



**Data Glacier**

Your Deep Learning Partner

# Exploratory Data Analysis

G2M Insights for Cab Investment Firm

**05/20/2022**

# Agenda

- Executive Summary
- Problem Statement
- Approach
- EDA
- EDA Summary
- Model
- Recommendations

# Executive Summary

- 'Yellow Cab' had more transactions and customers, especially male customers and customers in big cities.
- 'Yellow Cab' made higher profit and profit per kilometer.
- 'Yellow Cab' had higher user reach percentage and was good at retaining royal customers.
- It is recommended to invest in the 'Yellow Cab'.

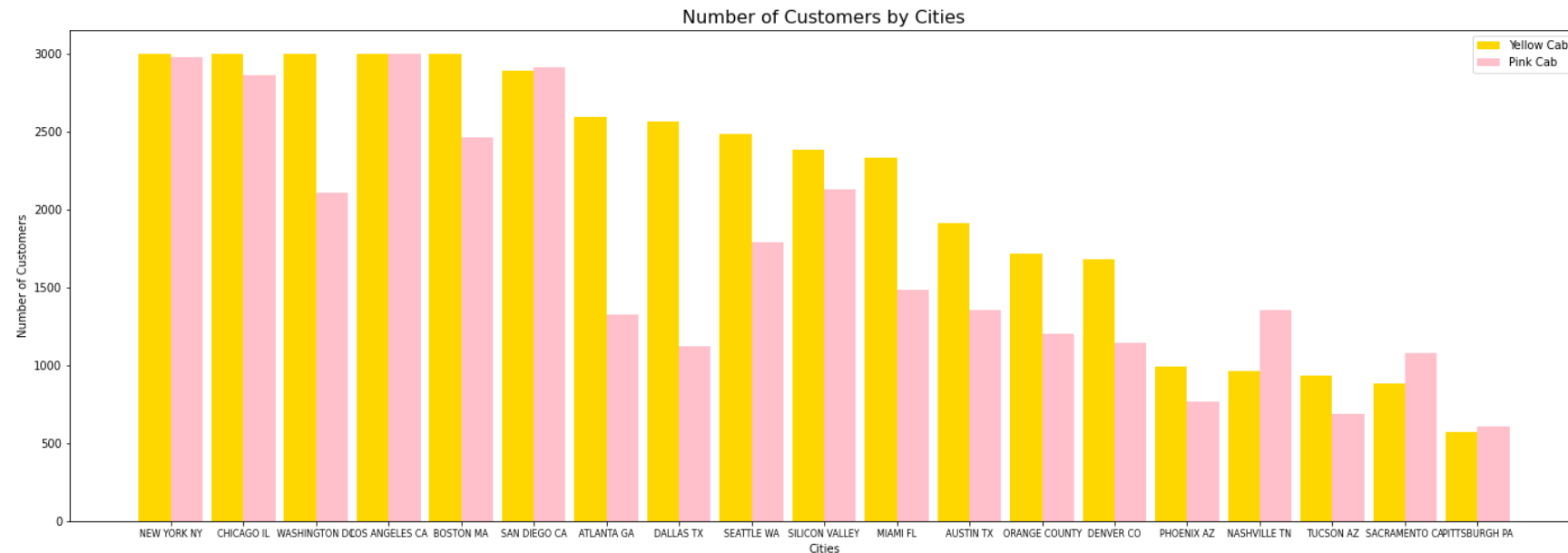
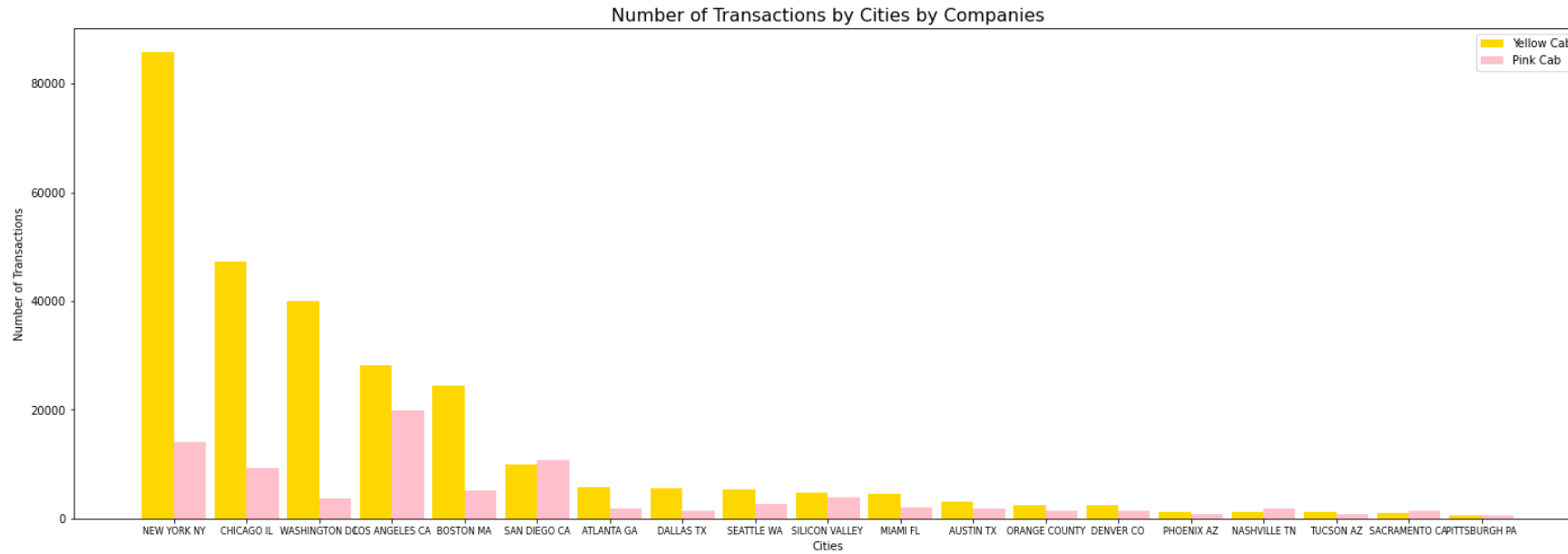
# Problem Statement

- Client:
  - XYZ is a private equity 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:
  - Help understand the cab industry
  - Provide Go-to-Market insights for identifying the right company to make investment.
- Business Problems:
  - Exploratory Data Analysis: Is there any seasonality in number of customers using the cab service? Which company has maximum cab users at a particular time period? Etc.
  - Modeling and forecasting profit and number of transactions for each cab company
  - Finding the most profitable cab company and making recommendations for investment

# Approach

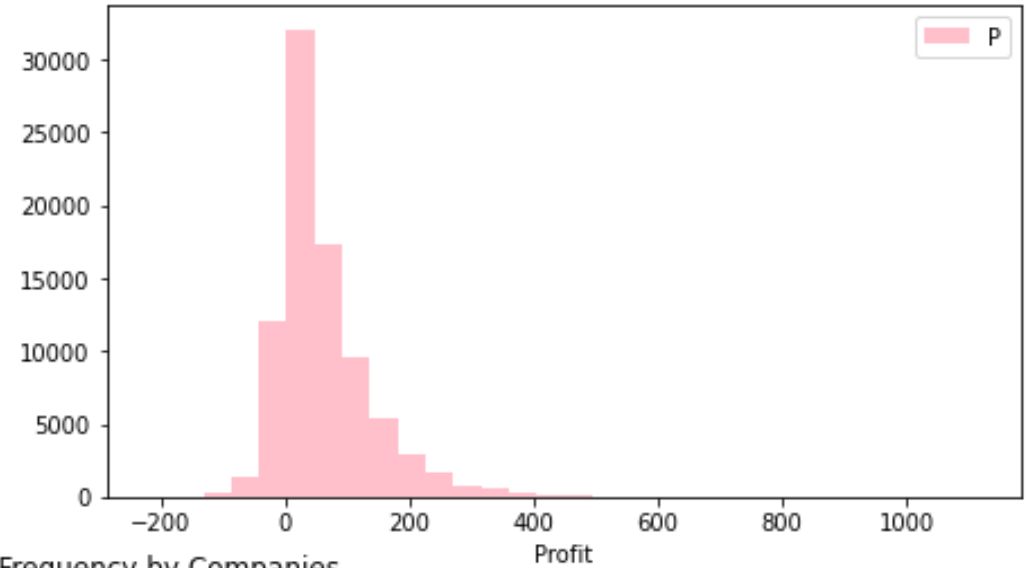
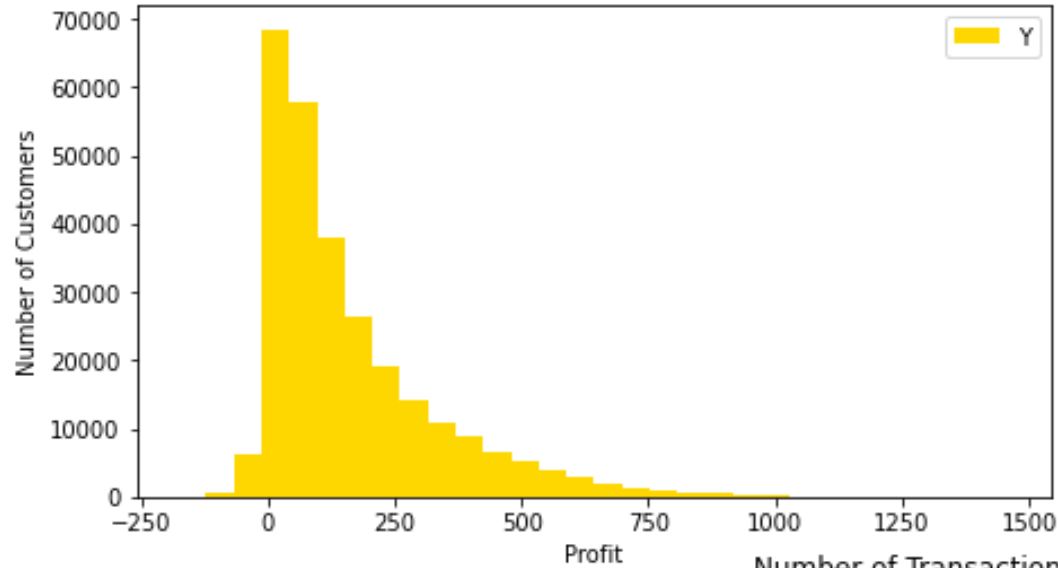
- Data:
  - Files: Cab\_Data.csv, Customer\_ID.csv, Transaction\_ID.csv, City.csv
  - Records: 359,392 (After merge)
  - Timeframe: 01/31/2016 - 12/31/2018
  - Features: 14 (After merge)
- Methods:
  - Dedupe Validation
  - Descriptive Analysis
  - Correlation Analysis
  - Contextual Analysis
  - Time Series Forecasting

# EDA - Descriptive Analysis

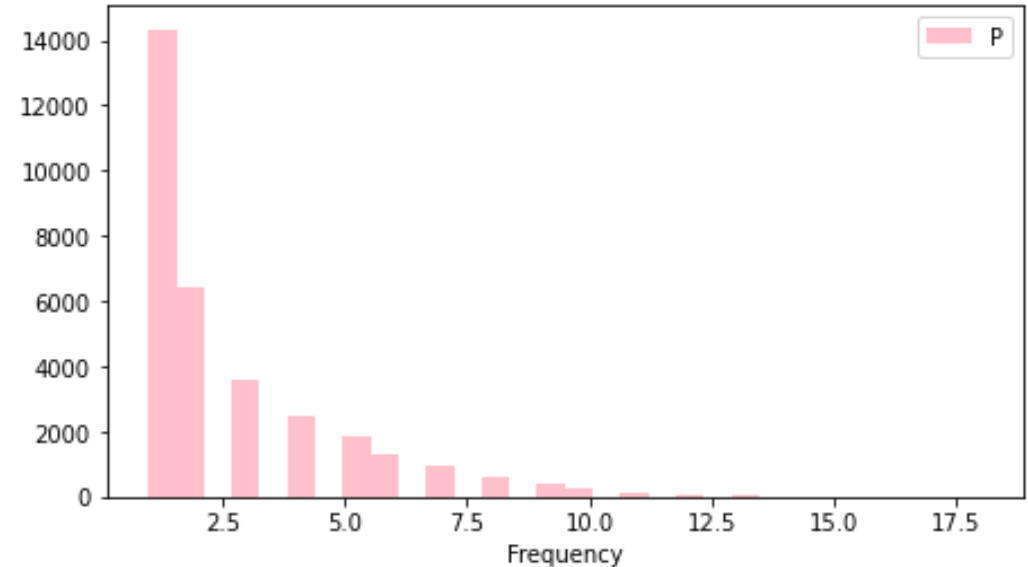
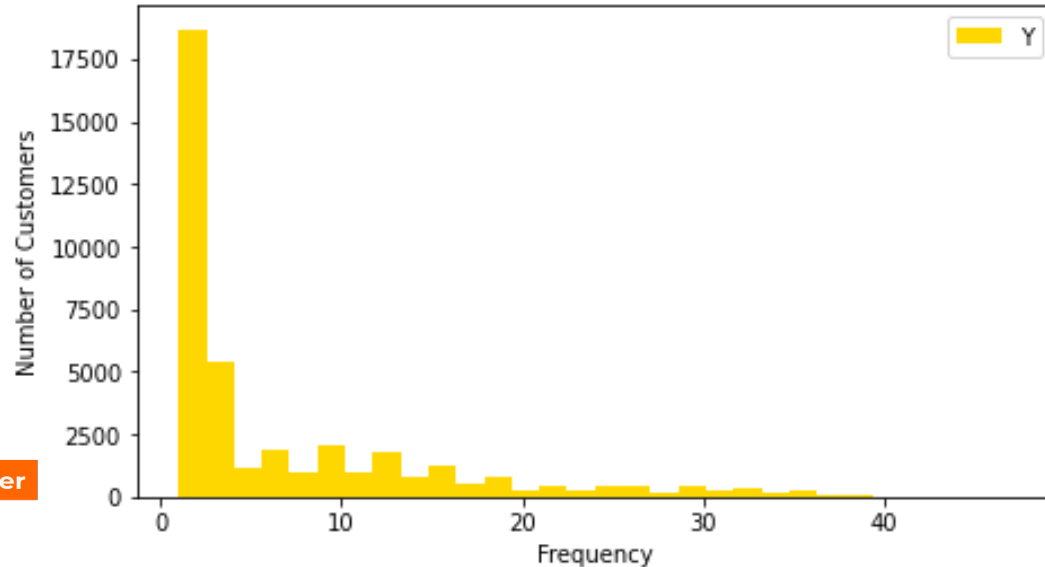


# EDA - Descriptive Analysis

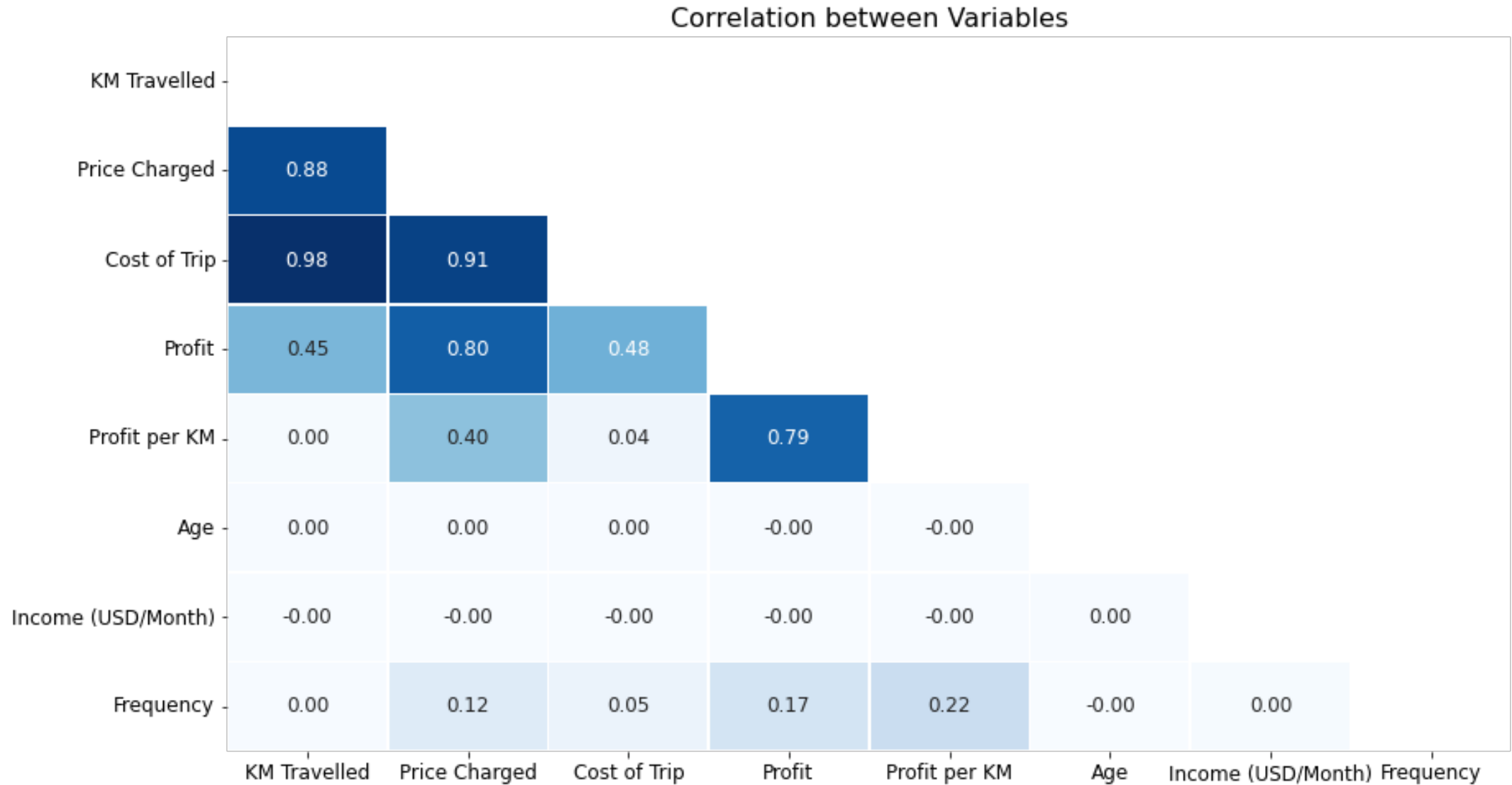
Number of Transactions by Profit by Companies



Number of Transactions by Frequency by Companies

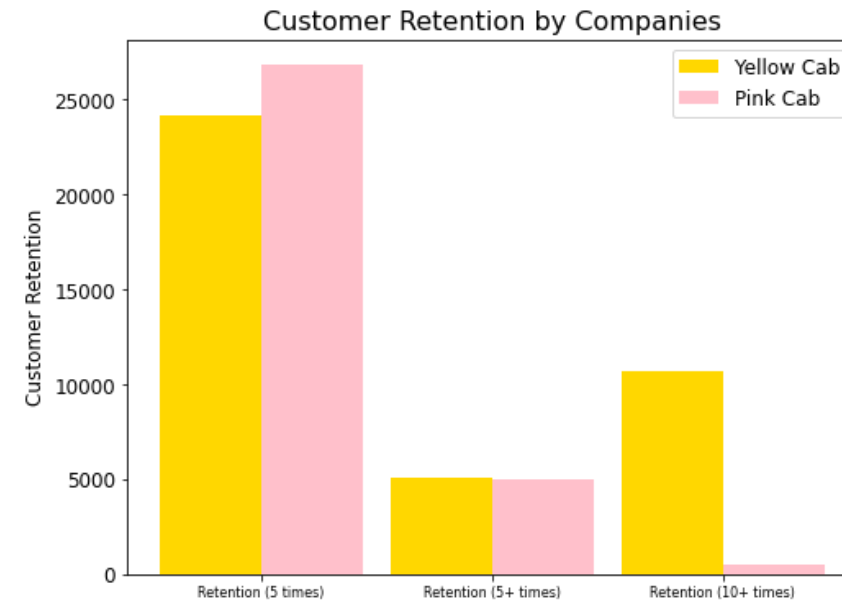
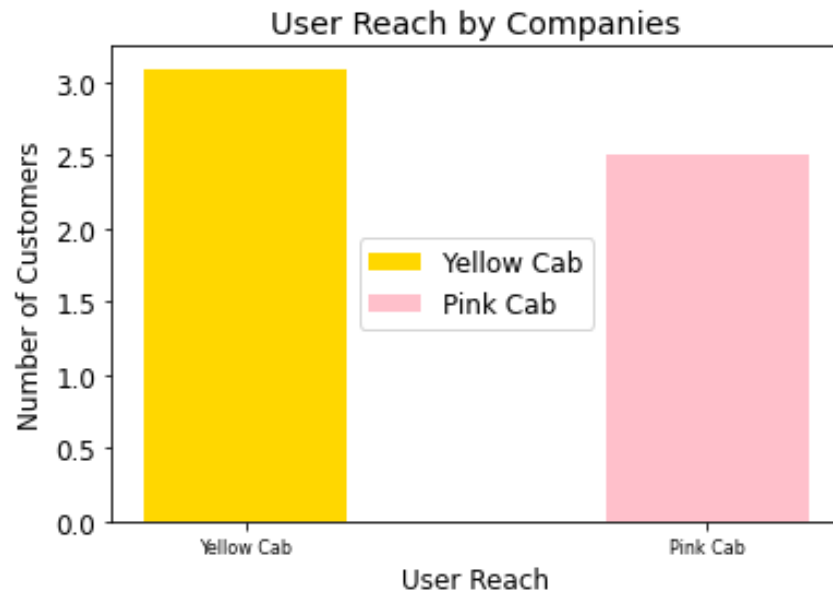
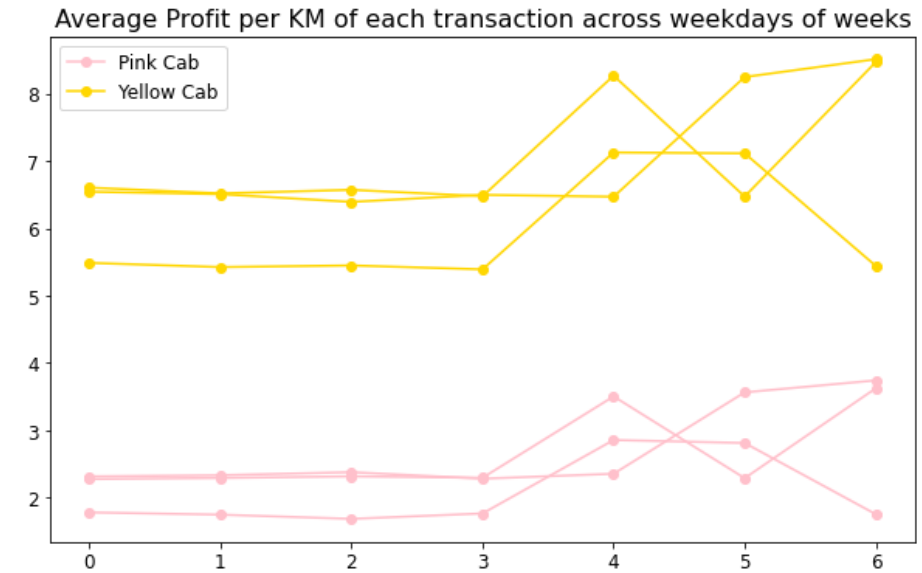
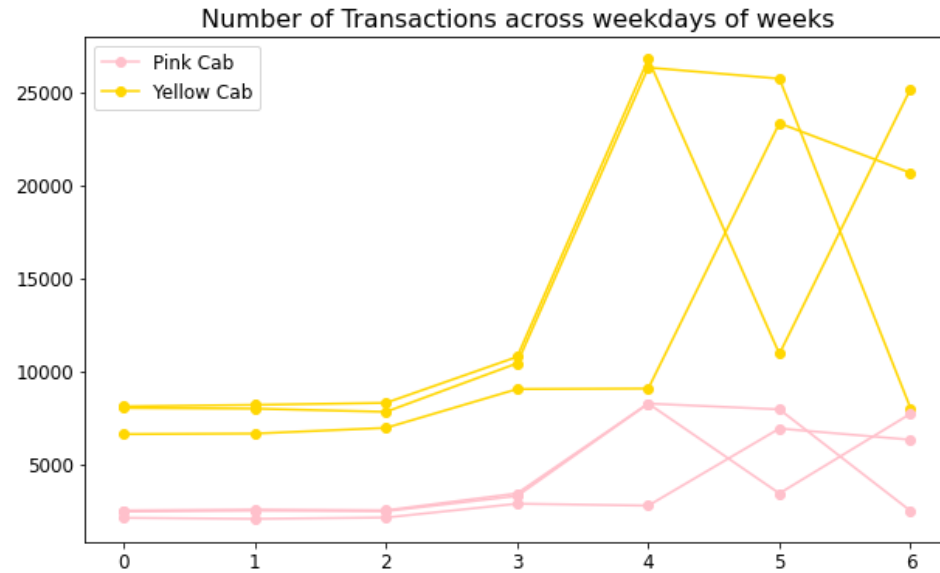


# EDA - Correlation Analysis





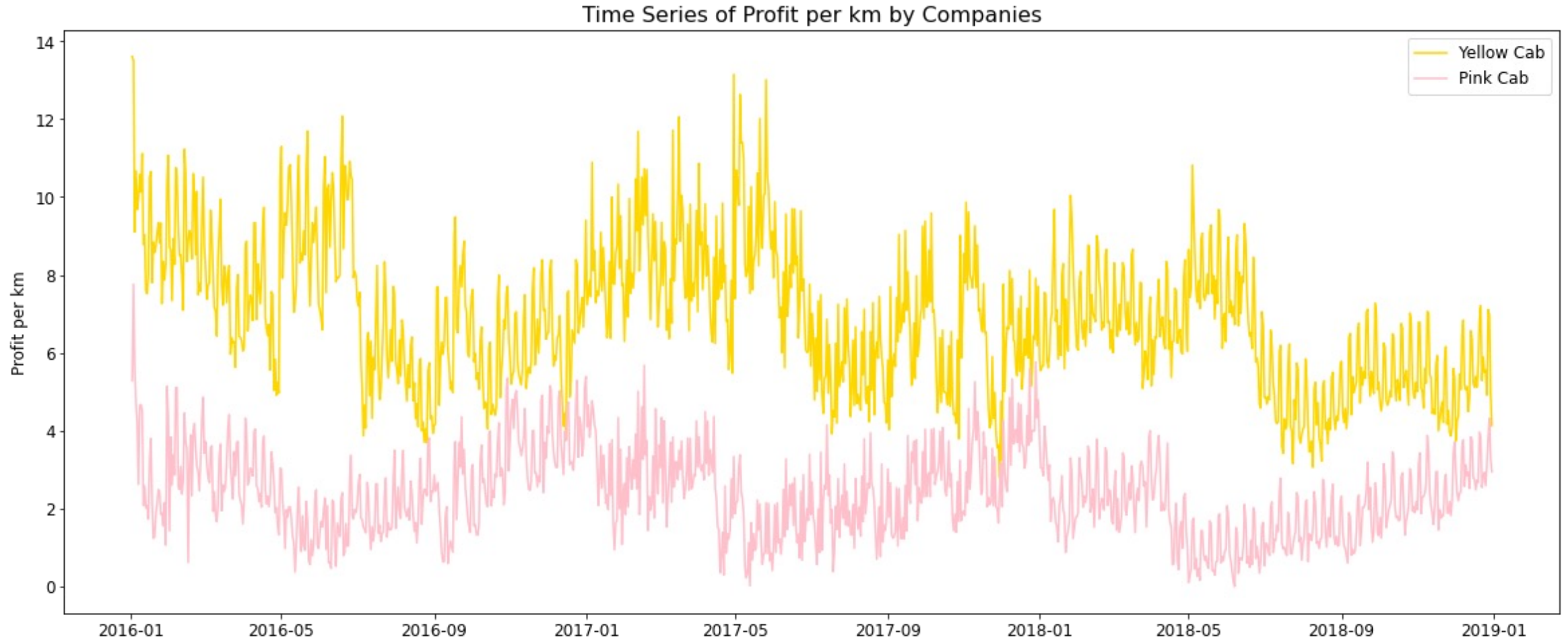
# EDA - Contextual Analysis



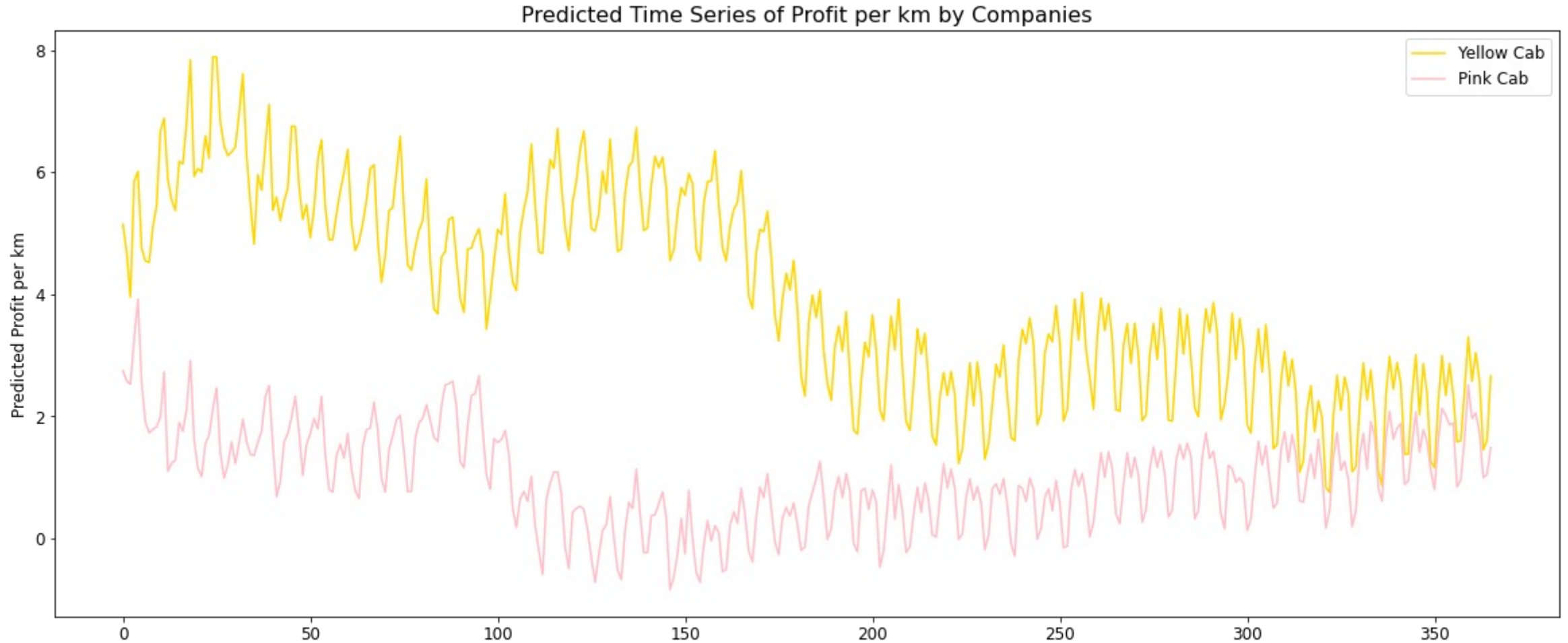
# EDA Summary

- Yellow Cab had both more transactions and more customers than Pink Cab.
- Yellow Cab made more businesses in bigger cities with higher frequency of taking cabs than Pink Cab.
- Yellow Cab had more customers with higher frequency of taking cab service.
- Yellow Cab made higher profit and profit per KM in total.
- Company & City are highly related categories; KM Travelled, Cost of Trip, Price Charged, Profit & Profit per KM are highly related; Company, City & Gender are related to Profit, Profit per KM & Frequency (based on both transactions and customer).
- Both number of transactions and customers increase from Friday to Sunday (Weekends), and average profit per km of each transaction also increase at Weekends.
- As for the user reach percentage, Yellow Cab has higher reach percentage than Pink Cab totally; As for the customer retention percentage, Yellow Cab has better customer retention situations than Pink Cab totally.

# Model – Time Series Forecasting



# Model – Time Series Forecasting



# Recommendations

In conclusion, we recommend 'Yellow Cab' for investment.

# Thank You