# Exploratory Data Analysis

G2M insight for Cab Investment firm

3-19-2022

- 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.

- The analysis has been divided into three parts:
  - Data Understanding
  - Finding the most profitable Cab company
  - Recommendations for investment

Data Exploration

Cab_Data.csv ustomer_ID.csv Transaction_ID.c... City...v USholiday.csv

Final cab data
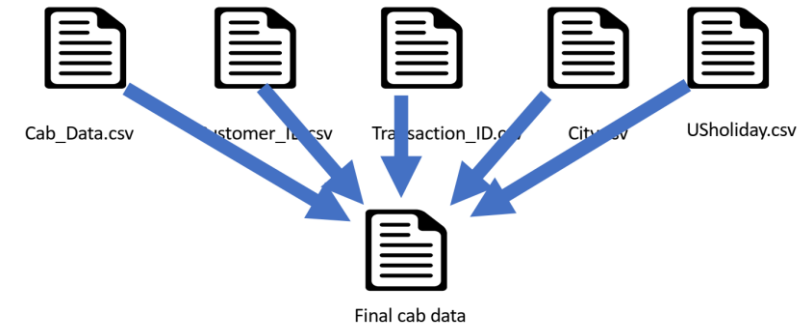
- 20 Features(including 7 derived features)
- Timeframe of the data: 2016-01-31 to 2018-12-31
- Total data points :354,746
- Assumptions:
  - Profit of rides are calculated keeping other factors constant and only Price_Charged and Cost_of_Trip features used to calculate profit.
- Hypothesis:
  - Is there any seasonality in number of customers using the cab service?
  - Is pink cab progress is better than yellow cab?
  - Is there any city that pink performed better in term of profit per user?
  - Did pink cab performed better in term of profit per kilometer?
 - Is age of customer affects our profit?

Data Exploration

- Date of Travel values range is from 42371 to 43465, Which is more than 1065 days.
- We need to remove extra 30 days data from our table. So we consider dates from 42401 to 43465.
- Then we map each date between (2016, 1, 31) and (2018, 12, 31) to values from 42401 to 43465.
- We extracted useful information from Date column:
  - Extracting week of year from date, varies from 1 to 52
  - Extracting month from date - 1 to 12
  - Extracting quarter from date variable - 1 to 4
  - We also extracted semester
  - Extracting year
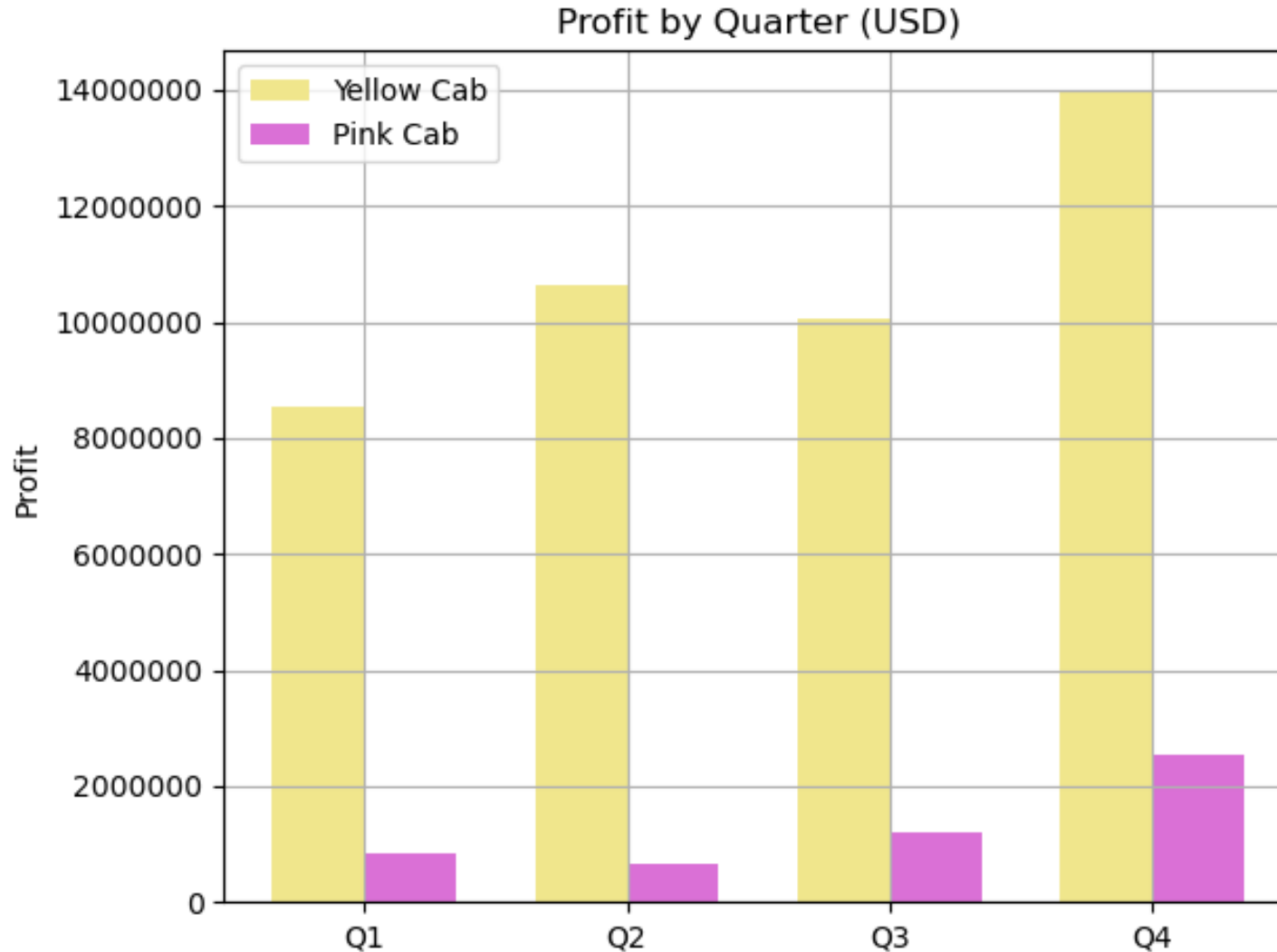 - Extracting day of the week - from 0 to 6
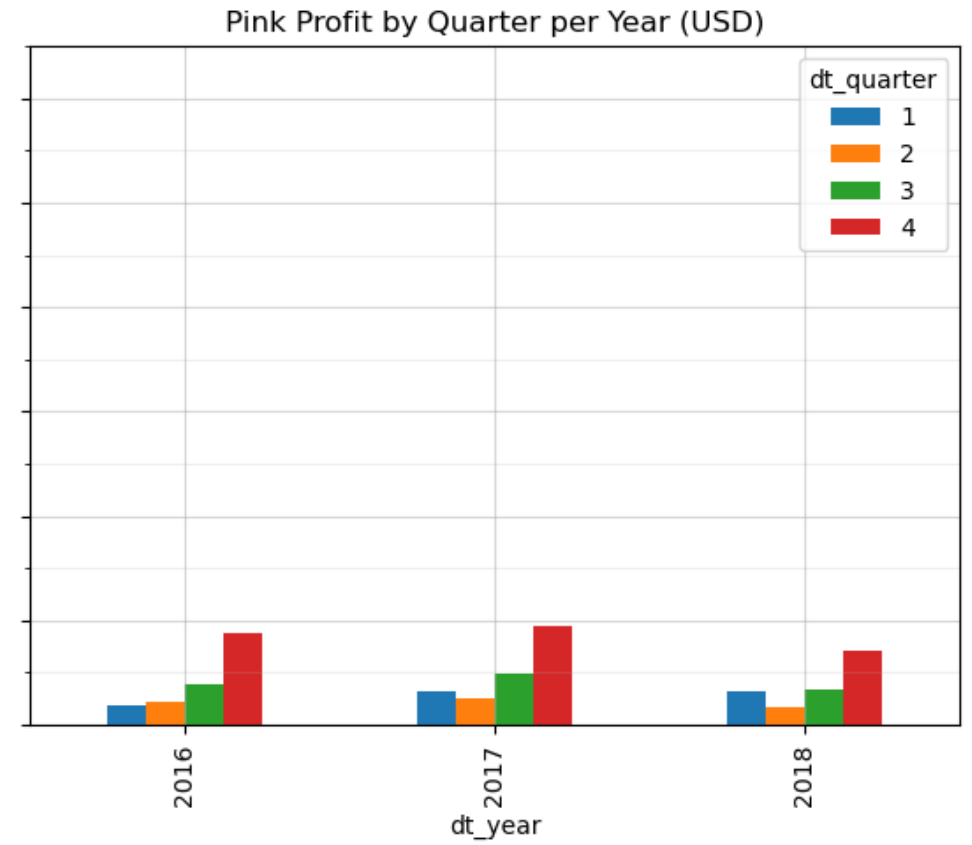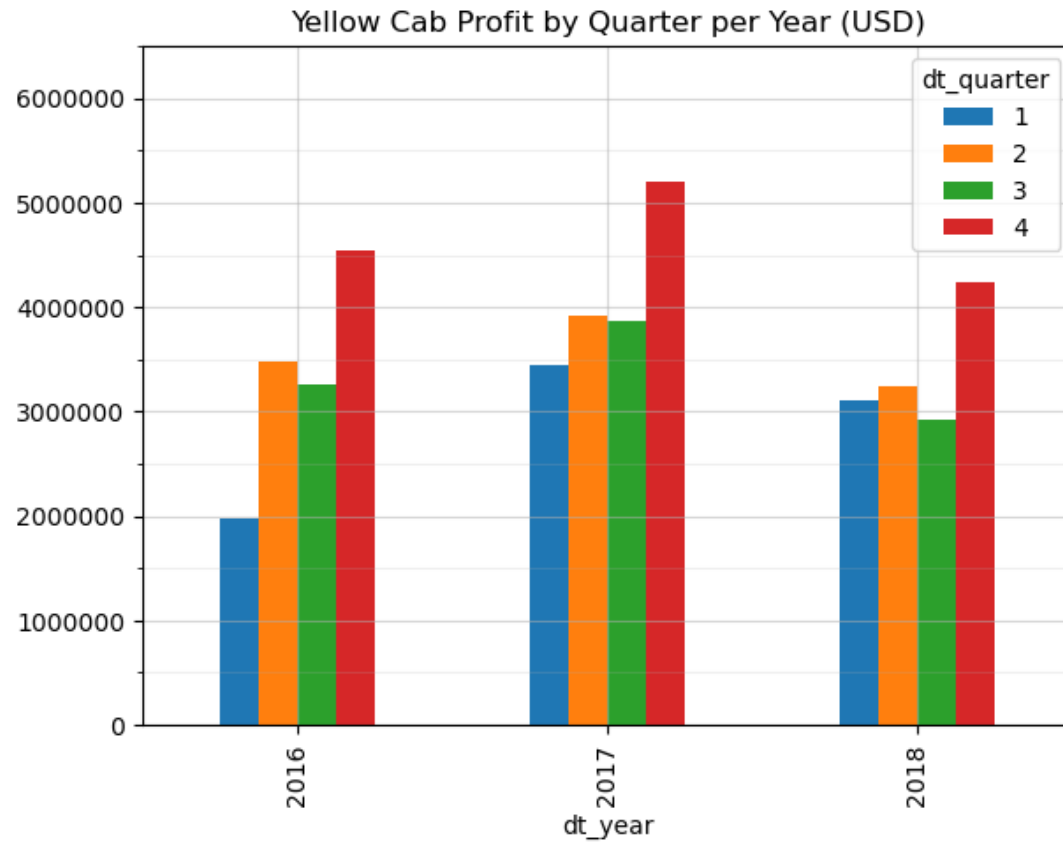
Profit of a company in each year
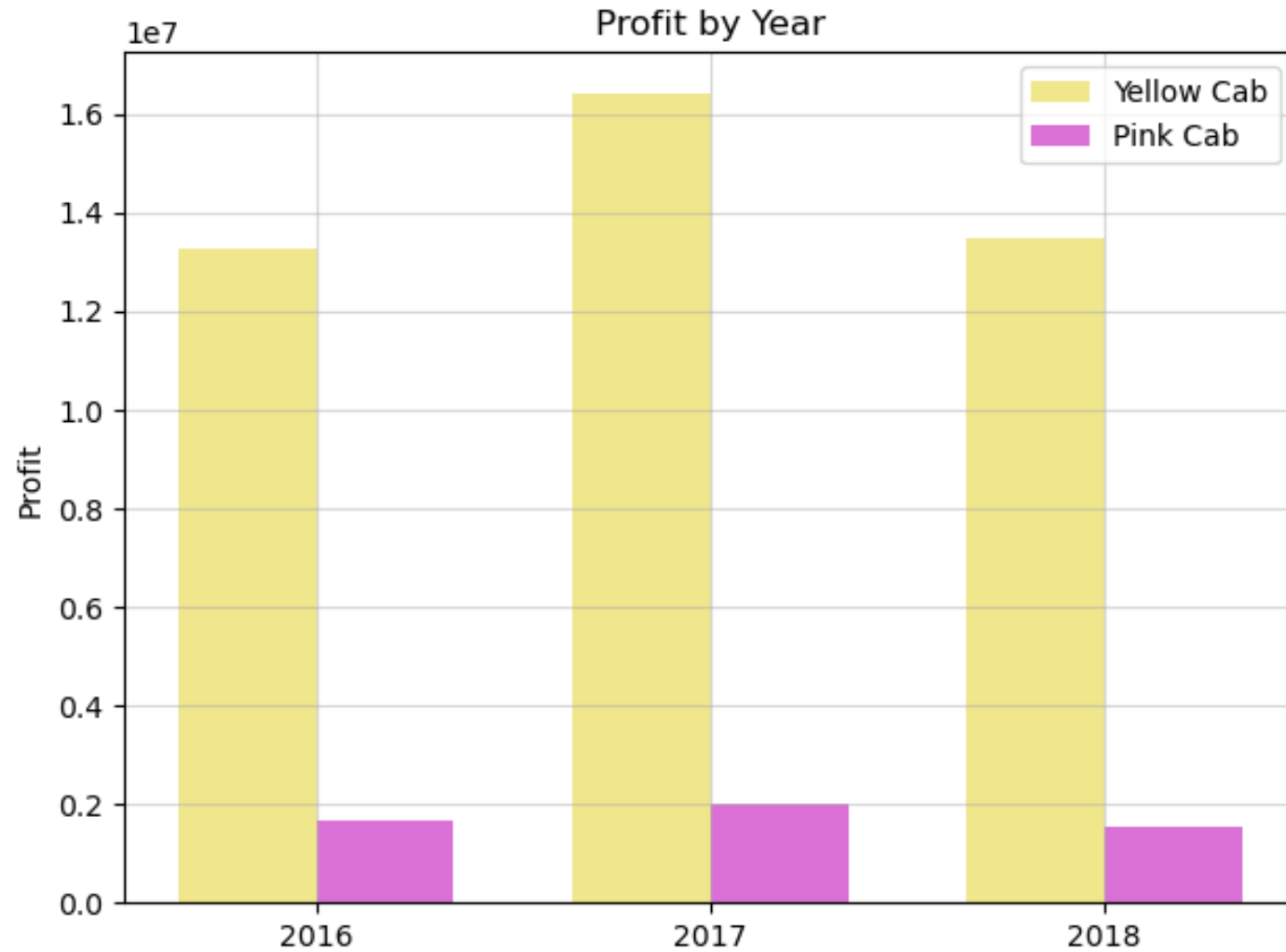
Profit percentage change by Year
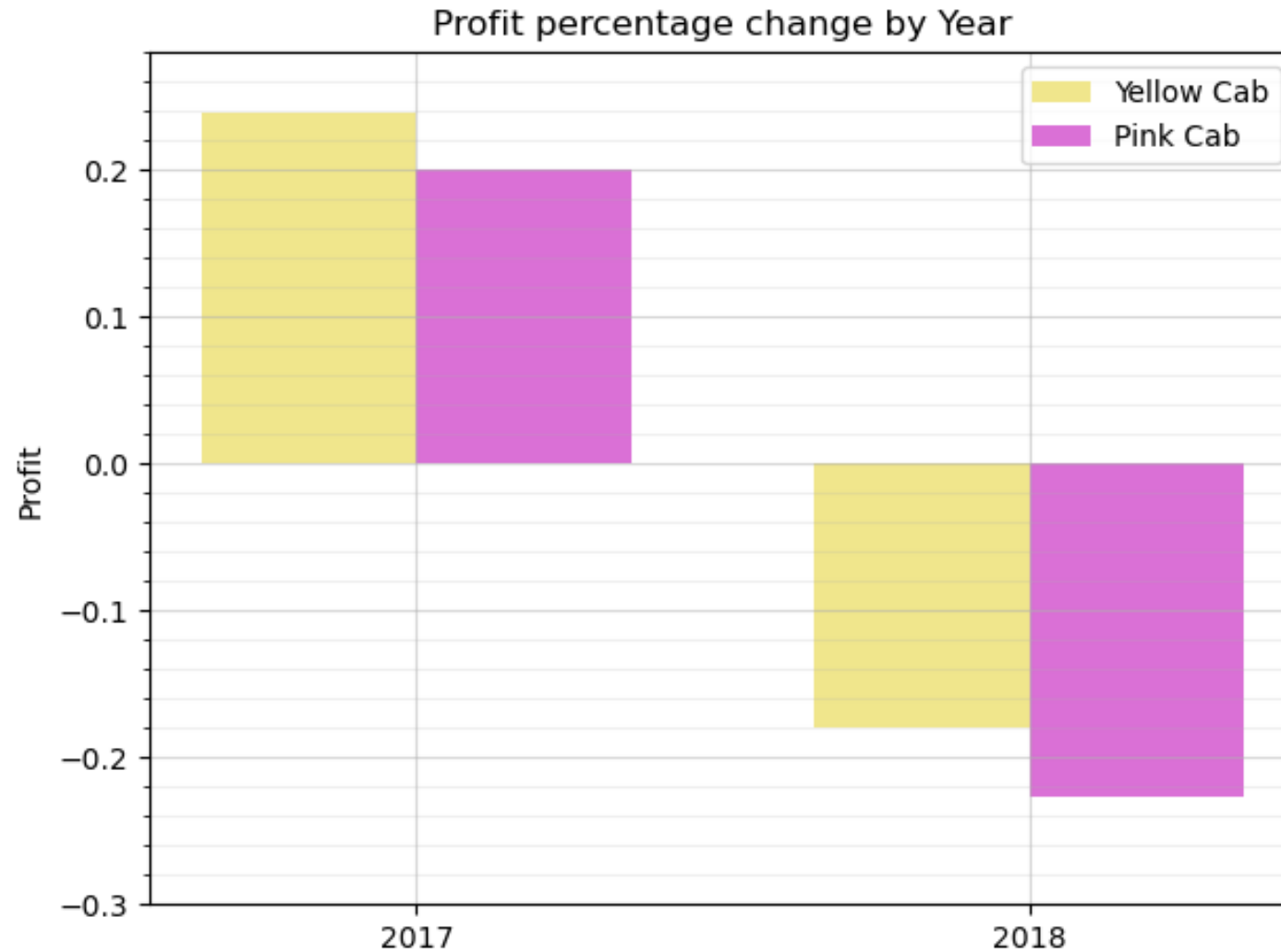
1. Is pink cab progress is better than yellow cab in quarters?
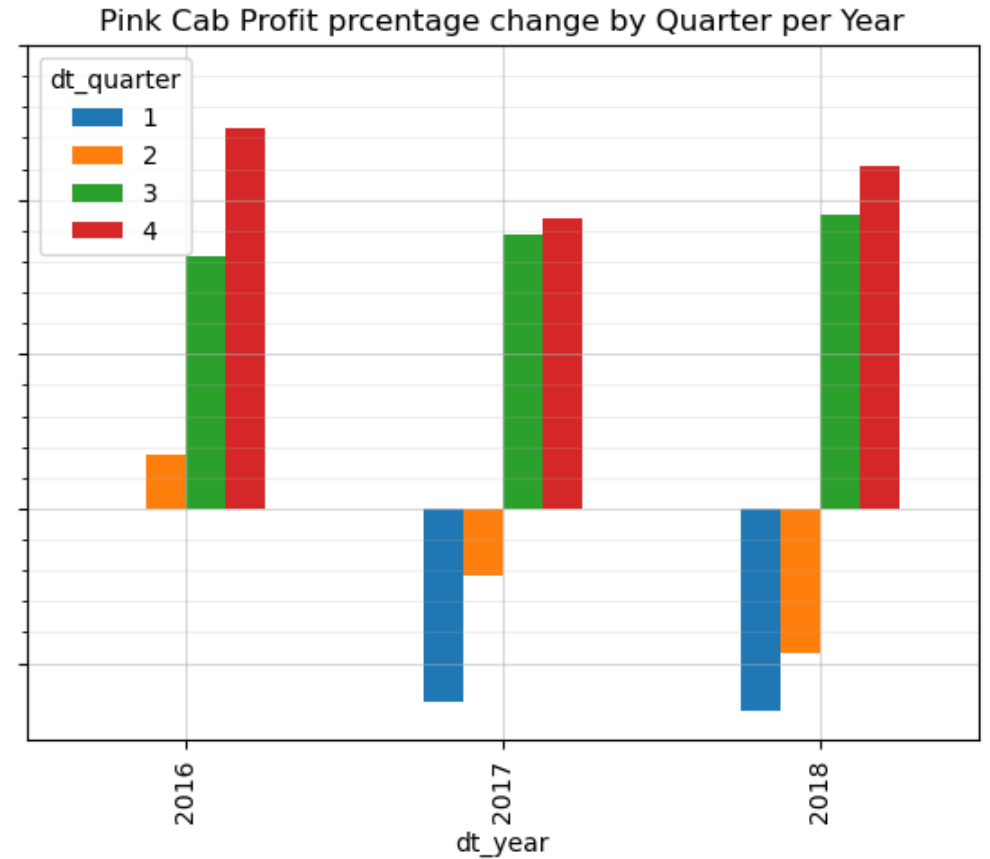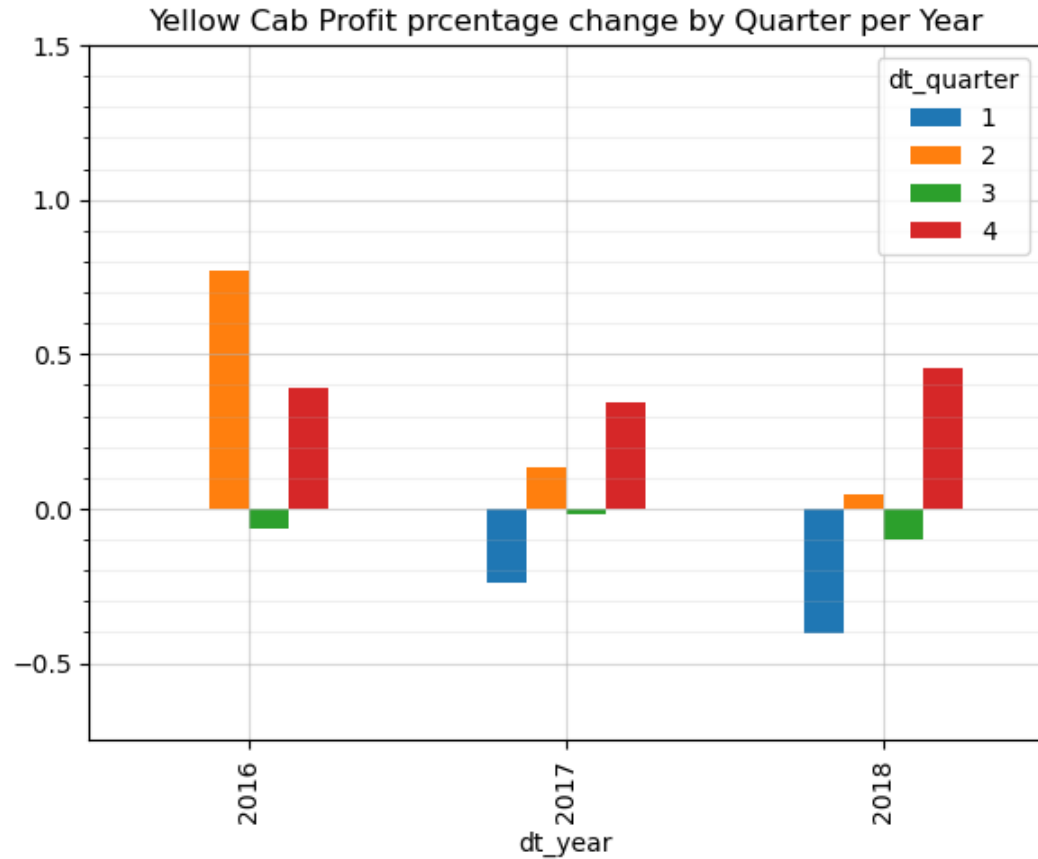

Profit by Quarter (USD)

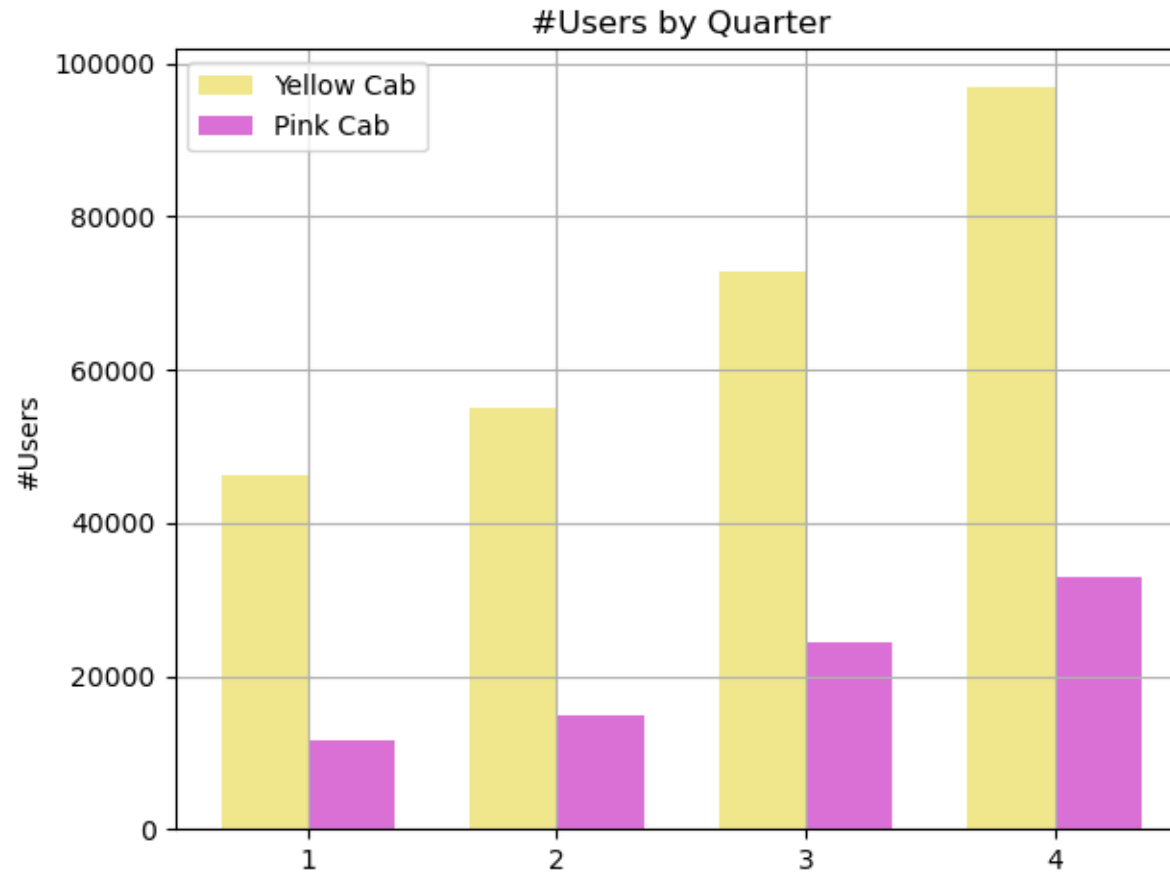# 1. Is pink cab progress is better than yellow cab in quarters?

1. Is pink cab progress is better than yellow cab in quarters?

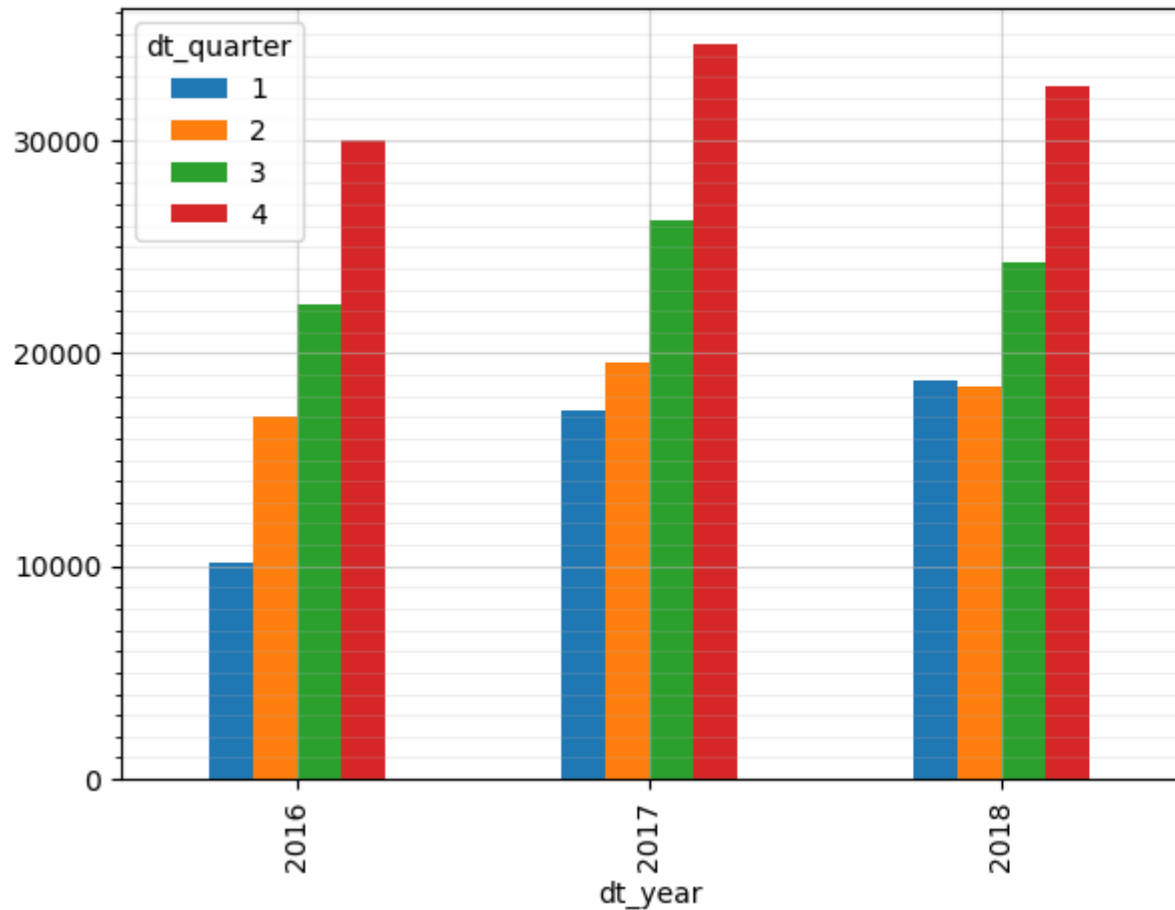# 1. Is pink cab progress is better than yellow cab in quarters?

- In some cases pink performed better but in total, yellow cap progress is more reliable and positive than pink cap.

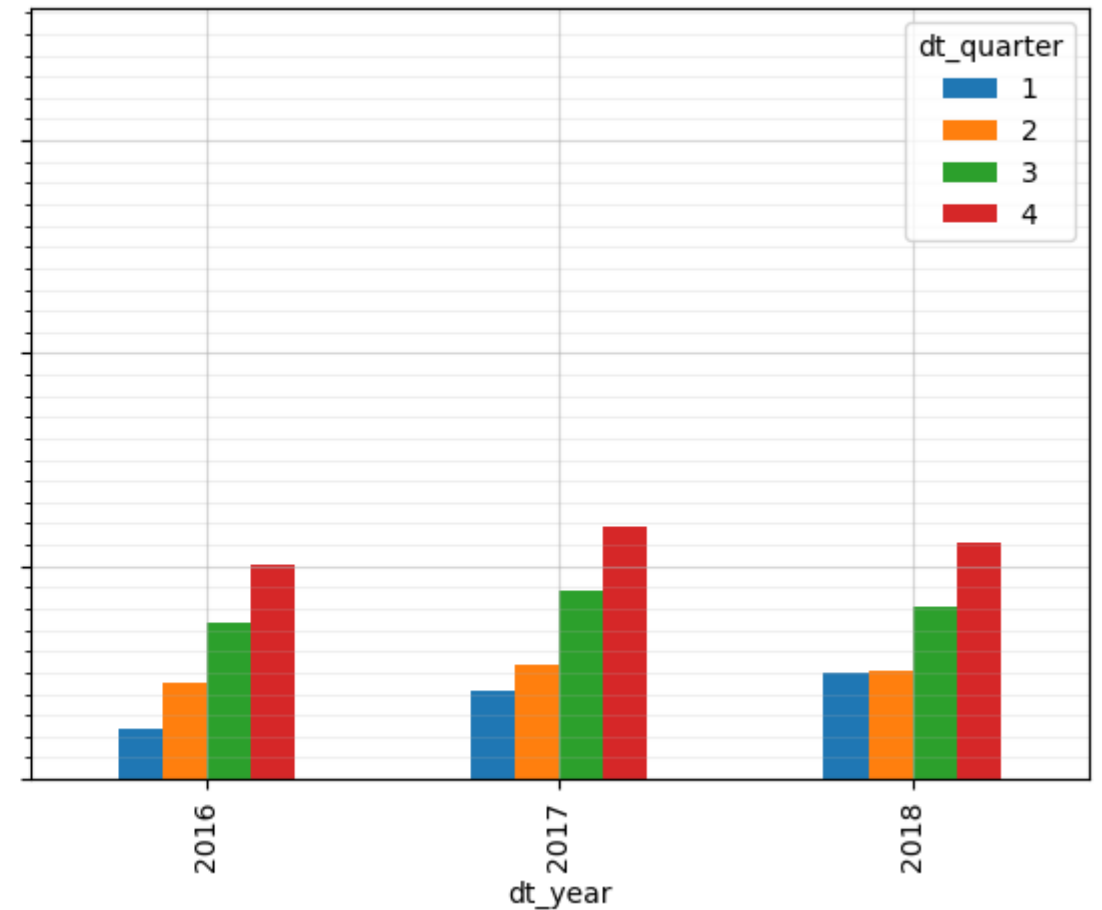# 2. Is there any seasonality in number of customers using the cab service?

2. Is there any seasonality in number of customers using the cab service?
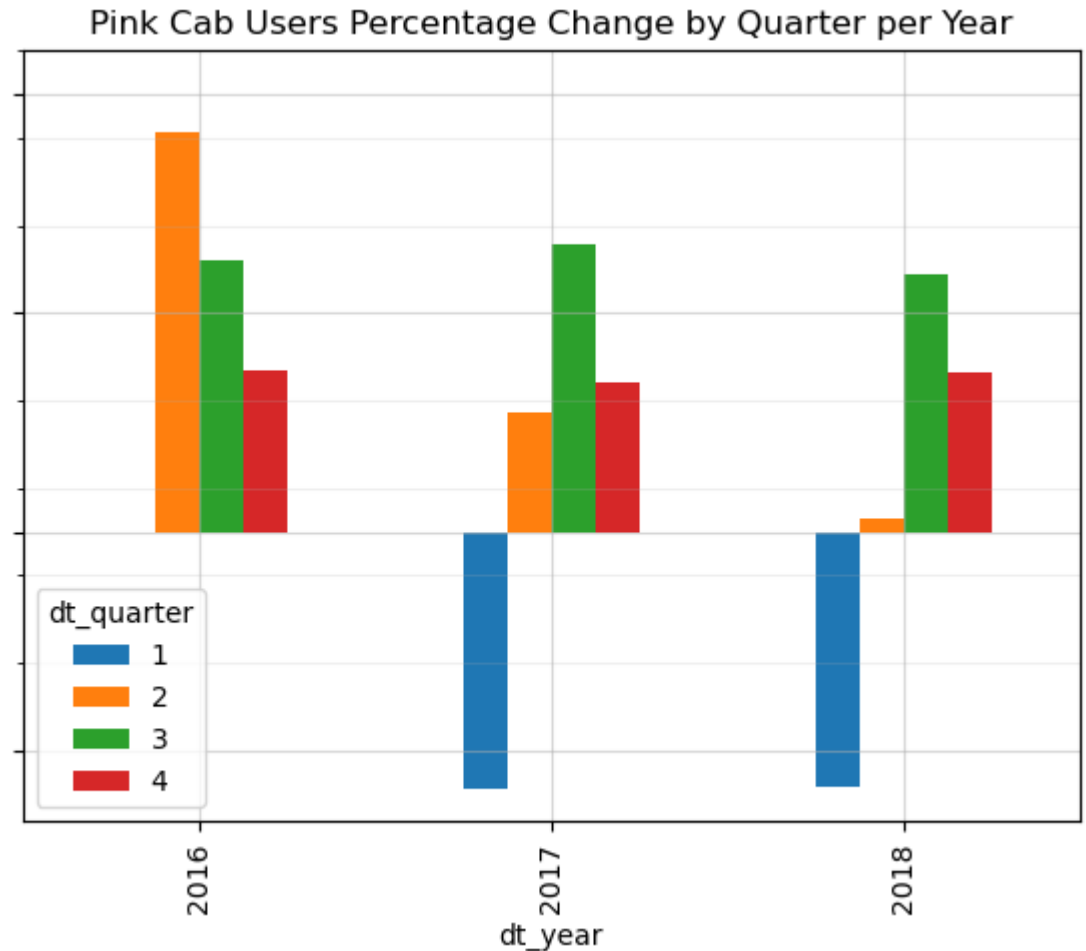
2. Is there any seasonality in number of customers using the cab service?

2. Is there any seasonality in number of customers using the cab service?
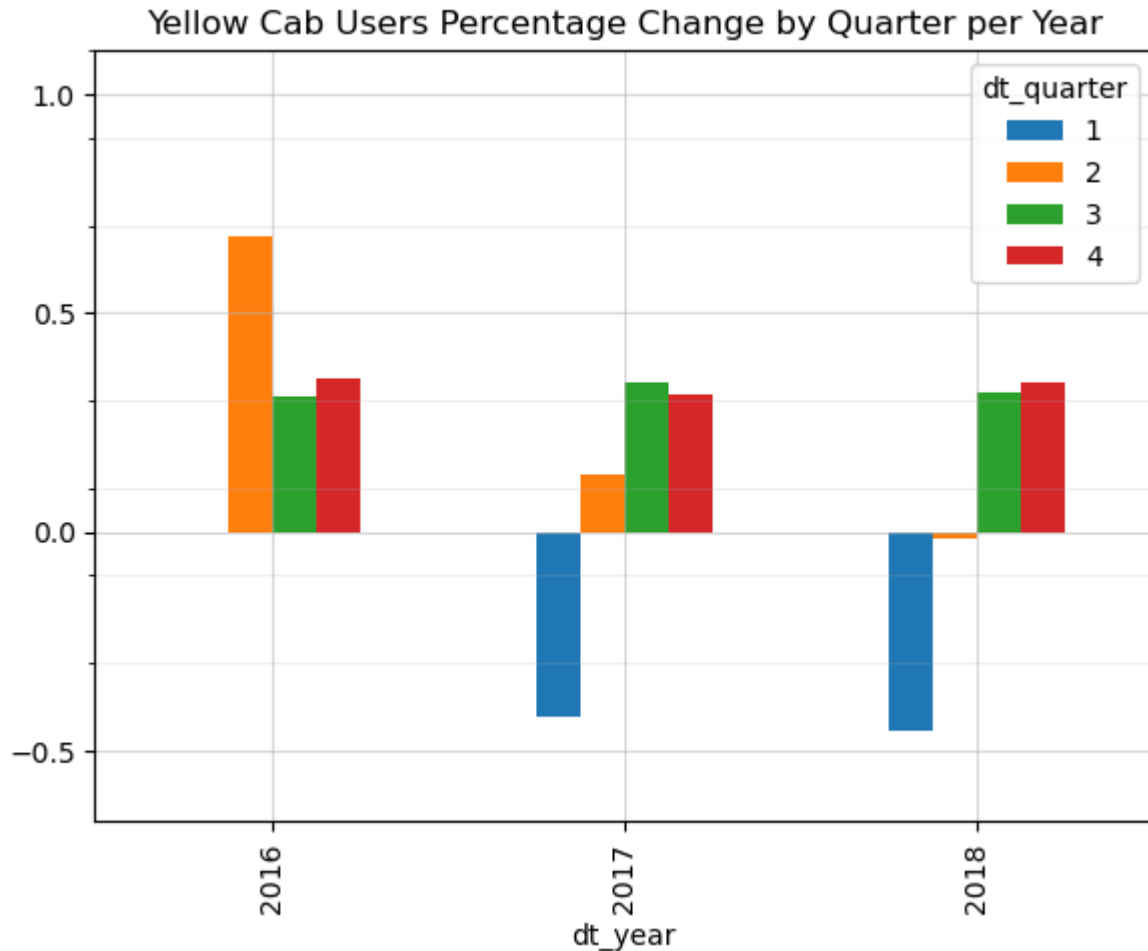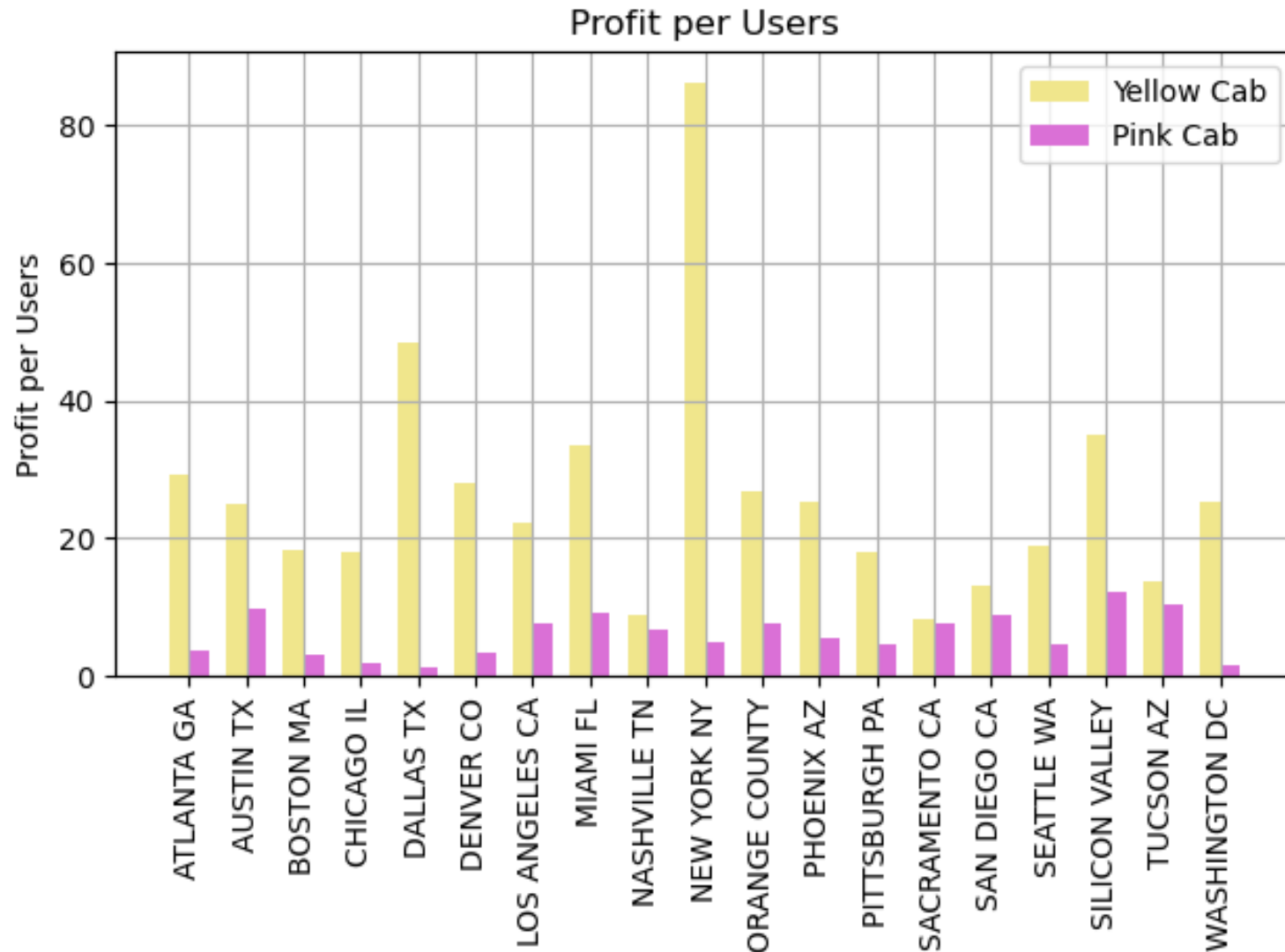
- Obviously pink cap has better performance in case of users percentage change and gaining more users, so we can say if pink cap has enough support, maybe in future it will pass yellow cap.
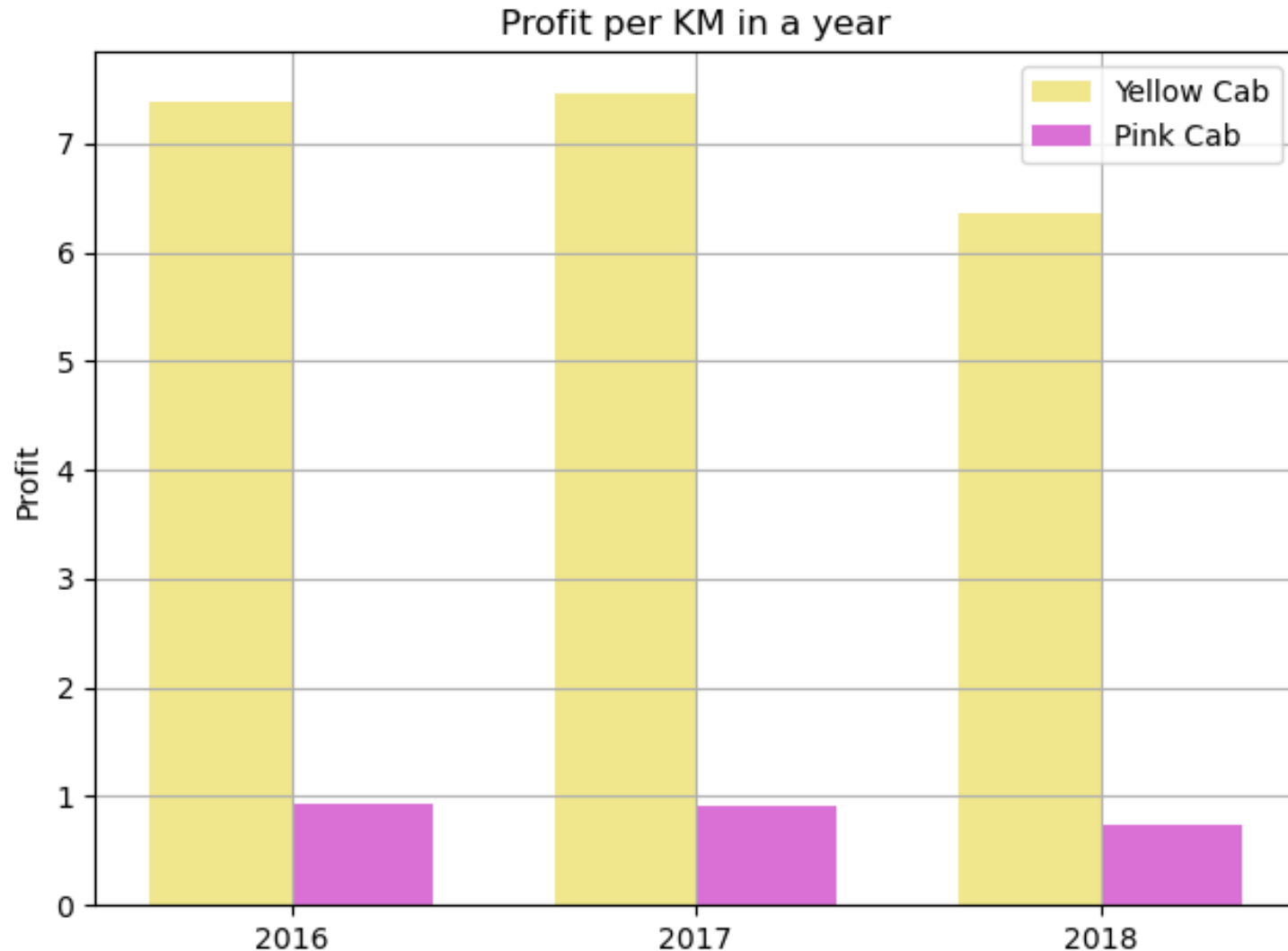
3. Is there any city that pink performed better in term of profit per user?

3. Is there any city that pink performed better in term of profit per user?

- No, but there are cities that pink cap has close enough competition with yellow cap. Which are:
 - Nashvile
  - Sacramento
  - San Diego
  - Tucsan

4. Did pink cab performed better in term of profit per kilometer?
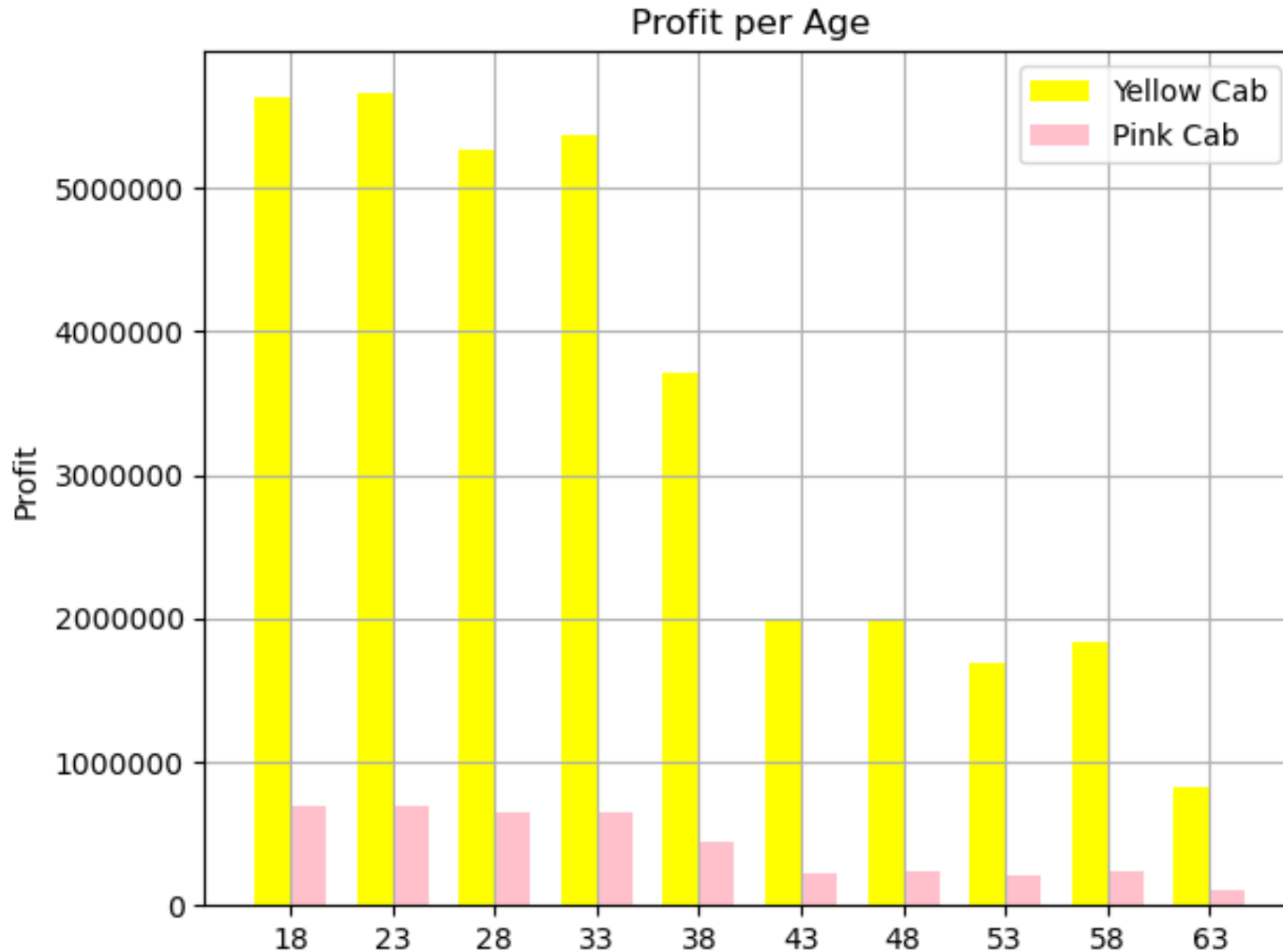

Profit per KM in a year

4. Did pink cab performed better in term of profit per kilometer?

- No, As we can see, pink cap profit is falling each year. But yellow cap profit is increased in second year.

5. Is age of customer affects our profit?

#Users per Age

5. Is age of customer affects our profit?

Profit per Age
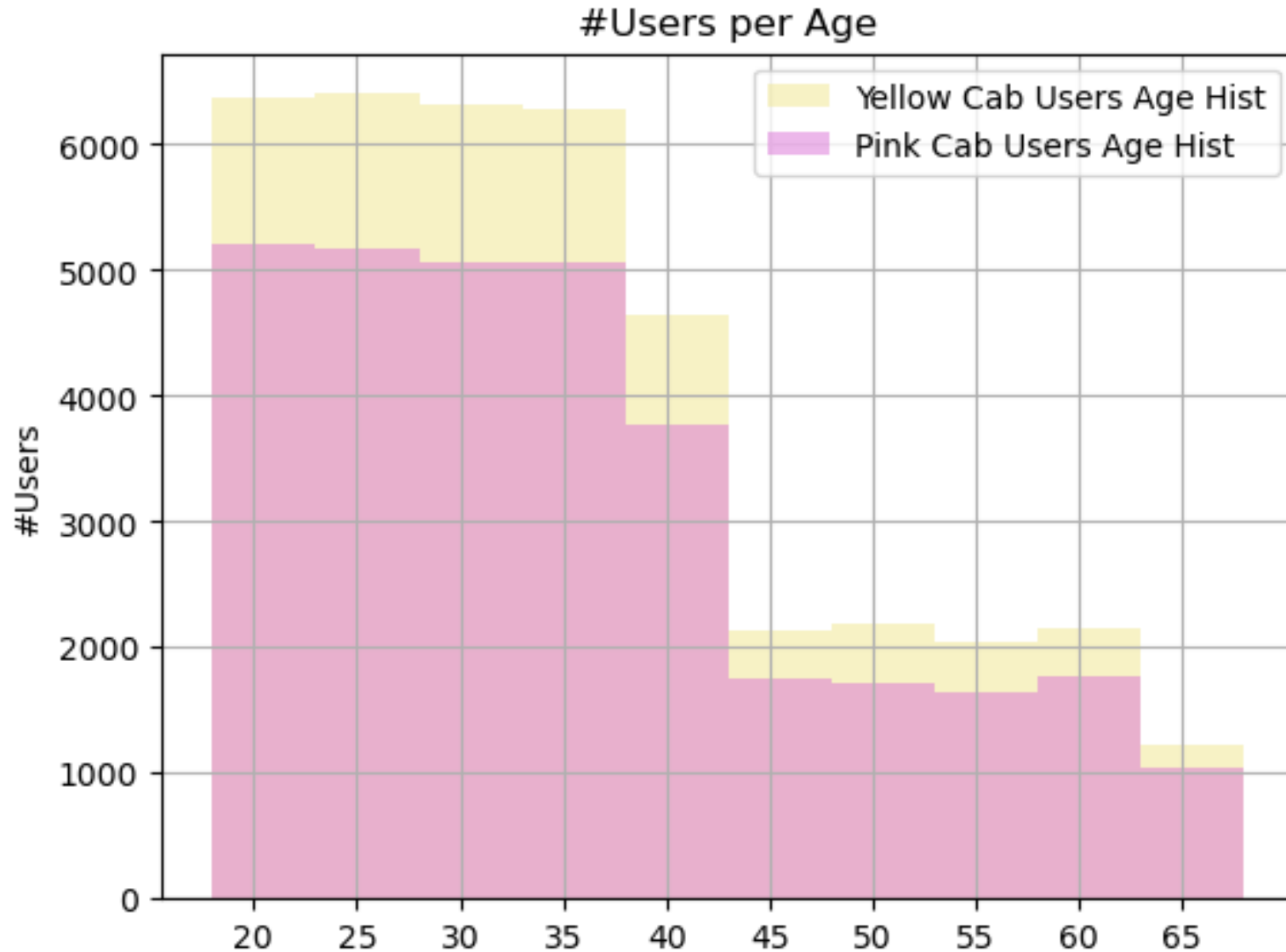
5. Is age of customer affects our profit?

- As we can see, Number of user at a specific age in both compaies has almost the same pattern. And also the profit they earn from them.

Conclusion

As we observed, both company almost have the same pattern in their data, but amount of them is different.

For now, we can say Yellow cap is doing better but with a little help and support, Pink Cap can grow faster and better than yellow cap.

Thank You