G2M Case Study

Name: XYZ

Location: Zimbabwe

Team: Data and Analytics

Date: 06-03-2021

Agenda

- Background
- Approach
- Data Overview
- EDA Summary
- Recommendations

Background

- A private firm in the US known as XYZ is planning to invest in the Cab industry due to the remarkable growth in the Cab Industry in the past years.
- The firm needs to gain an understanding of the market before making final decisions.
- The Objective of the presentation is to provide actionable insights to help the XYZ firm identify the right company to make their investment.

Approach

For this analysis we will look at four factors:

1. Find the most profitable cab company.
2. Find out which company has the best customer retention and coverage.
3. Find out the customer coverage in each city.
4. Find which company considers all age groups, income classes, genders and all distances

Data Overview

- Five data sets (Cab_data.csv, USHoliday.csv, Customer_ID.csv, Transaction_ID.csv and City.csv) were combined to one dataset.
- The full data set consists of 22 variables (7 derived features)
- The data was collected from 31/01/2016 to 31/12/2018

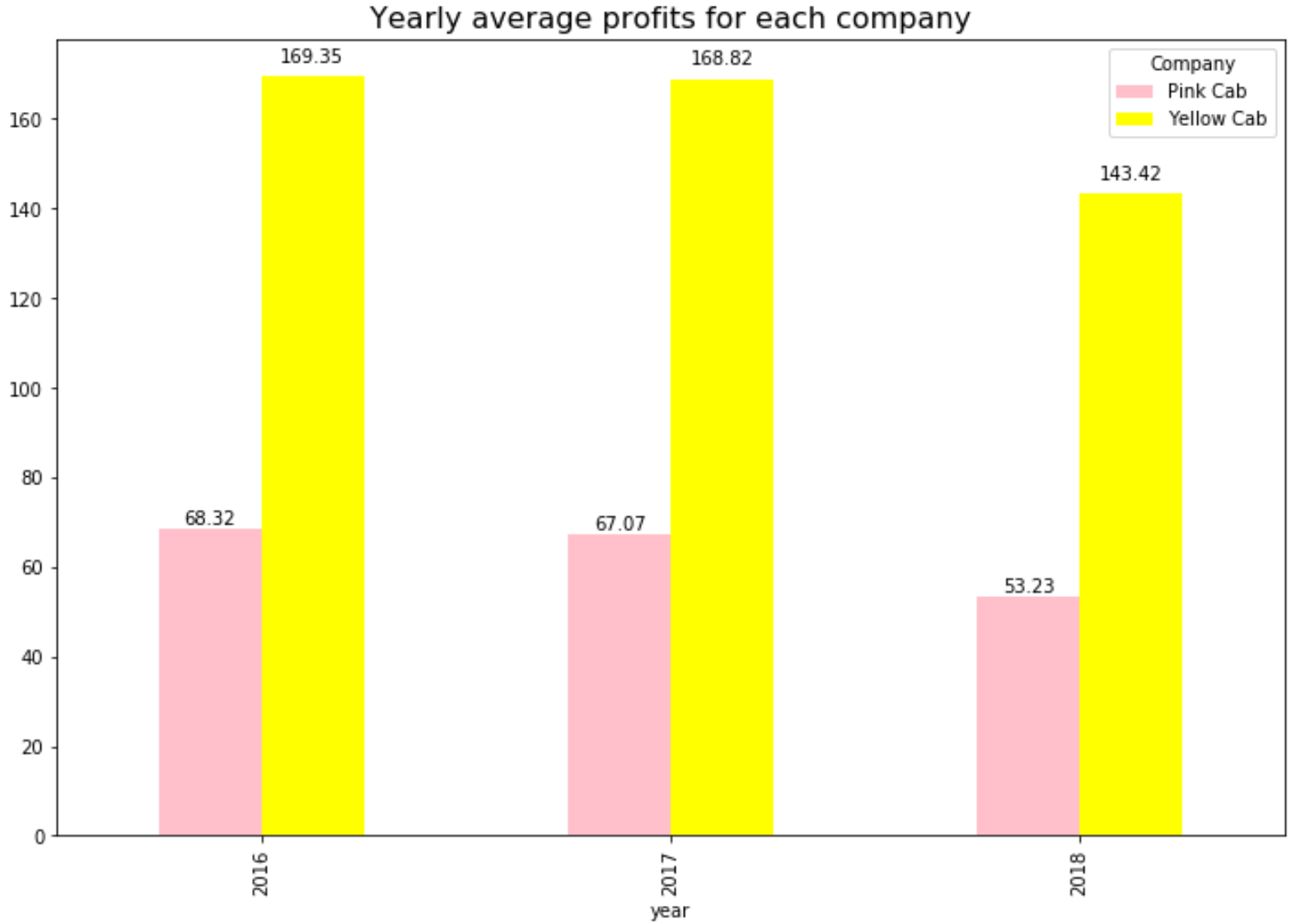
Assumptions:

- The difference between the profit_charged and cost_of_trip can be used to calculate profit.
- The customer IDs are the same for both companies
- The number of cab users is an approximate of the total number of all cab users in the city.

Manipulations:

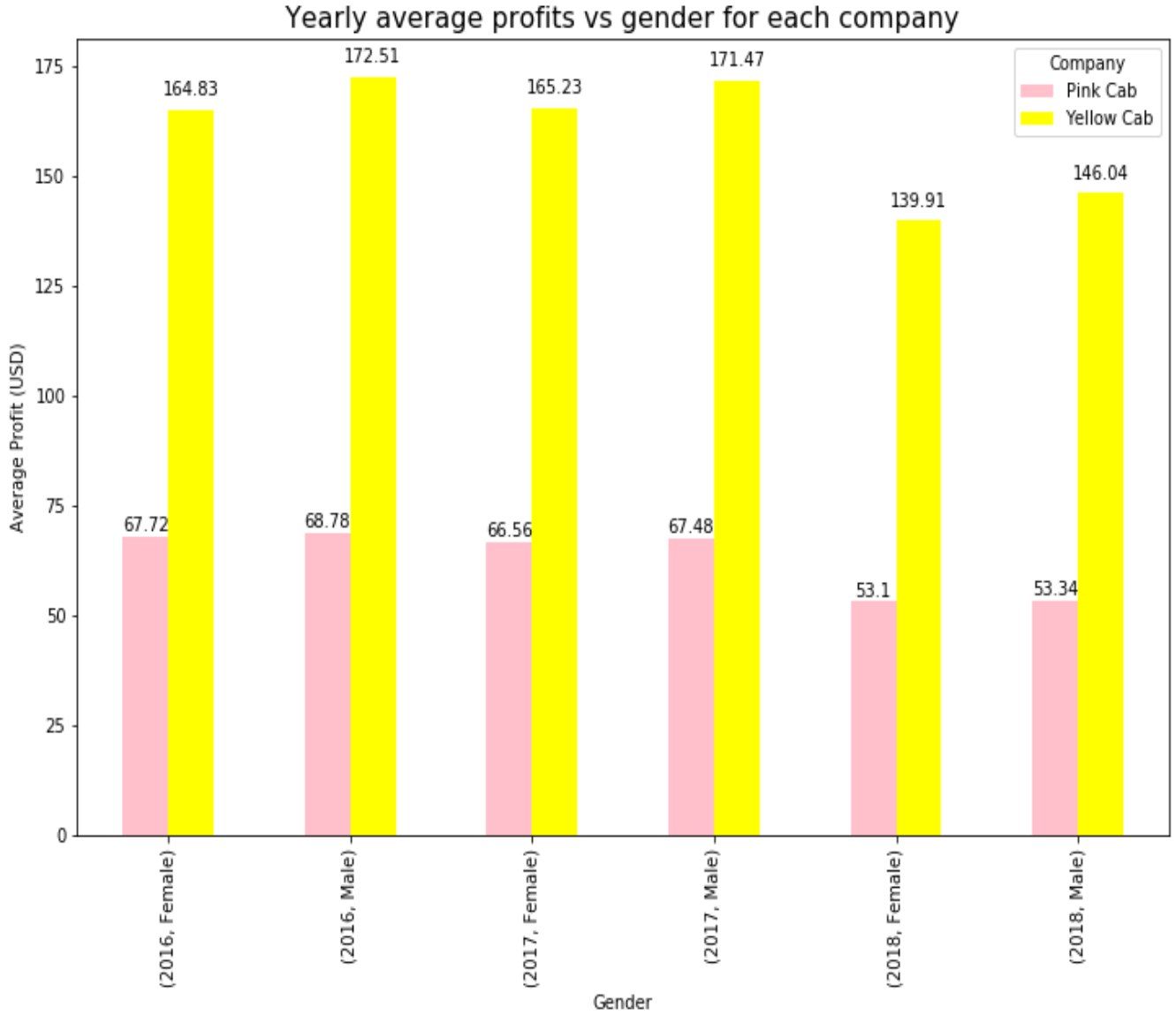
- Created an age group column for age groups (16-24, 25-34, 35-44 and 55+) using the age column.
- Created a holiday column which checks whether it is a holiday or a normal day.
- Created an income class column which uses the income column values to check if a customer is in low-income class or middle-income or high-income class.
    - Low-income class = [income < lower quartile]
    - Middle-income class = [lower quartile < income < mean]
    - High-income class = [income > mean]
- Created distance travelled column which places the distances in km to 3 categories (short [0-10 km], normal [10-30 km] and long [>30 km]).

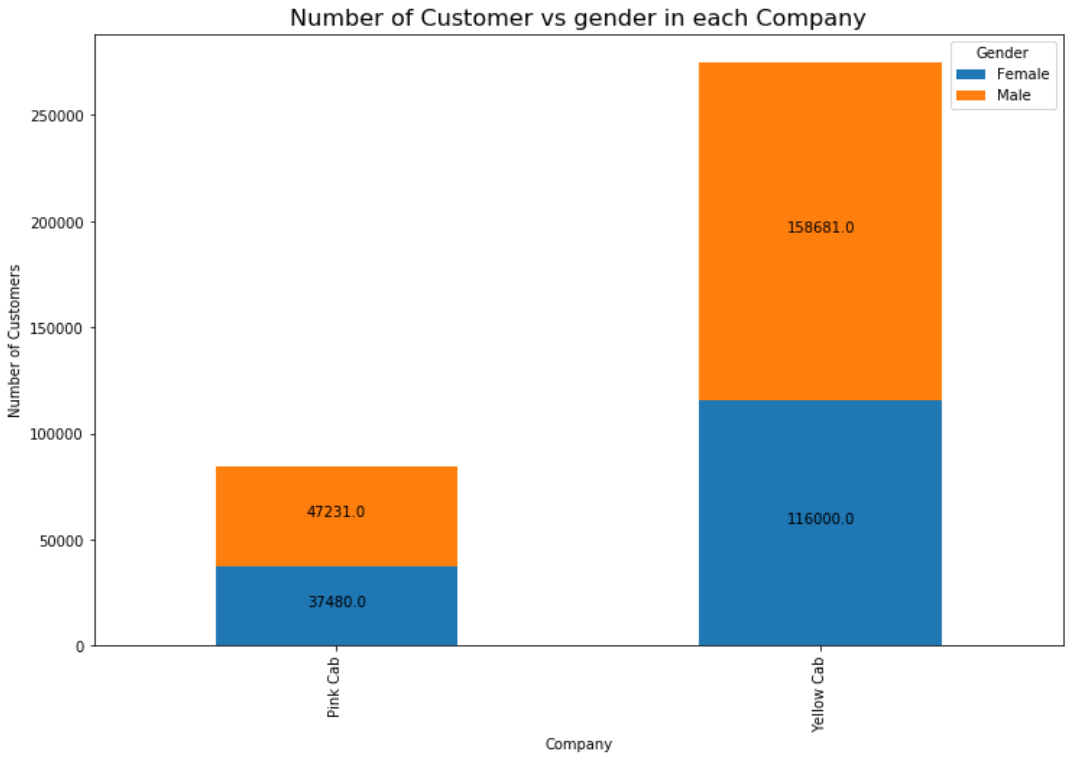
Which Cab has the highest average profit over the years?



- A similar trend over the 3 years is observed. The profits have decreased over the 3 years. Additionally, a serious decrease is observed from year 2017 to 2018.
- Overall, the yellow cab has the most average profit over the 3 years.

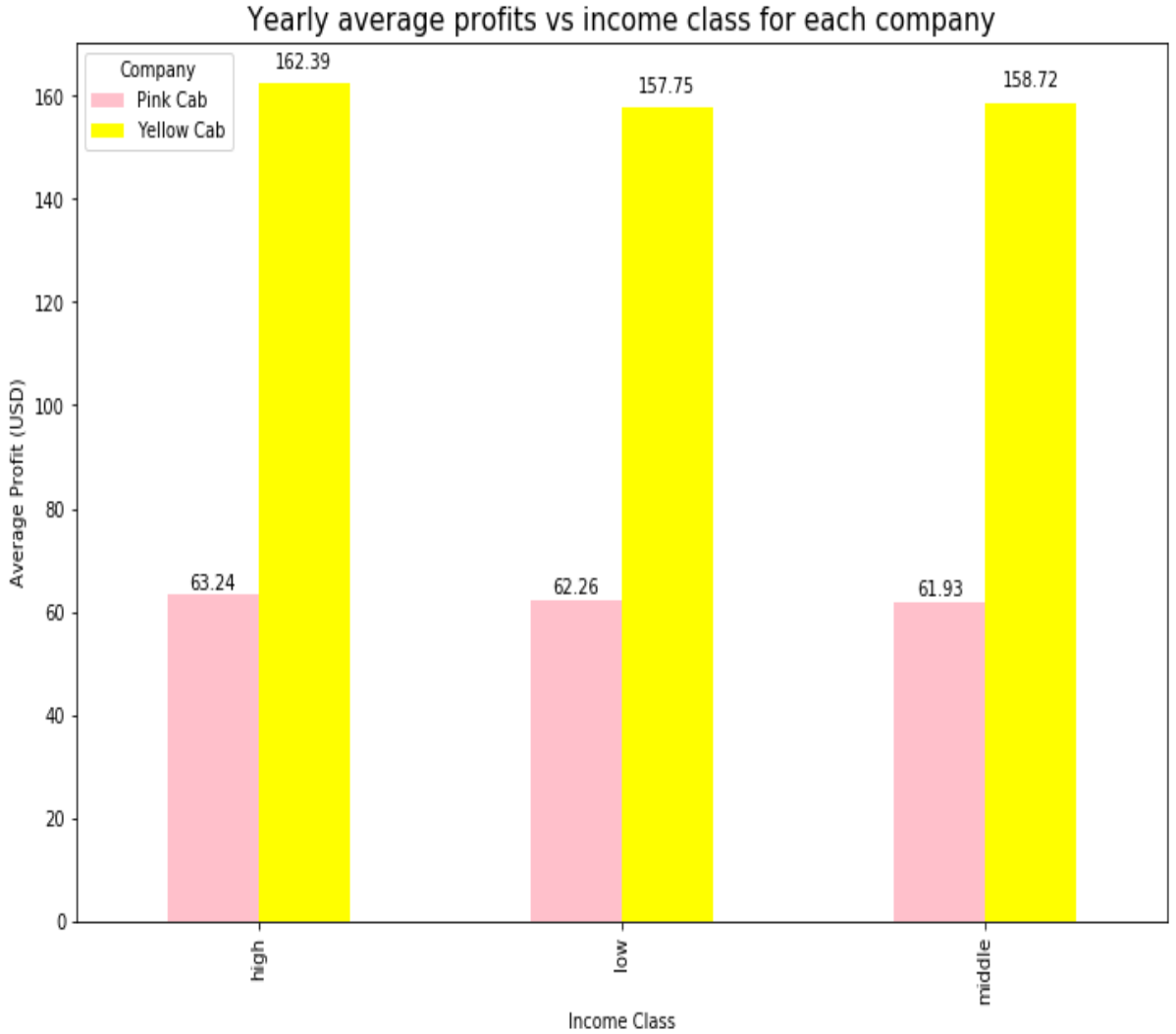
Do males contribute the most towards average yearly profits and customers sizes?

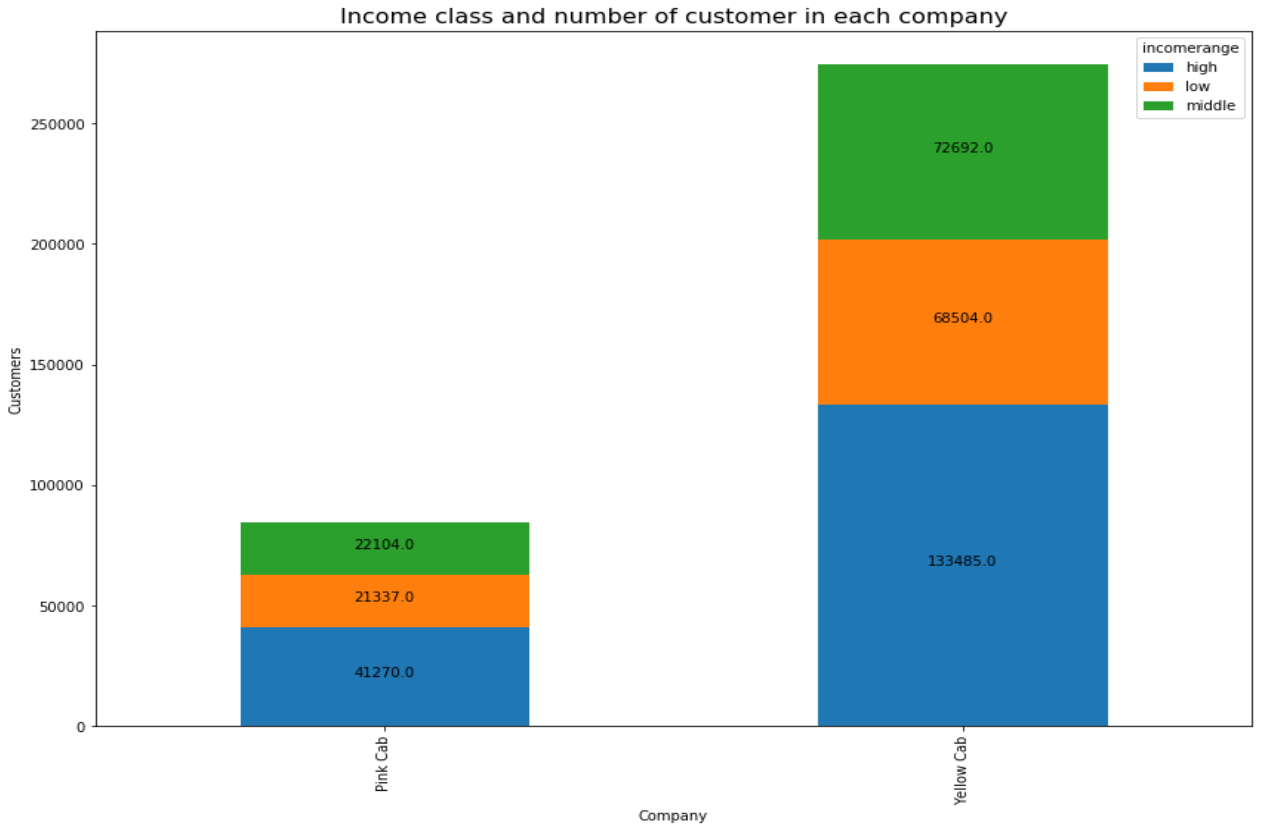




- There is an almost (49%/51%) equal contribution to profits and customer size for both genders for both cab companies.
- Thus, both companies have a balanced gender distribution in terms of profits and customer sizes. Both females have equal contributions towards customer size and profits.

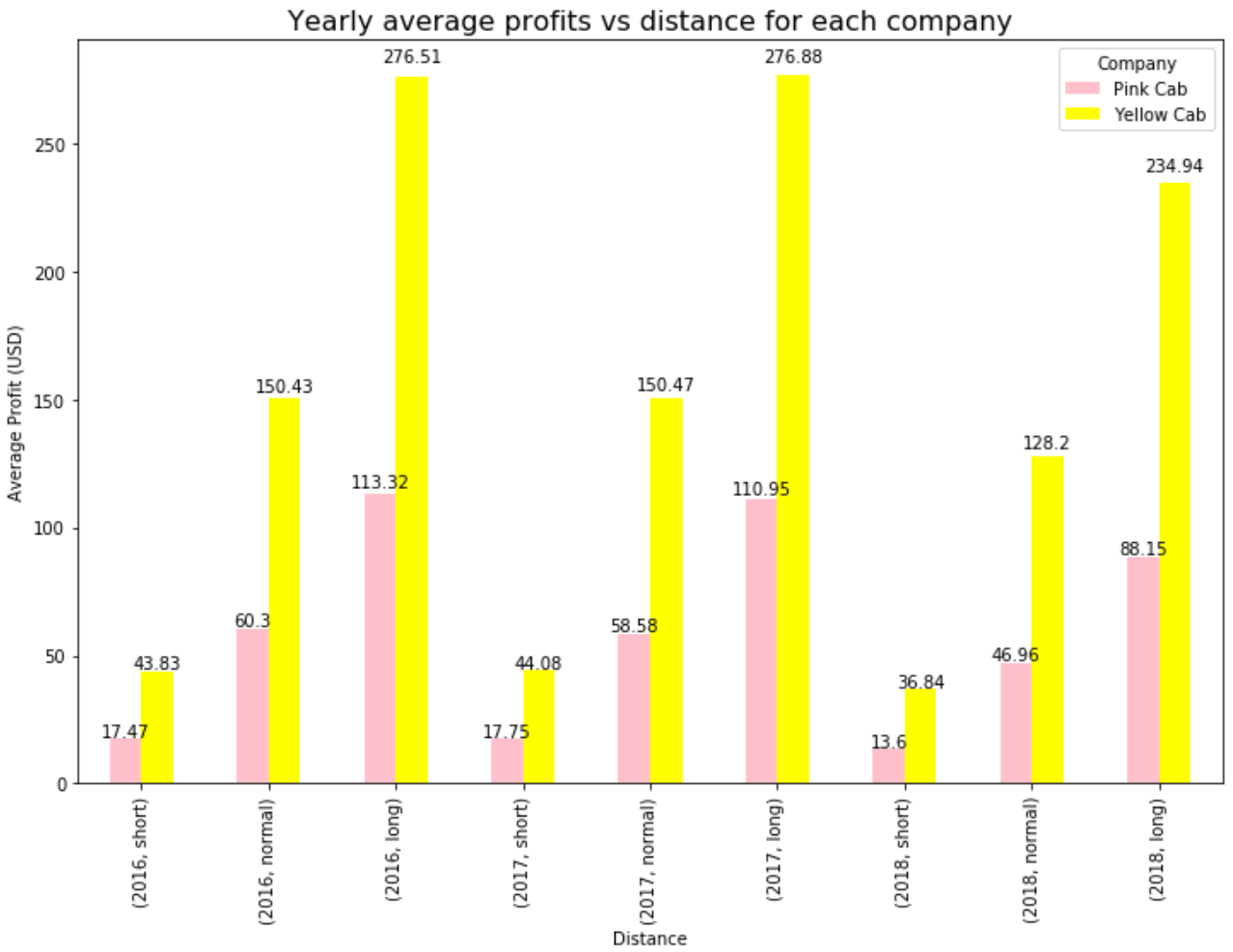
Do all income classes contribute equally towards average profits and customers sizes?

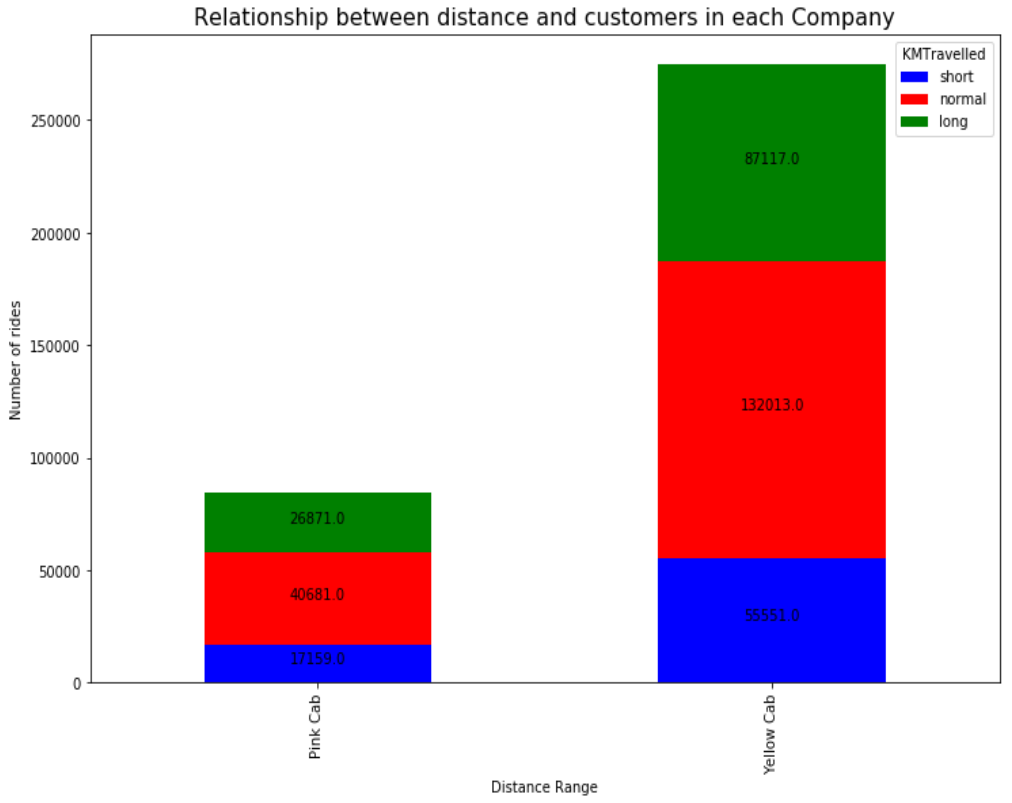




- The high-income class seems to be the highest contributor towards the profits and customer size.
- The other classes have a significant contributions too. However, the difference in number of customers between the 3 classes is too huge for the yellow cab.
- Both cab companies have an “okay” representation of all classes.

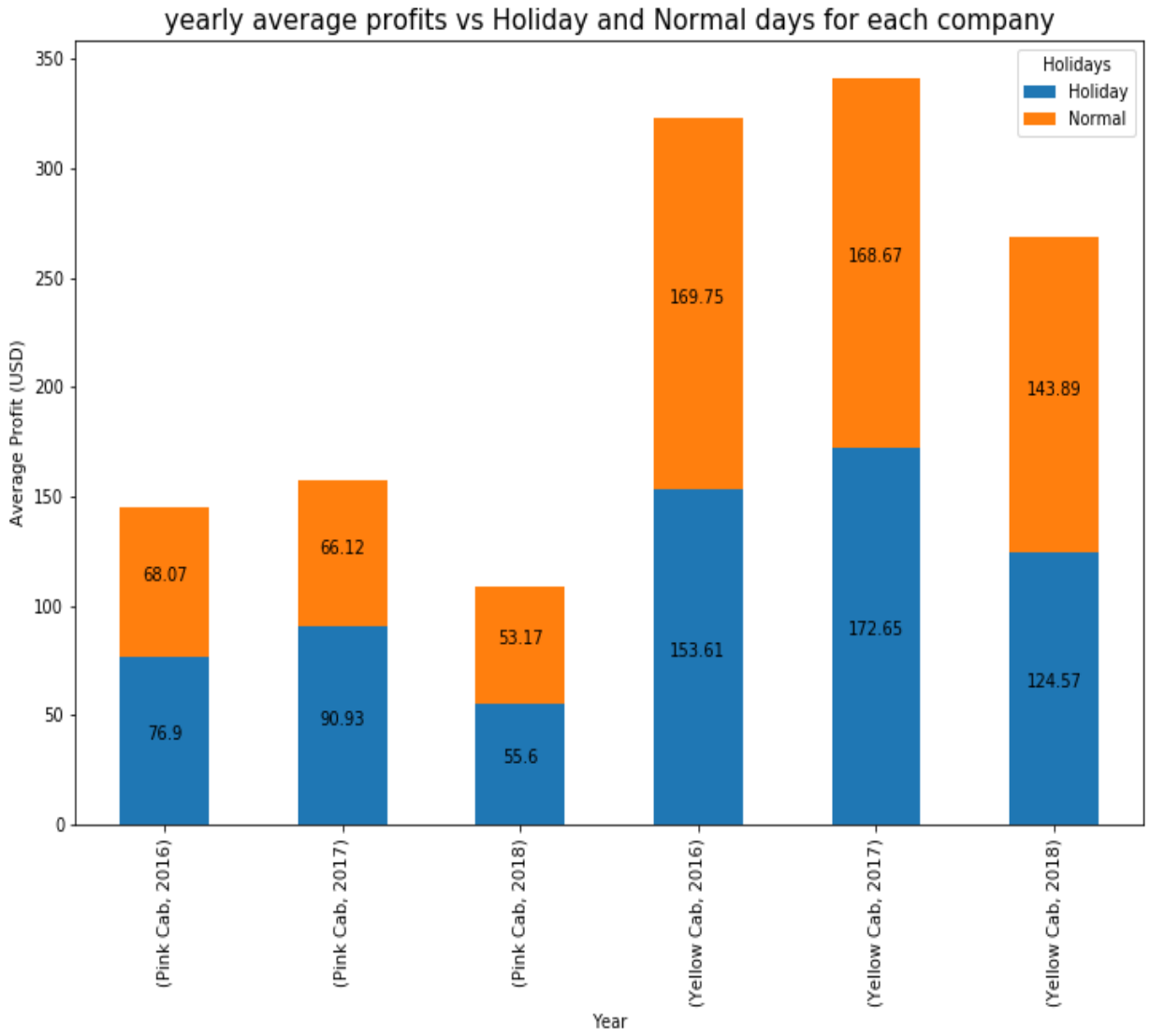
Do long distances contribute the most towards average profits and customer size?





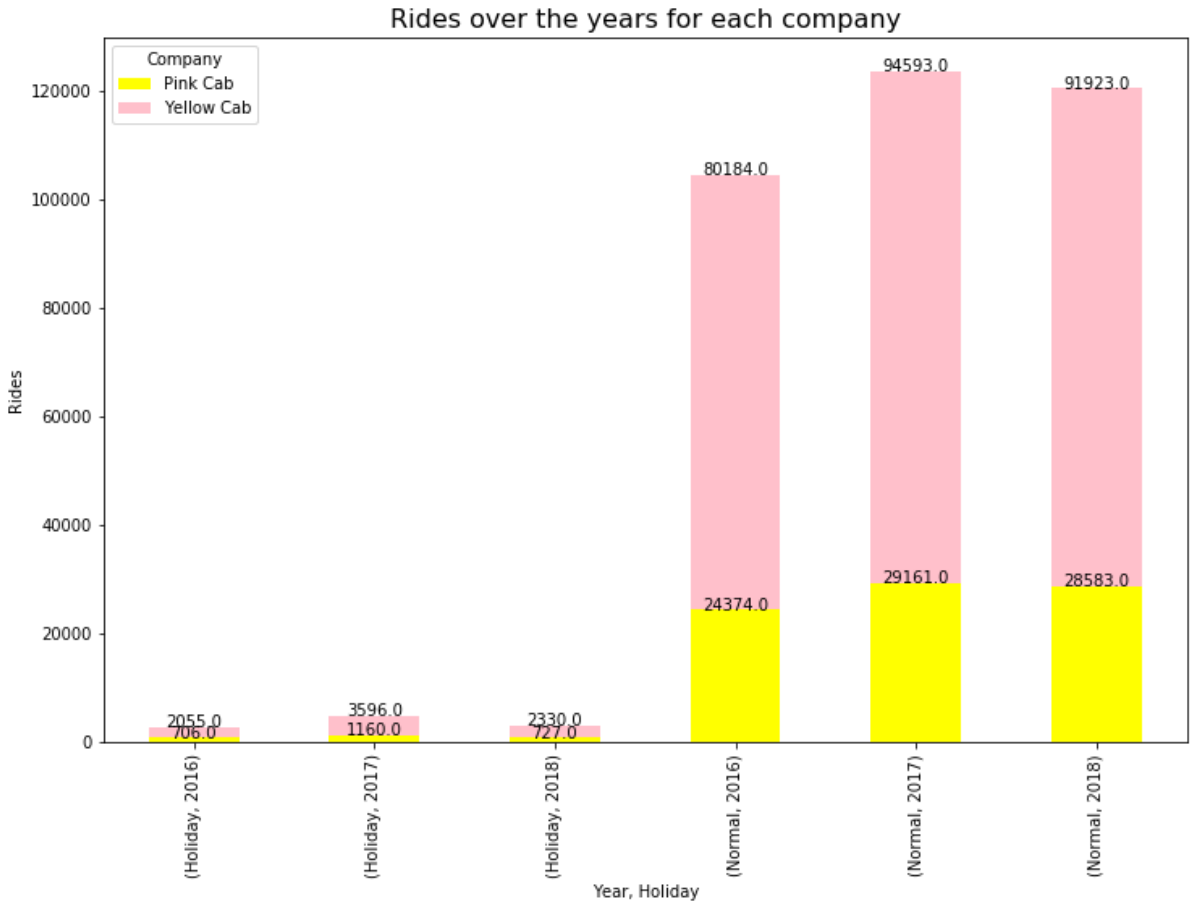
- For both companies, the longer the distance the higher the profit. However, most customers opt for middle distance rides.
- The yellow company seems to have the best distance coverage (covering all distances) than pink company.

Do normal days produce the most profit compared to holidays?



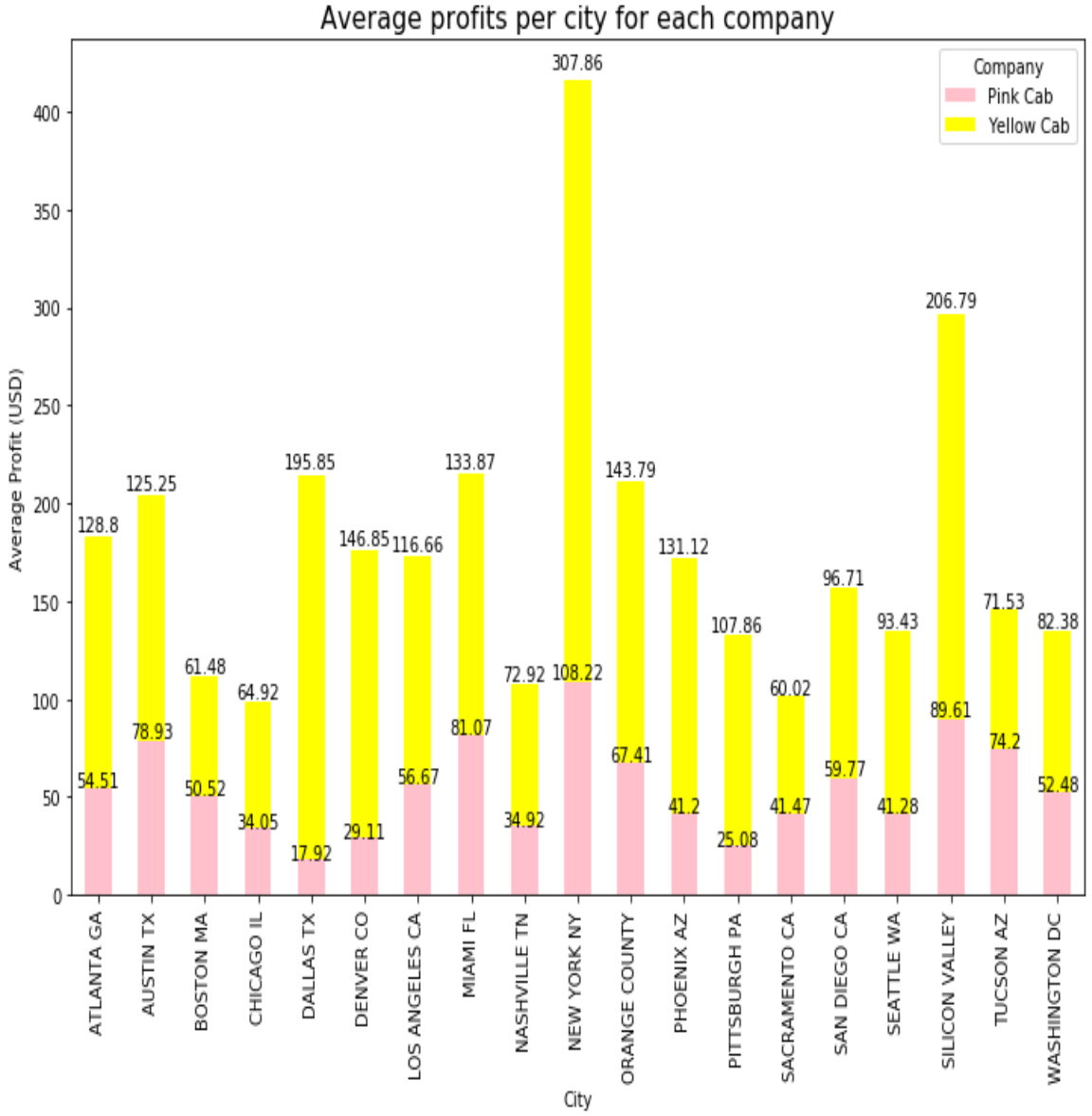
- For the yellow cab, the normal days have the most average profits compared to holidays.
- For the pink cab, holidays have the most average profits compared to normal days.
- The yellow cab company has a good profit balance for both normal days and holidays.

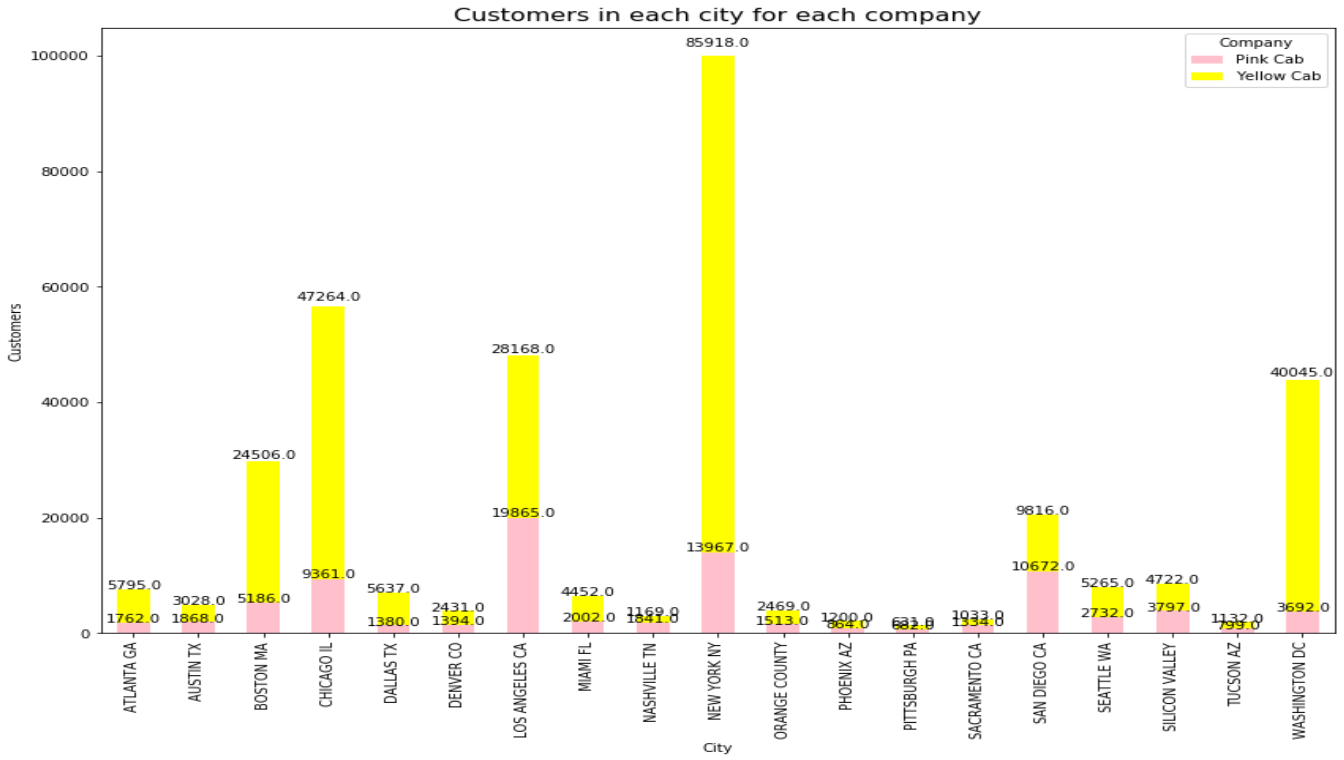
Do normal days have most rides compared to holidays?



- For both companies, the number of rides during normal days is higher than holidays.
- Thus, normal days have the most rides.

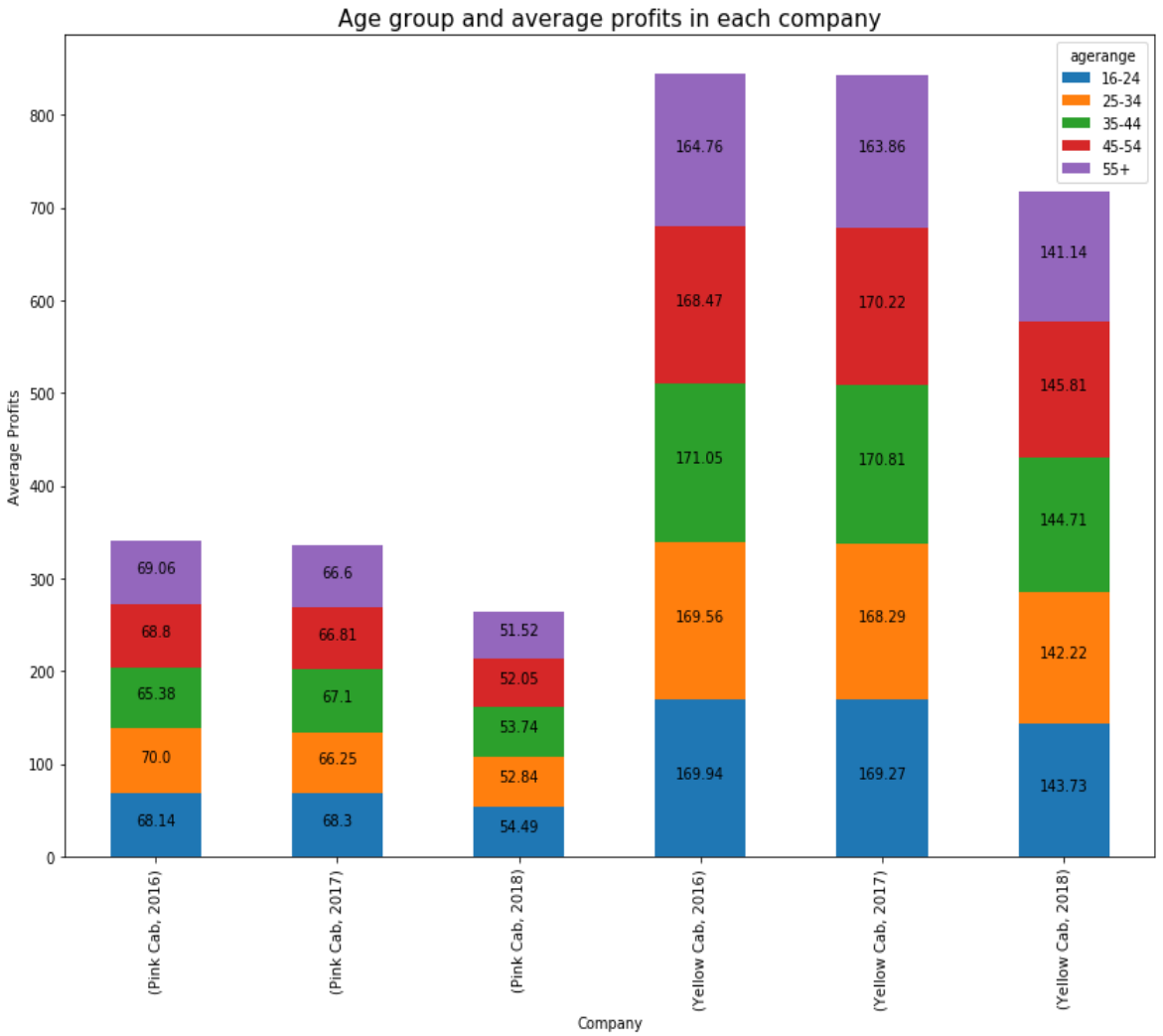
Do biggest and busiest cities contribute the most average profits and customer size?

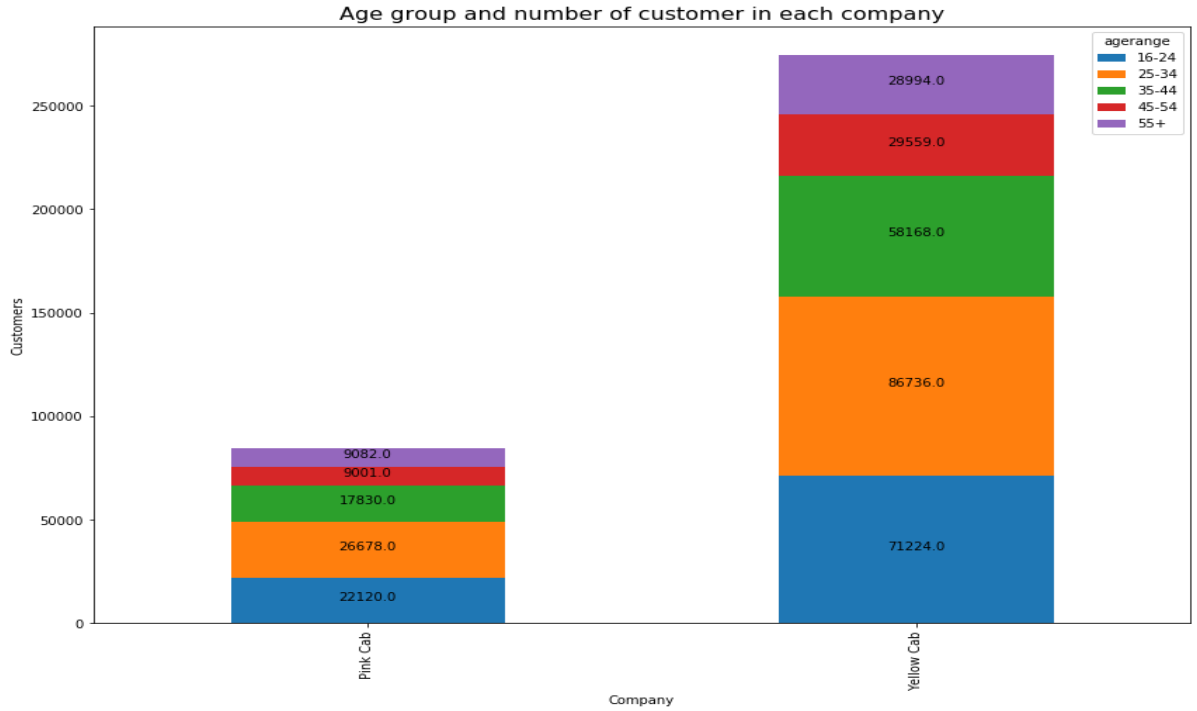




- The biggest and busiest cities like New York and Silicon Valley are the biggest profit contributors for both companies.
- As for customer size, the biggest cities contribute the most for the yellow cab company than the pink cab.
- The hypothesis holds for the yellow cab company. However, both companies are well represented in the US.

The working-class age group contributes the most profits and customer sizes?

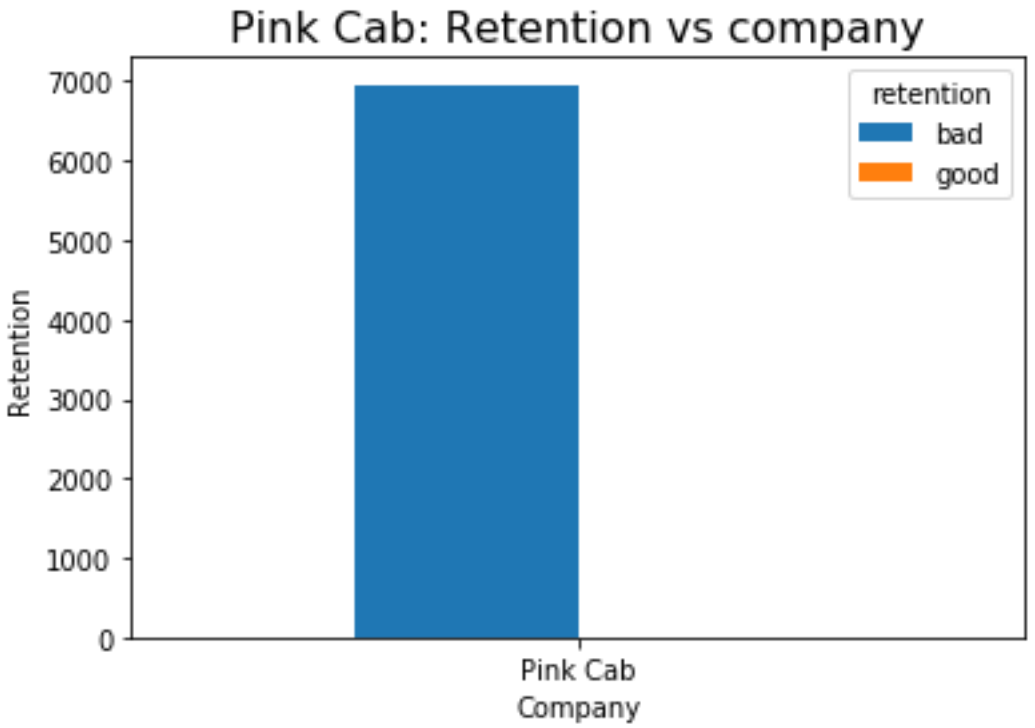


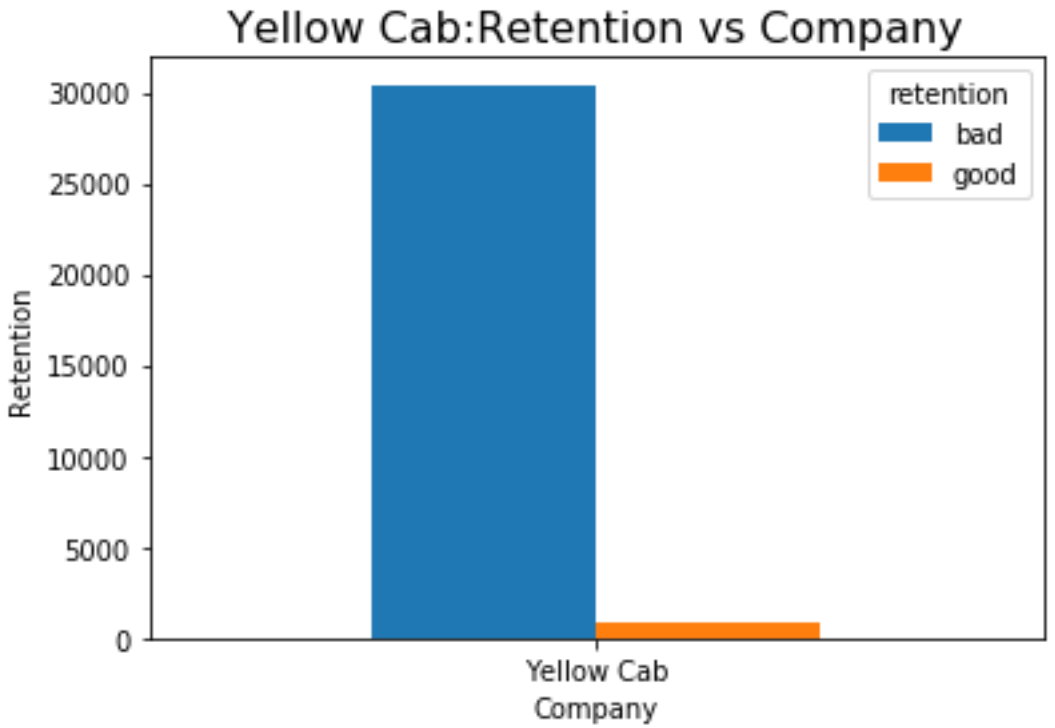


- There is an equal contribution of profits for all the age groups for both companies. The working-class age group is the biggest contributor of customers for both companies.
- However, the yellow cab has better distribution of age groups amongst customers.

Which cab company has the best customer retention?

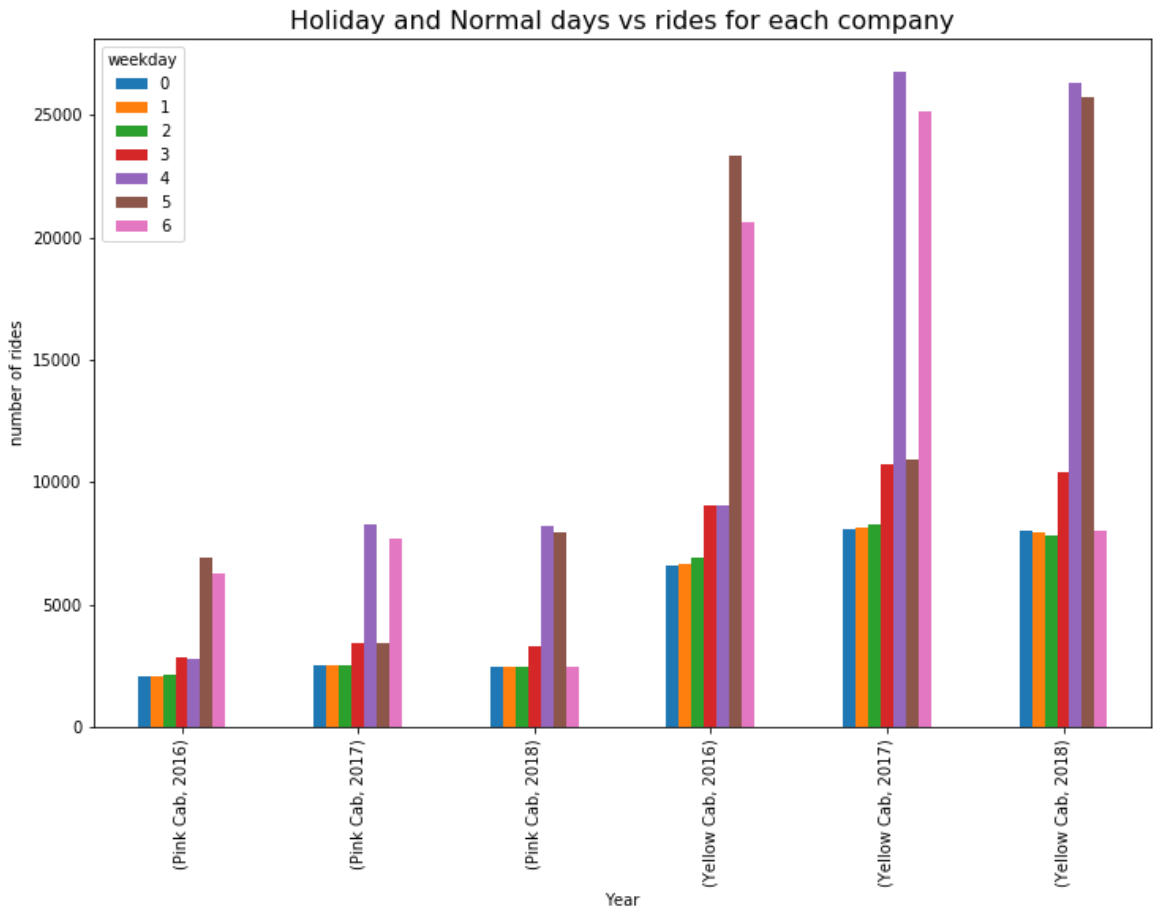
Years	2016	2017	2018
Yellow Cab customers	82239	98189	94253
Pink Cab Customers	25080	30321	29310





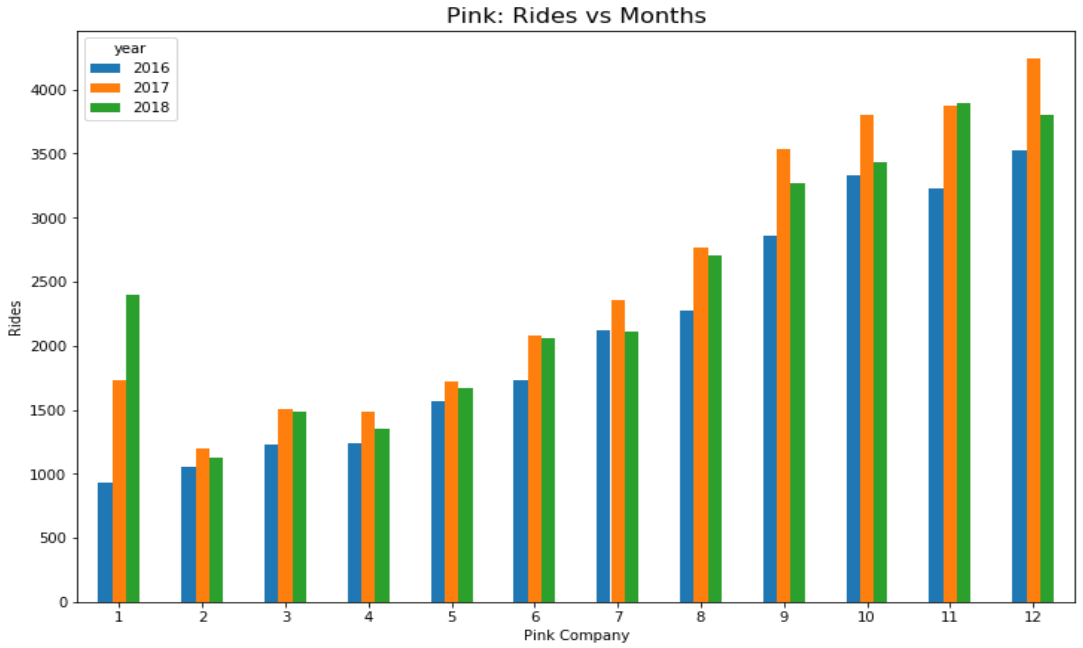
- The Yellow cab company has the best customer retention. On average customers ride 10 rides per year which is not that good. However, it is better than the Pink Cab which on average has customers taking 2 rides per year.

Are week-days the busiest days?



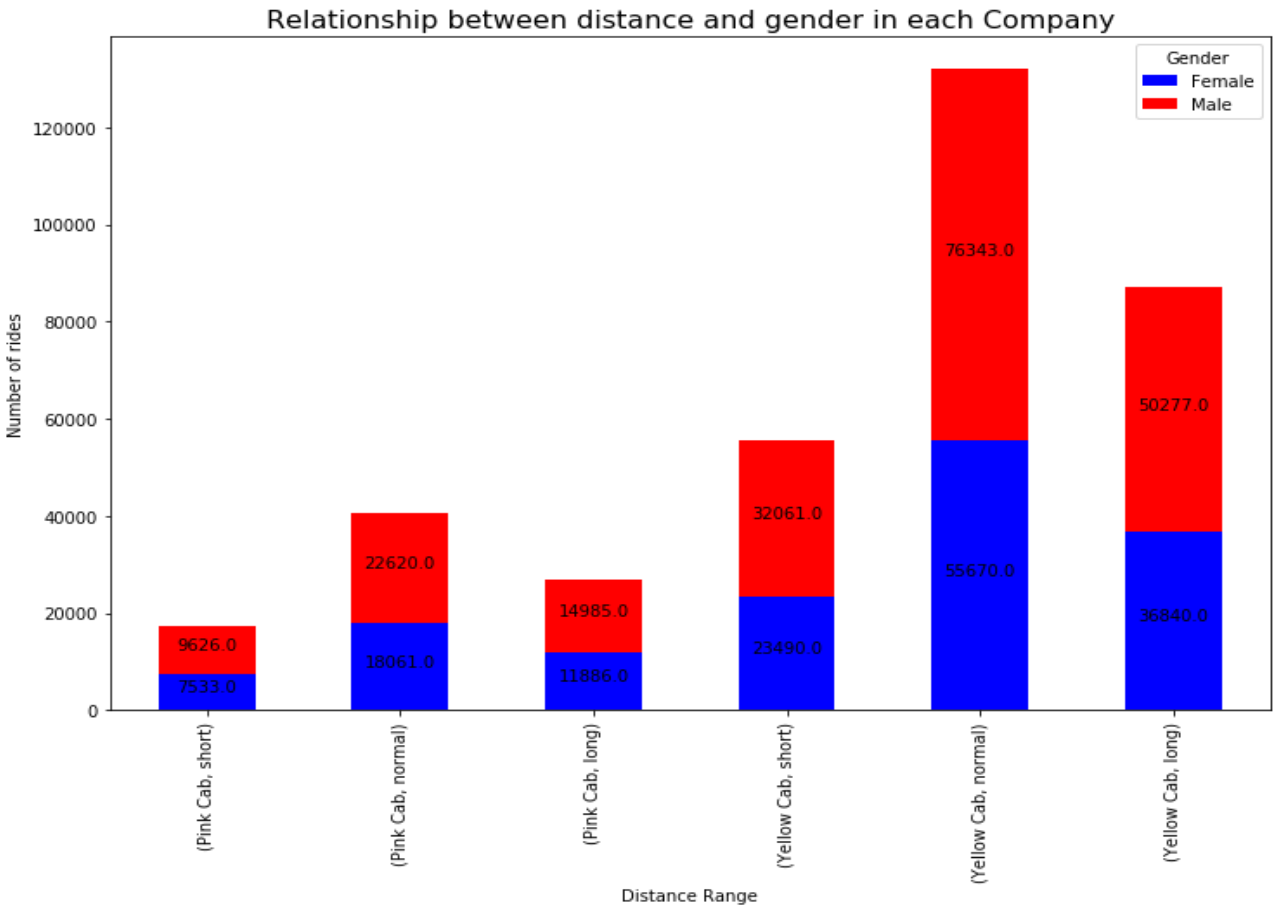
- Fridays to Sundays are the busiest days over the 3 years. There is a similar trend in both companies.
- However, there is a reasonable number of rides over the different days of the week for both companies.

Is there a trend in terms of rides used yearly?



- There is a similar trend in both companies. The busiest months over the 3 years has been October-December.

Is there a relationship between distance and gender?



- Both genders prefer “normal distance” most. Both cabs seem to be safe for both genders because a good number of female customers use both cabs for long distance rides.

Recommendations

Data Glacier
Your Deep Learning Partner

Comparison of both cab companies:

Profit Analysis: The yellow cab has the most consistent average profit gains over the three years. This shows that the Yellow cab is more stable than pink cab. Additionally, the yellow cab company has a good profit balance for normal days and holidays.

Customer retention: The yellow cab customer retention is better than the pink cab. The yellow cab has 10 rides per year for most customers while pink cab has 2 rides per year for most customers.

Customer Coverage: Both companies have a good balance in terms of gender (almost equal number of females and males use the cabs) and income classes. However, the yellow cab company has the best distance coverage compared to the pink company. Additionally, the yellow company has the best age-group coverage. The best cab is one that can be used by every gender, age group and income class.

City Popularity: Both companies are present in most cities in the US. However, the Yellow cab company is popular in most of the biggest and busiest cities in the US (it has more than 4000 customers for 10 cities) and the pink cab company has more than 4000 customers for 5 cities.

Safety: Both cab companies are safe for both genders and all ages.

The yellow cab has the best potential to grow. It is most stable and has the best customer coverage and retention.

Based on the analysis, Yellow cab is the best for the investment.

Thank You