Exploratory Data Analysis

Week 2: EDA

2021-03-07

Executive Summary

- Exploratory Data Analysis was performed on a data obtained from two different Cab service companies, operating across multiple cities in the USA.

- The "Yellow Cab" company proved to be a more promising investment and a thriving business as opposed to the "Pink Cab" company.

Problem Statement

- 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-Marketstrategy they want to understand the market before taking final decision.

- Multiple data sets containing information on 2 cab companies, within the time period from 31/01/2016 to 31/12/2018 were provided and to be investigated.

- **Data Sets:**
  - *Cab_Data:* Details of Transactions.
  - *Customer_ID:* Mapping each customer to their Demographic Details
  - *Transaction_ID:* Maps each transaction to customer and payment mode.
  - *City:* List of US cities, their population and number of cab users.

Approach

- Investment strategies involves:
  - Profit Analysis
  - Exploring company's activity through
    - Total distance travelled yearly
    - Number of cab rides on a yearly basis
    - Number of cab rides on a monthly basis
    - Number of cab rides on a weekly basis
  - Measuring company's reach across the USA
  - Exploring the customer basis
    - Gender Wise
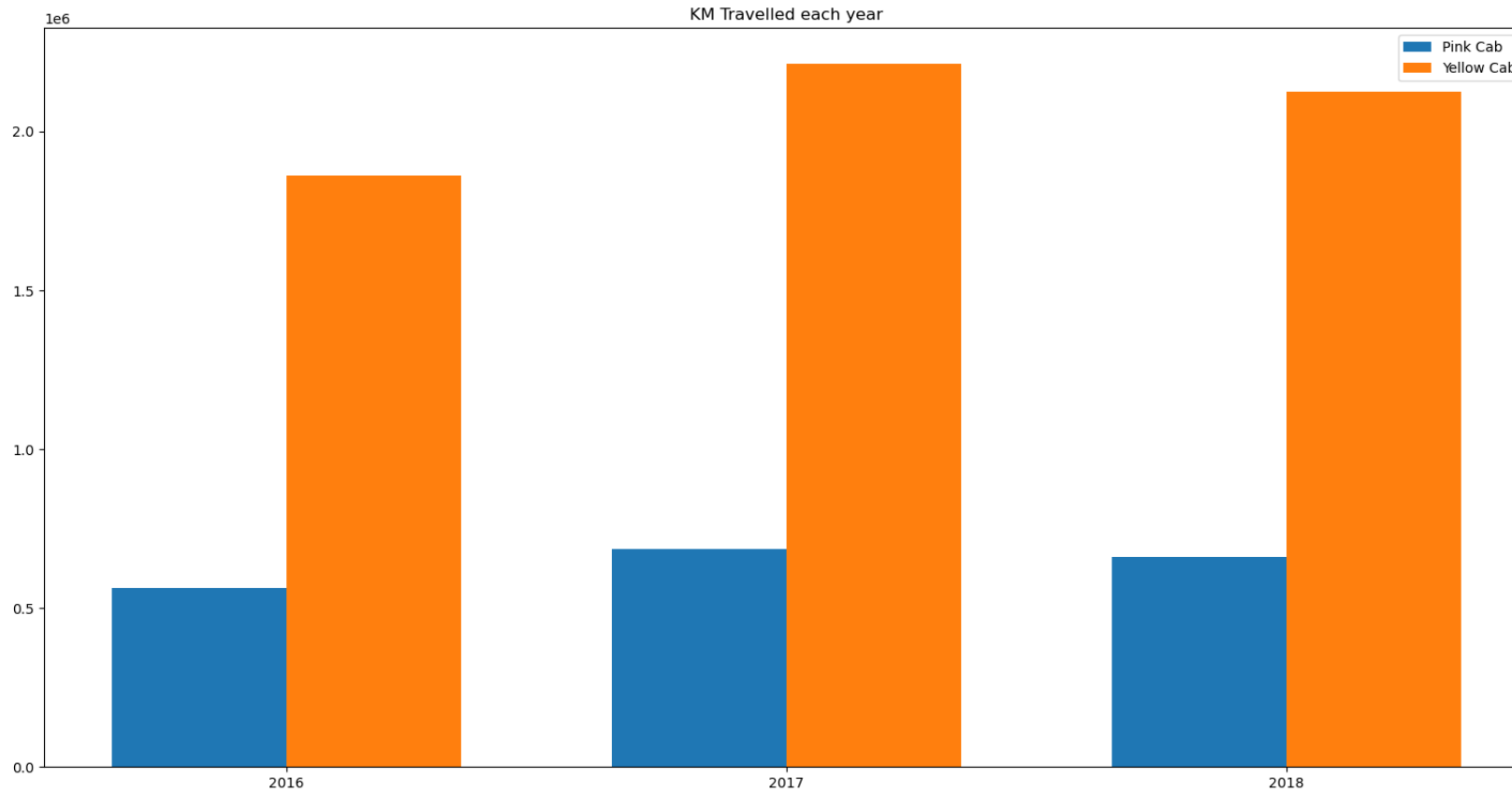   - Customer preference
    - Customer Repeated Purchases

Profit Analysis

"Yellow Cab'' achieved an average profit 4.8 times larger than "Pink Cab" over three years.
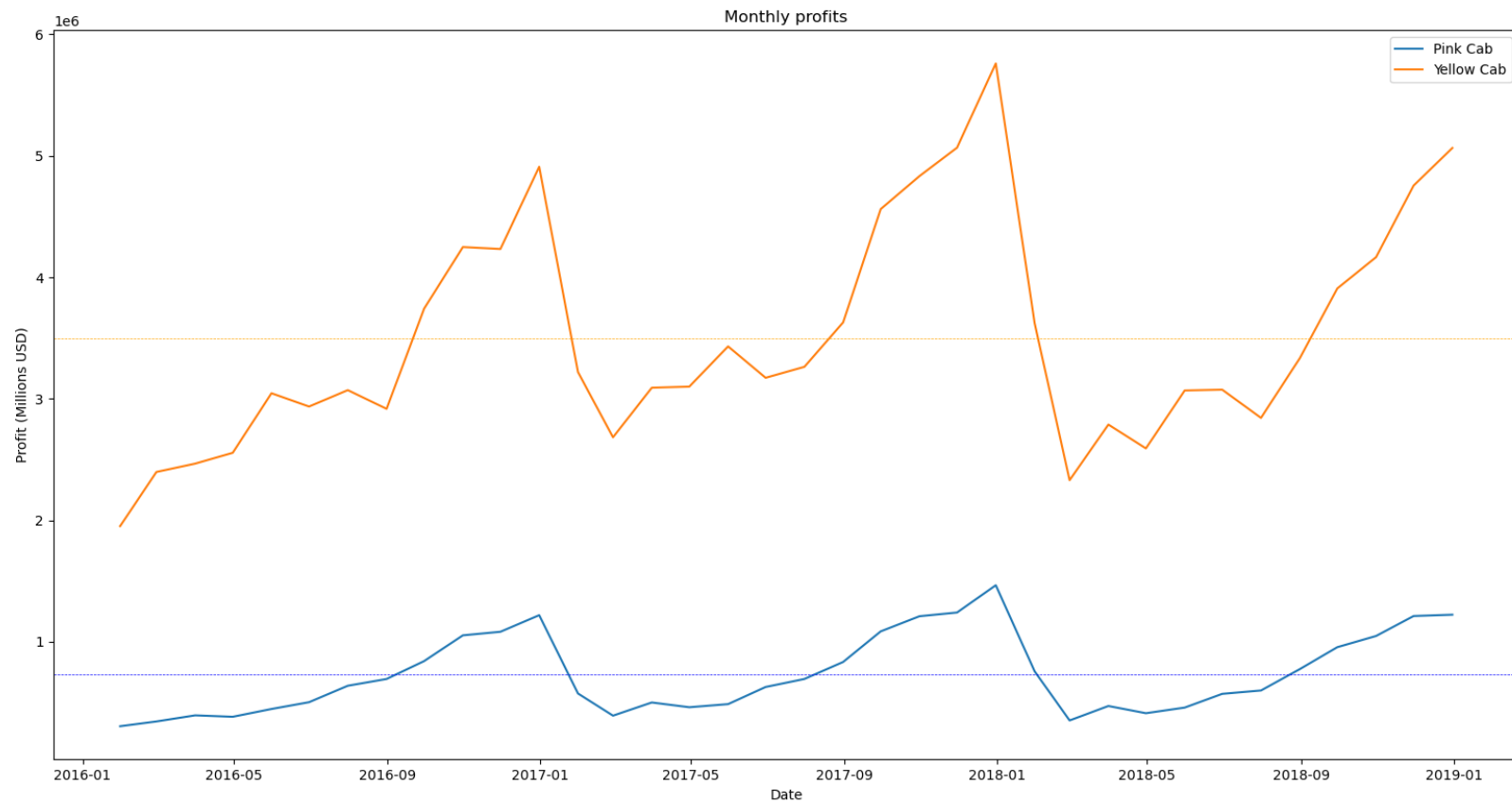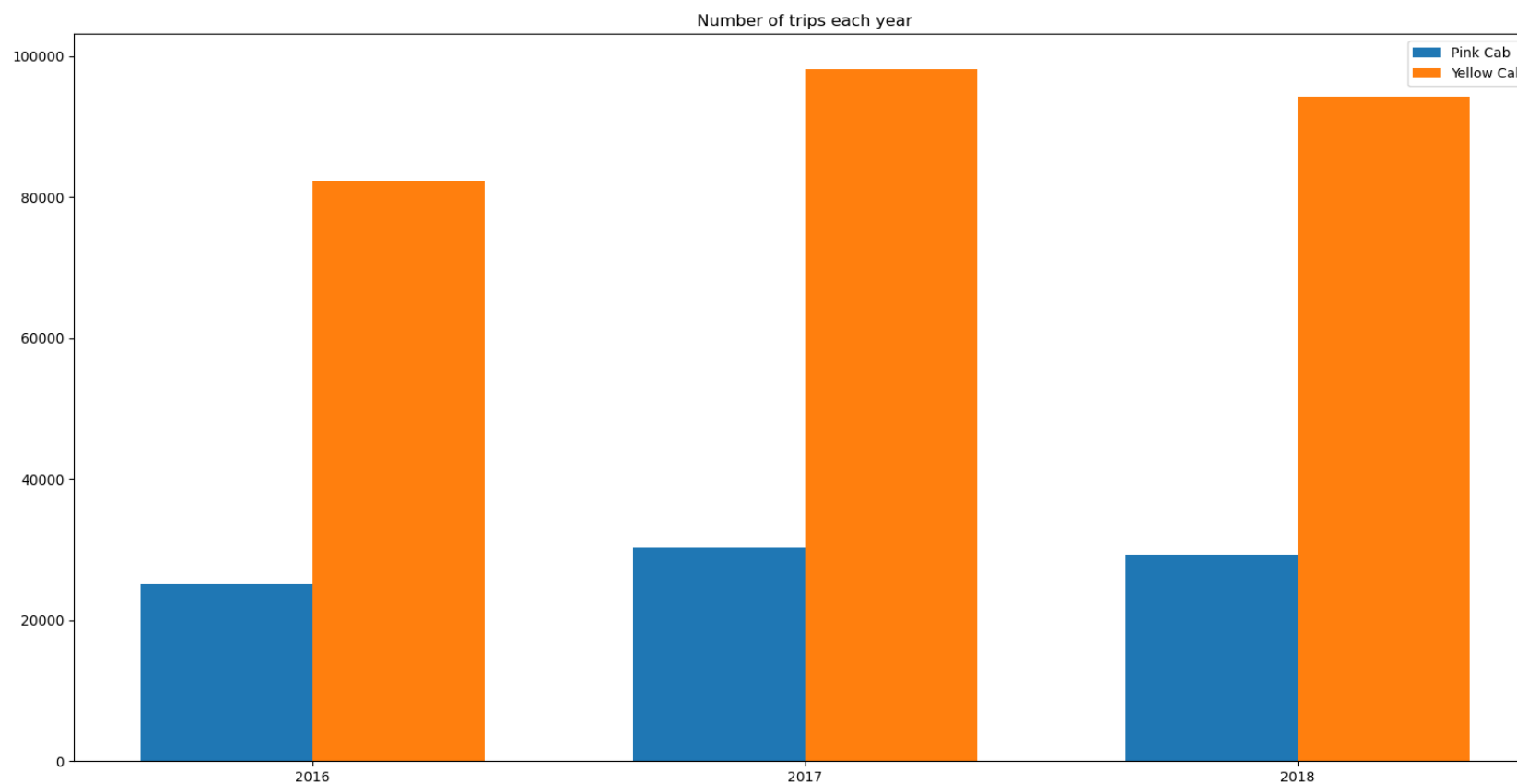
Activity by distance Travelled

By 2017 "Yellow Cab" was able to exceed 2 Million KMs in distance travelled during cab rides
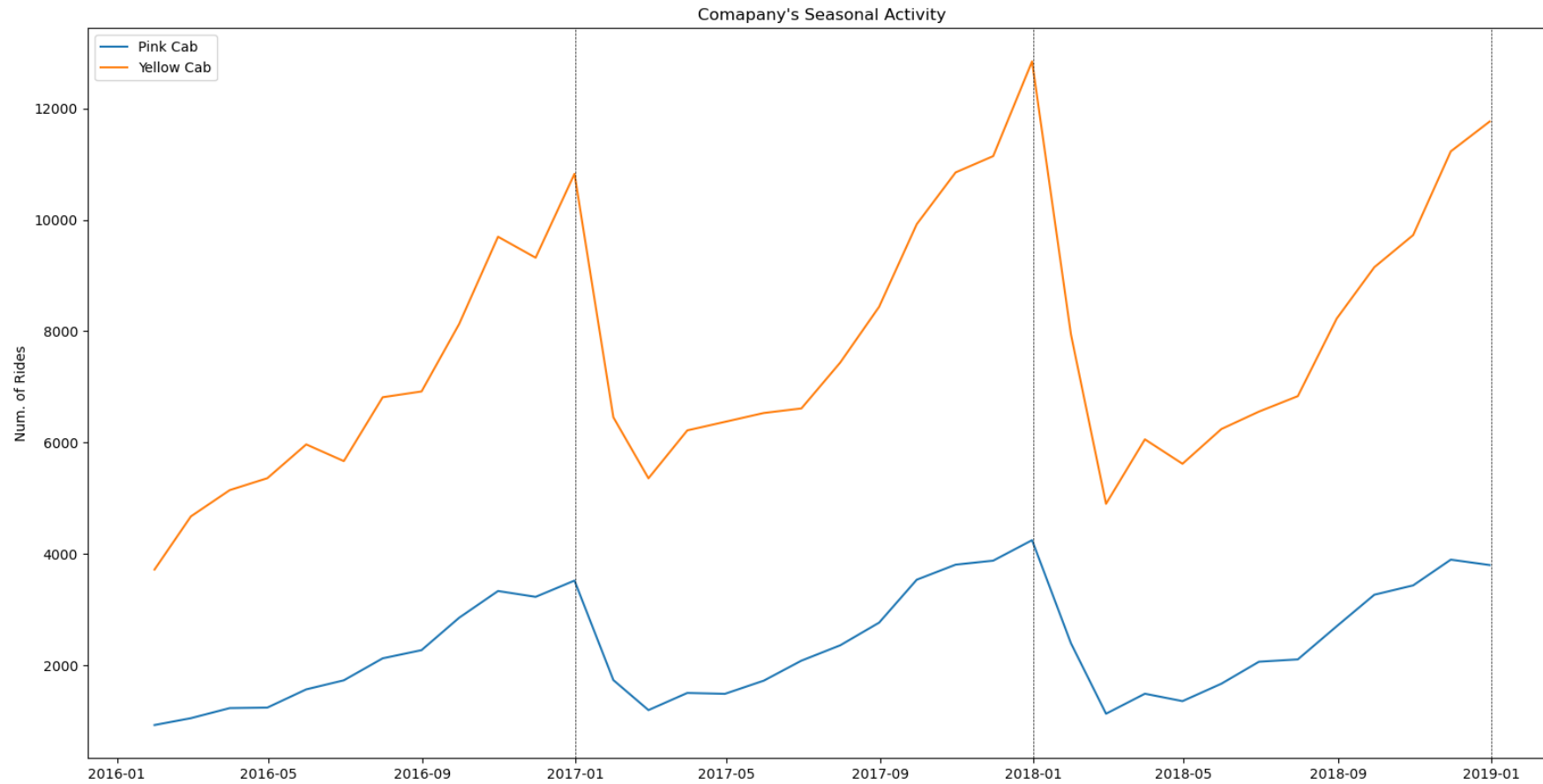
Activity by Number of cab rides

Number of rides achieved yearly by "Yellow Cab" is significantly higher than "Pink Cab"
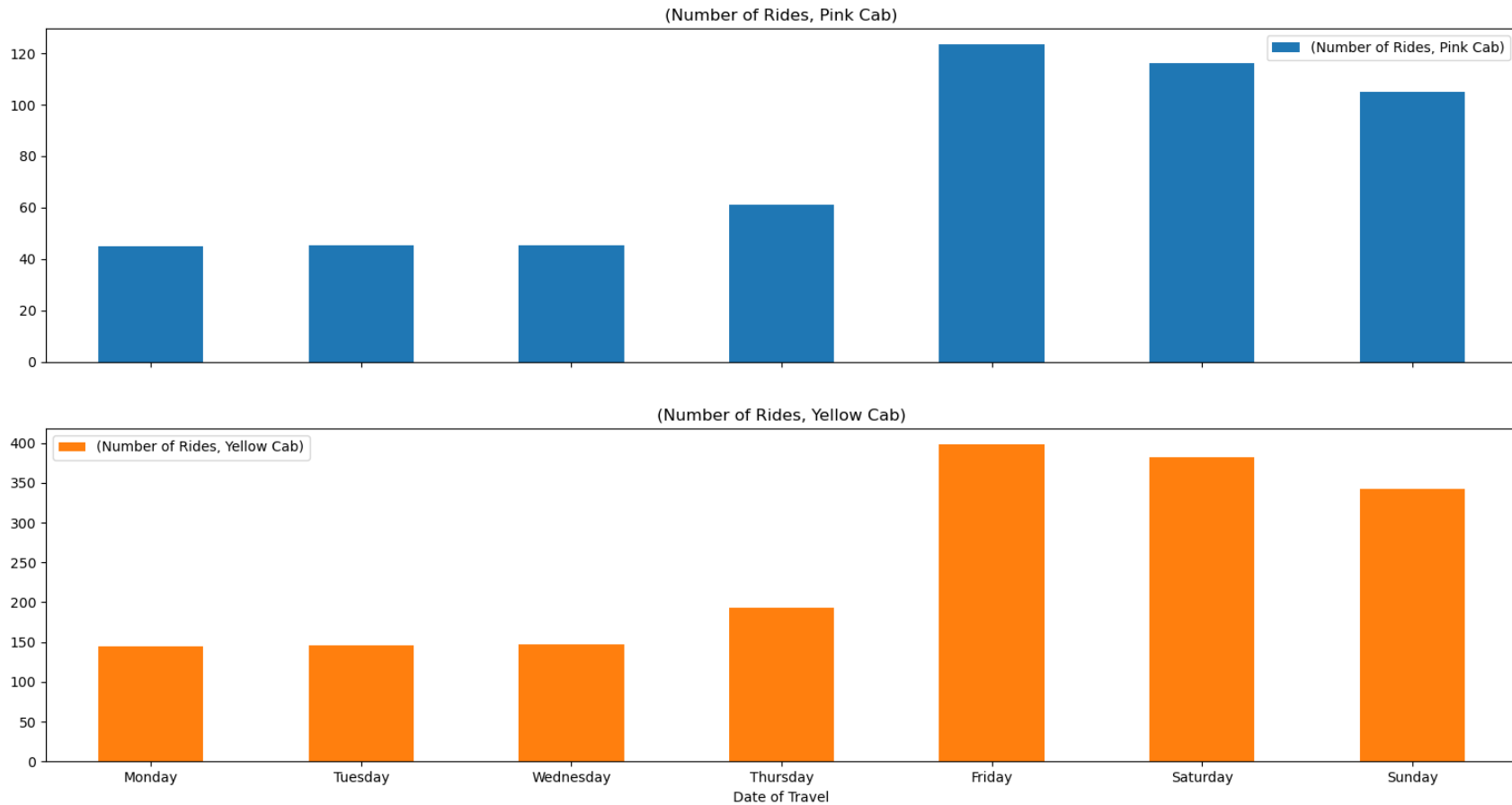
Activity: Monthly

Both companies are most active around late December and least active from January to March
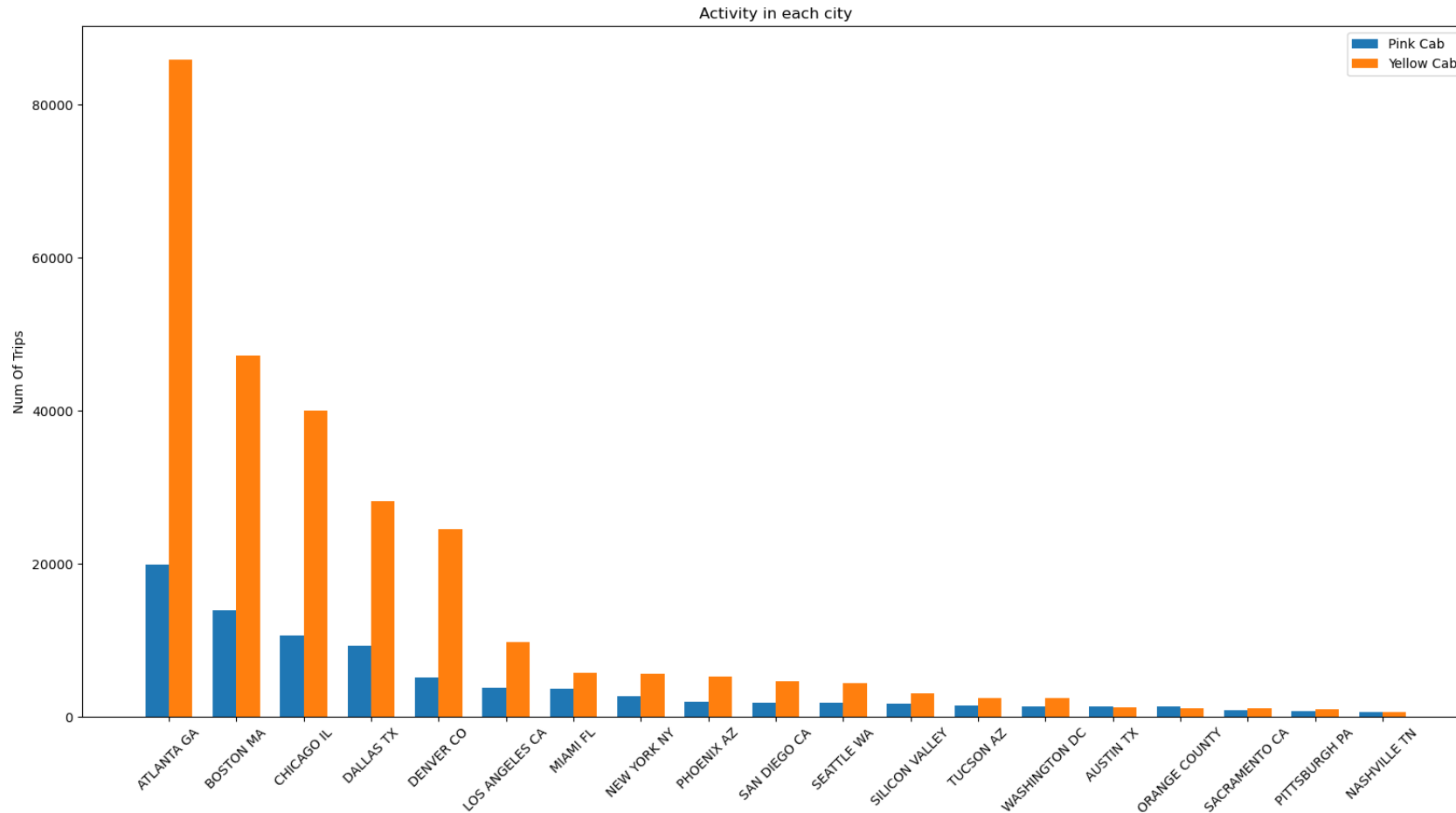

Comapany's Seasonal Activity

Activity: Weekly


Activity is highest on Weekends

Company's reach across the USA

"Yellow Cab" is more active in the majority of the cities

Customer preference and gender

- Rise in total number of customers of "Yellow Cab" after 2016, indicates that customers are switching over from "Pink Cab"
- Both Men and Women are equally using both cab services

Number of Customers Classified by Gender (Bottom: Male, Top: Female)
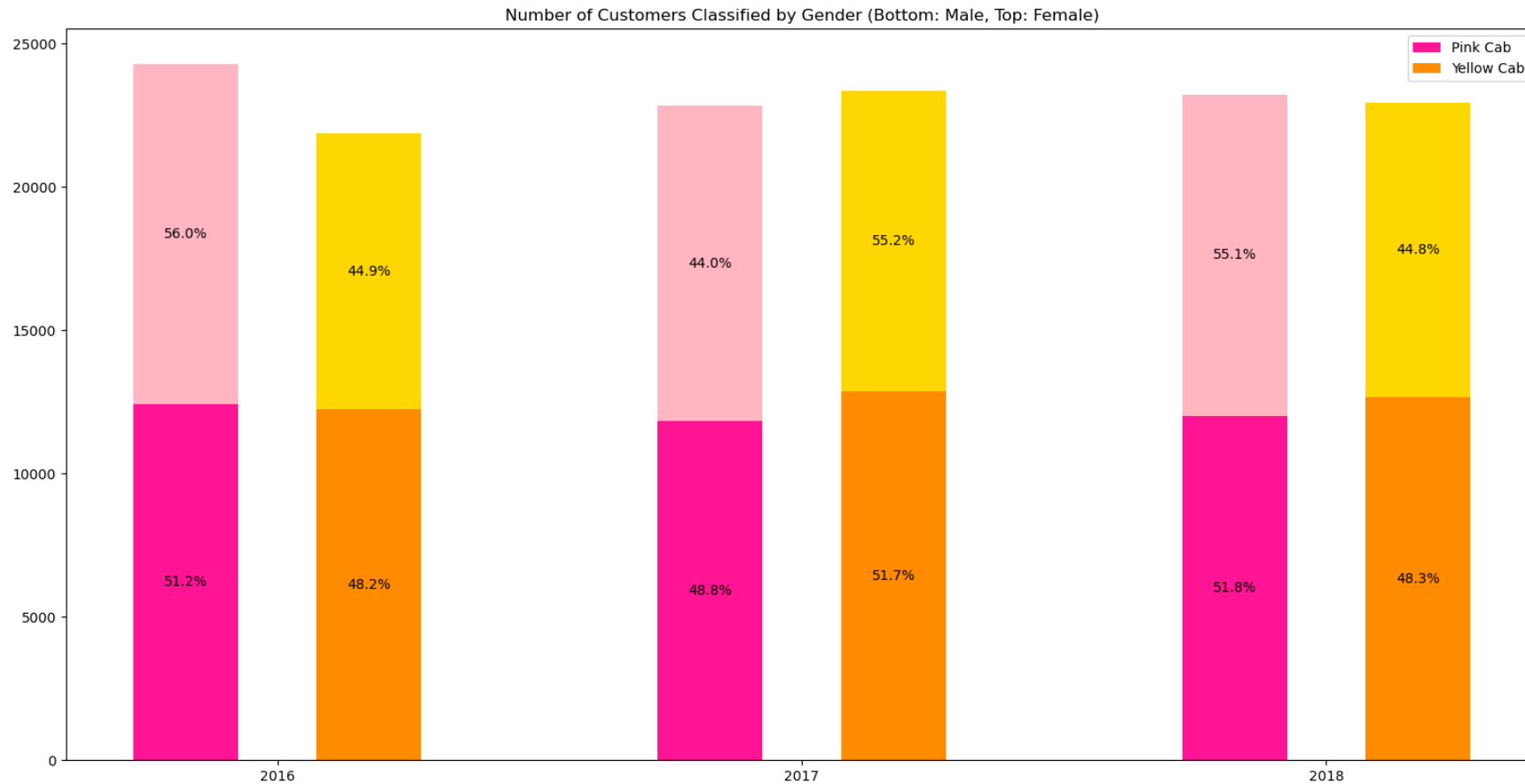
Customer Repeated Purchases

Customers of are more likely to reuse services offered by "Yellow Cab" than by "Pink Cab"
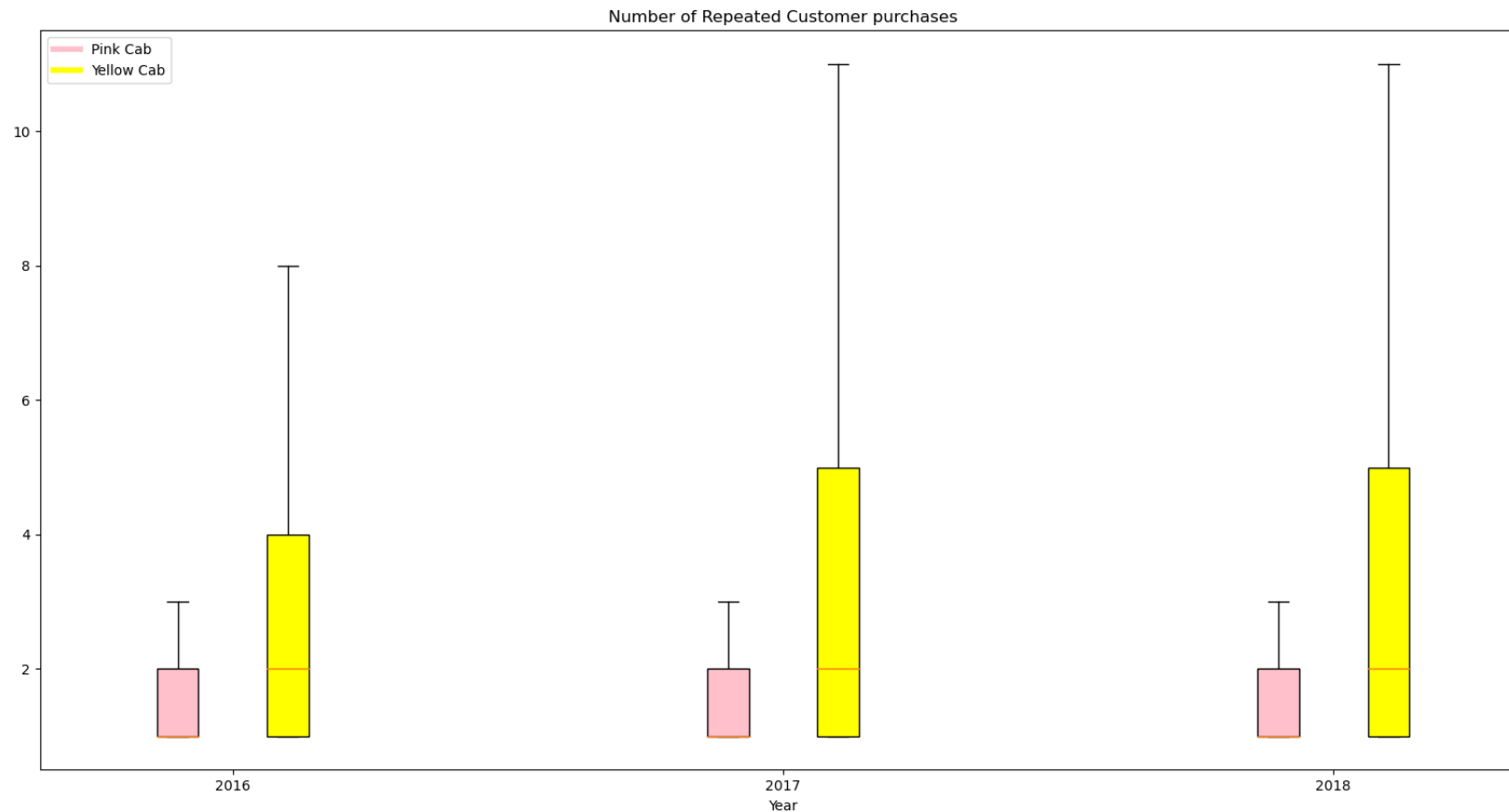
Summary and Recommendations

- *Profit analysis:*
  - "Yellow Cab" achieves higher profits than "Pink Cab"; 4.8 folds larger on average.

- *Activity:*
  - "Yellow Cab" is more active, with a mileage approximately 5 times larger than "Pink Cab", and 330% higher number of trips.

- *Company's reach across the USA:*
  - "Yellow Cab" is more popular in most of the larger cities, especially: Atlanta, Boston, Chicago, Dallas and Denver.

- *Customer basis*
  - No gender preference was observed in the data.
  - Data shows that a number of customers have switched over to using "Yellow Cab" after 2016.
  - Customers using "Yellow Cab" are more likely to use it more often in the same year.

- **Recommendation:**
  - EDA shows that investing in "Yellow Cab" is much more promising than in "Pink Cab"

Week 2: EDA

Thank You