Data Glacier
Your Deep Learning Partner

Exploratory Data Analysis
G2M Case Study

1/19/2023

Agenda

Data Glacier
Your Deep Learning Partner

Problem Statement

- 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 :** Provide actionable insights to help XYZ firm in identifying the right company for making investment.

Approach

The analysis has been divided into four parts:

- Data Understanding

- Forecasting profit and number of rides for each cab type

- Finding the most profitable Cab company

- Recommendations for investment

Data Understanding

- We have been provided 4 individual data sets for the analysis:
  - 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
- We have total of 22 Features (including 7 derived features)
- Timeframe of the data: 2016-01-31 to 2018-12-31
- Total data points :359392

Assumptions

- Outliers are present in Price_Charged feature but due to unavailability of trip duration details, we are not treating this as outlier.

- Profit of rides are calculated keeping other factors constant and only Price_Charged and Cost_of_Trip features used to calculate profit.

- Users feature of city dataset is treated as number of cab users in the city, we have assumed that this can be other cab users as well (including Yellow and Pink cab)
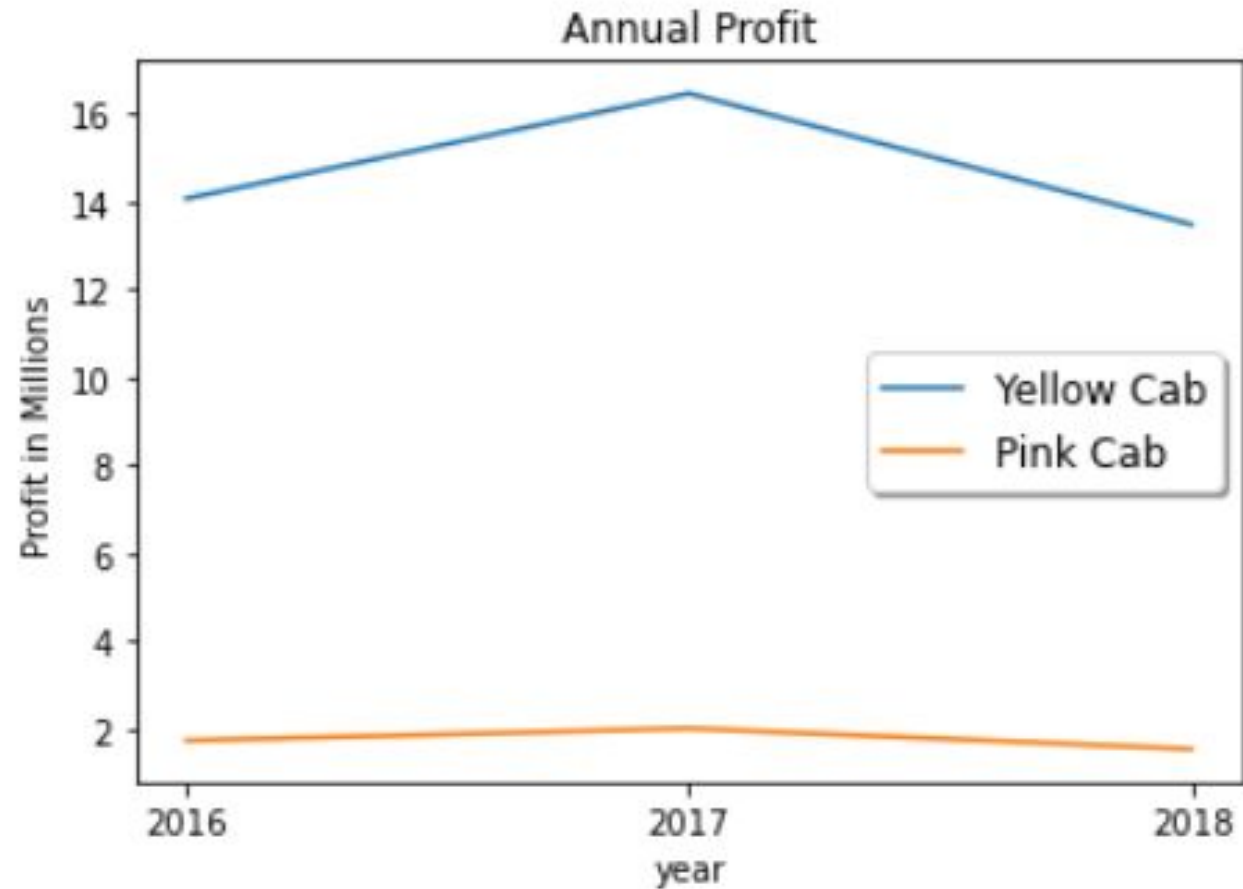
Profit Analysis

Annual profit rate per KM
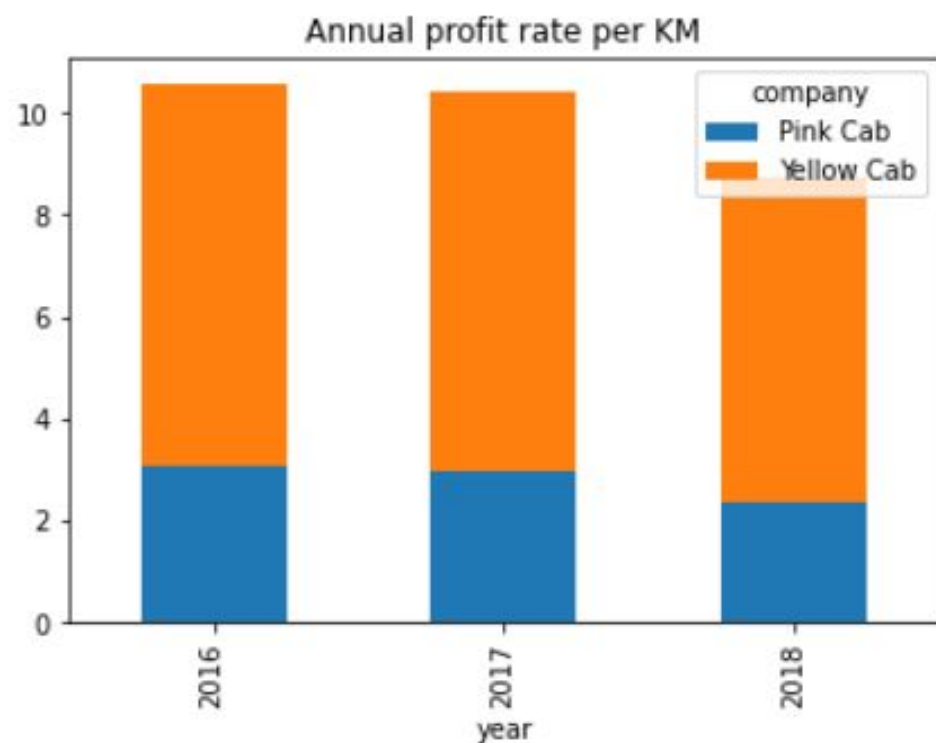
company year	Pink Cab	Yellow Cab
2016	3.062027	7.503564
2017	2.949786	7.465487
2018	2.353015	6.364322

Annual profit rate

Annual Profit Rate

company year	Pink Cab	Yellow Cab
2016	27.854227	56.841811
2017	26.812112	56.542007
2018	21.385401	48.228080

Profit Analysis
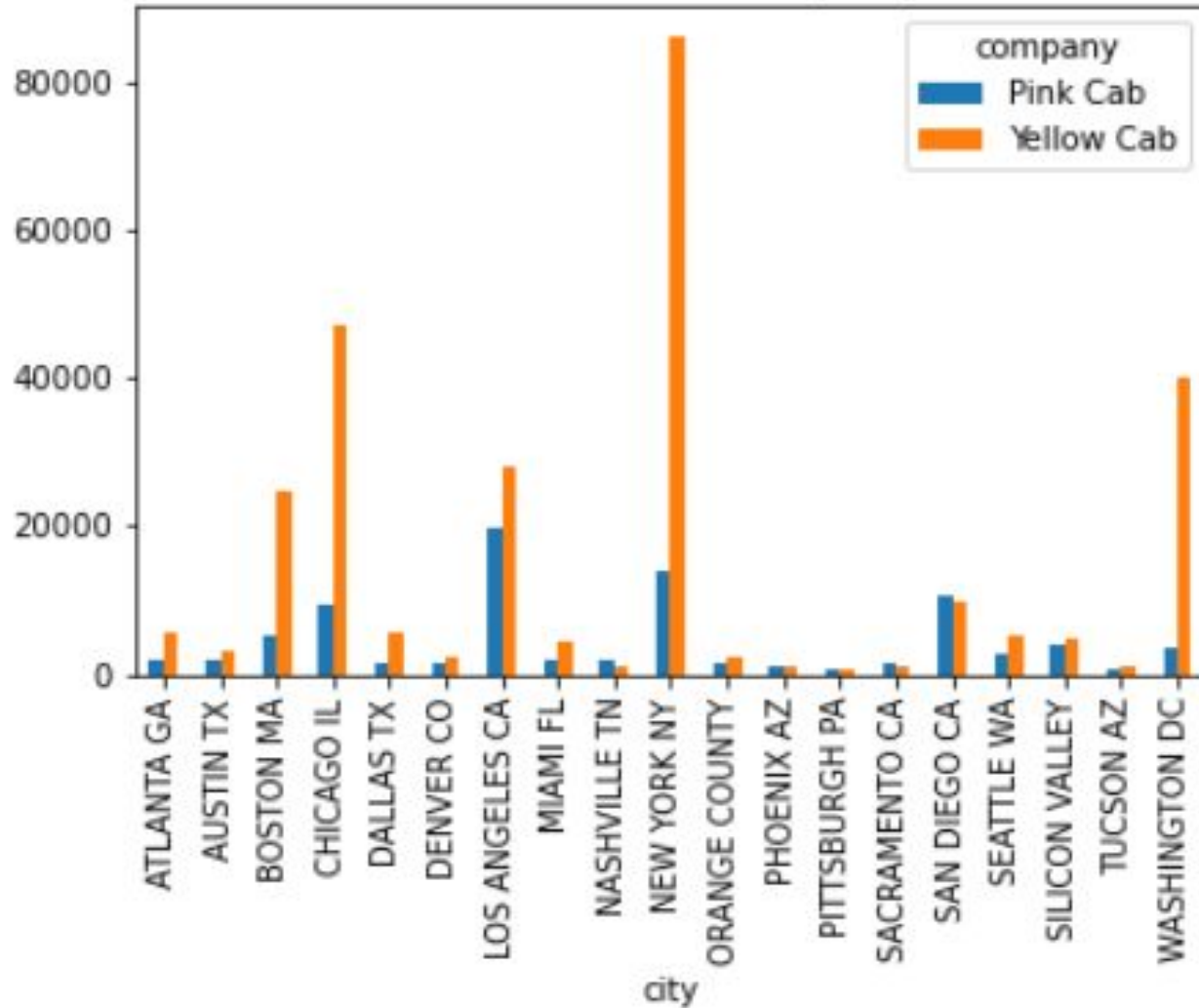
Profit Analysis By City

Customer base Analysis Gender wise
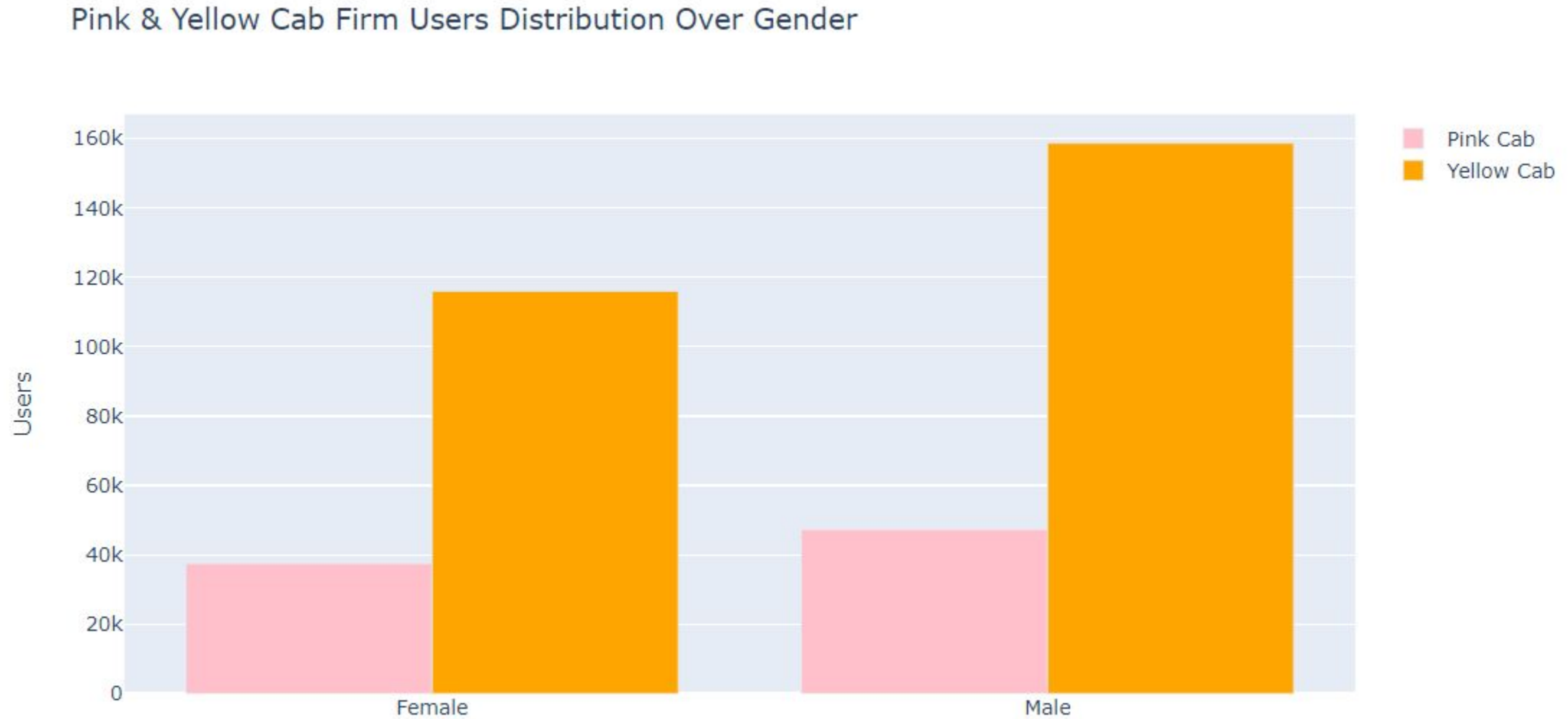
Customer base Analysis Gender wise

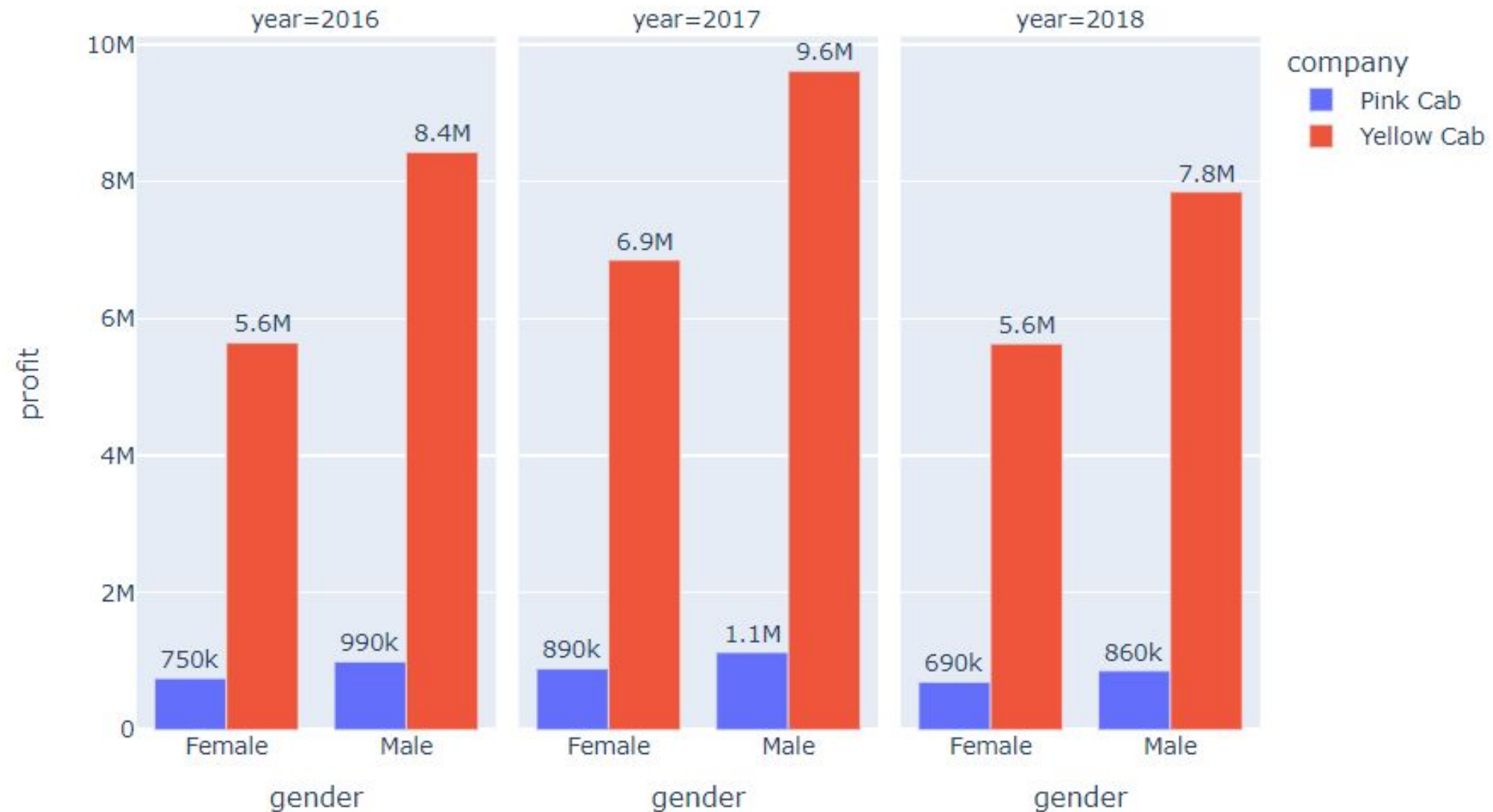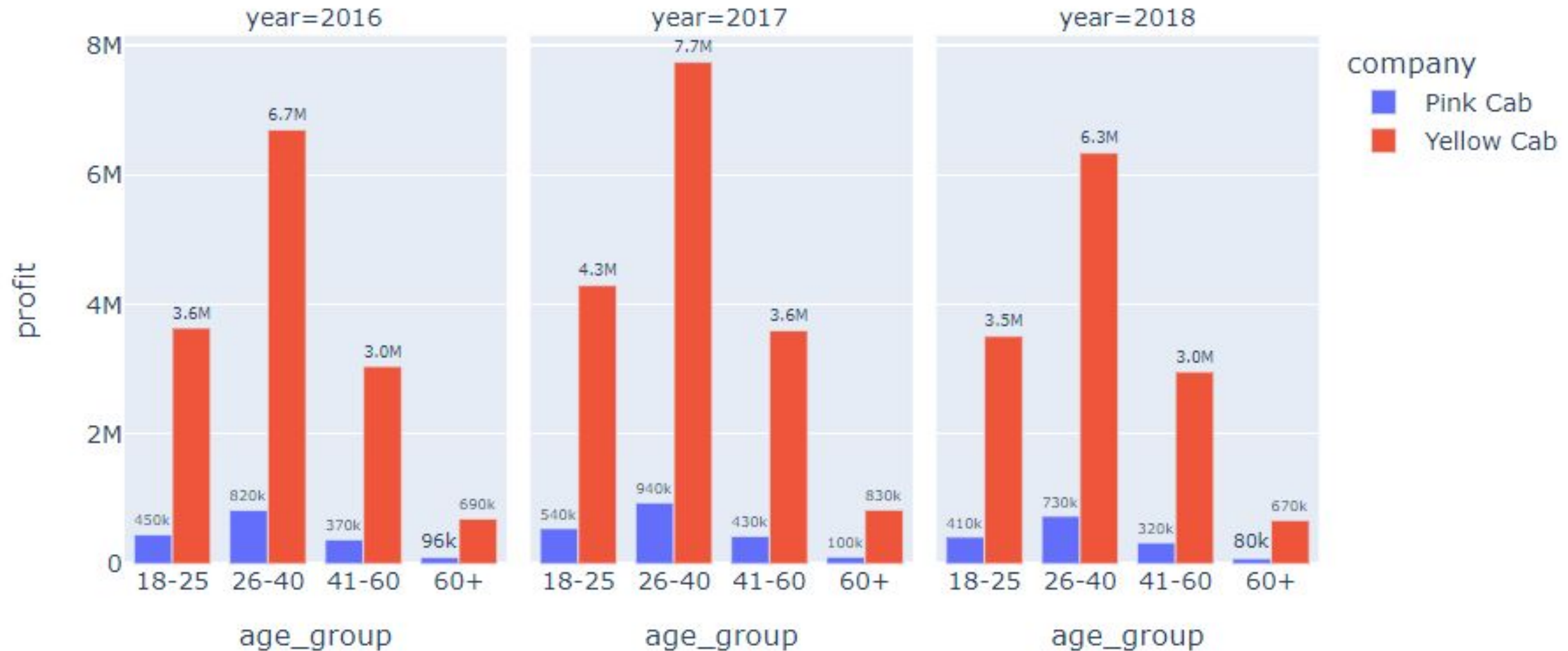Customer base per yer per Cab Company

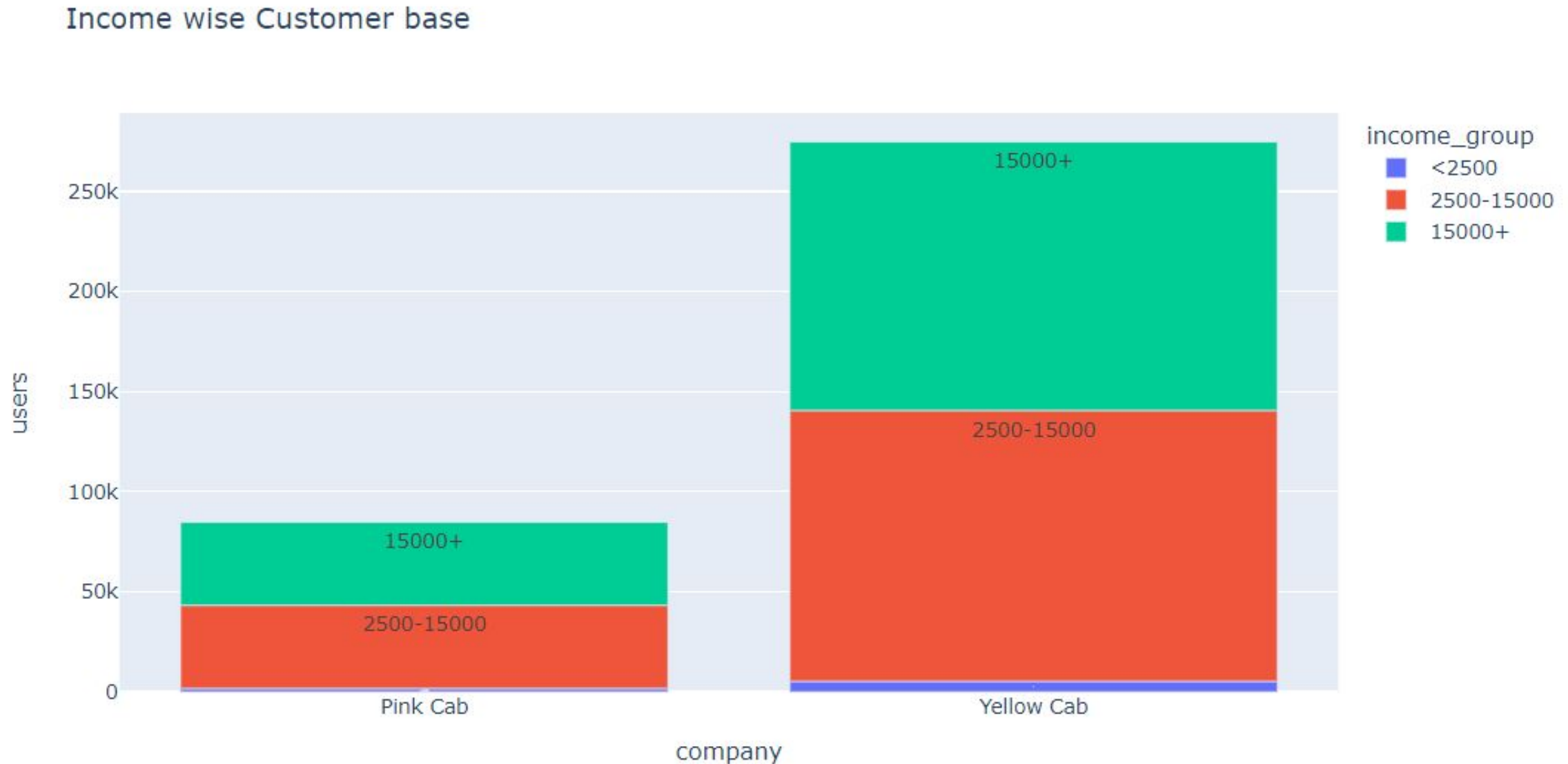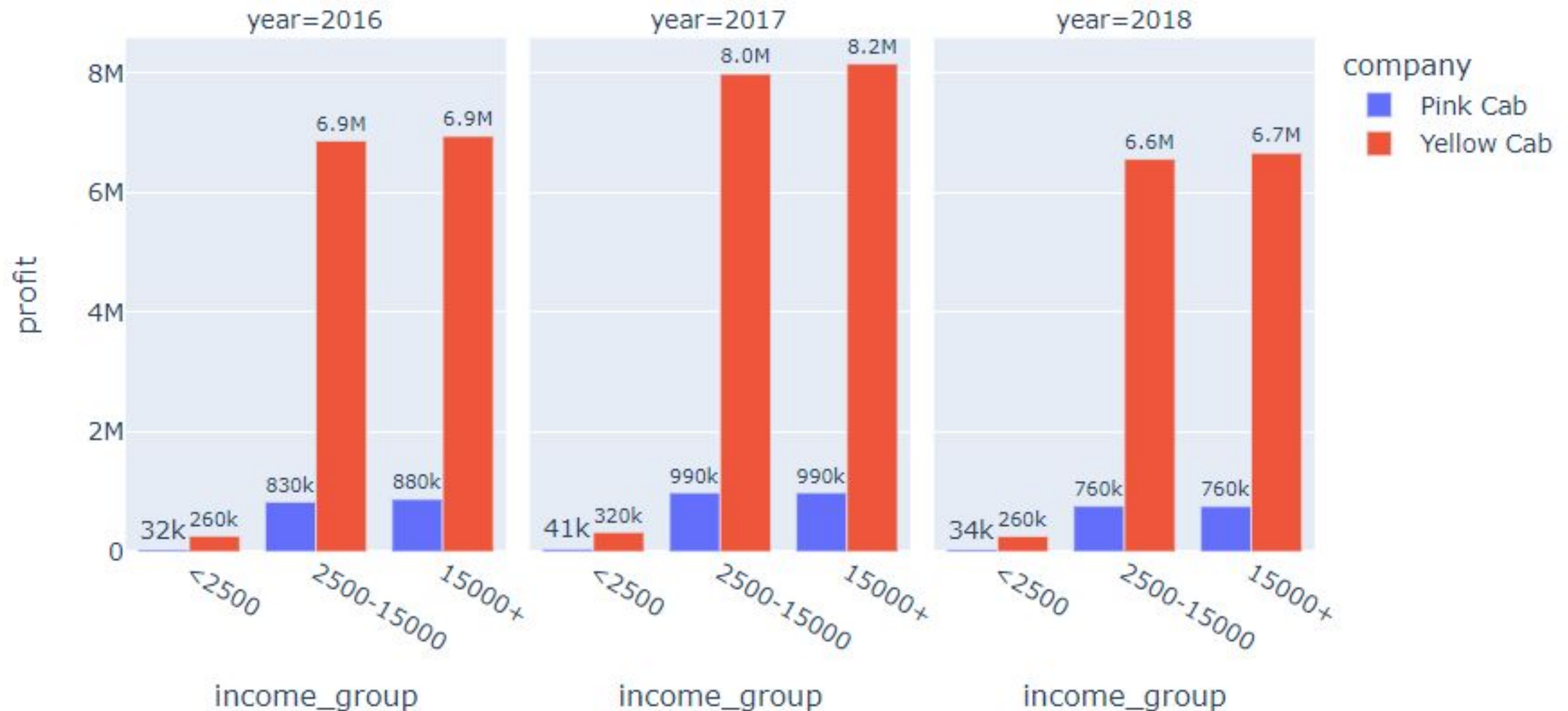# Profit and Customer base Analysis Gender wise

Customer base Analysis Age wise

Profit and Customer base Analysis Age wise

Profit and Customer base Analysis Income wise

Income wise Customer base

Profit and Customer base Analysis Income wise

Observations and Recommendations

- **Yellow cab's average profit per KM** is almost **three times** the average profit per KM of the Pink cab.

- Most profit earned from users for both Cab companies fall in between 26-40 age group. **Yellow cab has more customer in all age group** compared to that of Pink Cab resulting more profit

- **Yellow cab has higher customer reach in 25 citie**s while Pink cab has higher customer reach in 4 cities.

- Both the cabs are very popular in high and medium income class but here also **Yellow cab is performing better than Pink cab in offering their services to all the three income class group** (low, medium and high)

On the basis of above point , we will recommend Yellow cab for investment.

Thank You

Data Glacier
Your Deep Learning Partner