Exploratory Data Analysis

G2M insight for Cab Investment firm

20 July 2022

Agenda

Executive Summary

Problem Statement

Approach

EDA

EDA Summary

Recommendations

Data Glacier
Your Deep Learning Partner

Executive Summary

Target Audience

Investment Stakeholders, Company, Government

Problem Statement

XYZ is a private firm in US, planning to invest in Cab industry and decide their Go-to-Market(G2M) strategy based on the growth in the Cab Industry in last few years.

Project Delivery

Provide investment recommendations to XYZ private firm based on analysis of Cab Industry growth

Problem Statement

XYZ is a private firm in US who is planning to invest in Cab industry and decide their Go-to-Market(G2M) strategy based on the growth in the Cab Industry in 2016 to 2018.

Project Objectives

Provide investment recommendations to XYZ private firm based on analysis of Cab Industry growth based on:
- Customer Reach
- Customer Retention
- Age wise Reach
- Gender wise Reach
- Average profit per KM
- Average profit per day
- Overall profit from 2016 to 2018

The analysis has been divided into **four** parts:

- Data Preprocessing and Understanding
- Exploratory Data Analysis (EDA) on factors of profit for each cab type
- Finding the most profitable Cab company
- Recommendations for investment

Approach

Data Preprocessing and Understanding

12 Features (after merging and dropping meaningless columns)
- Timeframe of the data: 2016-01-31 to 2018-12-31
- Total data points :359392

Assumptions
- No null or NA values present
- No outliers present
- No duplicated rows

Transaction_ID.csv   City.csv

Cab_Data.csv   Customer_ID.csv

Final cab data
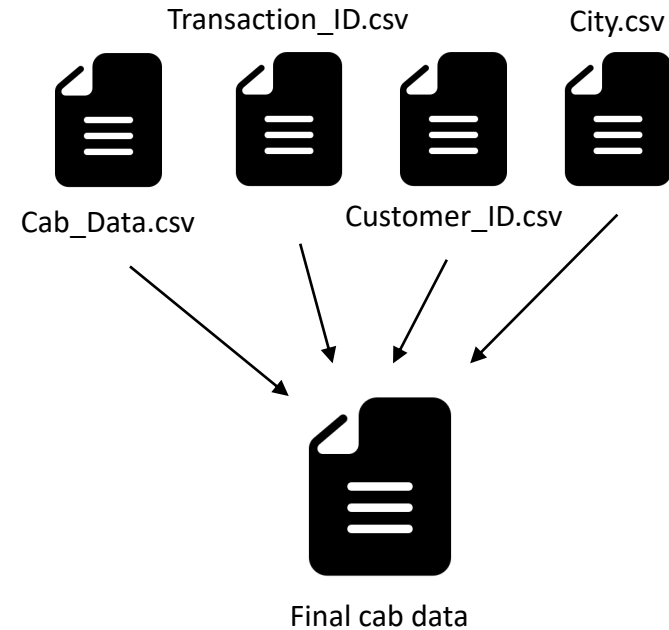
Approach

Data Preprocessing and Understanding
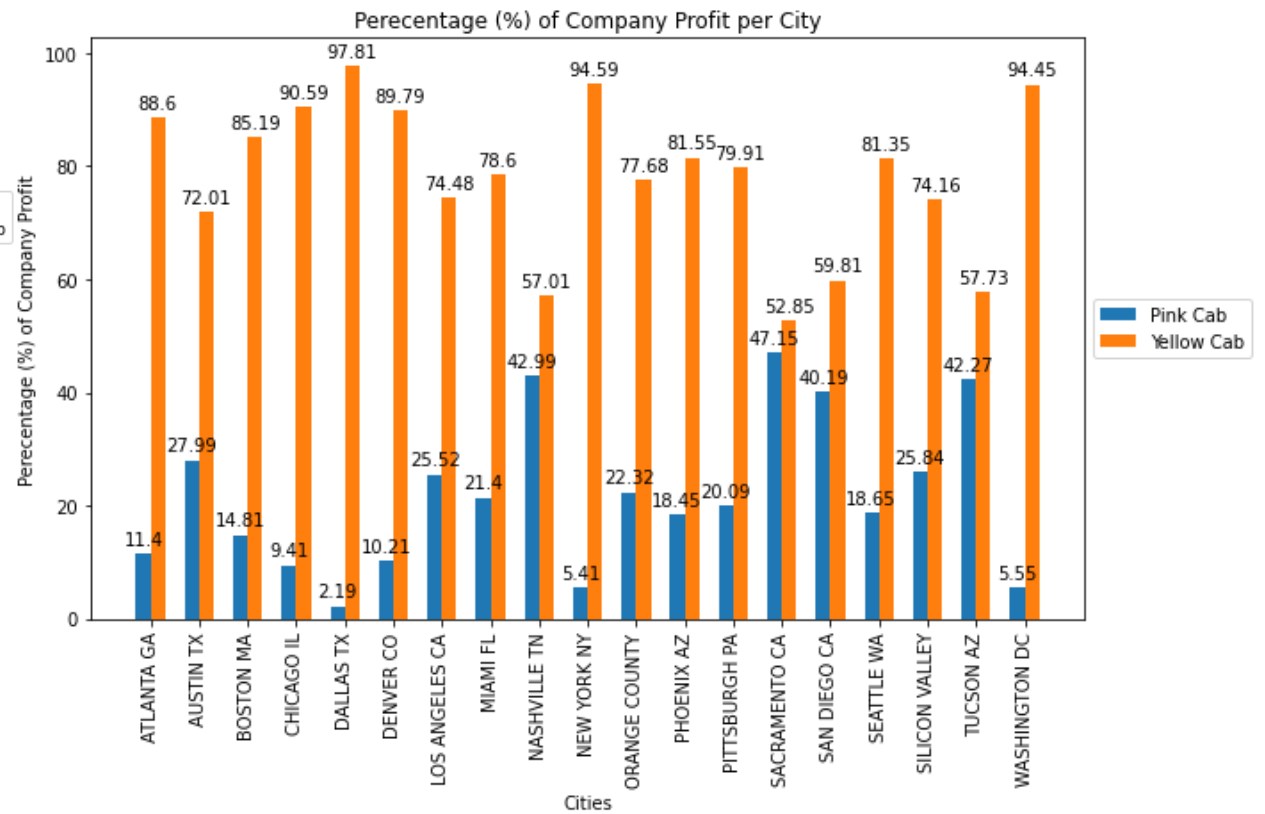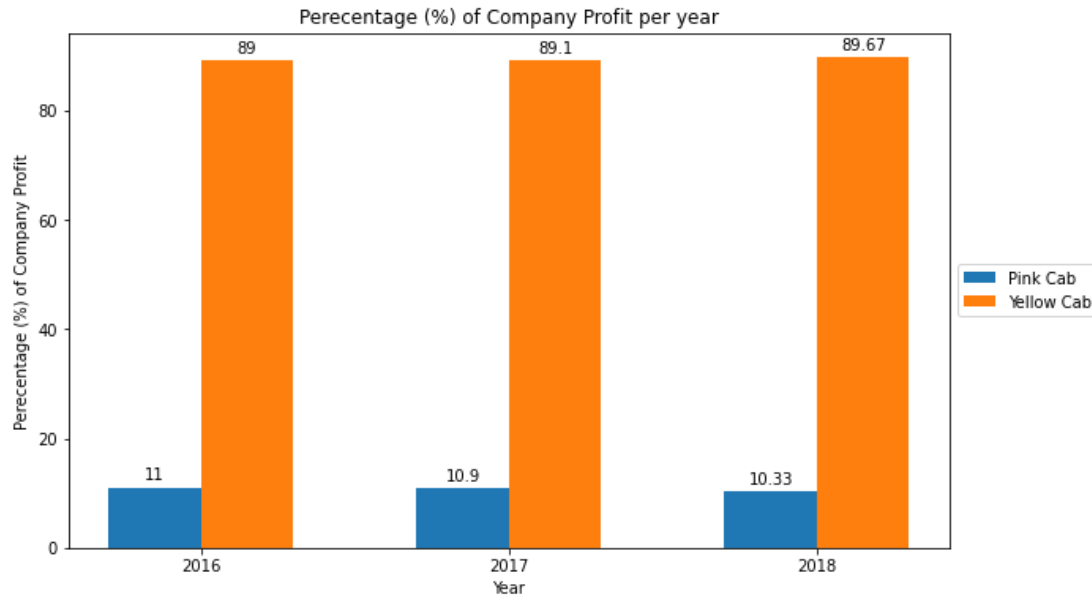
Observations:

- Customer Preference on Cab Company reflects directly on the profit of each company.

- Customer Gender and Residential Cities influences company profits.

Customer Background
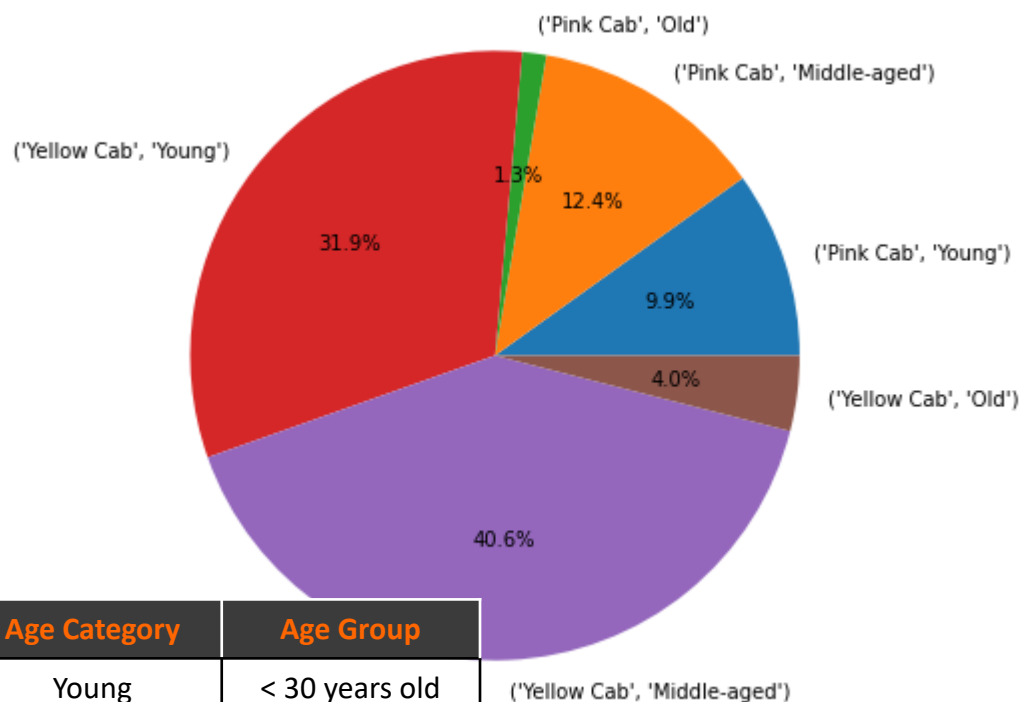

Company Users' Age Group

('Pink Cab', 'Old')

('Pink Cab', 'Middle-aged')

('Yellow Cab', 'Young')

1.3%

12.4%

31.9%

('Pink Cab', 'Young')

9.9%

('Yellow Cab', 'Old')

4.0%

40.6%

('Yellow Cab', 'Middle-aged')

Customer Users' Gender

('Pink Cab', 'Male')

('Yellow Cab', 'Female')

13.1%

32.3%

('Pink Cab', 'Female')

10.4%

44.2%

('Yellow Cab', 'Male')

Age Category	Age Group
Young	< 30 years old
Middle-aged	30 – 60 years old
Old	> 60 years old
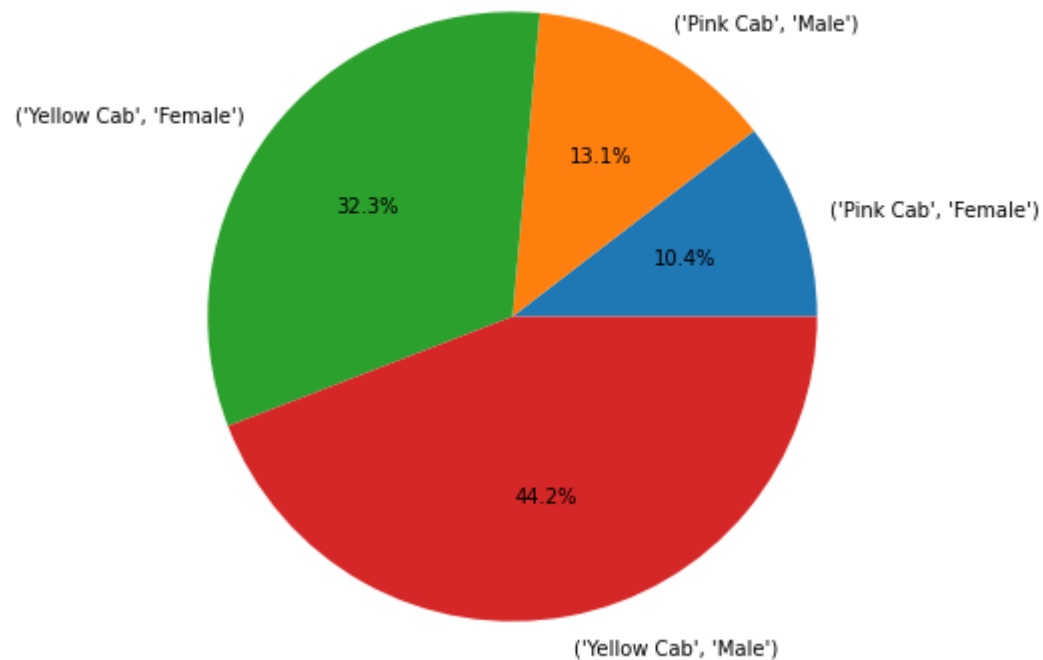
Profit based on Customer Background
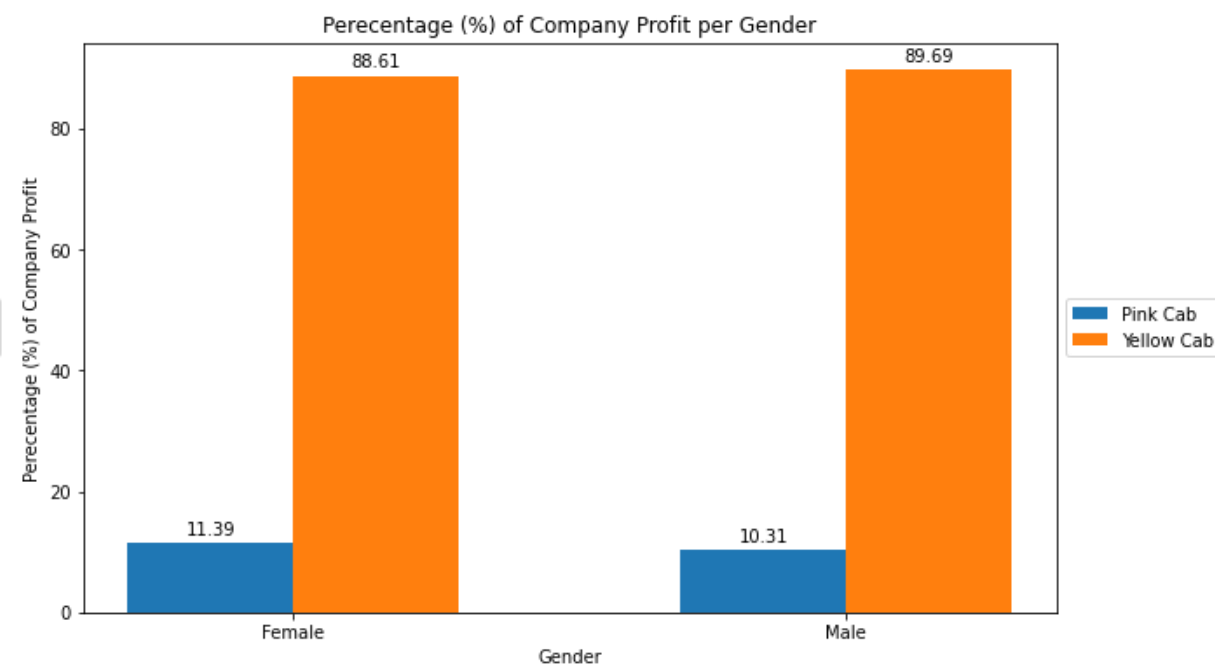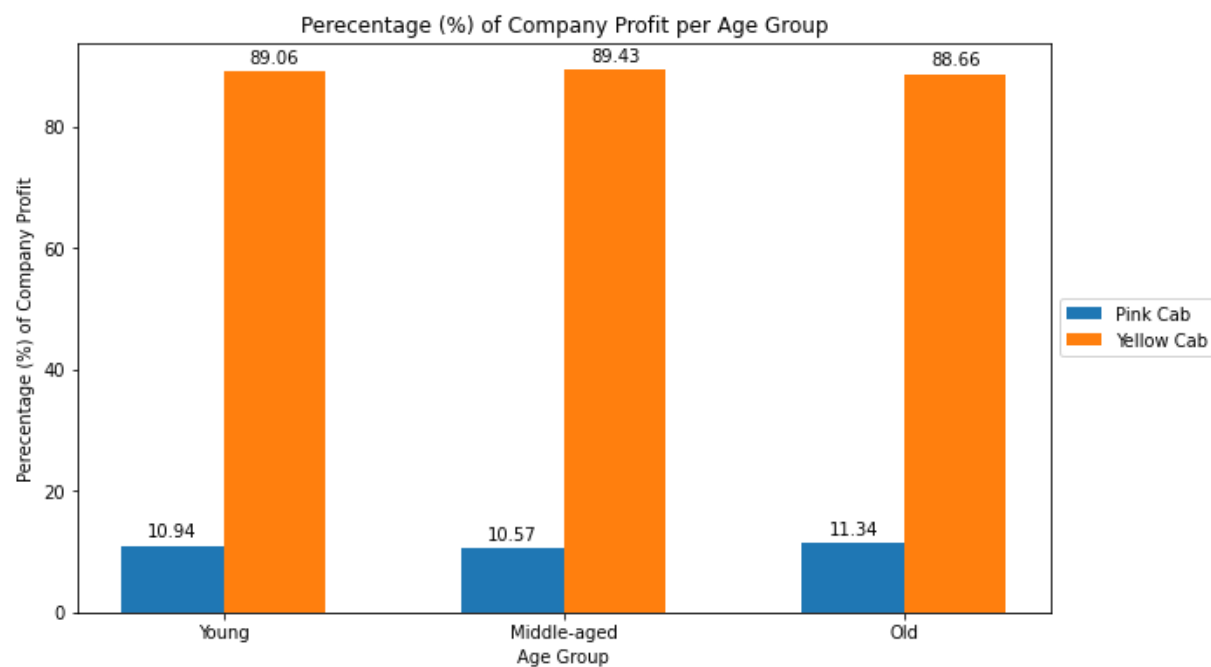
EDA Summary

1. Profit Coverage per year:
   - Yellow Cab covers **almost 90%** of the profit from 2016 to 2018.
2. Profit Coverage based on city:
   - **Yellow cab** covers **more than 50%** of profit over **all 19 cities**
   - Pink Cab takes 40% to 45% of the profit **only across 4 cities**
3. Average profit per day:
   - Average profit per day of **Yellow Cab** is **almost 10 times** the average profit per day of the Pink cab.
4. Average profit per KM:
   - Average profit per KM of **Yellow Cab** is **more than 3 times** the average profit per KM of the Pink cab.
5. Customer preference over companies:
   - **76.5% of users prefer Yellow Cab** over Pink Cab
   - **Users** who choose **Yellow Cab gradually increase around 1%** over the years from 2016 to 2018.
   - **Middle-aged users (53%)** travel in Cab **the most**, followed by Young (41.8%) and Old (5.3%) users.
     - Yellow Cab: 40.6% middle-aged, 31.9% young, 4.0% old users
     - Pink Cab: 12.4% middle-aged, 9.9% young, 1.3% old users
   - **57.3% of users** who prefer to travel in Cab are **Male**.
     - Yellow Cab: 44.2% Male, 32.3% Female
     - Pink Cab: 13.1% Male, 10.4% Female

EDA Summary

6. Profit Coverage based on Age Group:
   - Young: 89.06% from Yellow Cab, 10.94% from Pink Cab
   - Middle-aged: 89.43% from Yellow Cab, 10.57% from Pink Cab
   - Old: 88.66% from Yellow Cab, 11.34% from Pink Cab
7. Profit Coverage based on Gender:
   - Male: 88.61% from Yellow Cab, 11.39% from Pink Cab
   - Female: 88.69% from Yellow Cab, 10.31% from Pink Cab

Recommendations

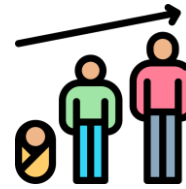Highly recommended to invest in **Yellow Cab**.

3x average profit per KM higher than Pink Cab

10x average profit per day higher than Pink Cab

Customer Reach

Covers **more than 50%** of **profit** over **all 19 cities**

Customer Retention

Users gradually increase about 1% over the years from 2016 to 2018

Age Wise Reach

- **Middle-aged users (53%)** travel in Cab **the most**, followed by **Young users (41.8%)**
- Yellow cab occupies **75.75%** of **Middle-aged** users and **76.32%** of **Young** users

Gender Wise Reach

- **57.3% of users** who prefer to travel in Cab are **Male**
- Yellow cab occupies **77.13%** of **Male** users

Thank You