Exploratory Data Analysis

G2M Insight for Cab Investment Firm

June 9, 2022

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Data Glacier

Your Deep Learning Partner

Executive Summary

- There are two cab companies that's data has been provided from 01/31/2016 to 12/31/2018

- The goal is to create an analysis so that the company knows which cab company to invest in based on this data

- There will be analysis of all aspects of both the companies so that XYZ will be able to put their money in the right place

Problem Statement

- Which cab company should XYZ invest in based on the trends in the data?

Approach

- Look at the data to see the different unique and foreign keys so that we are able to join all of the tables together

- Look sort through each of the tables to see if there are NULL values so that we can decide what to replace them with.

- Go through the data and find if there are any duplicate values

- Aggregate data based on different categories to see any trends between cab companies and these groups

- Analyze profits for each company based on this aggregated data to find which would be the better investment

Transactions

- If we look at the breakdown of the transactions had between each of the companies by city, month and year we can clearly see that Yellow Cab company is destroying the market in every fashion
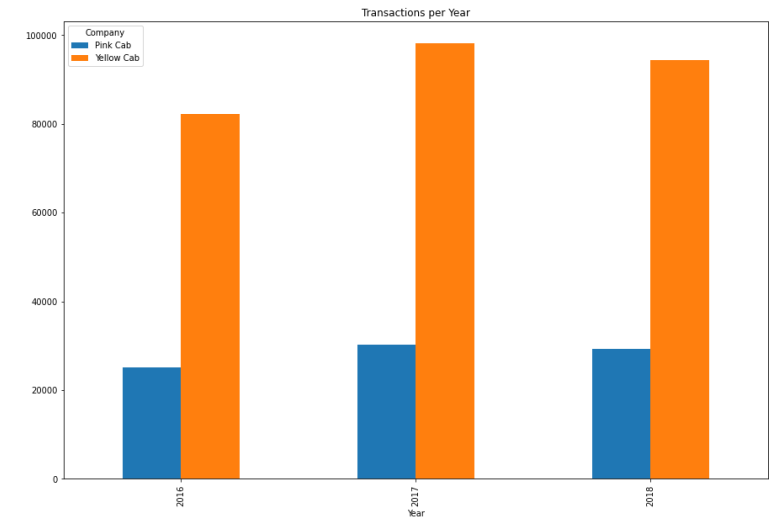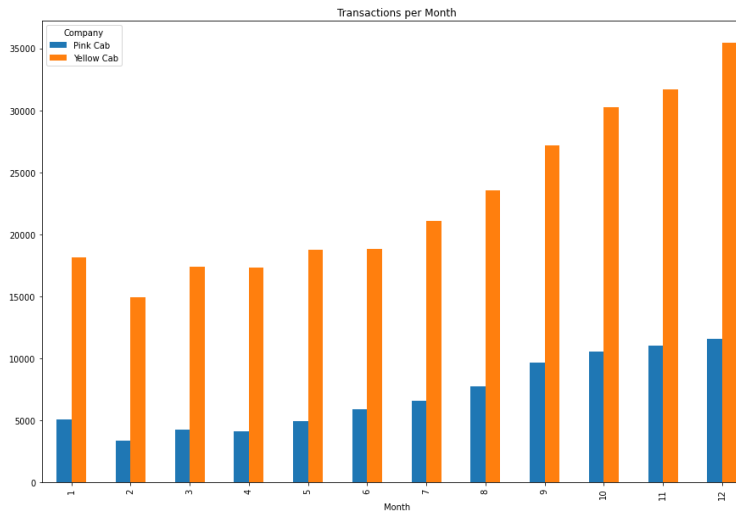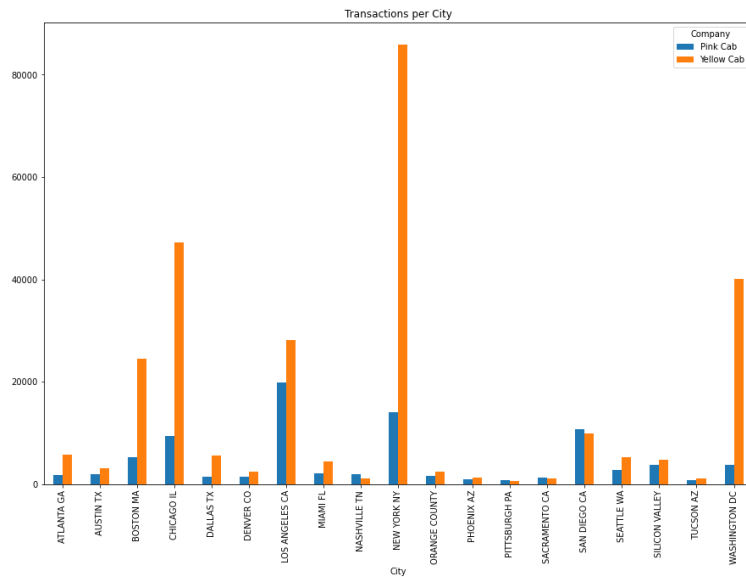
Transactions By City Each Year

- Looking at the transaction amounts per city by each of the years we can see that there is not really an upward trend of Pink Cab coming up on Yellow Cab as Yellow Cab is still dominating the market in almost all cities through the three years
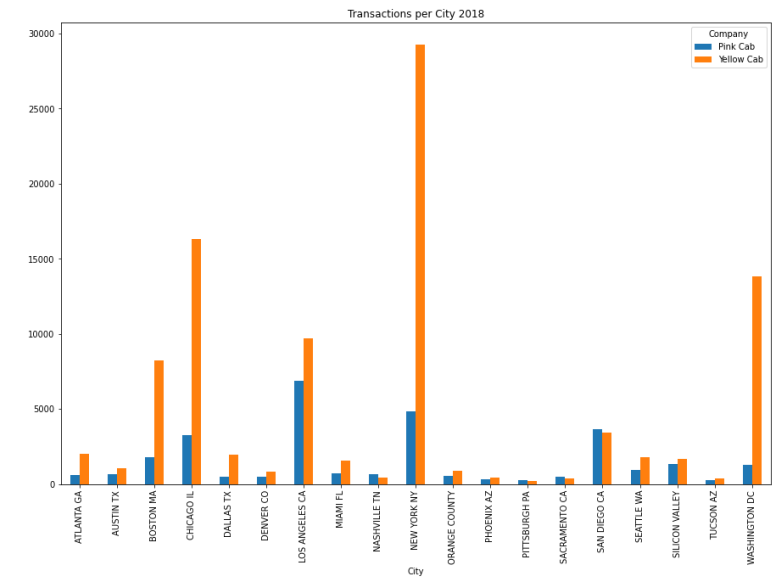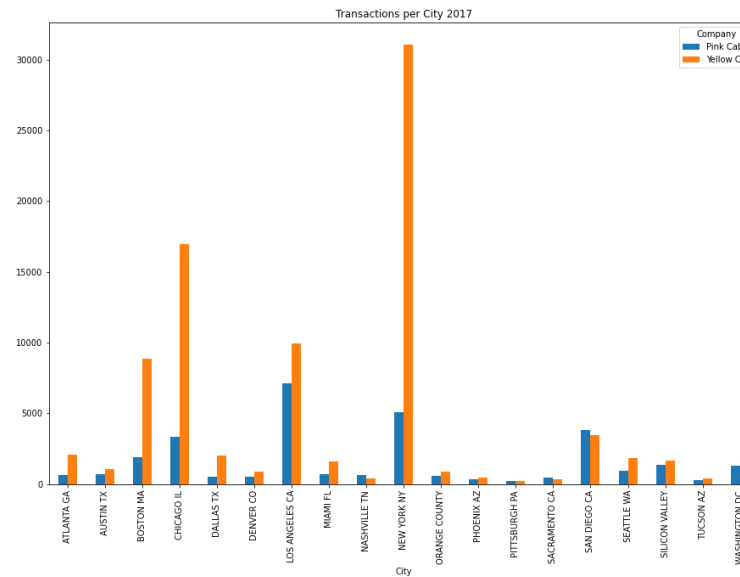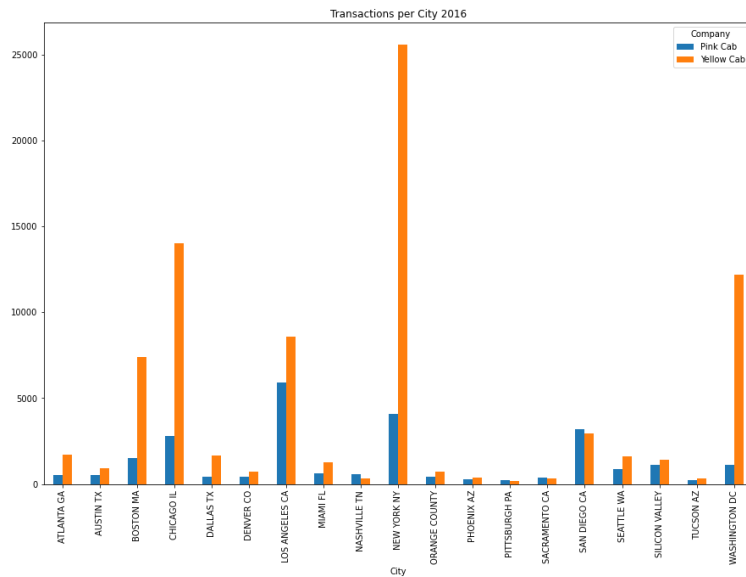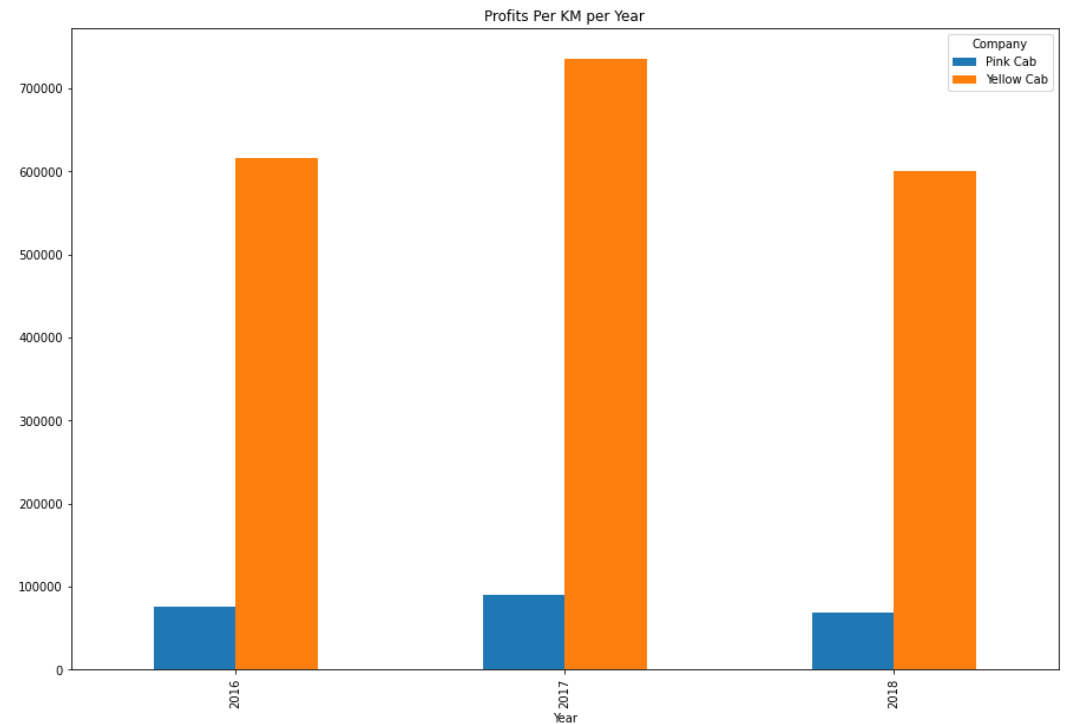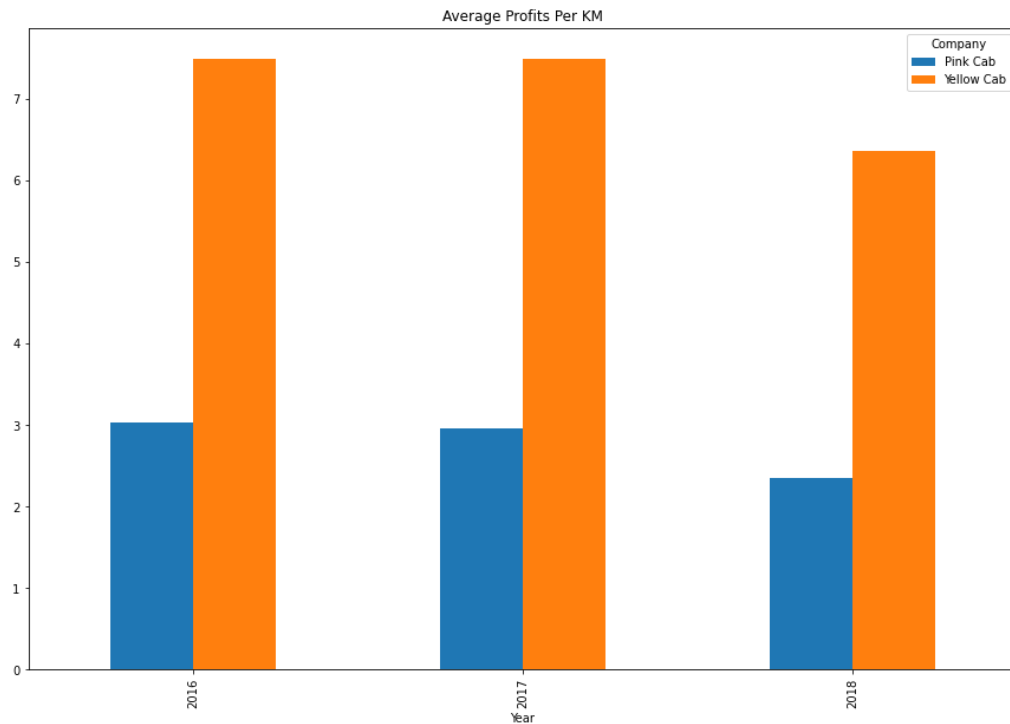
Profits

- Looking at the profits broken down by year and month we can clearly see that Yellow Cab is destroying Pink Cab in profits
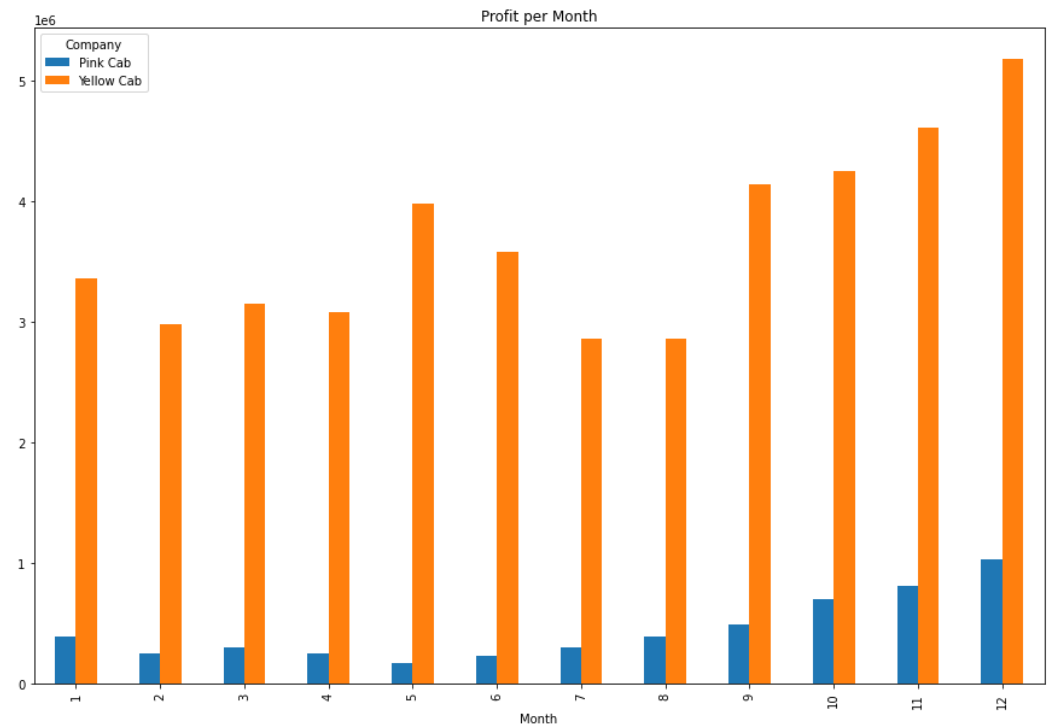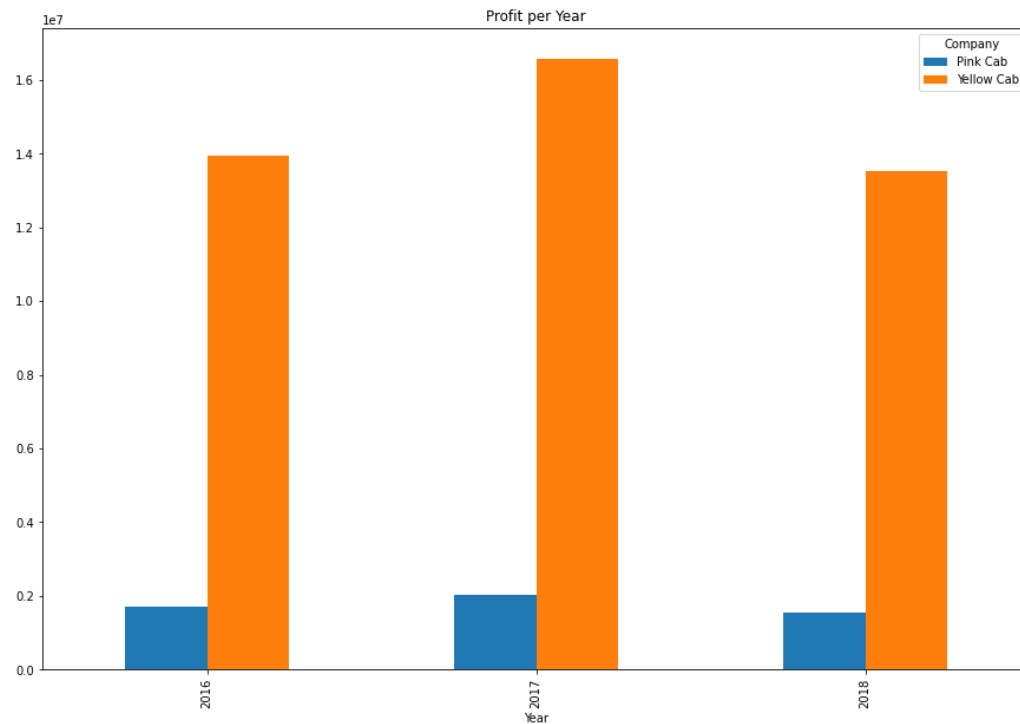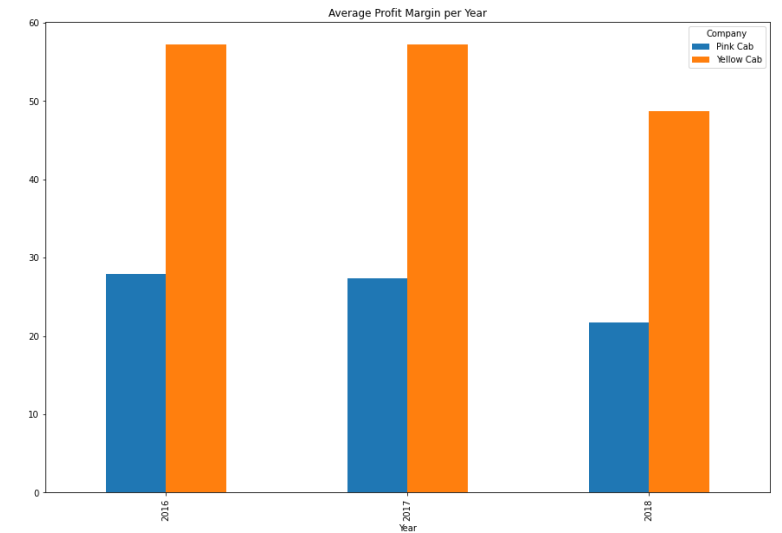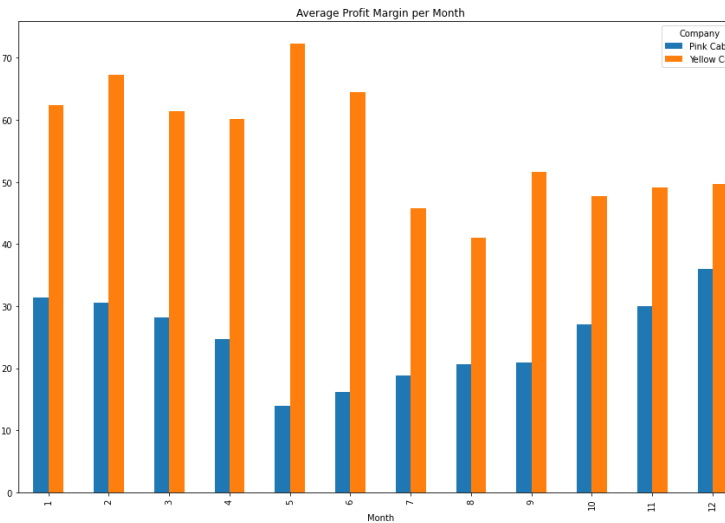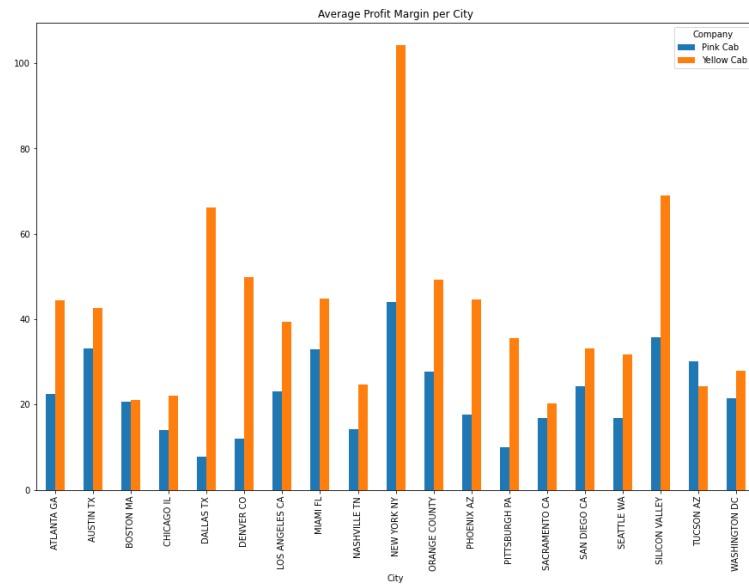
Profits Per KM

- Then taking the average profits per KM and the profits per KM Yellow Cab destroyed Pink Cab every year to the point it is barely comparable

Profit Margins

- Then taking a look at the average profit margin broken down by city month and year, we can clearly see that Yellow Cab's profit margins are extremely higher than Pink Cab's profit margins

Customer Breakdown

- Finally breaking down the transactions by the customers we can see that Yellow cab is dominating every category whether it be every age that are taking cab trips to overall male and female to male and female breakdown over the years.
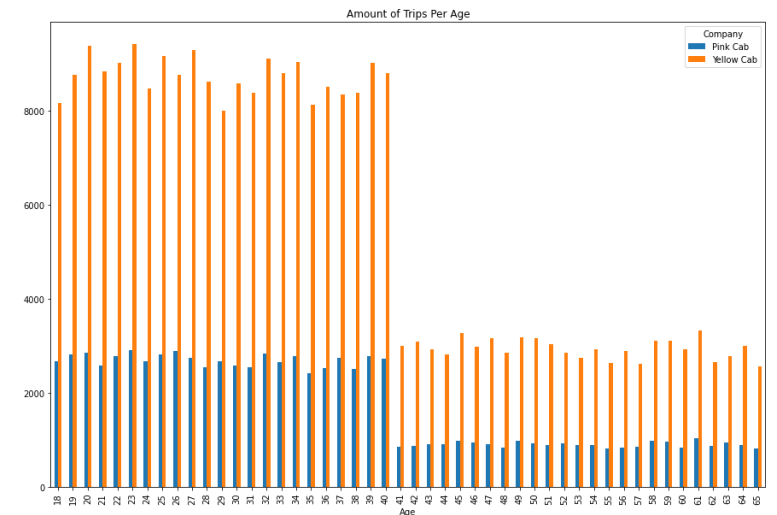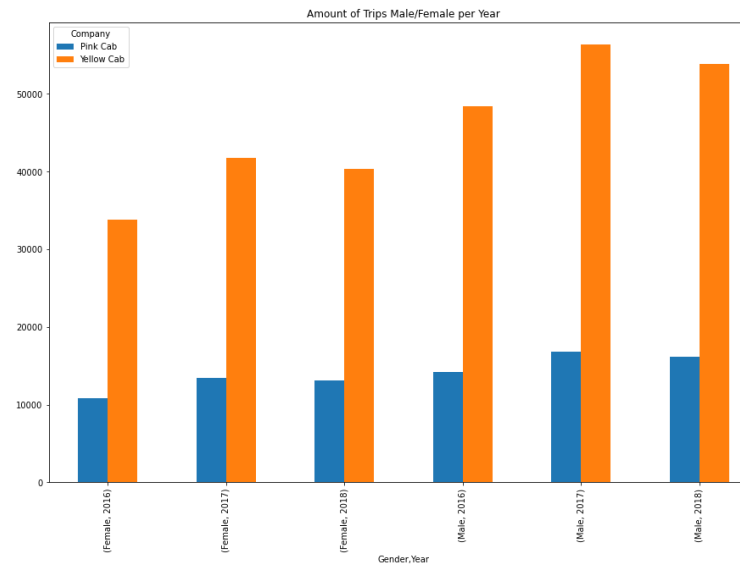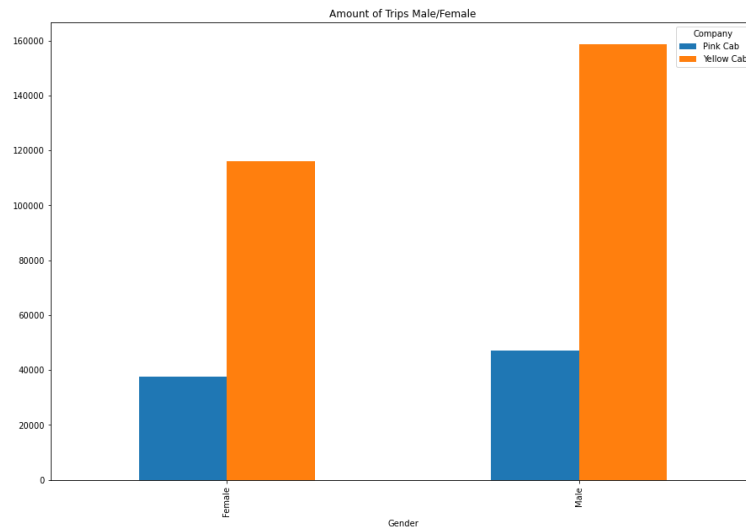
EDA Summary

- After breaking down each of the companies by categories we can see that Yellow Cab is dominant in many categories

- They dominate in the profits they make in every way along with the amount of trips they are getting

- Yellow Cab also dominates the market in every age group along with male and female customers

Recommendations

We have evaluated the cab companies in the following ways:

- Amount of Transactions broken down by city, month and year: Yellow Cab dominated the amount of transactions they had per month year and had more transactions in almost every city even when broken down by year
- Profits: Yellow Cab had way more profits whether it be the total amount or the average profits per KM travelled
- Profit Margin: Yellow Cab had better profit margins in almost every city and destroyed Pink Cab's profit margins over every month and year
- Transaction Breakdown: When looking at the transactions, Yellow Cab dominated in both the female and male markets over the years and overall. They also have significantly more transactions for every age group

Based on all of the information provided by looking at the data, we recommend that XYZ invests in Yellow Cab

Thank You

Data Glacier
Your Deep Learning Partner