Exploratory Data Analysis

<G2M Case Study>

<2024-7-17>

Agenda

Background

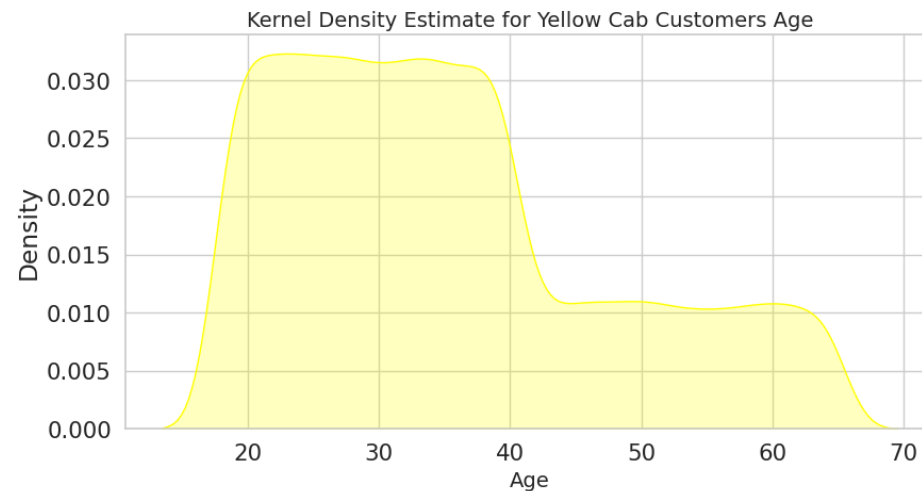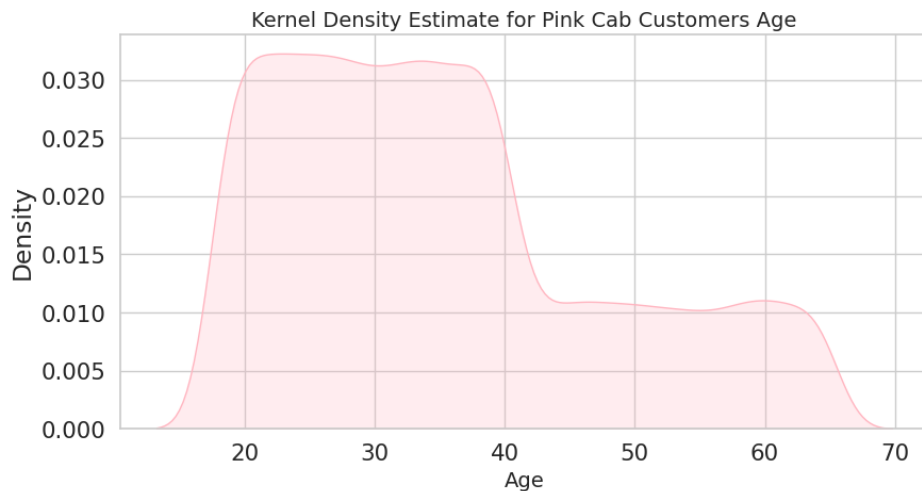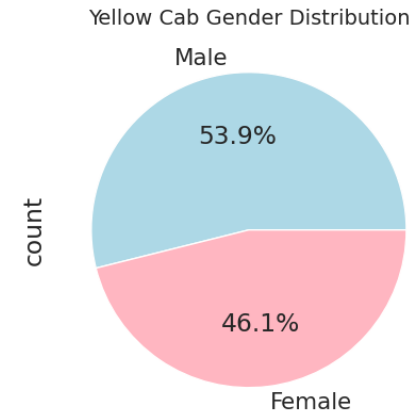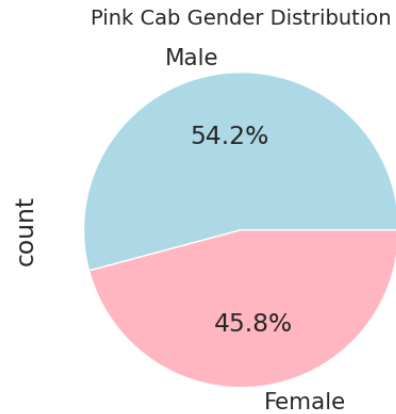Problem Statement

EDA

Recommendations

Background Information: G2M Case Study

- 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 and as per their Go-to-Market(G2M) strategy they want to understand the market before taking final decision.

- Objective: identify which company is performing better and is a better investment opportunity for XYZ.

- Data Features:
  - Total 4 individual data sets merged into 1 final_cab_data data set.
    - 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
  - 14 features, and a total of 35, 9391 data points
  - Timeframe: 31/01/2016 - 31/12/2018

  - The profit generated from each ride is calculated while holding other variables constant. Only the Price_Charged and Cost_of_Trip features are utilized to determine the profit.

Problem Statement

- Some questions that we are looking into during analysis:
    - Is there any difference in customer demographics between the 2 companies?
    - Is there any difference in company's operations (eg: which company travels more, which company has higher price or cost etc.)?
    - Which company has higher customer reach?
    - Which company has higher city coverage?
    - Which company has better customer retention?
    - Which company makes more profit?
    - How does the customer attributes contributes to the company's profit?
    - Which company has better future prospects?

Is there any difference in customer demographics between the 2 companies?

Pink Cab Gender Distribution
Male 54.2%
Female 45.8%

Yellow Cab Gender Distribution
Male 53.9%
Female 46.1%

Kernel Density Estimate for Pink Cab Customers Age

Kernel Density Estimate for Yellow Cab Customers Age

- Gender distribution for both companies are similar
  - About 54% male customers and 46% female customers

- Age distribution for both companies are similar:
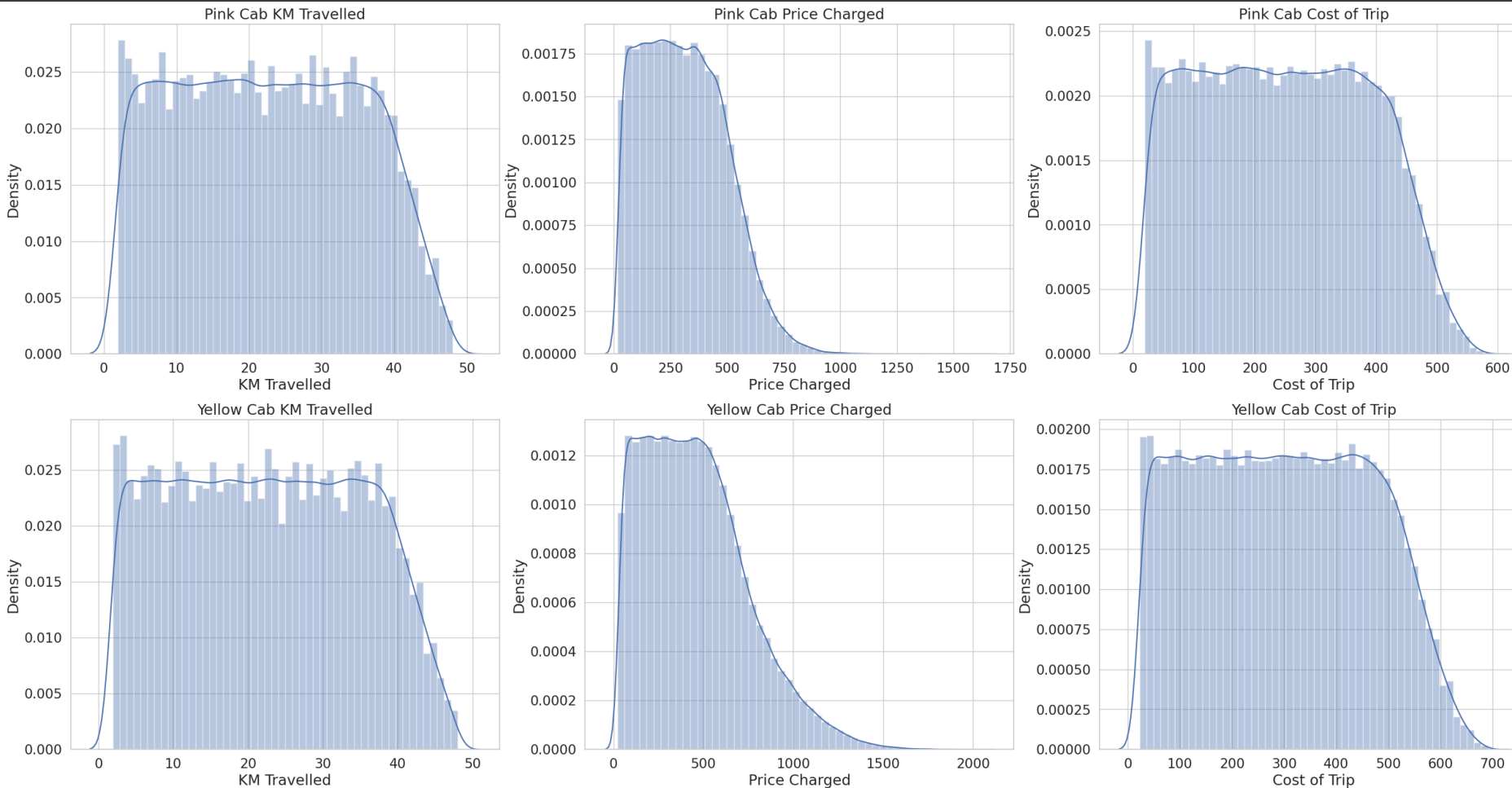 - Mainly concentrated in the 20 – 40 age group

Data Glacier

Is there any difference in customer demographics between the 2 companies?
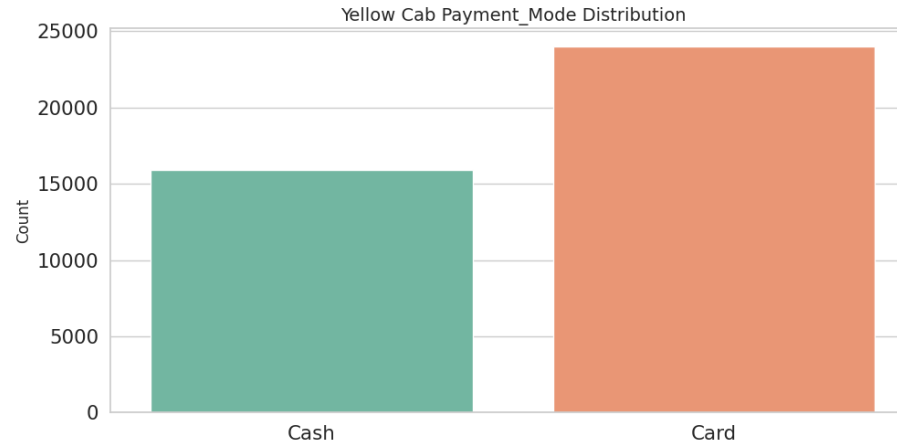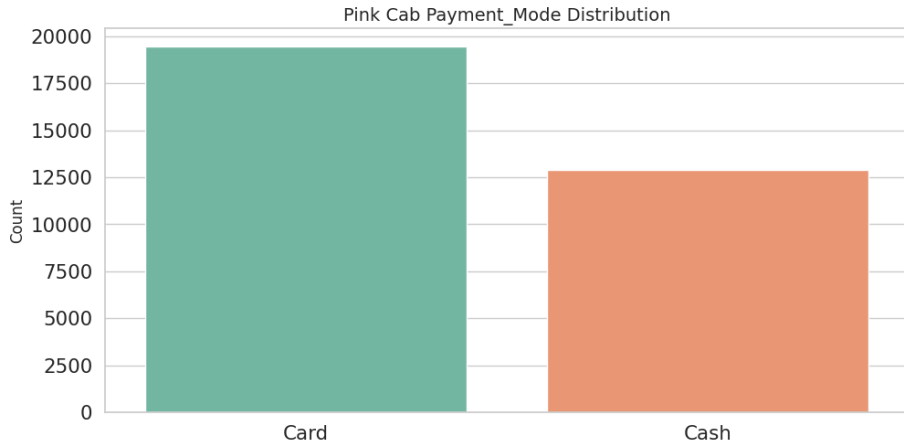
- Payment mode are similar:
  - Both are mainly card payment
- Customer's income level distribution are quite average:
  - Mainly concentrated below 25,000 USD/month

Data Glacier
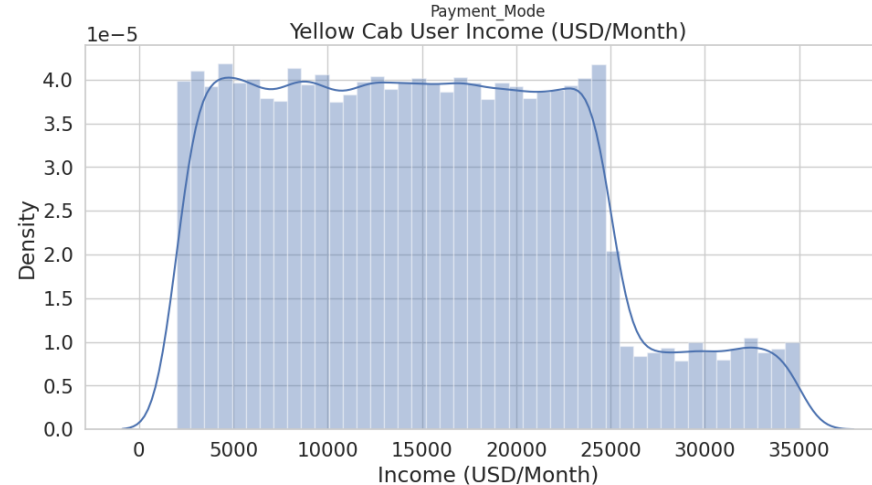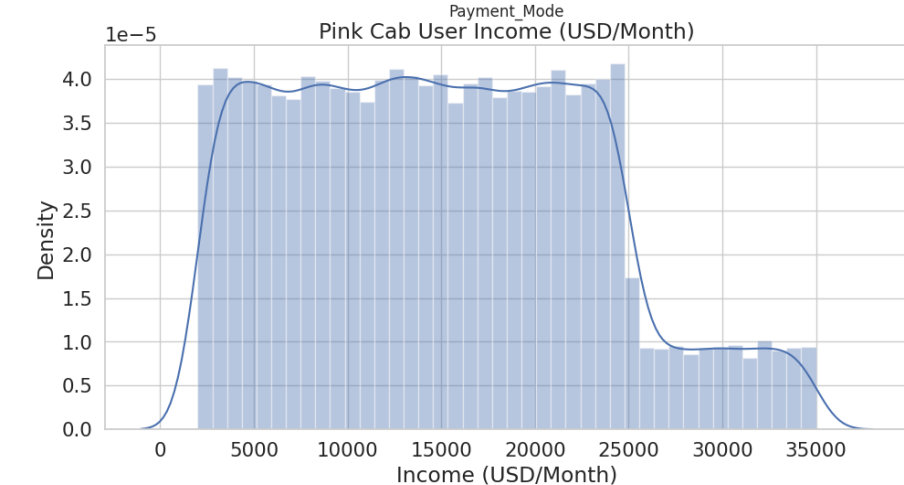
# Is there any difference in company's operations?

- Yellow cab company seems to have higher price charged compared to the pink cab company.

Which company has higher customer reach?

- Yellow cab company obviously has higher customer reach compared to pink cab company, reach to almost twice the amount of unique customers of the pink cab, covers 2.4% of the entire cab customer market while pink only covers 1.1%.

- Both companies faced a drop from 2017 to 2018, but pink cab decreased 1.61%, more than yellow cab's 1.36%.
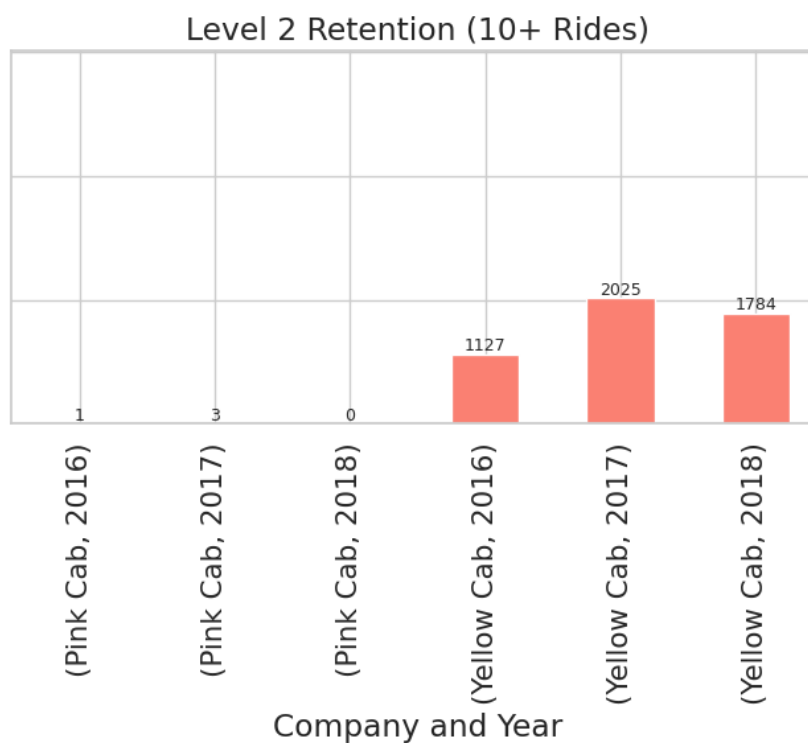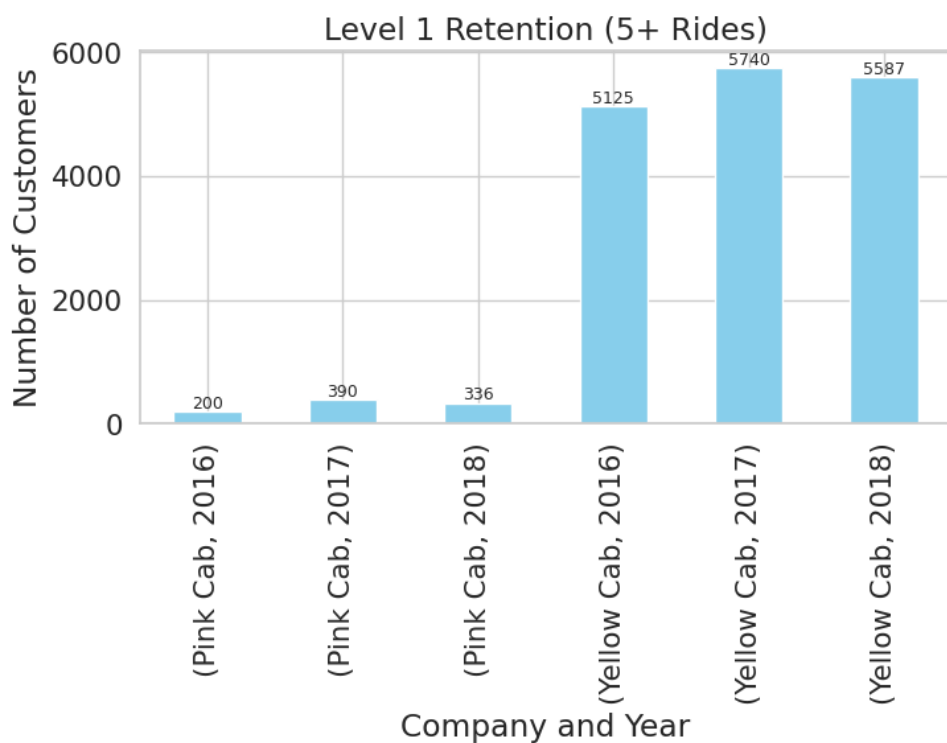
Which company has higher city coverage?

Total Number of Users by City

Percentage of Users by City
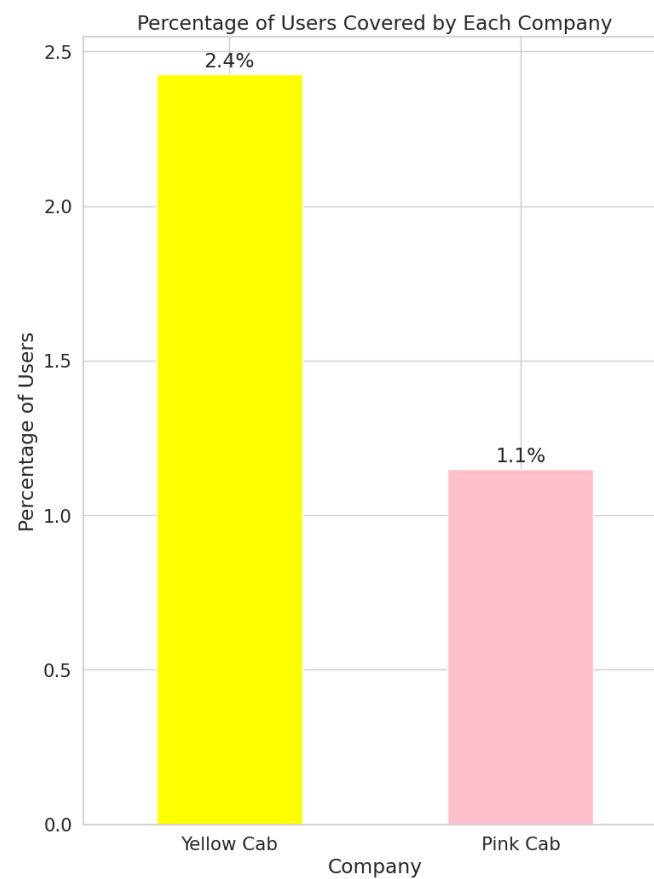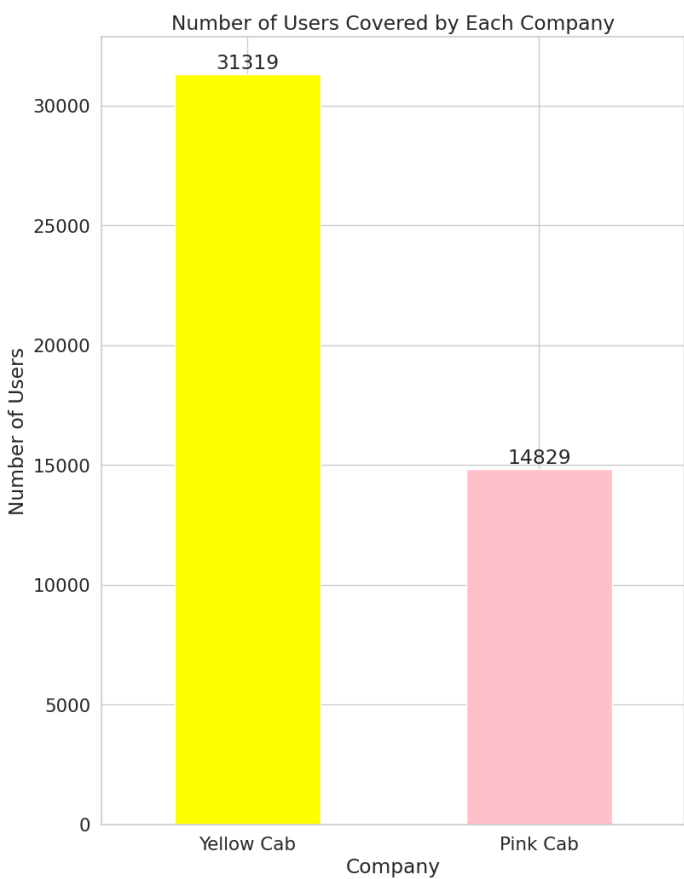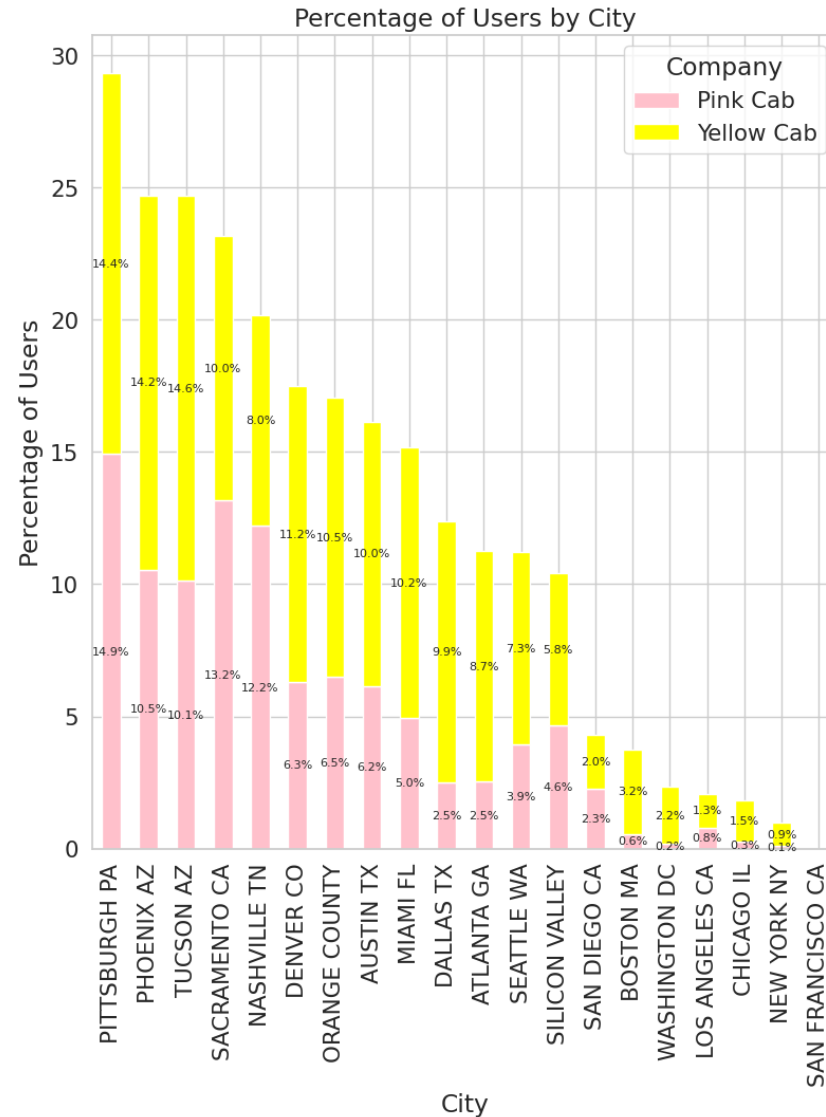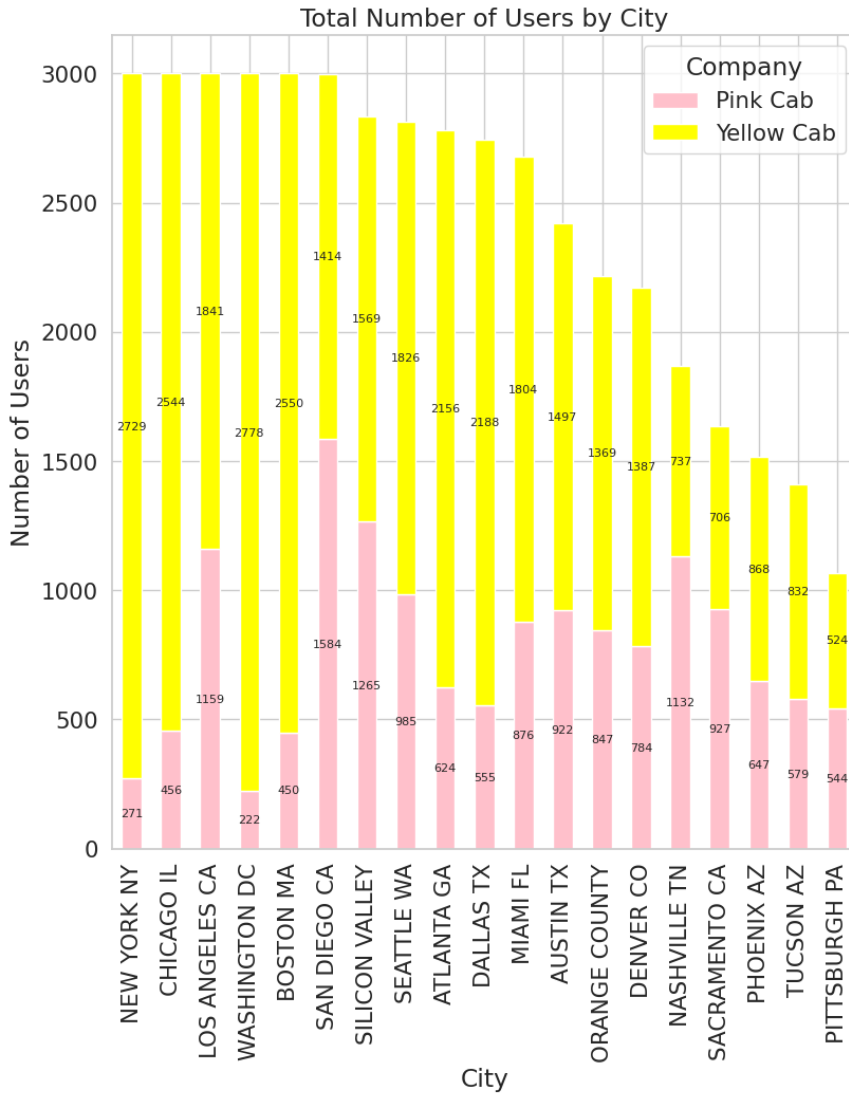
- Yellow cab has significant higher coverage compared to pink cab (leading in 16 cities in a total of 19 cities).

Which company has better customer retention?

- We analyze customer retention by segmenting customers into 2 segments: 1 is customers who took at least 5 drive with the same cab company, and 1 is at least 10 drive with the same company, indicating more loyalty.

- It is evident that Yellow Cab outperforms Pink Cab in both segments.
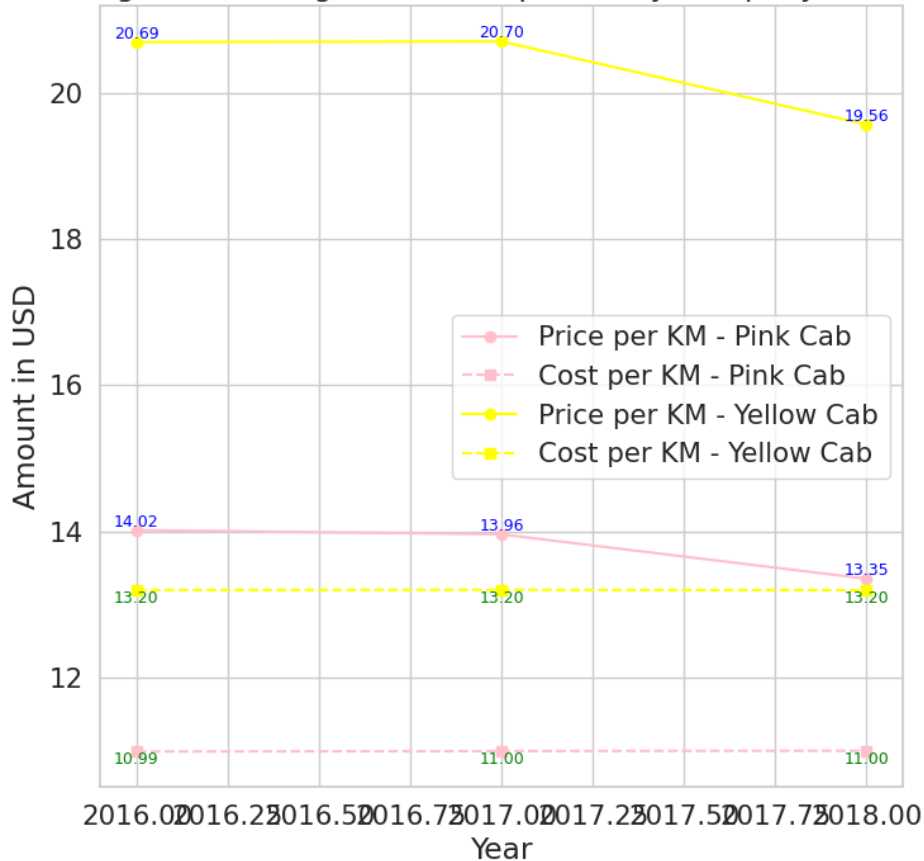
Data Glacier

Which company makes more profit?
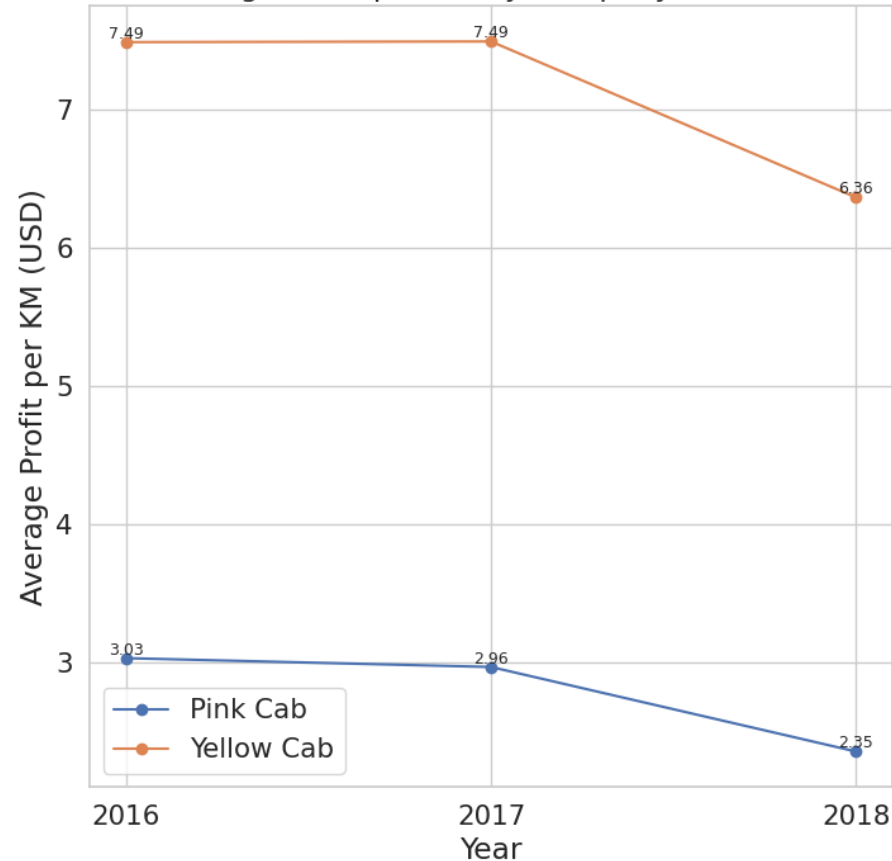

Annual Profit Made by Each Company
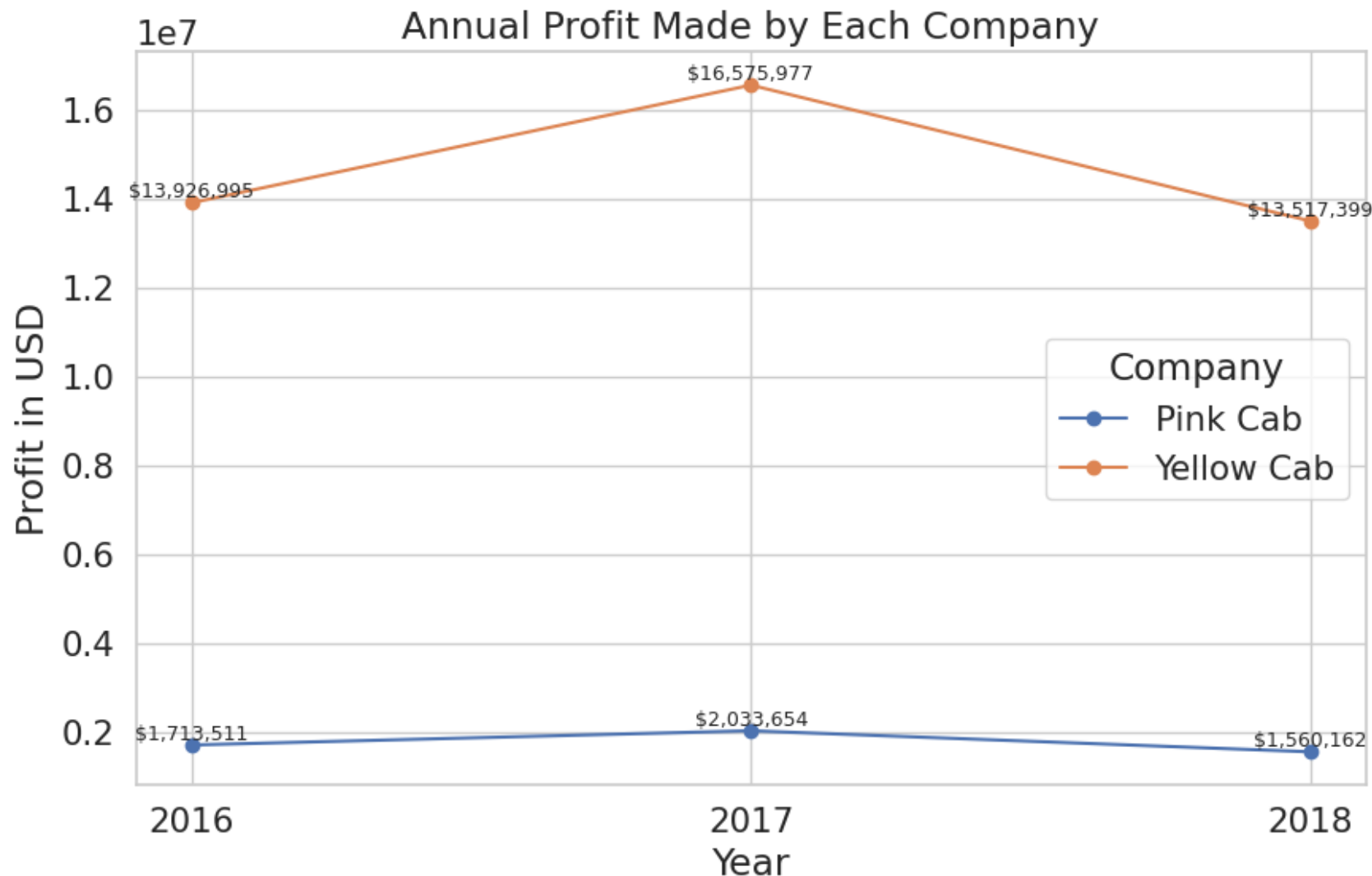
- yellow cab company is significantly making more profit, around 11-14 million more than the pink cab company.

- In 2017-2018, pink cab dropped 23% while yellow cab dropped 18%.

Data Glacier

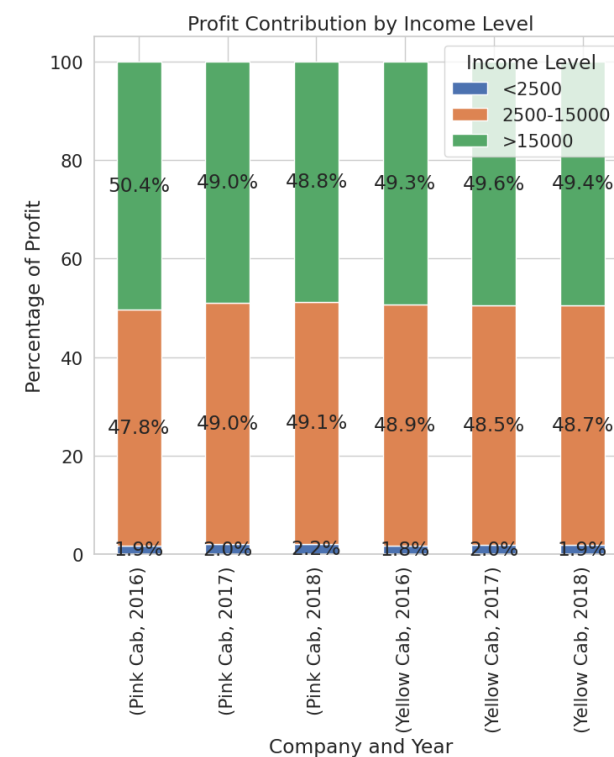# Which company makes more profit?

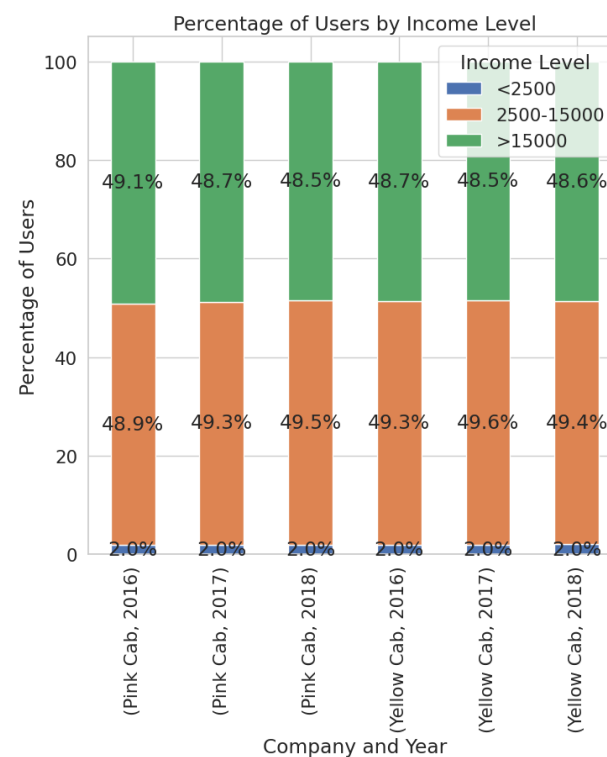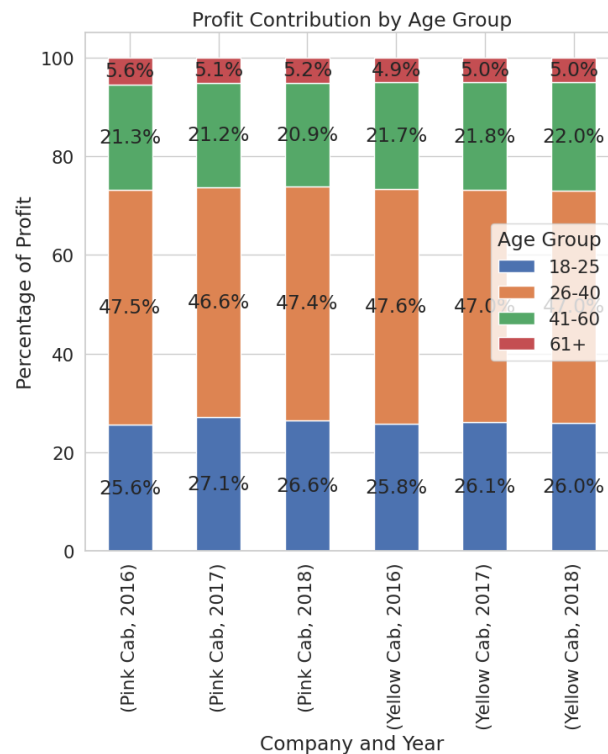Average Price Charged and Cost per KM by Company Each Year

Average Profit per KM by Company Each Year

- Yellow cab has an average of about 4 dollars more profit per KM than the pink cab, stronger profitability.

How does the customer attributes contributes to the company's profit?
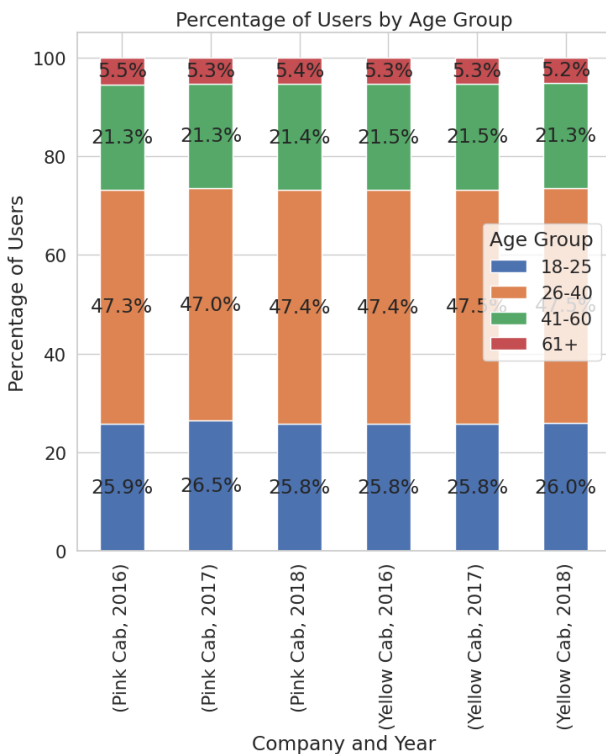
How does the customer attributes contributes to the company's profit?
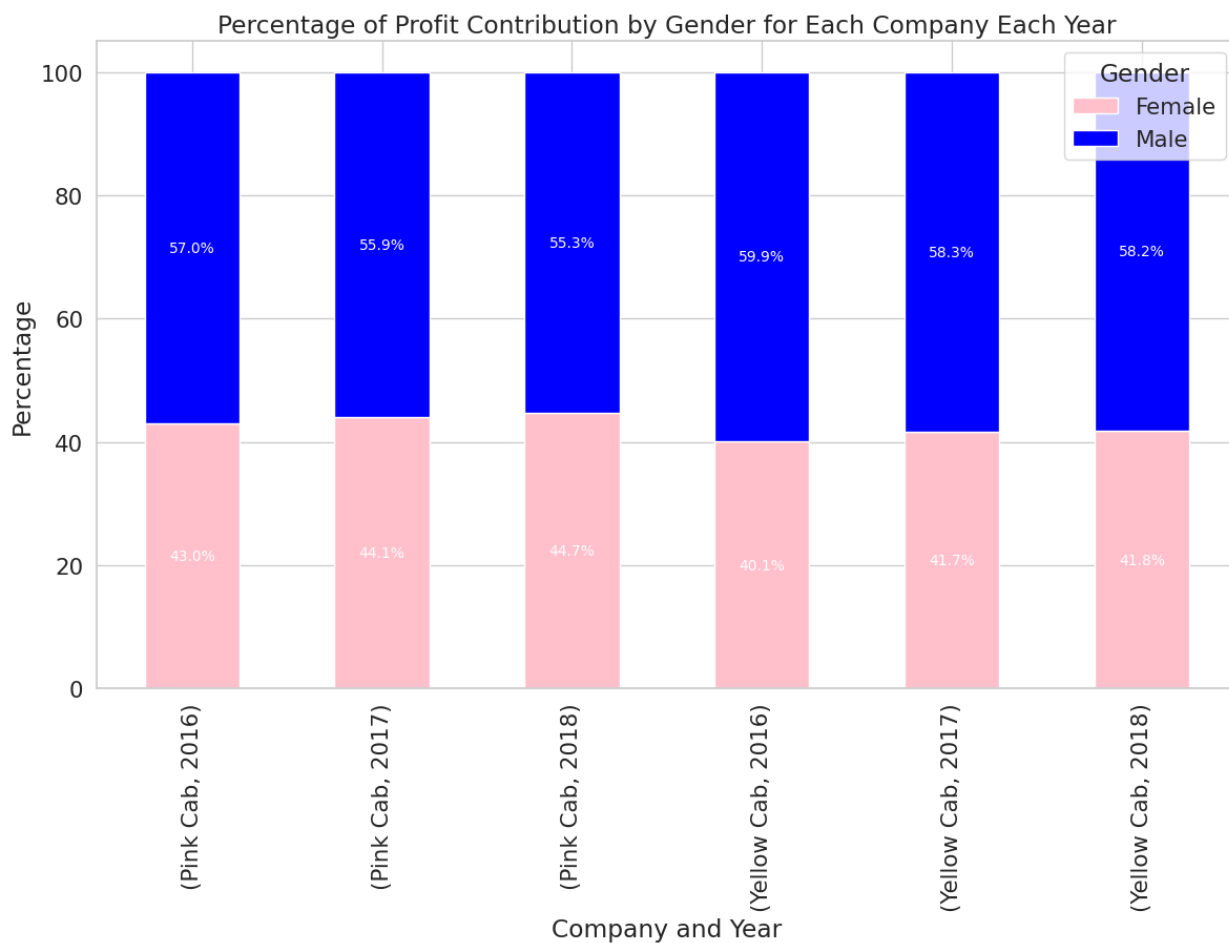
Percentage of Profit Contribution by Gender for Each Company Each Year

- As predicted, gender, income level, and age group contribution to profit have no significant difference between the 2 companies all genders between the ages of 18 to 40 that are at the middle and high-income level contribute most to companies' profit.

How does the customer attributes contributes to the company's profit?

Profit Contribution per City for Pink Cab
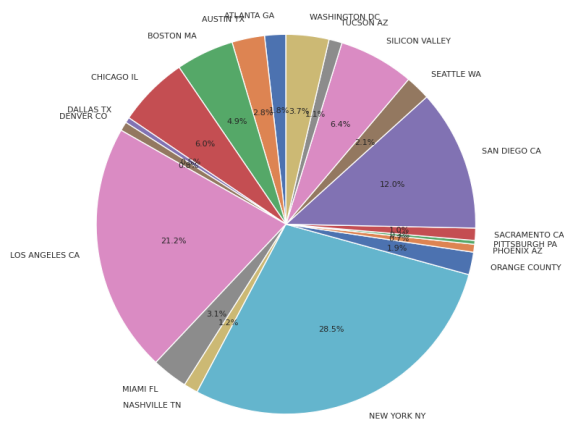
Profit Contribution per City for Yellow Cab
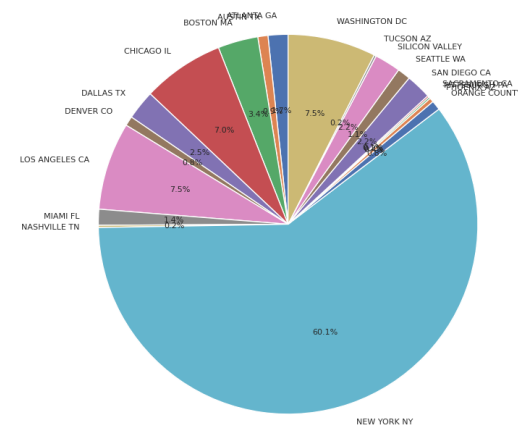
- New York cab users contributed 60.1% of the profit for the yellow cab company while only 28.5% for the pink cab company. Yellow cab has a stronger presence in more developed cities with higher demands for cabs, and have higher profitable rides percentage in basically all cities compared to pink cabs.

Profitable Rides Percentage per City for Pink Cab
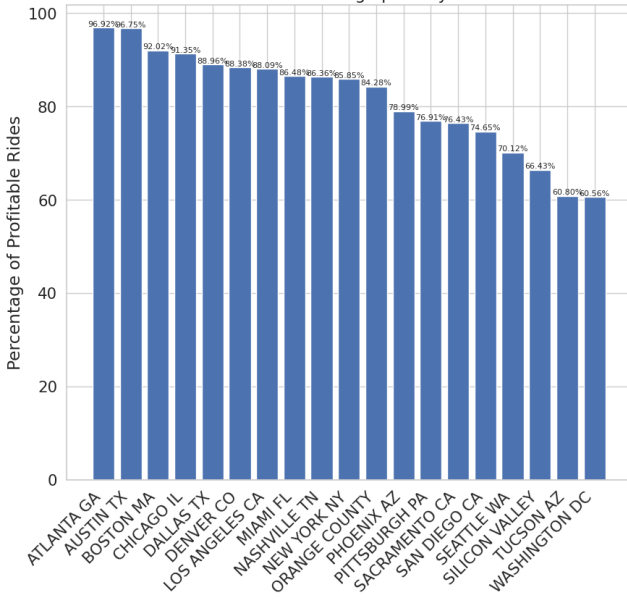
Profitable Rides Percentage per City for Yellow Cab

Which company has better future prospects?
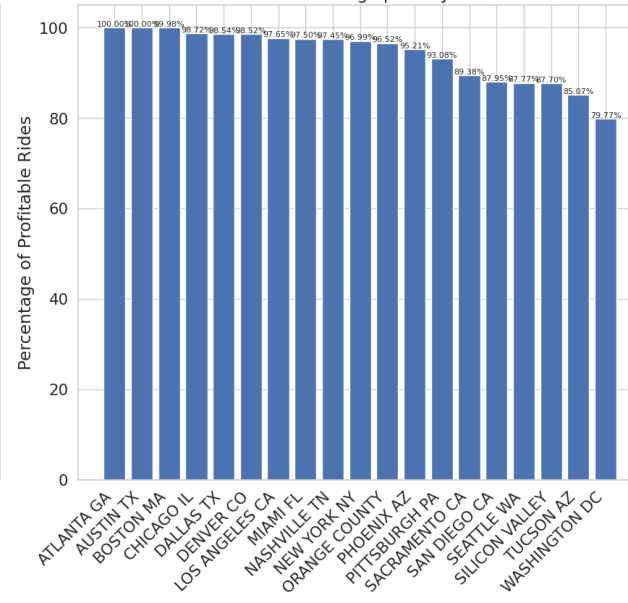
Profit Forecasting for Future Years

- Pink cab has lesser increase in 2019 than yellow cab and higher decrease from 2019 to 2021 than yellow cab.

Recommendations

Demographics:

- User demographics for the yellow and pink companies are very similar. Both companies' users have almost identical gender distribution, age distribution, and income level distribution. The main age group for both cab companies is 18-40, and both are very popular in high and medium-income classes, indicating strong future development potential and steady revenue.
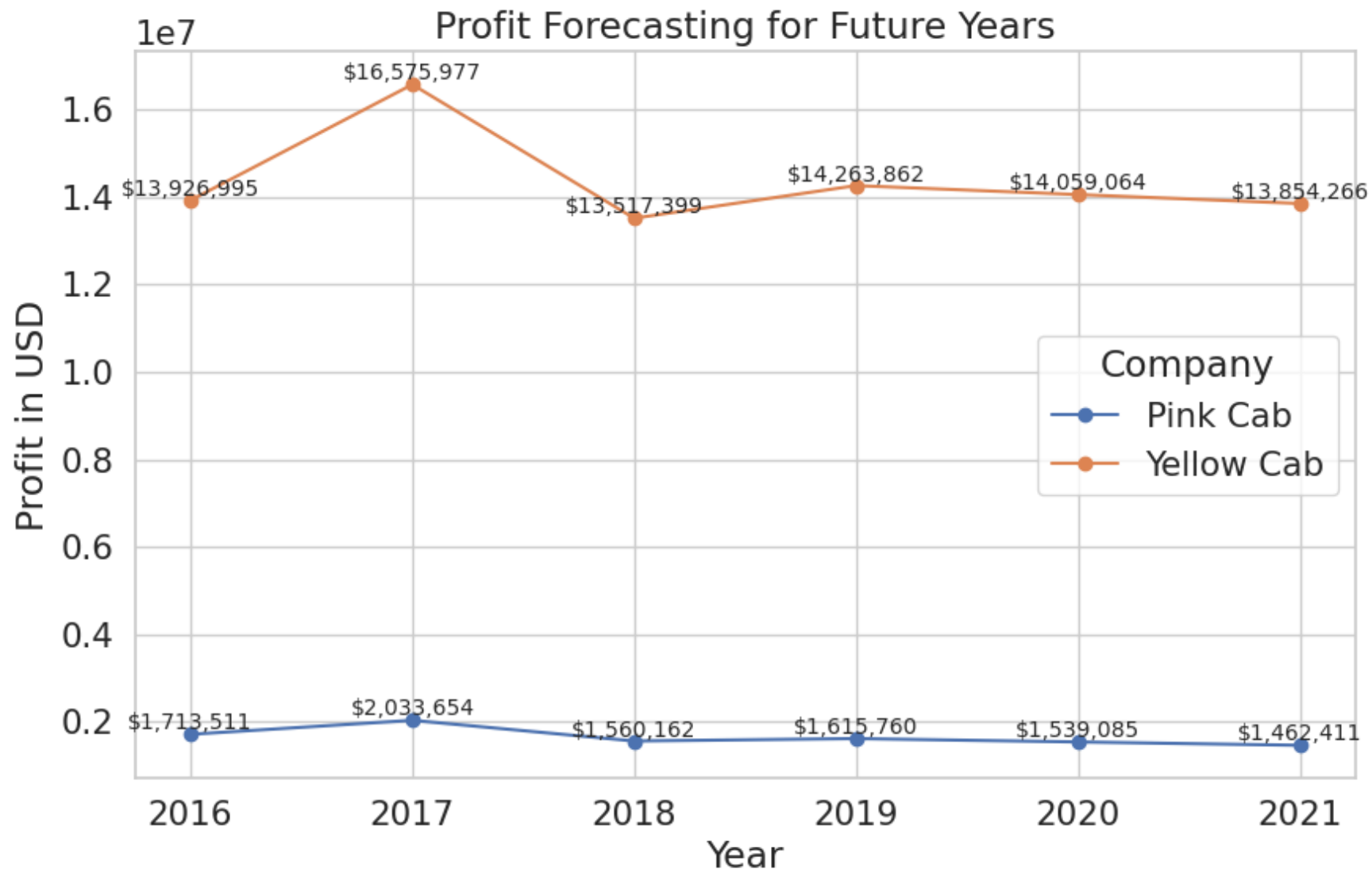
User Reach:

- Yellow cab company has 31,319 unique users while Pink Cab company has 14,829, about twice the amount of users. Yellow cab covers 2.4% of total cab users compared to pink cab's 1.1%, a significantly larger market. Both companies' users amount have faced a drop in users from 2017 to 2018, yellow cab decreased by about 1.36% while pink cab decreased by 1.61%.

City Coverage:

- Out of 19 cities, pink cab only reached more users, and has a higher user percentage in 3 cities, while yellow cab has more users in 16 others.

User Retention:

- After analyzing this in two segments: at least 5 drive and at least 10 drive with the same cab company, it is evident that Yellow Cab outperforms Pink Cab in both segments. This indicates that Yellow Cab is more effective in retaining its customers than Pink Cab.

Recommendations

Profit:

- We can see that the yellow cab company is significantly making more profit, around 11-14 million more than the pink cab company. Both faces decreased from 2017 to 2018, pink cab dropped 23% while yellow cab dropped 18%, indicating better stability for yellow cab company.

- Even though the yellow cab has a higher cost per KM compared to the pink cab, with a higher price per KM yellow cab has an average of about 4 dollars more profit per KM than the pink cab, stronger profitability.

- As predicted, gender, income level, and age group contribution to profit have no significant difference between the 2 companies all genders between the ages of 18 to 40 that are at the middle and high-income level contribute most to companies' profit.

- New York cab users contributed 60.1% of the profit for the yellow cab company while only 28.5% for the pink cab company. We can see that yellow cab has a stronger presence in more developed cities with higher demands for cabs, and have higher profitable rides percentage in basically all cities compared to pink cabs, indicating stronger future potential.

- Yellow cab has a 5.5% predicted increase in 2019 and a 2.87% decrease in the future 3 years from 2019 to 2021, while pink cab has a 3.6% predicted increase in 2019 and a 9.49% decrease in the future 3 years from 2019 to 2021. There is a lesser increase and more decrease for the pink cab compared to the yellow cab.

Conclusion:

- Based on the above points, I recommend the **yellow cab company** for investment.

Thank You