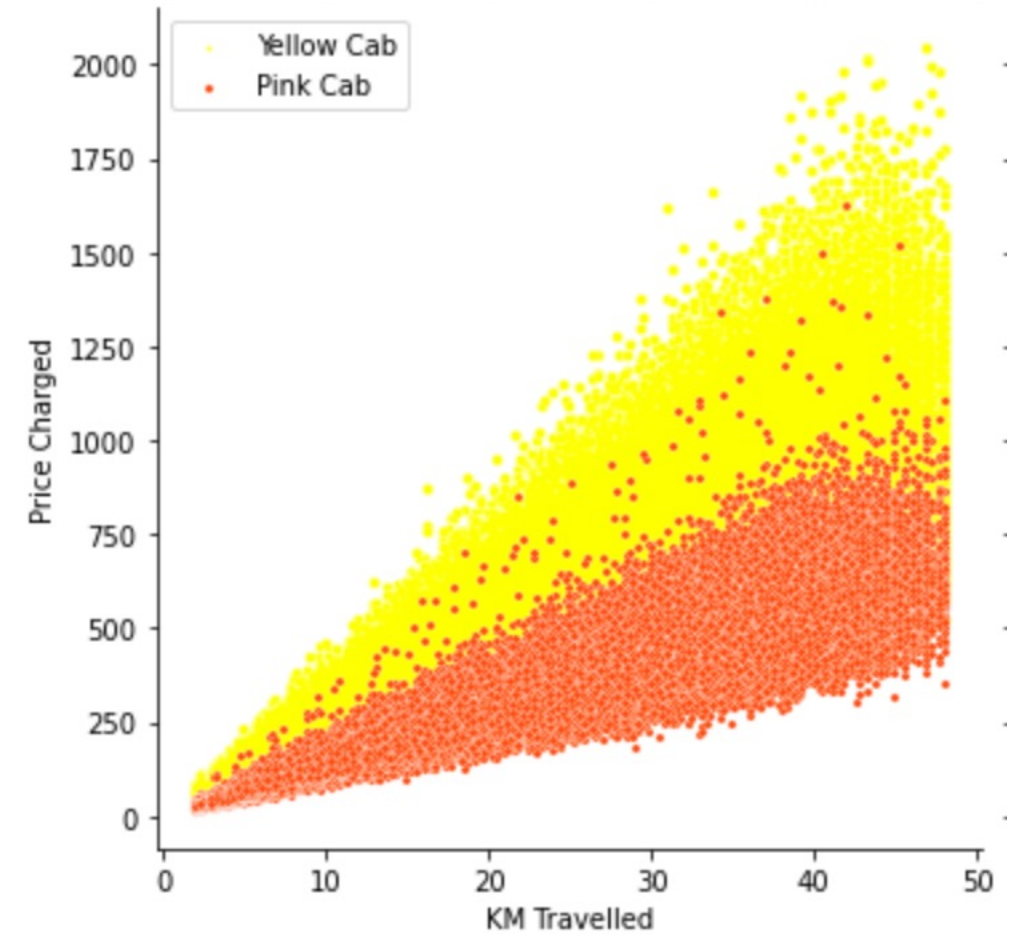
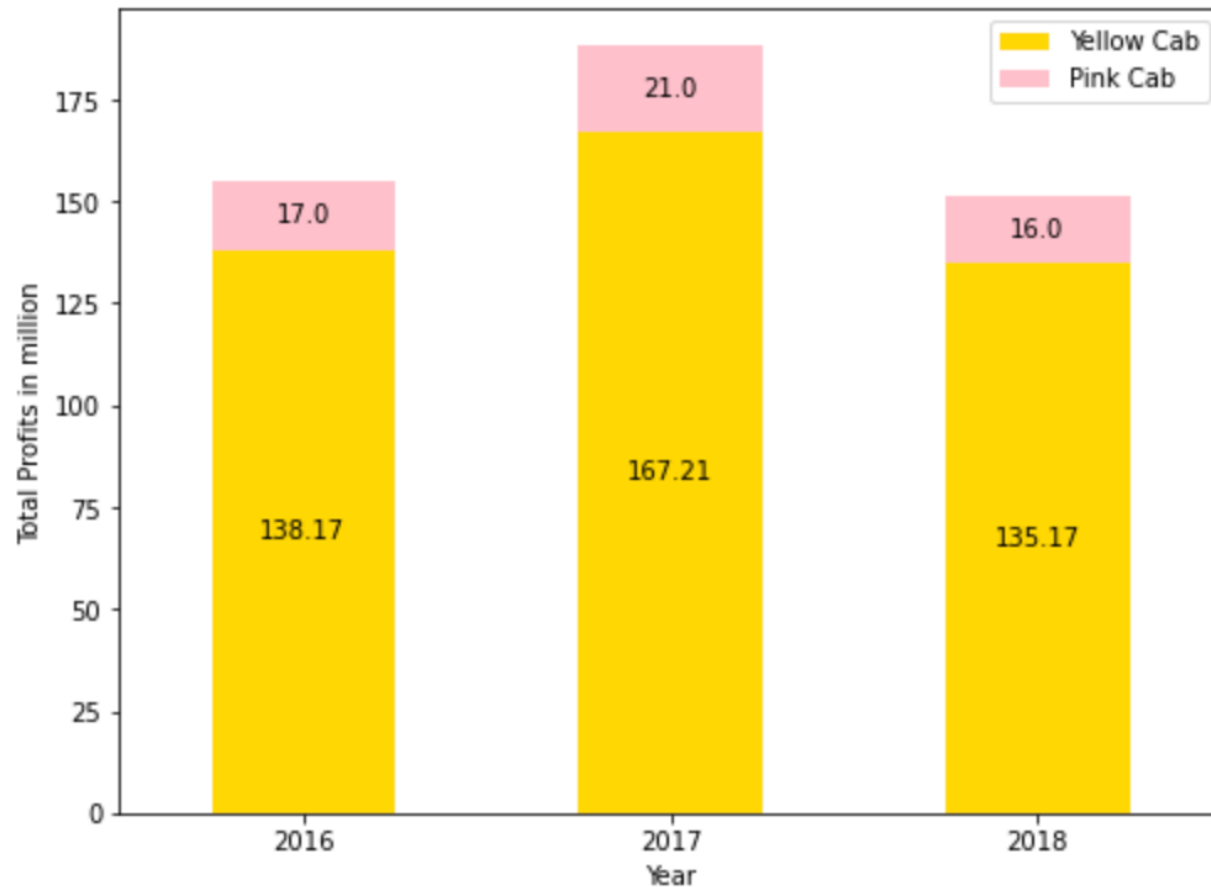


Figure.2

Comparison of Total Profits by Year

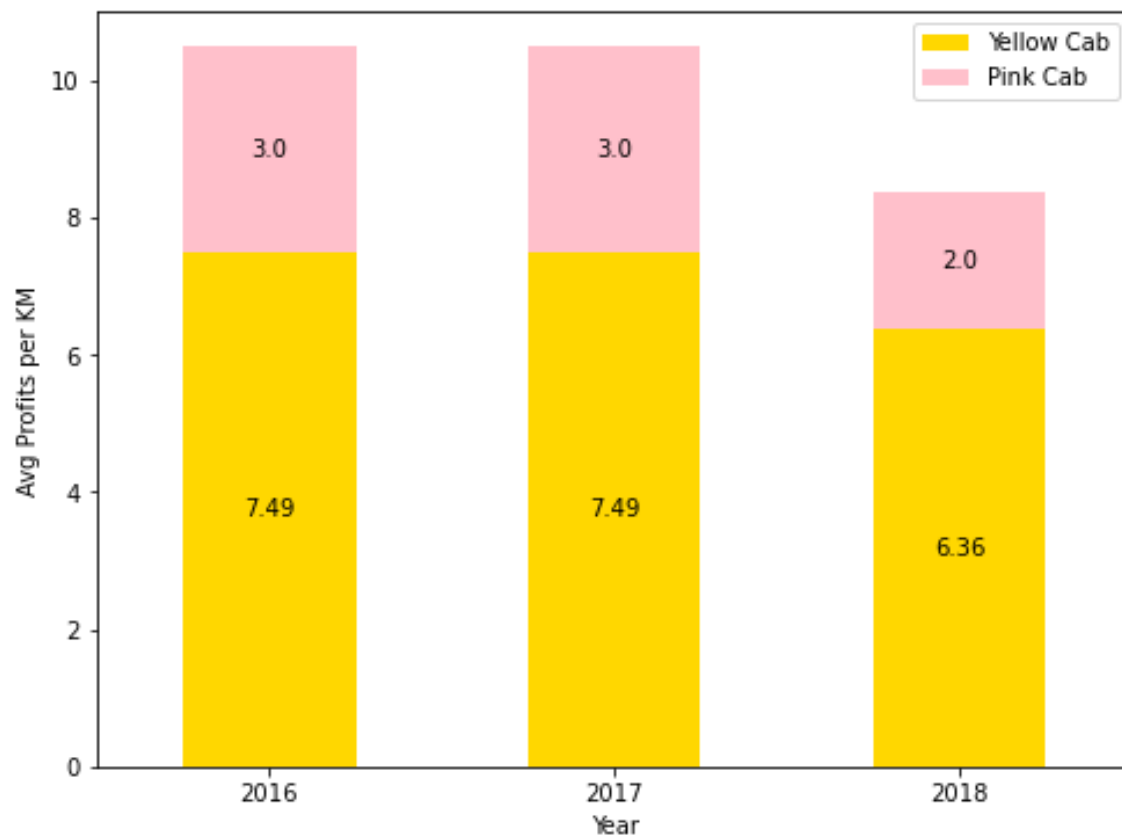


Comapny	Total Profits	Total transactions
Yellow Cab	44,020,373	274681
Pink cab	5,307,328	84711

Average profit per KM by Year

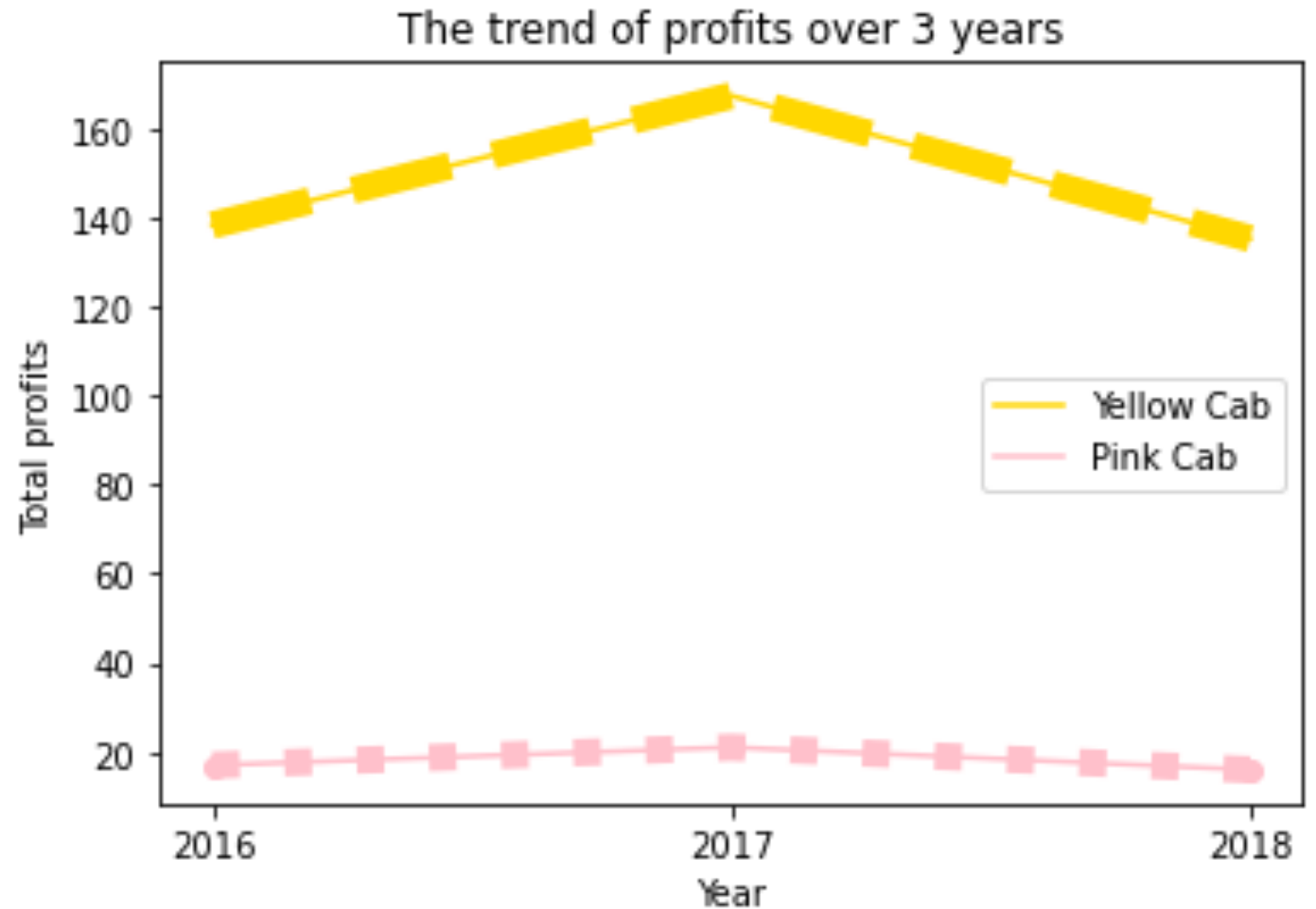
Comapny	Avg profits per transaction
Yellow Cab	160.25998
Pink Cab	62.65217

It can be seen that Yellow Cab company has a profit per transaction that's 60% higher than Pink Cab company.



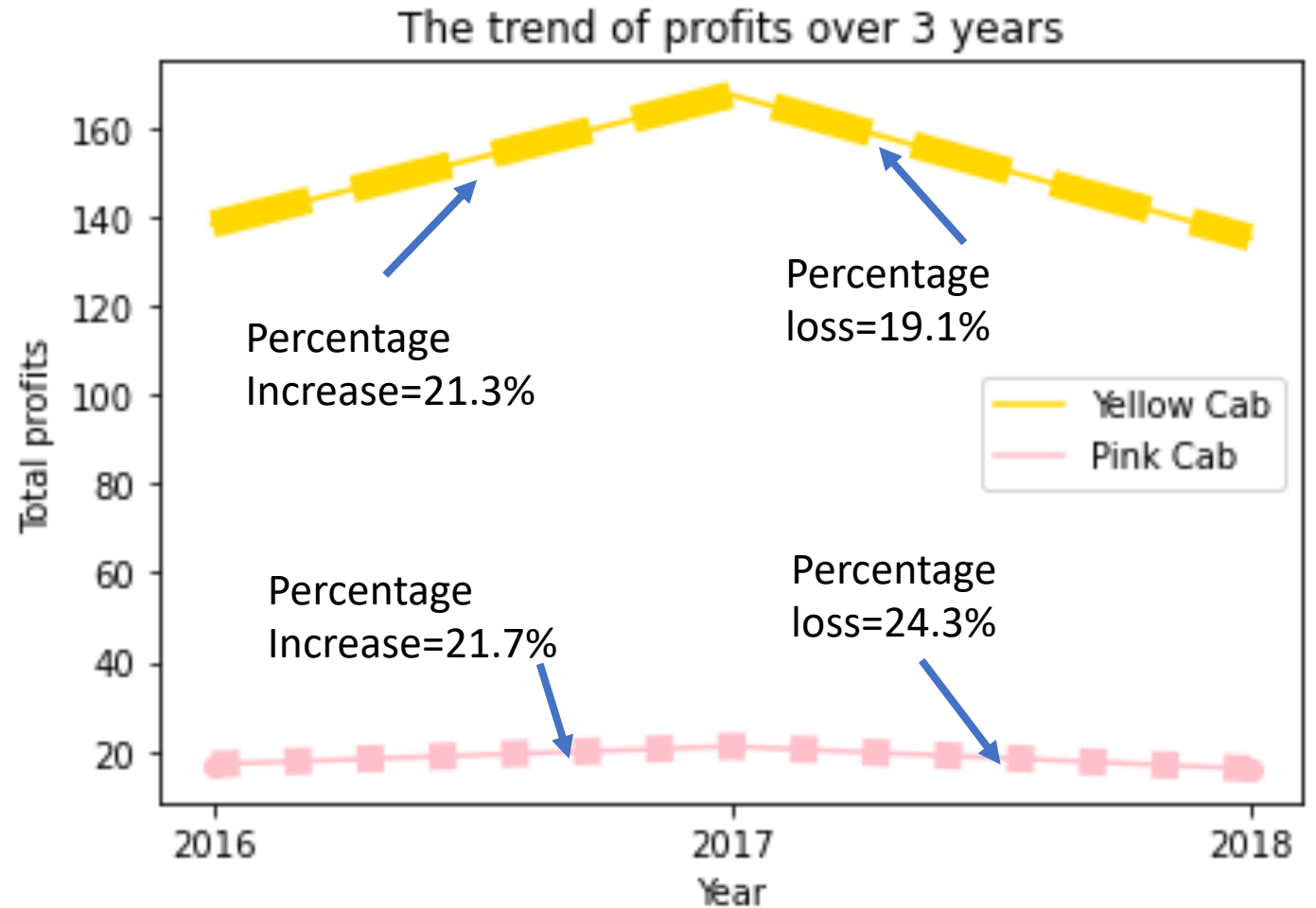
Total profits by Year

Comapny	Profit in year (2016)	Profit in Year (2017)	Profit in year (2018)
Yellow Cab	13817365	16720787	13517399
Pink cab	1693009	2061973	1560162

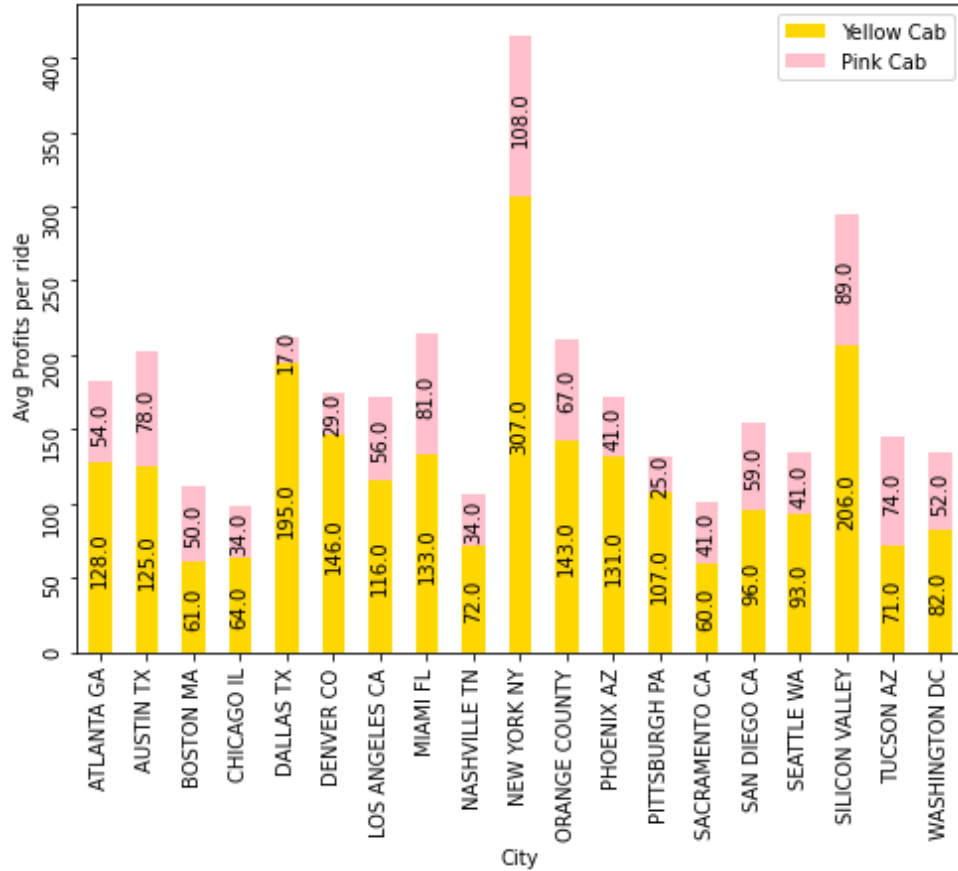


Total profits by year

Company	Profit in year (2016)	Profit in Year (2017)	Profit in year (2018)
Yellow Cab	13817365	16720787	13517399
Pink cab	1693009	2061973	1560162



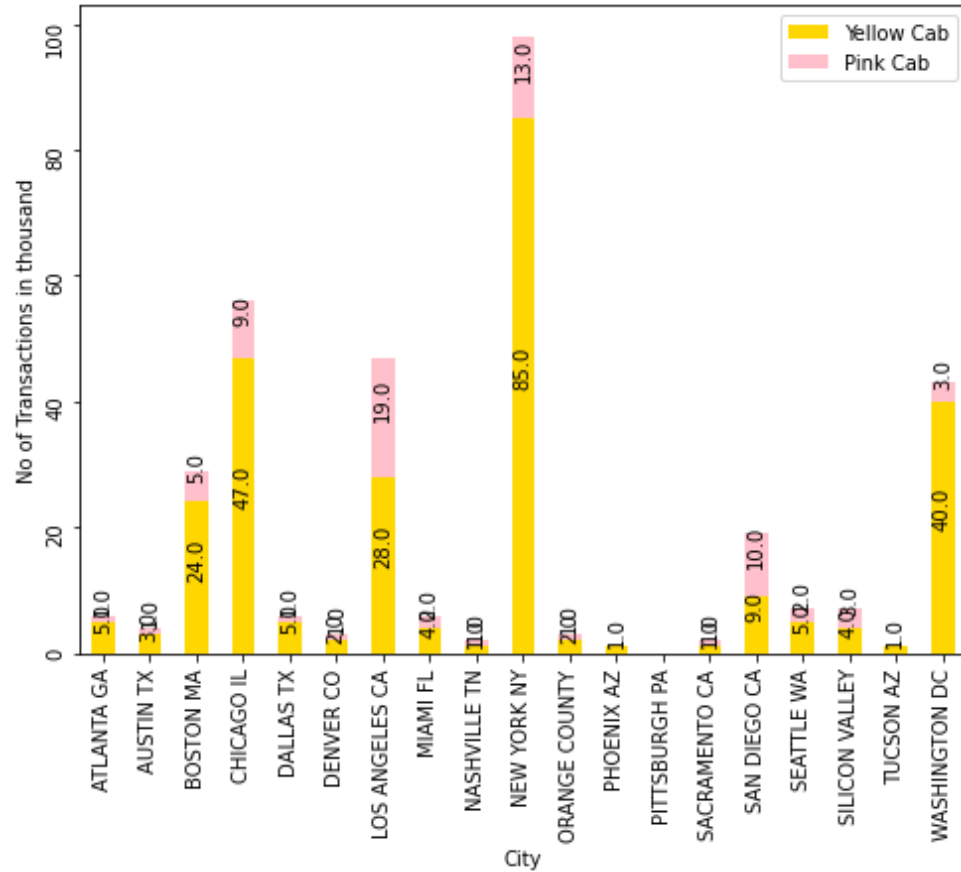
Avg Profits by City



It can be seen that New York City is the city where most profitable transactions are taking place.

Company	Top 1 (profitable)	Top 2 (profitable)	Top 3 (profitable)
Yellow Cab	New York City	Silicon Valley	Dallas TX
Pink cab	New York City	Silicon Valley	Tuscon AZ

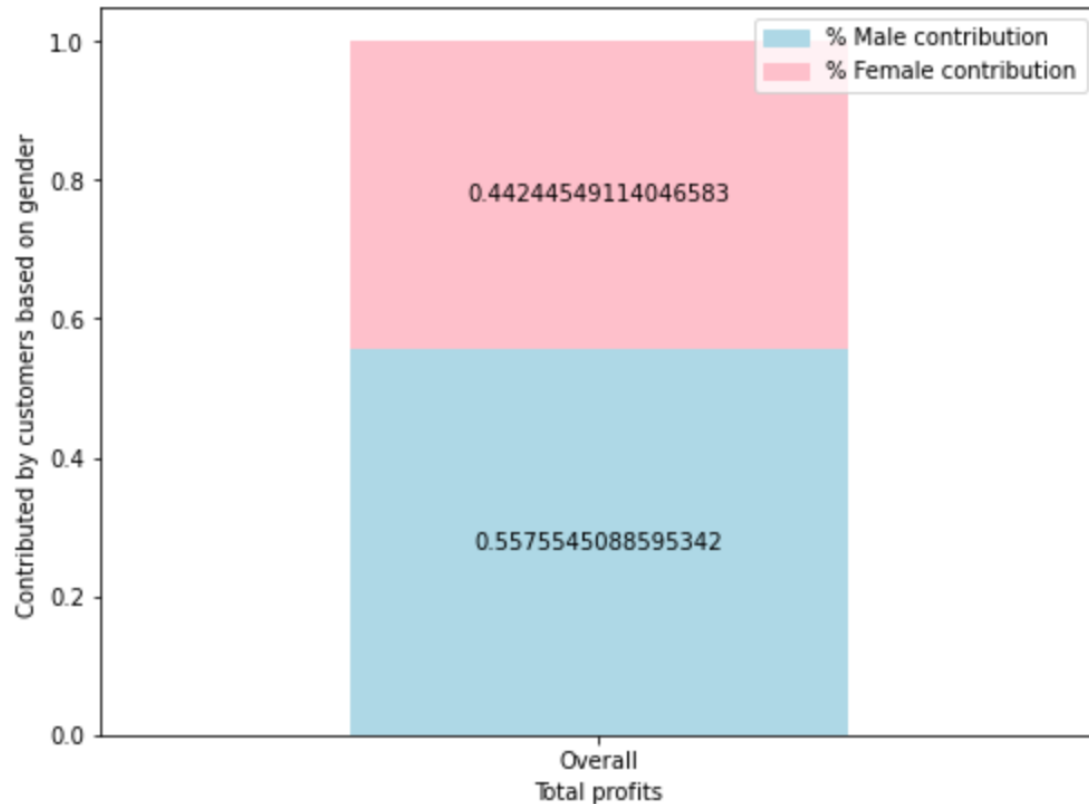
Transaction Frequency by City



It can be seen that yellow cab company has more frequent transactions than pink cab company in most cities.

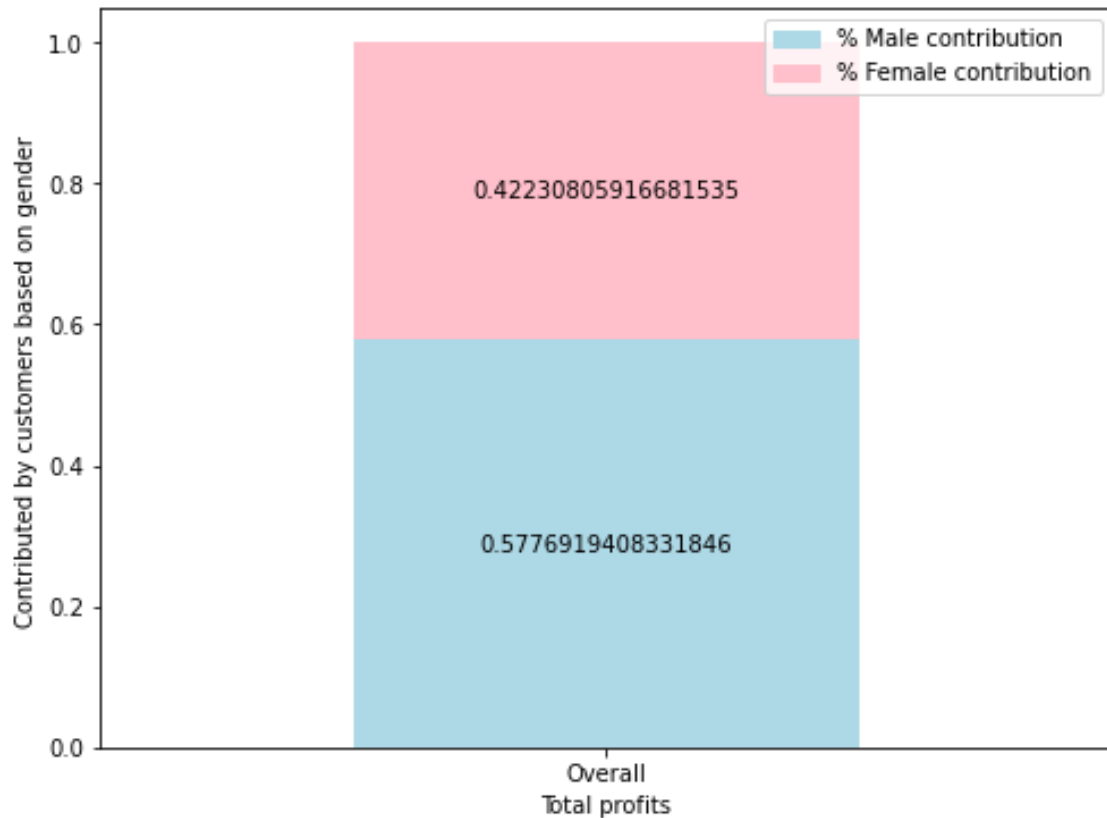
Company	Top 1 (frequent)	Top 2	Top 3
Yellow Cab	New York City	Chicago IL	Washington DC
Pink cab	Los Angeles LA	New York City	San Diego CA

Overall Profits for the pink cab company contributed by gender



	Total number of transactions made	Transactions made by female customers	Transactions made by male customers
Yellow Cab			
Pink cab	84711	37480	47231

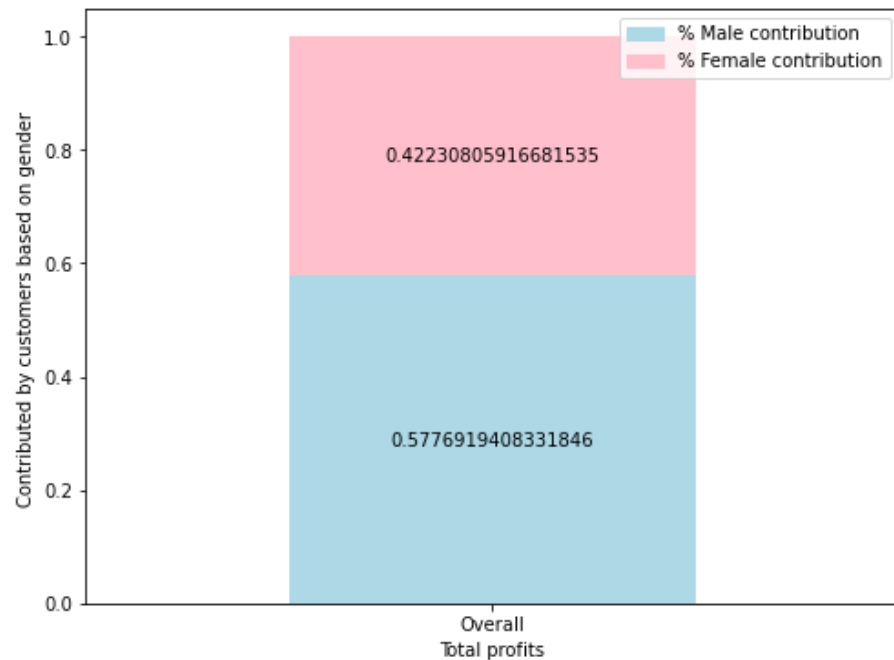
Overall Profits for the yellow cab company contributed by gender



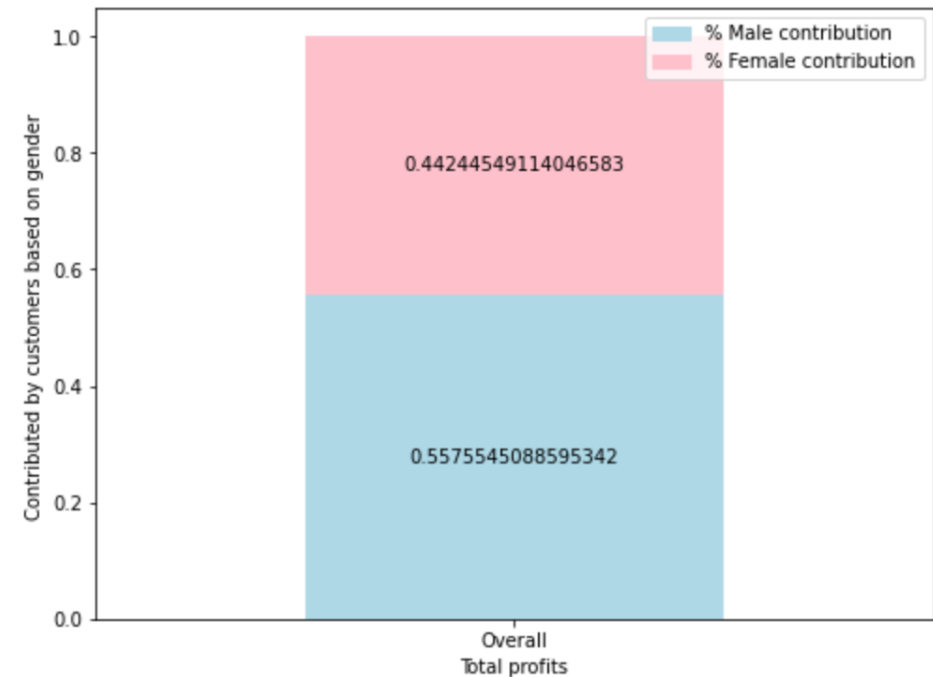
	Total number of transactions made	Transactions made by female customers	Transactions made by male customers
Yellow Cab	274681	116000	158681
Pink cab	84711	37480	47231

Customer Analysis Part 1

Overall Profits for the yellow cab company contributed by gender

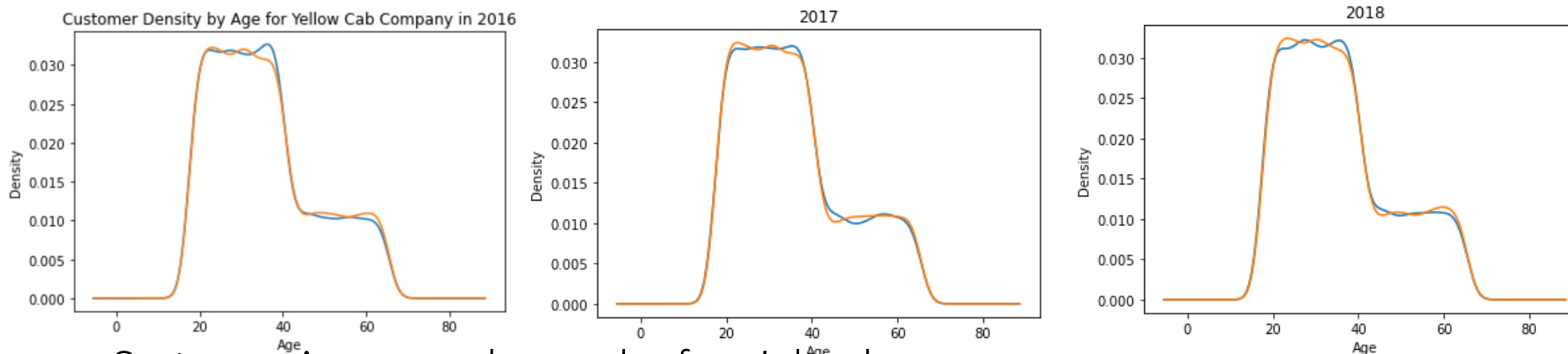


Overall Profits for the pink cab company contributed by gender

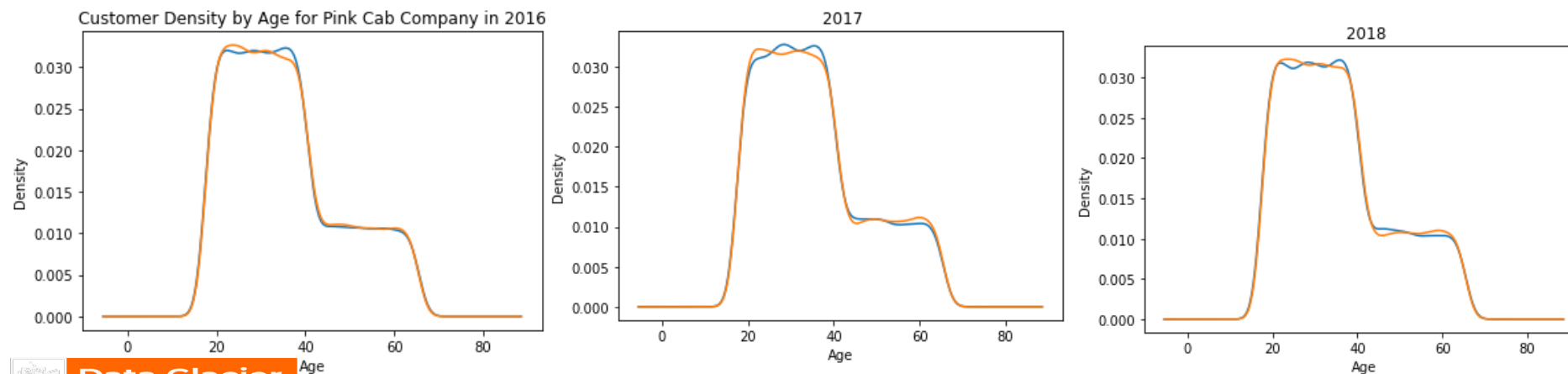


EDA SUMMARY: Both yellow and pink cab companies have more transactions from male customers.

Customer Age group by gender for yellow cab company



Customer Age group by gender for pink cab company



Recommendation

The logo for Yellow Cab Company is centered within a yellow, ribbon-like banner that has a 3D effect with shadows. The banner is set against a dark gray background. The text "Yellow Cab Company" is written in a bold, orange, sans-serif font.

Yellow Cab
Company

After analyzing different data sets about two companies, I have concluded that Yellow Cab company is more worthwhile of investing in for the following reasons:

Profit-wise advantages:

- Yellow Cab Company's profit per transaction is 60% higher than that of Pink Cab Company.
- Yellow Cab Company's profit loss rate is lower than that of Pink Cab Company.
- Yellow Cab Company has more customers in most cities whereas Pink cab company only has more customers in 4 cities.

Customer-wise advantages:

- Yellow Cab Company has more customers overall.

The end

Thank You 😊