



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

<G2M Cab Investment>

Name: Özgür Ateş

Team: Data Science

Date: 21 December 2022

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Executive Summary

- 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.
- Objective:
Make an analysis and present your analysis to the firm and make a recommendation about which company to invest in.

This analysis consists of:

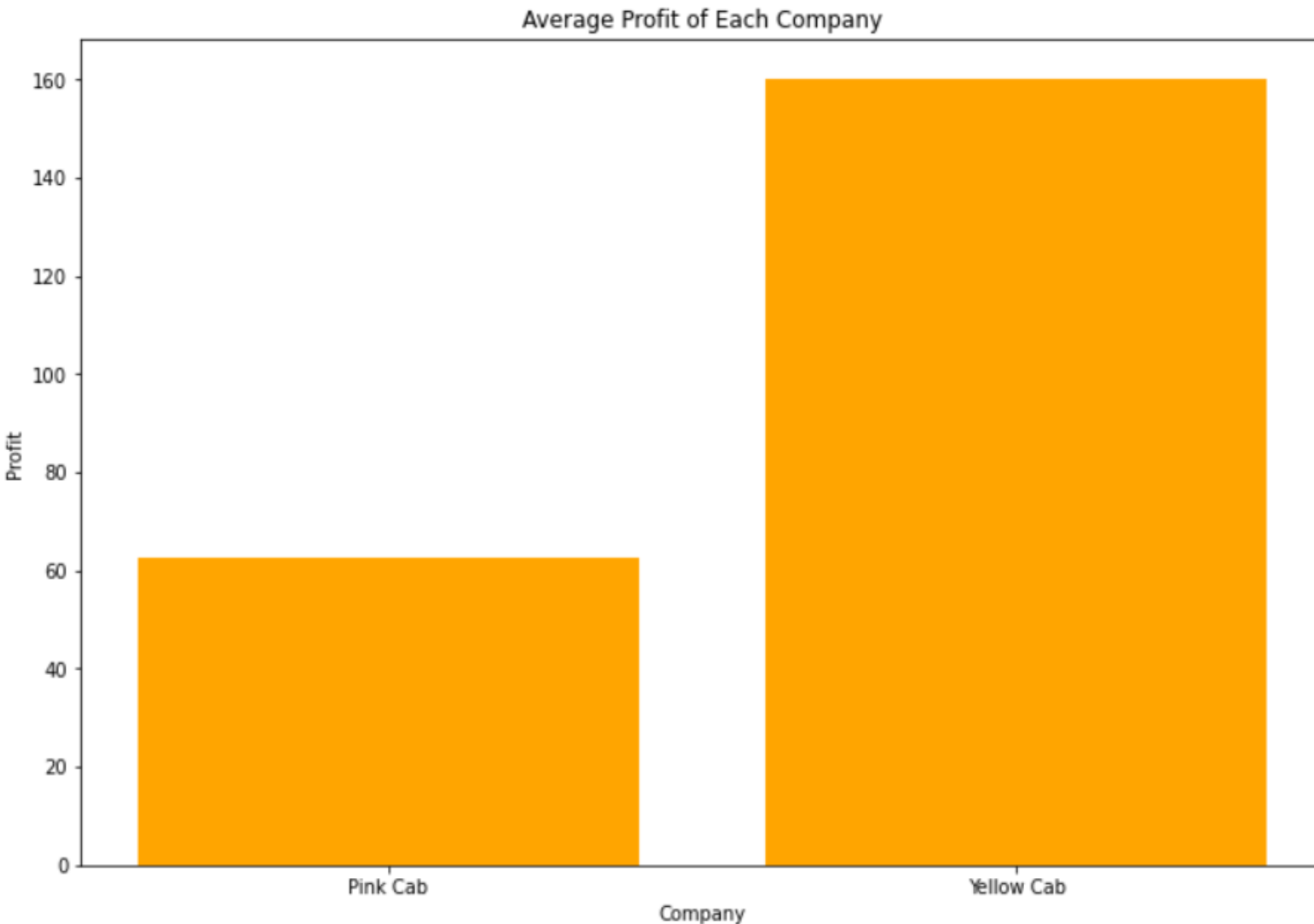
- Data analysis
- Data visualisation
- Hypothesis testing

Approach

Four Datasets have been provided to us:

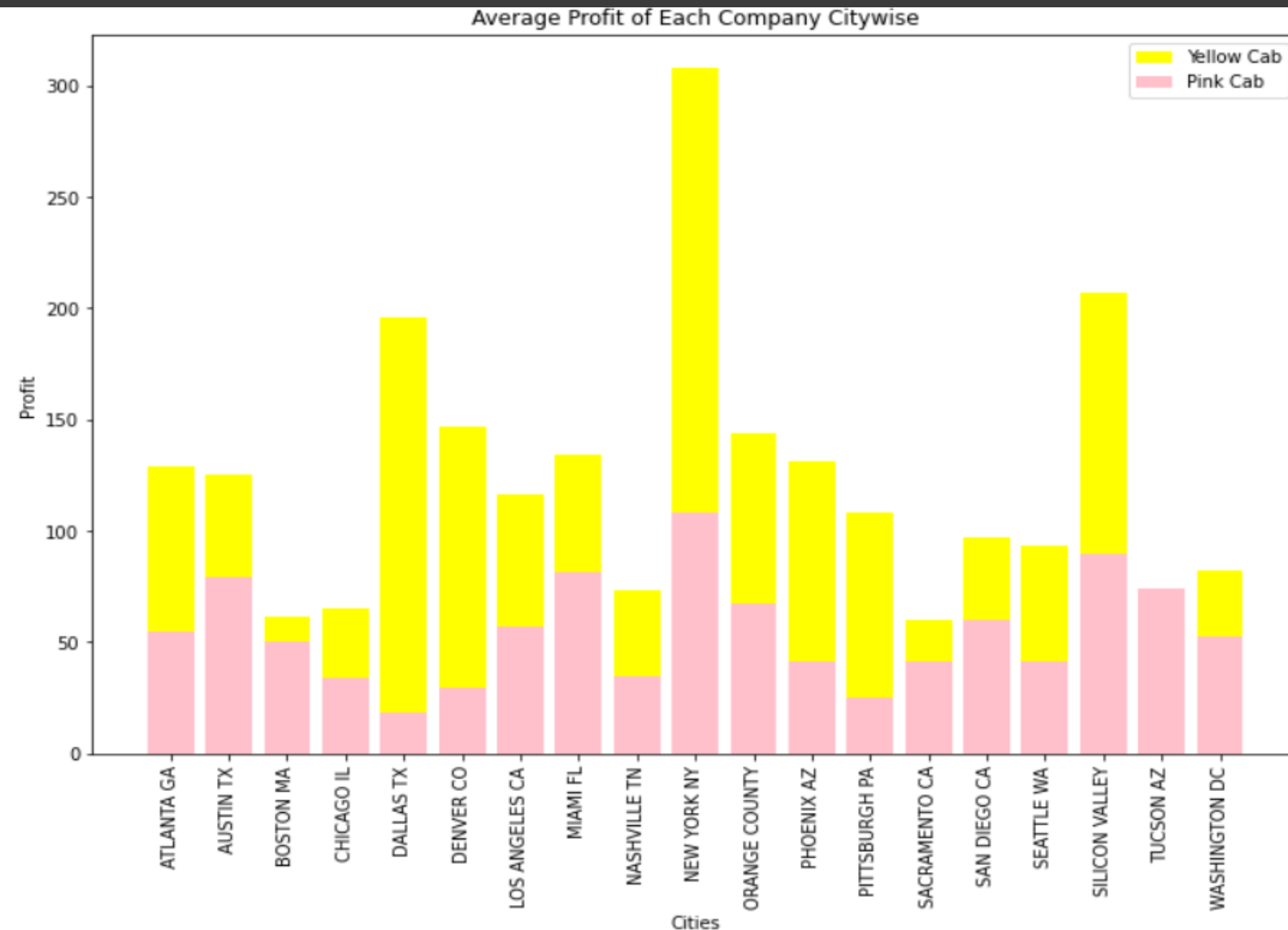
- **Cab_data.csv** - this file includes details of transaction for 2 cab companies
- **Customer_ID.csv** – this is a mapping table that contains a unique identifier which links the customer's demographic details
- **Transaction_ID.csv** – this is a mapping table that contains transaction to customer mapping and payment mode
- **City.csv** – this file contains list of US cities, their population and number of cab users

EDA – Company Profits



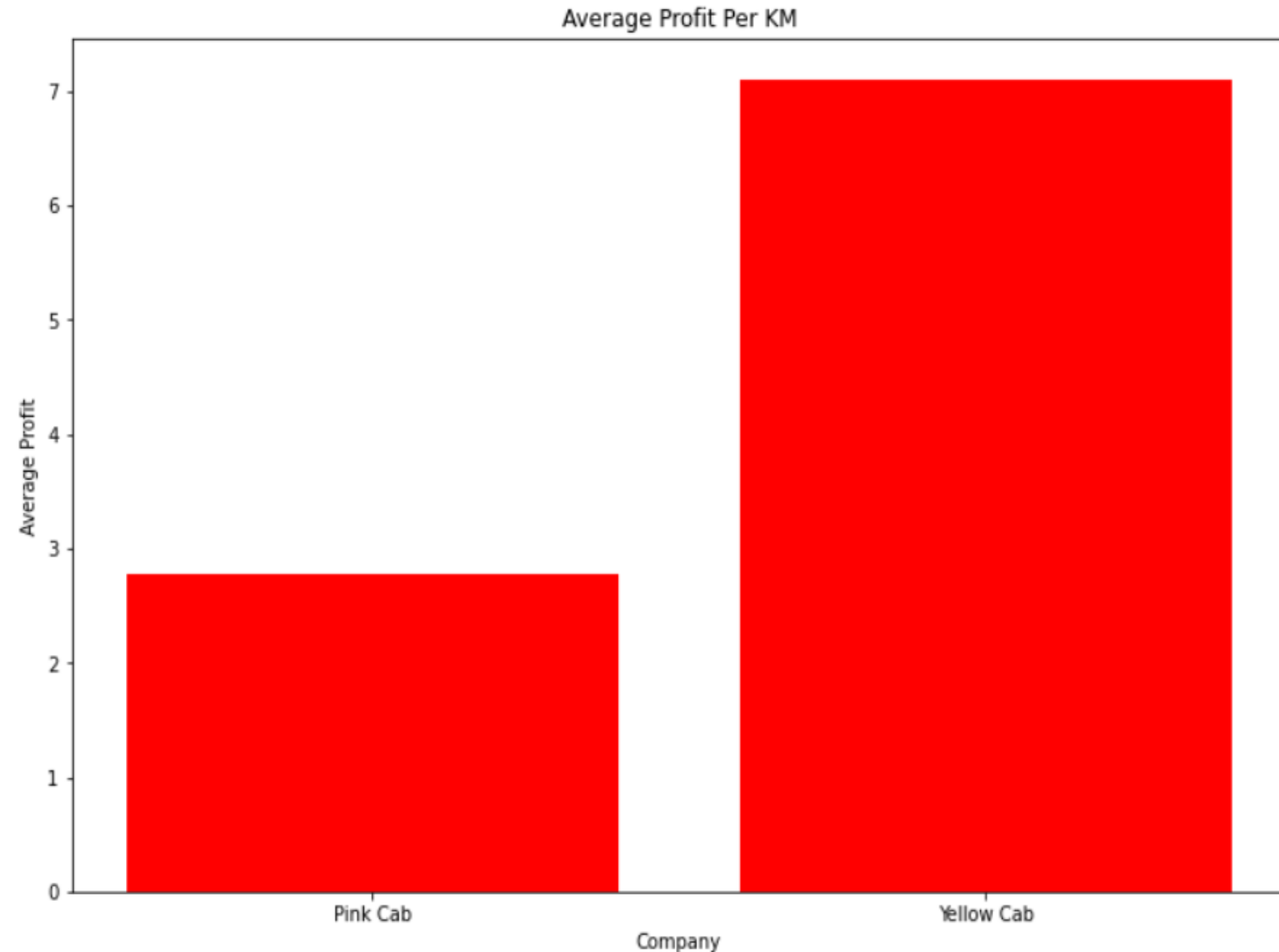
- It is understood from here that Yellow Cab Company gains way more profits compared to the Pink Cab Company. Since Yellow Cab Company Transaction number was way many compared to the Pink Cab Company, we took the mean of the profits.

EDA – Company Profits Citywise



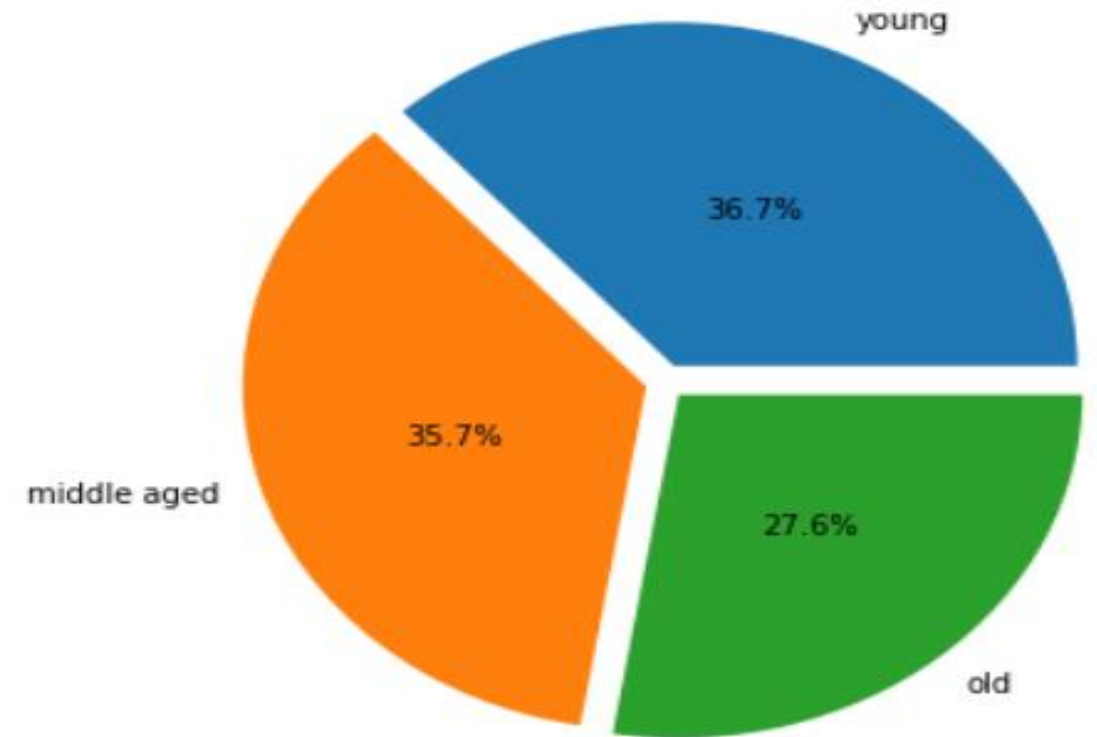
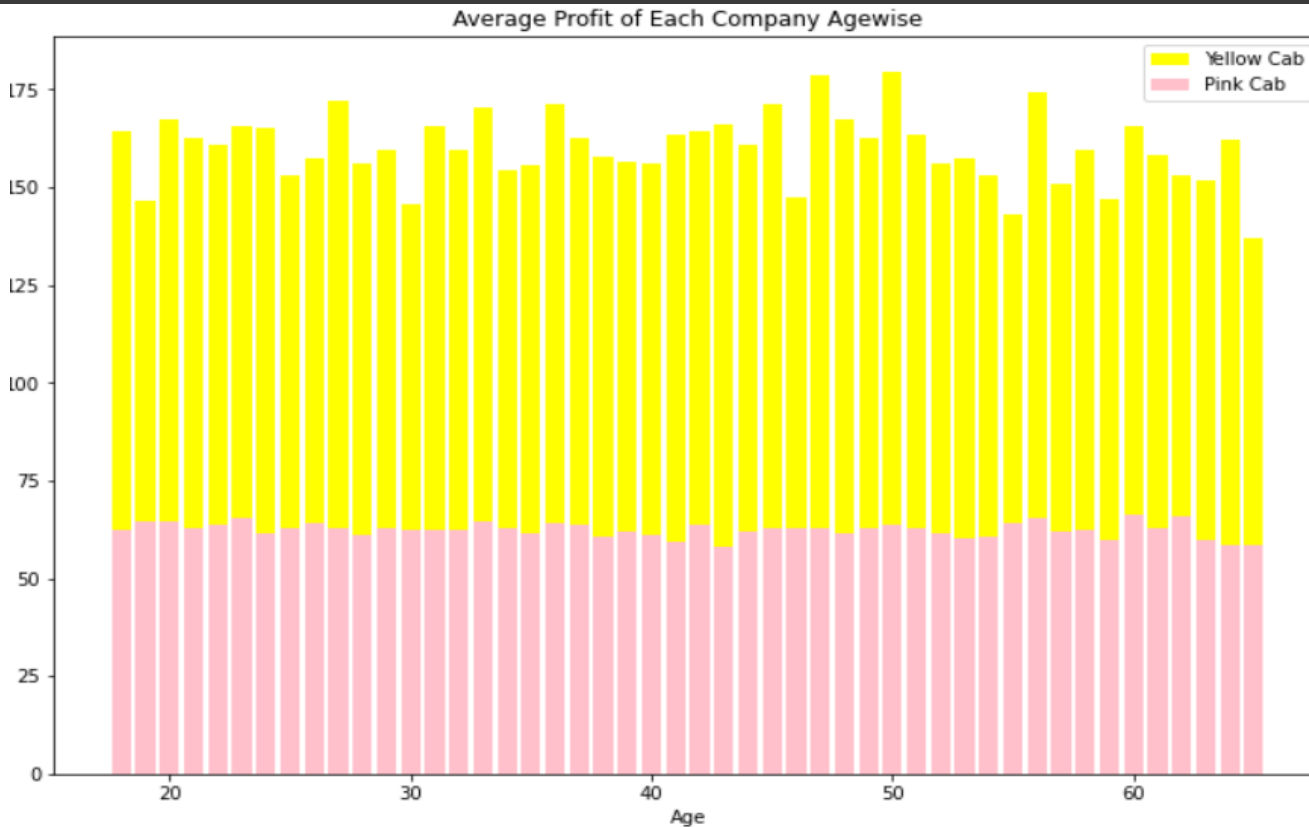
- It can be seen from this figure that the city of New York is the most active city in terms of making drives.
- We can also see that for Pink Cab Company Dallas TX city is the least drives made city and for Yellow Cab Company that city is Sacramento CA.
- We can see the amount of profit differences between the two companies Yellow Cab Company has an insane amount of advantage in this field.

EDA – Average Profits Per KM



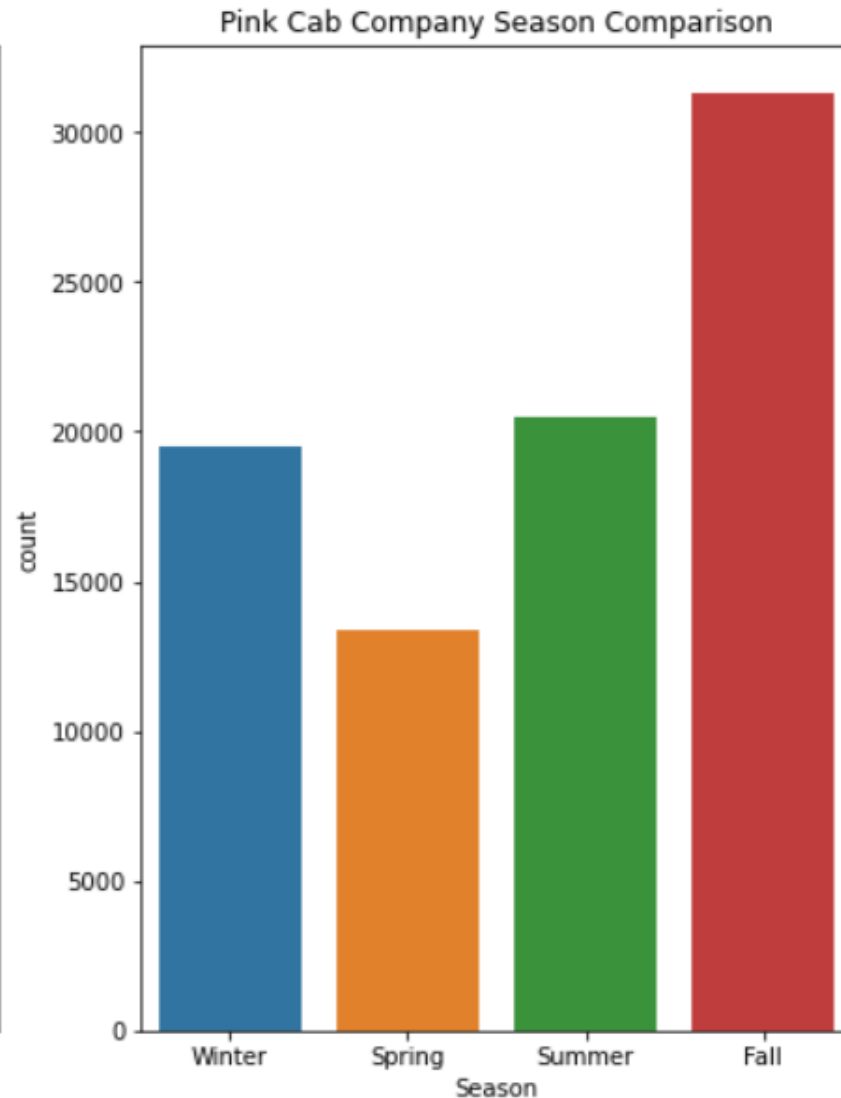
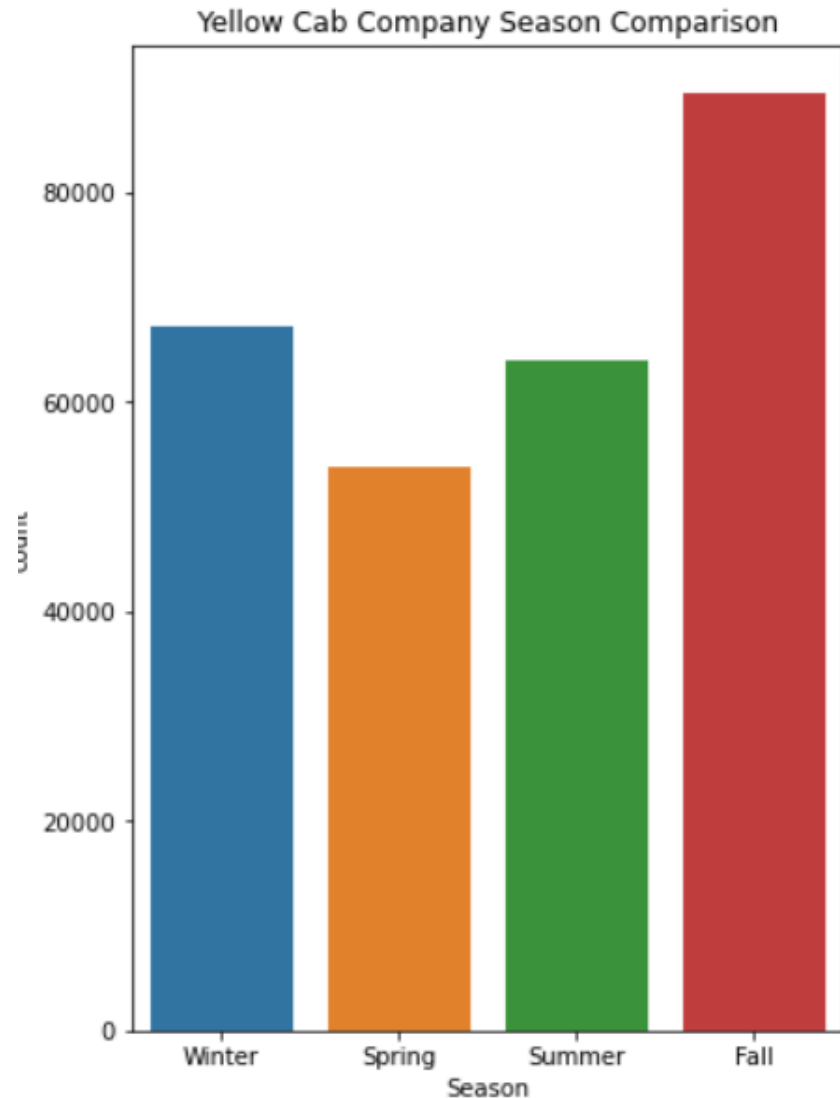
- We saw that Yellow Cab Company makes more profits but what if we check the average profits made per KM taken.
- Well the result doesn't change, still Yellow Cab Company makes about 2.3 times more profits per KM taken.

EDA – Average Profit Agewise



From these graphs, we can see that young people and middle aged people contribute almost same to the profits while young people contribute just a little more. Old people are not far behind but they contribute less compared to young and middle aged people.

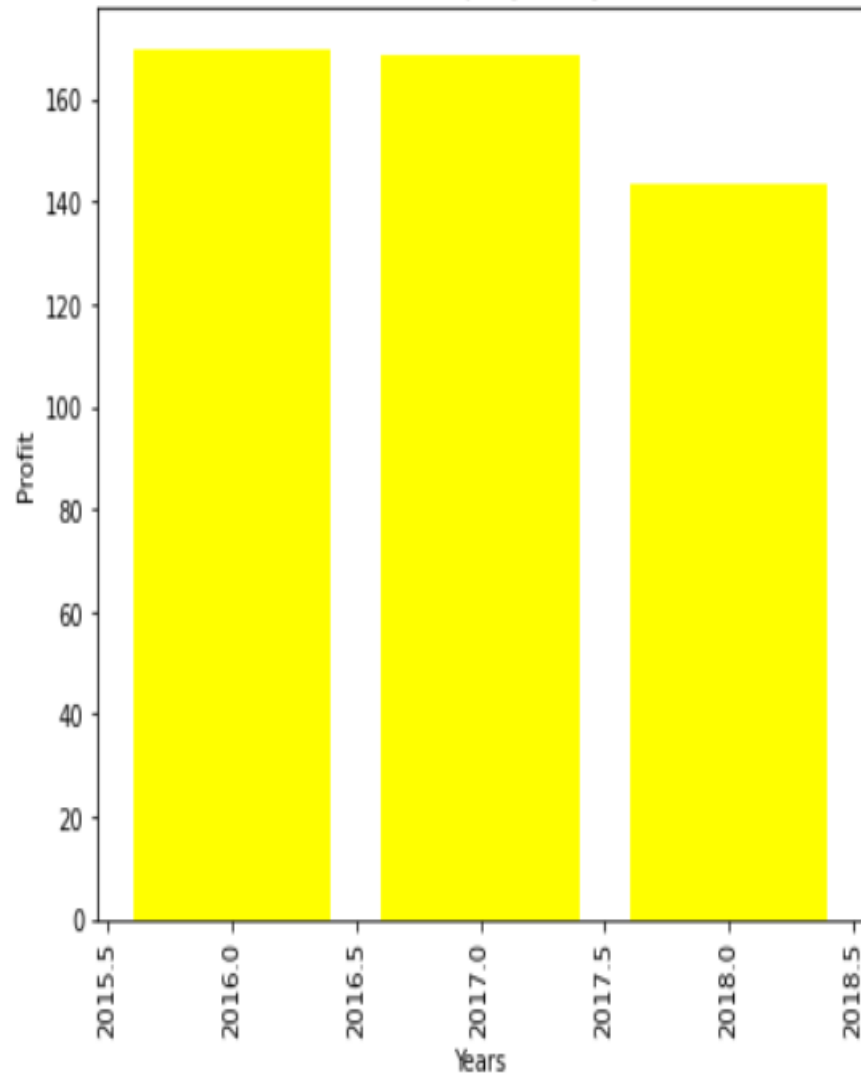
EDA – Number of drive changes from Season



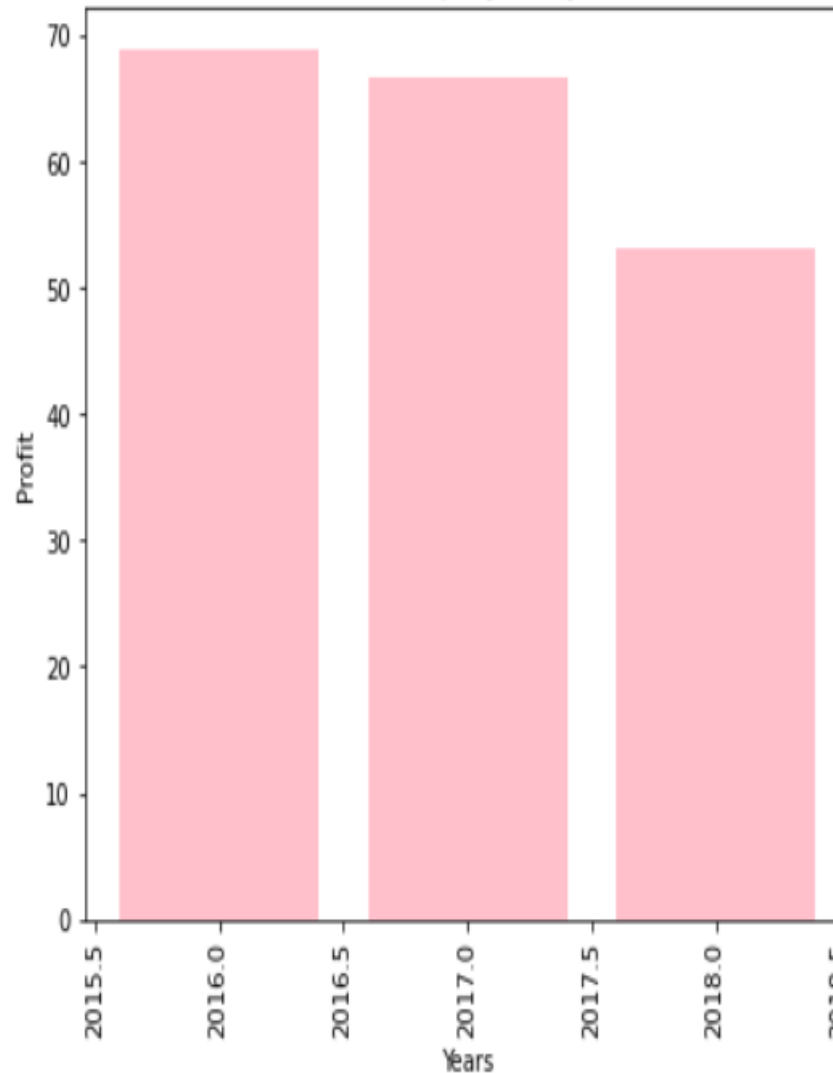
- From these graphs, we can see that Fall season is the most drives made for each companies and Spring is the least drives made for each companies. In Pink Cab Company Summer Season drives are just a little more than the Winter Season while in the Yellow Cab Company it's the other way around.

EDA – Profit changes throughout years.

Yellow Cab Company Yearly Profits

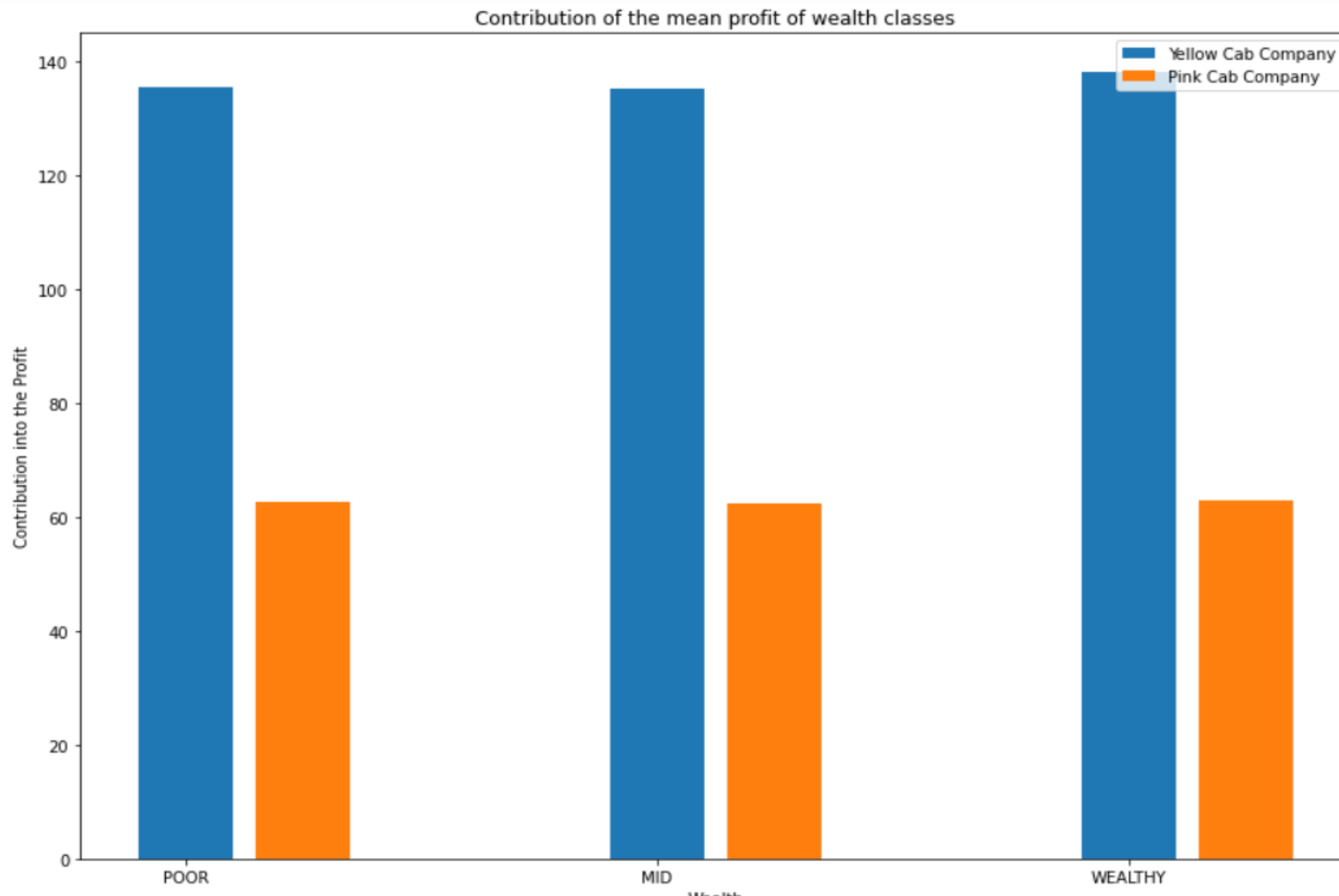


Pink Cab Company Yearly Profits



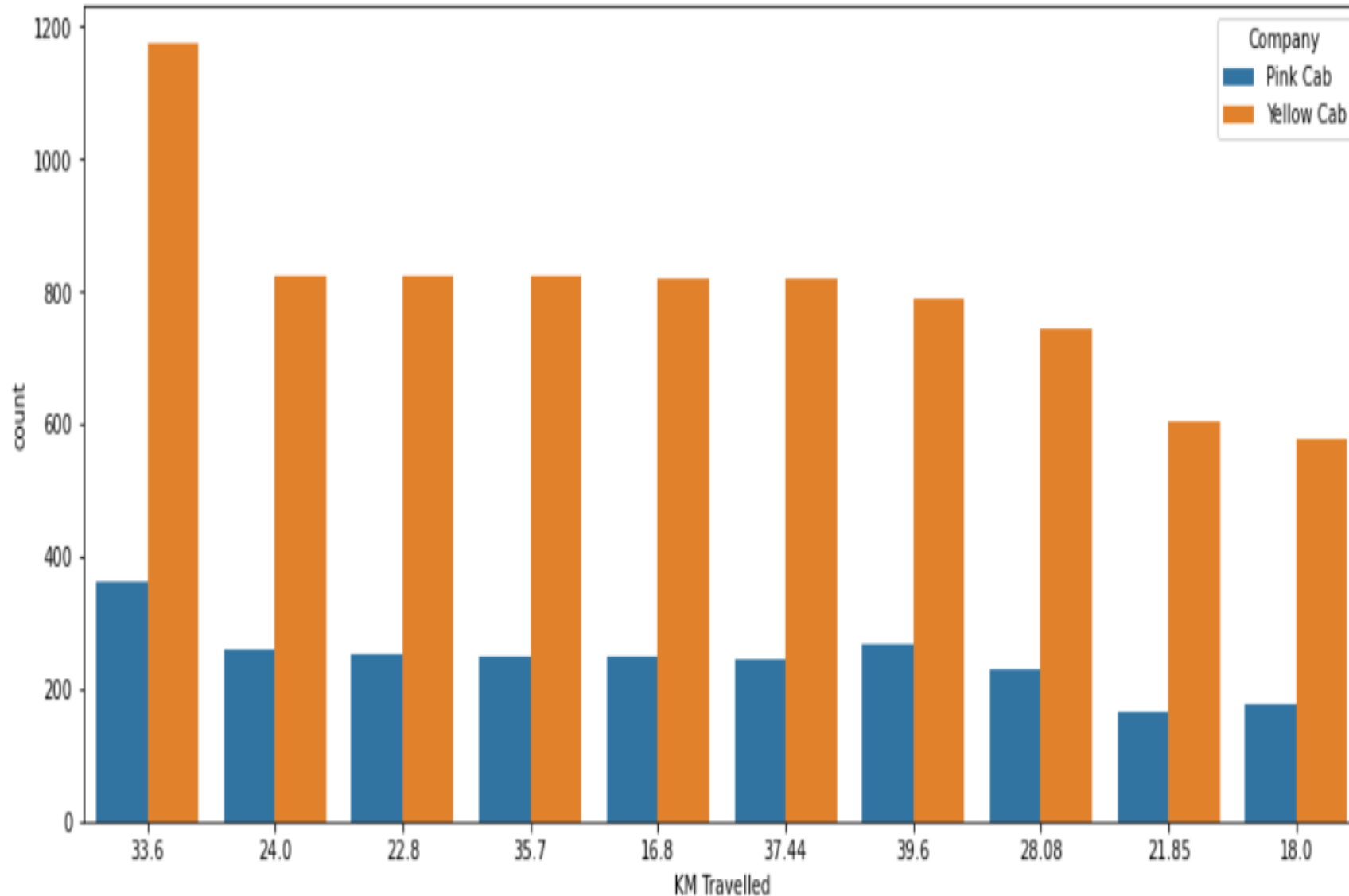
- From these graphs, we can see that from 2016 to 2017 both companies suffered a small profit loss while in Pink Cab Company that margin is a little bigger.
- From 2017 to 2018 both companies suffered a big amount of profit loss.
- But still, Yellow Cab Company profits are almost always 2 times more than the Pink Cab Company.

EDA – Profit contribution of social classes



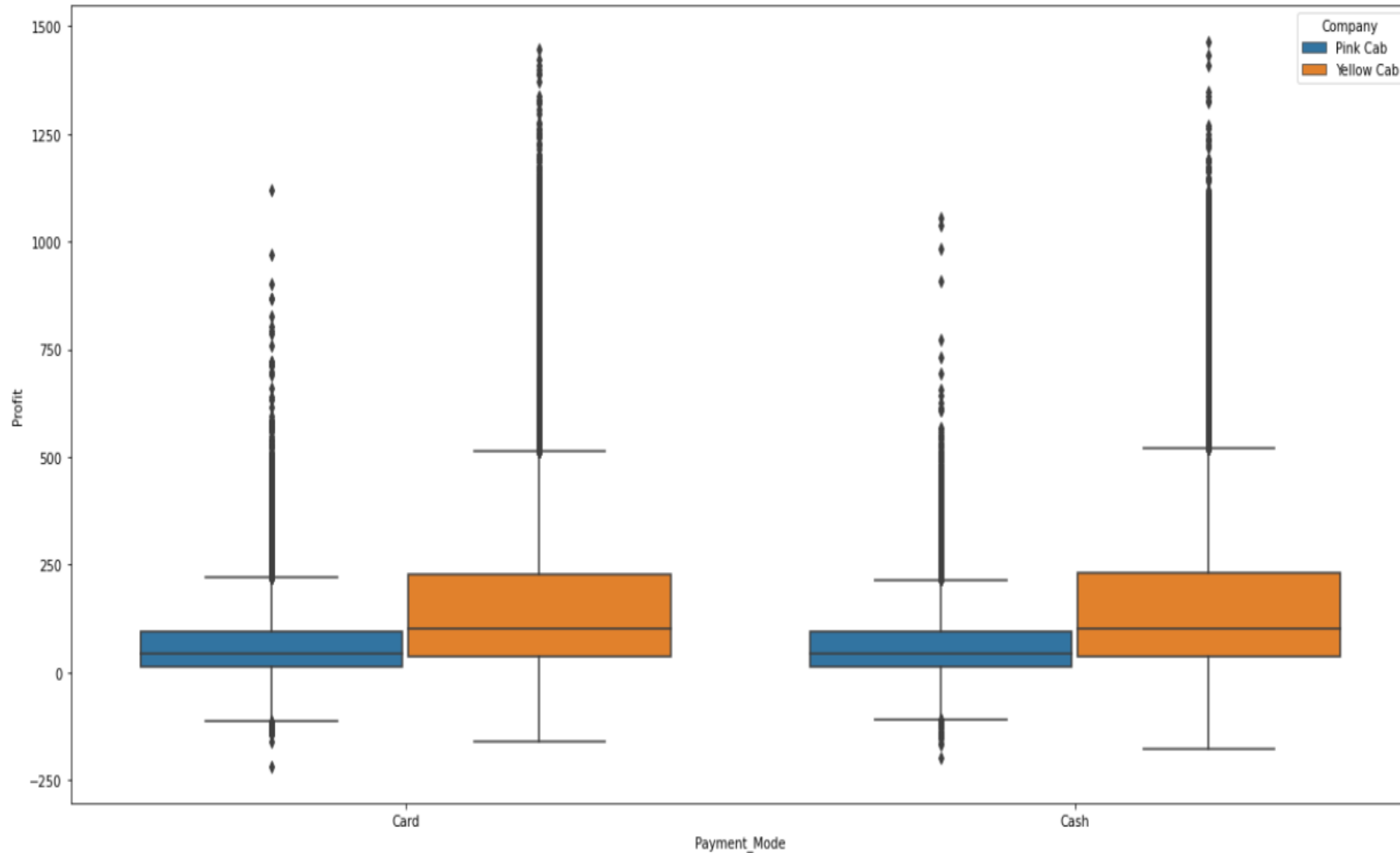
- We can see that Poor Mid and Wealthy classes are all contribute the same amount to the companies profits for each companies.

EDA – Number of drive changes from Season



- We can see that both companies travel around 20 to 40KM mostly.

EDA – Number of drive changes from Season



- From this graph, we can see that both card and cash payments contribute almost the same amount to the companies profits while Yellow Cab Company profits are bigger than the Pink Cab Company.

Hyphotesis Testing

Hyphotesis One:

Females bring less profit than Man

Conclusion = For Yellow Cab Company the hyphotesis is false while for Pink Cab Company, the hyphotesis is true.

Hyphotesis Two:

Does profit/margins change depending on gender of customer?

Conclusion = For Yellow Cab Company the hyphotesis is false while for Pink Cab Company, the hyphotesis is true.

Hyphotesis Three:

The mean Profit for the different Age groups for Cab Companies are equal.

Conclusion = For both companies, this hyphotesis is False.

Recommendations

- **Transaction Analysis:** Yellow Cab has domination over Pink Cab Company in terms of profits made.
- **Gender:** Yellow Cab Company has the most transactions made for both male and female.
- **Age:** Yellow Cab Company has the most transactions made for all age classes and ages.
- Even though both companies lost a profit in 2018 Yellow Cab Company lost less.

Therefore, I recommend to invest in Yellow Cab Company.

Thank You