G2M Case Study

Name: Laith Adi
Location: Toronto, ON
Team: Data Science
Date: 11-March-2021

Agenda

Brief Background

Introduction to The Data

EDA

Recommendations

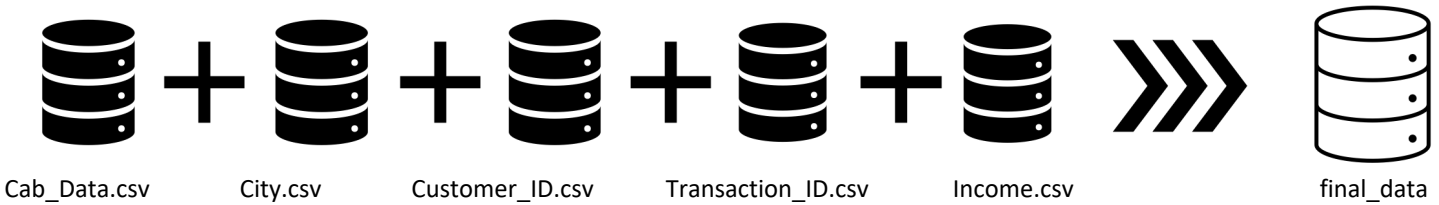
Brief background

- Our objective is to analyze the data about the cab industry and provide actionable insights to you, the XYZ Executive team.
- We are looking at at two companies:
  1. Pink Cab Company
  2. Yellow Cab Company
- The layout of this presentation will be in the following order…

Intro. to data → Analyzing the data using visuals → Final recommendations

Introduction to the data

- 359,392 instances
- 16 features



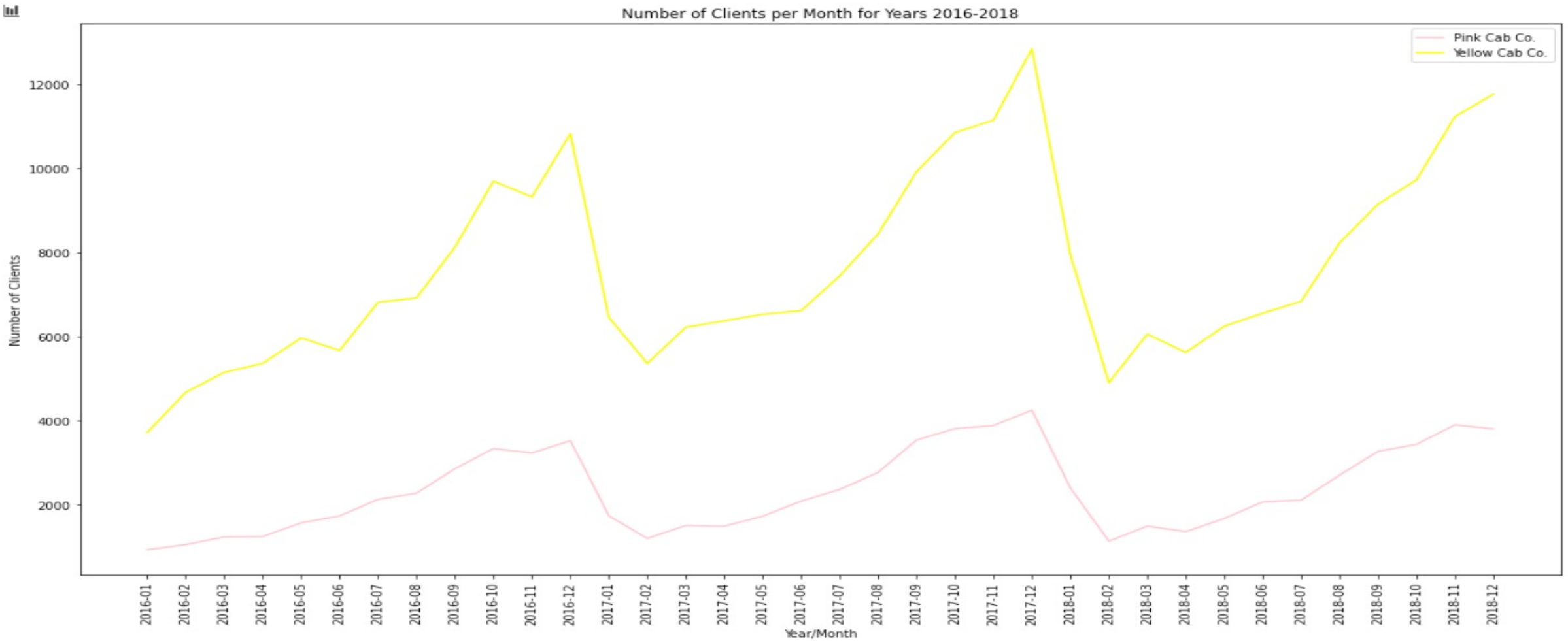
- Assumptions about the data:
  - Income.csv data is from 2012-2016. Here we will assume the same numbers are similar to the years 2017-2018.
  - Profit of rides are calculated using each respective instance by subtracting the cost of the trip by the price charged. We are not taking consideration of other features.

Analyzing the data using visuals

- In order, we will look at the following:
  1. Number of rides each company gave every month.
  2. Characteristics of the customers. Sex and age.
  3. Table of prices, costs, and profits. Prices charged vs. costs of trips. Compare yearly profits.
  4. Relationship and correlation between distance travelled and prices, costs, profits.
  5. Customers income vs. the average income in their city. For the top 3 cities for each company.
  6. Users in city vs. actual customers.
  7. Payment methods.
  8. Number of repeated customers.

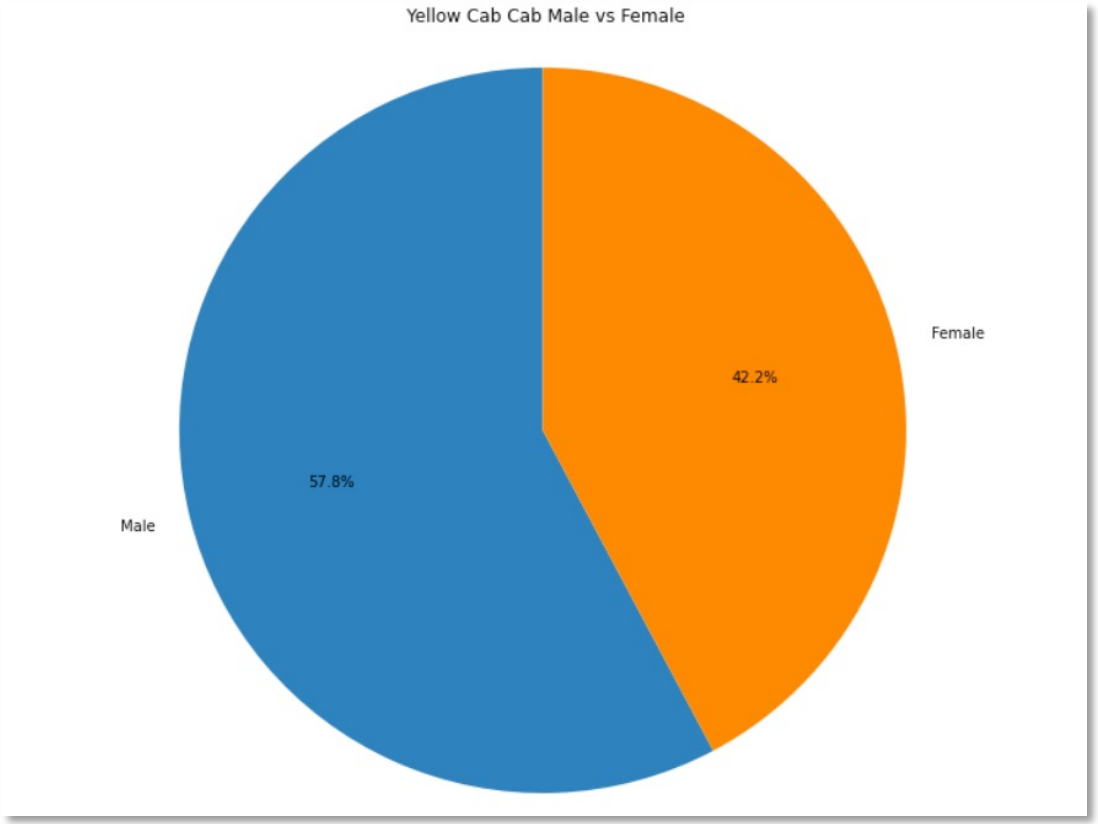
1

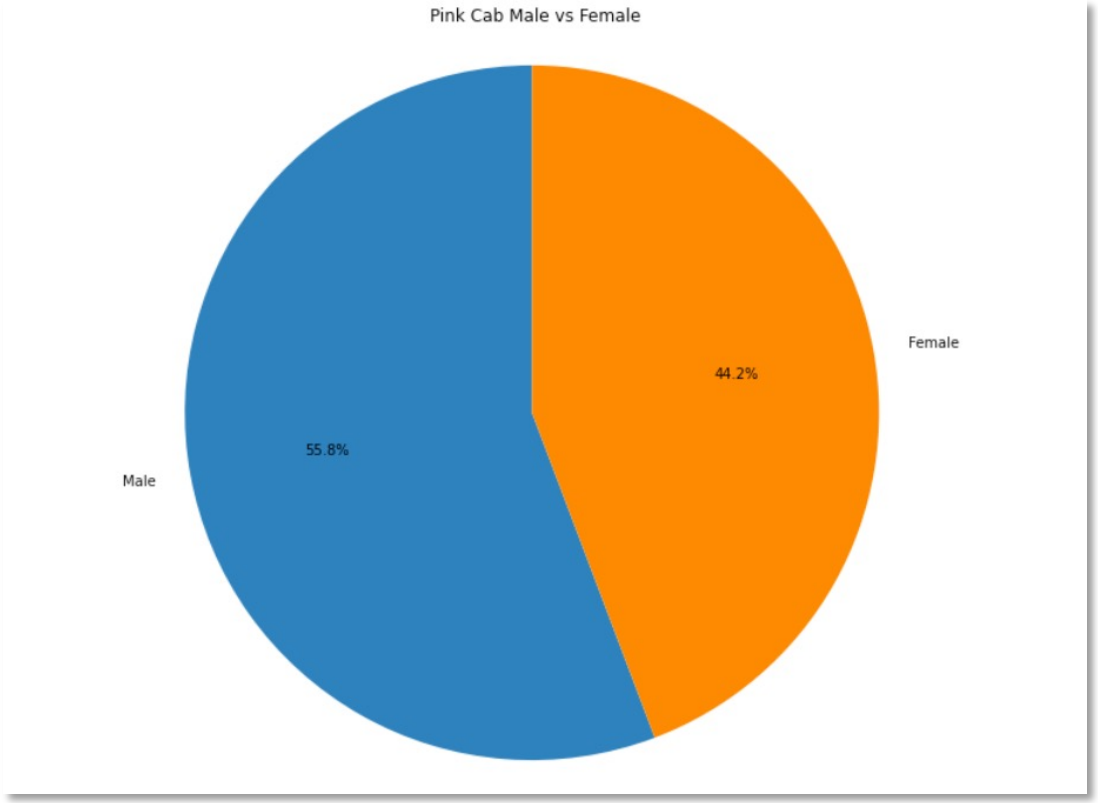
Here we see the number of customers per month, each company had. It is clear there are similarities in trends when each company either loses or gains customers. Yellow does provide its services more frequently. Which does not mean that the Yellow Cab company has more customers. We will get to that later at slide 18. From this graph we can safely infer that any increase or decrease in customers will likely apply to both cab companies.



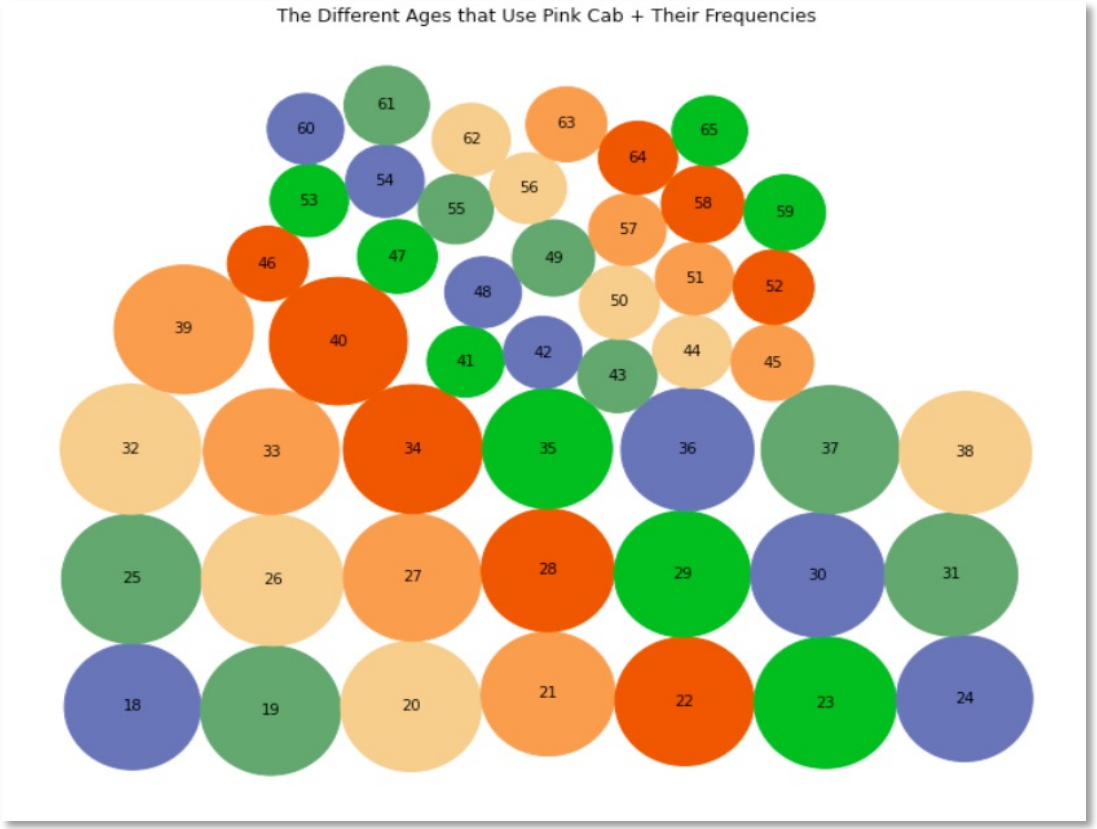
2

These are two pie charts showing the percentage of male and female customers in each company.





Here is the distribution of ages that use each cab company. The larger the amount of age group, the bigger and lower the bubble is location in the chart.



3

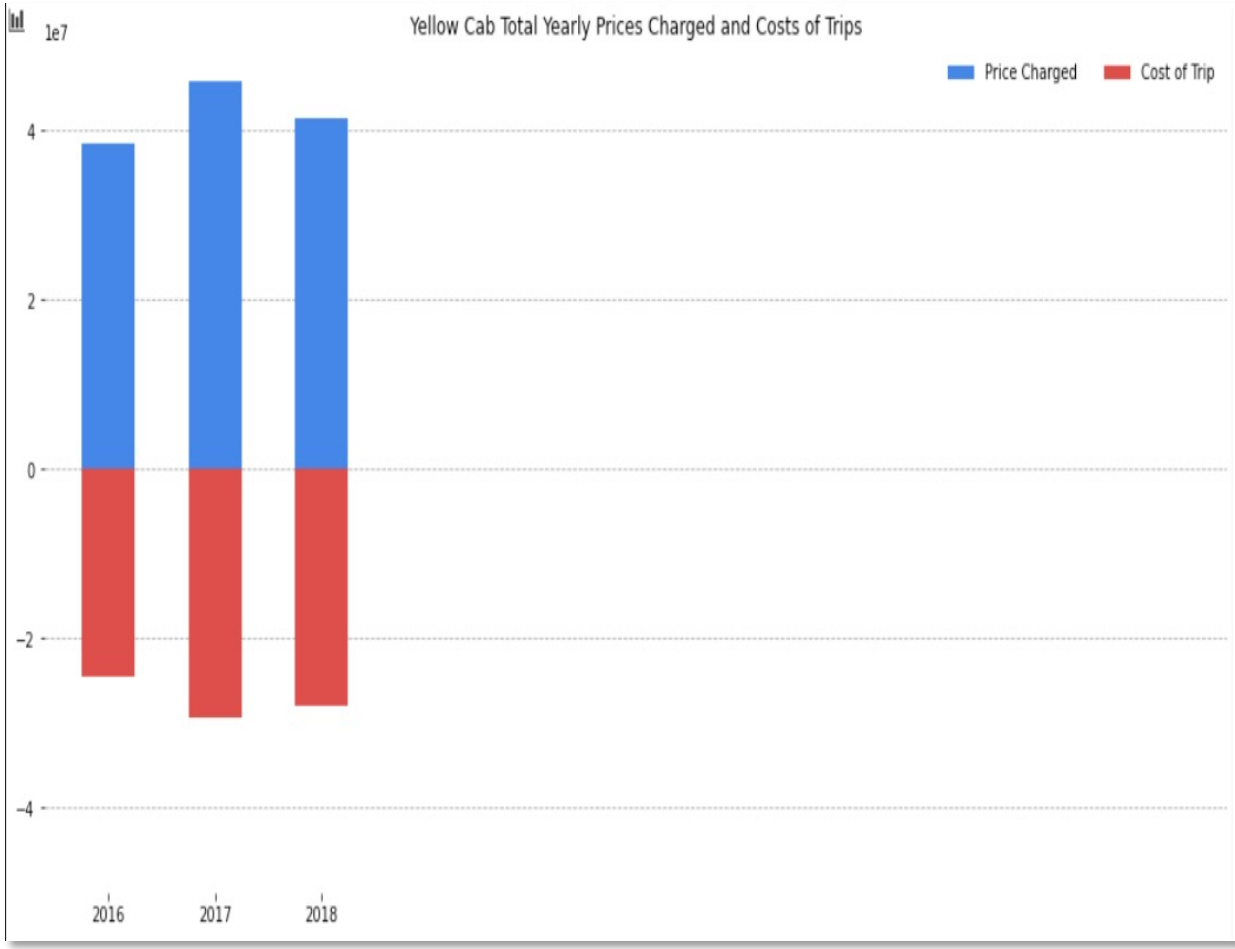
We can see that the Yellow Cab company yearly profits are significantly greater than the Pink Cab company. The costs for the Pink Cab are increasing at a slower amount but the Yellow Cab is making up for it by increasing the total yearly prices charged faster than the Pink Cab. Yearly profits seem to increase then decrease at a similar rate. Yet the Yellow Cab company did not suffer as much losses from 2017 to 2018 as the Pink Cab company.

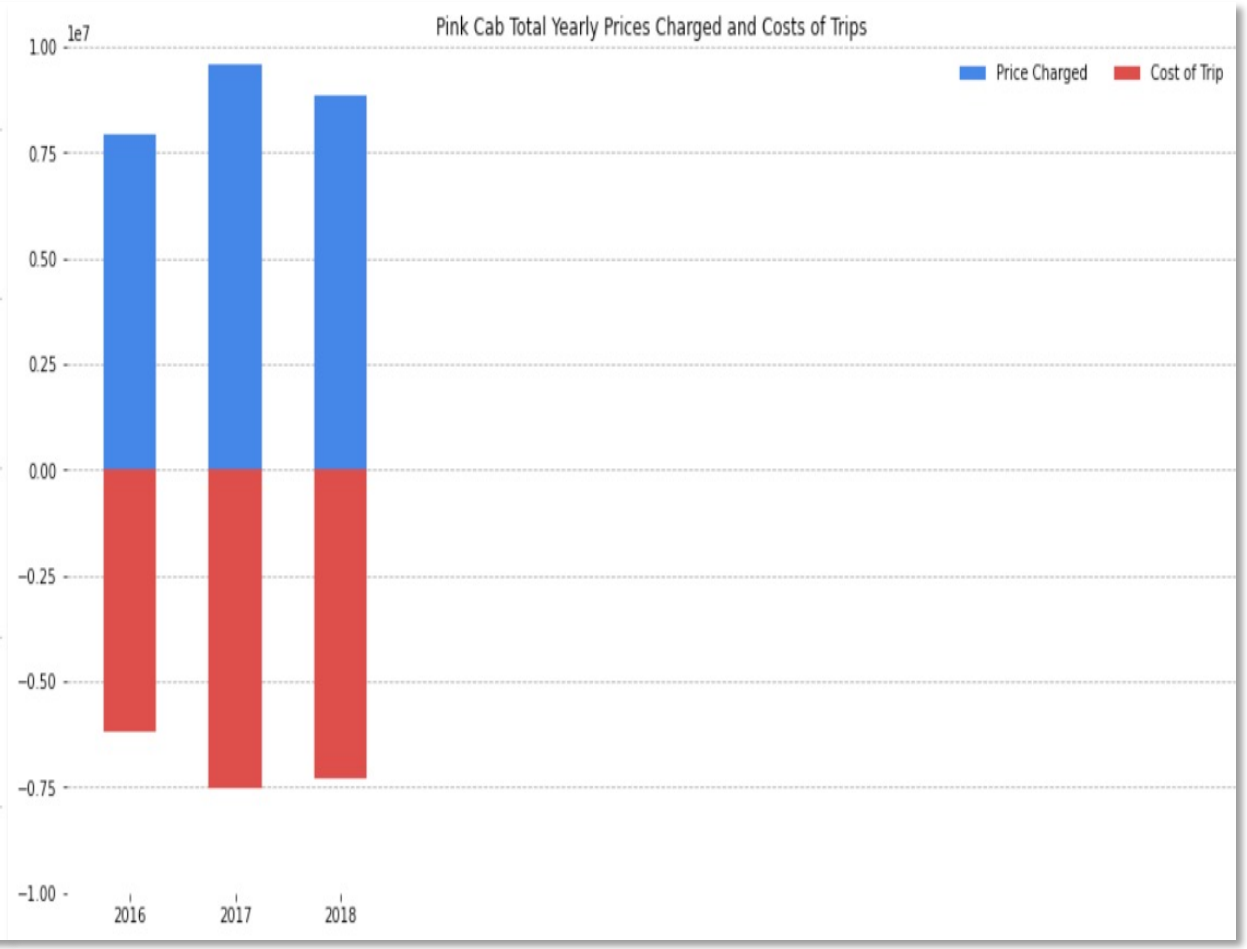
Pink Cab

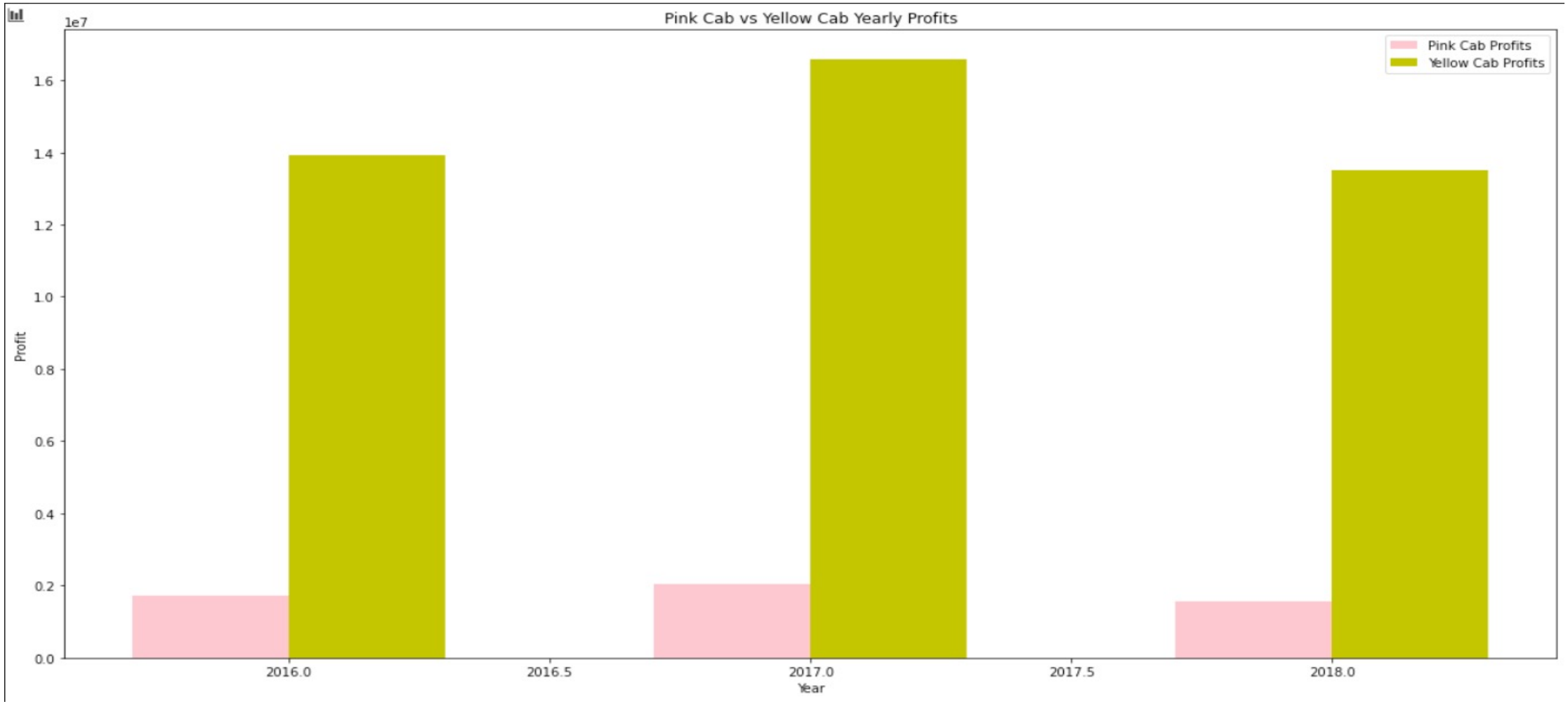
	Date of Travel	Price Charged	Cost of Trip	Yearly Profits
0	2016	7908479.23	6194968.01	1713511.22
1	2017	9578629.54	7544974.63	2033654.91
2	2018	8841142.56	7280980.37	1560162.19

Yellow Cab

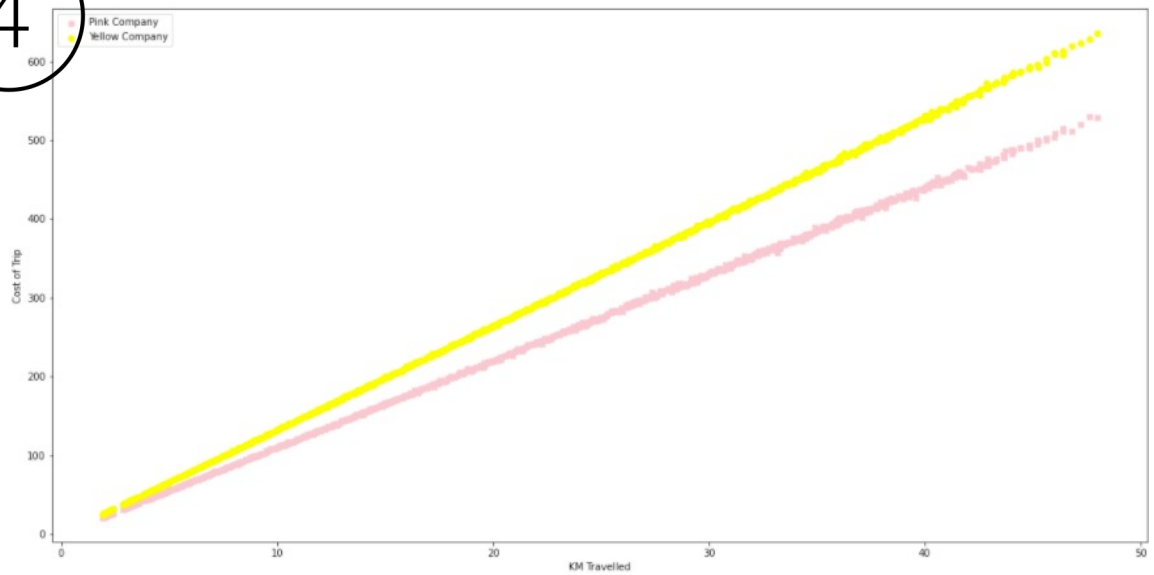
	Date of Travel	Price Charged	Cost of Trip	Yearly Profits
0	2016	38481133.18	24554137.75	13926995.43
1	2017	45818910.04	29242932.07	16575977.97
2	2018	41553843.97	28036444.20	13517399.77

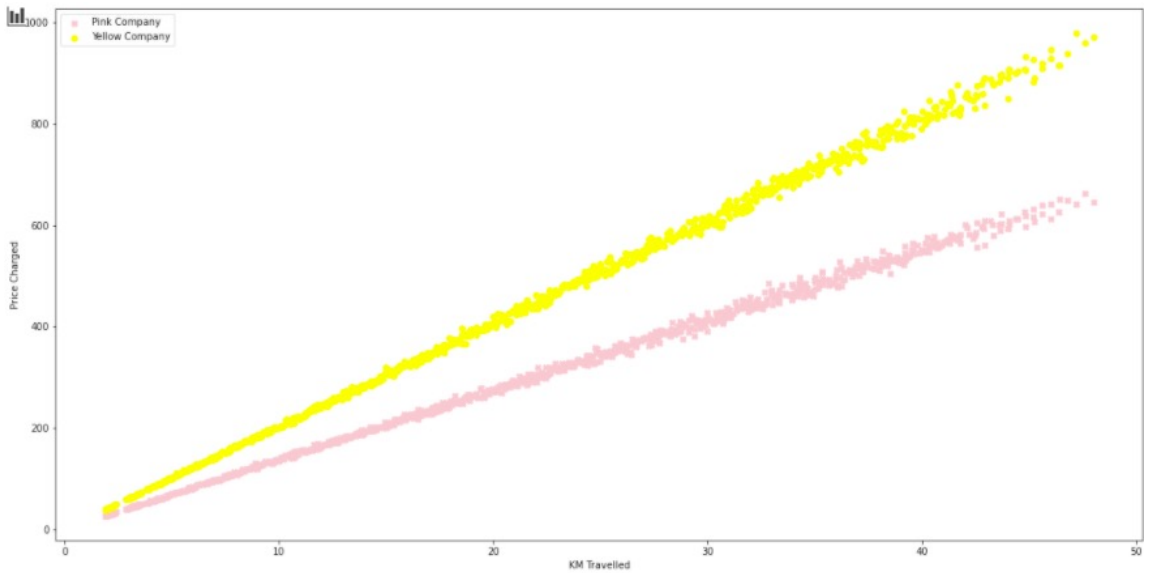

Yellow Cab Total Yearly Prices Charged and Costs of Trips
1e7
Price Charged
Cost of Trip
4
2
0
−2
−4
2016
2017
2018

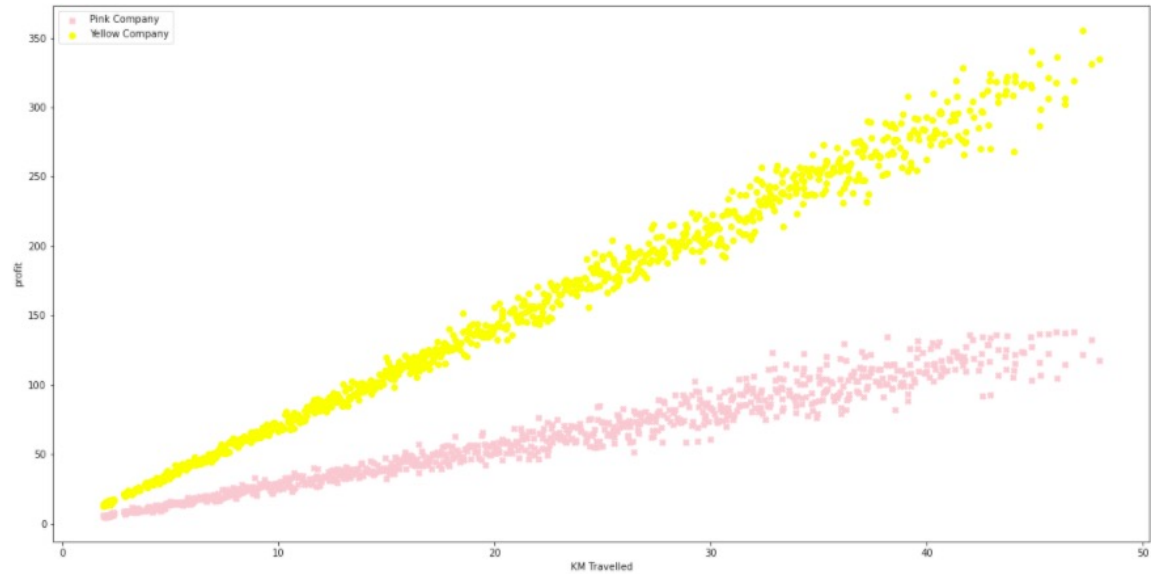

Pink Cab Total Yearly Prices Charged and Costs of Trips
1e7
Price Charged
Cost of Trip
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
2016
2017
2018


Pink Cab vs Yellow Cab Yearly Profits
1e7
Pink Cab Profits
Yellow Cab Profits
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Profit
2016.0
2016.5
2017.0
2017.5
2018.0
Year

4


Pink Company
Yellow Company
Cost of Trip
KM Travelled
0
100
200
300
400
500
600
0
10
20
30
40
50


Pink Company
Yellow Company
Price Charged
KM Travelled
0
200
400
600
800
1000
0
10
20
30
40
50


Pink Company
Yellow Company
profit
KM Travelled
0
50
100
150
200
250
300
350
0
10
20
30
40
50

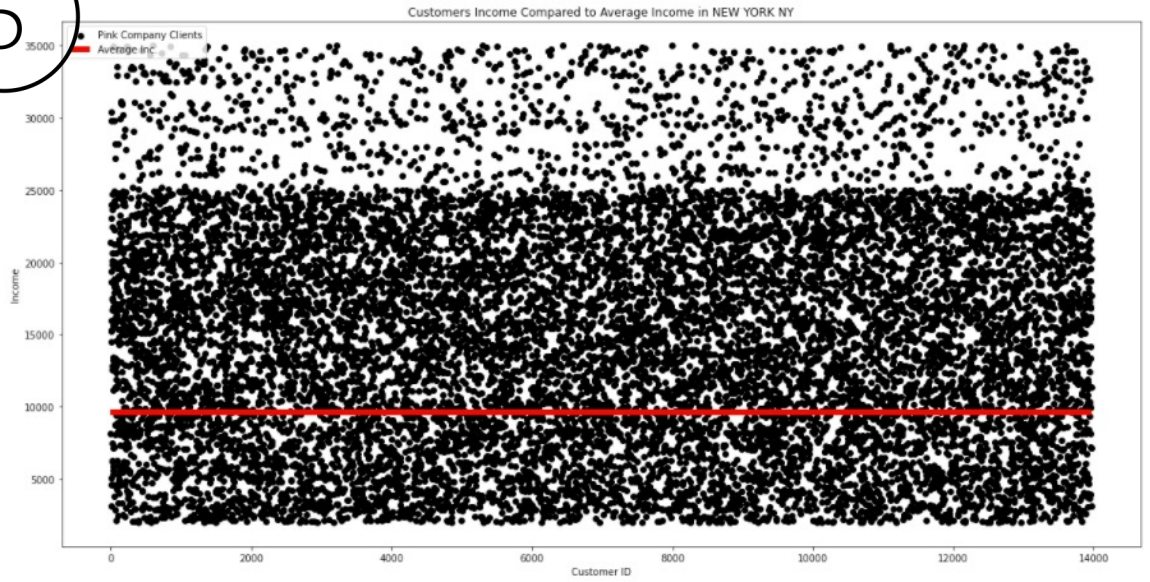
For the previous slide – inference

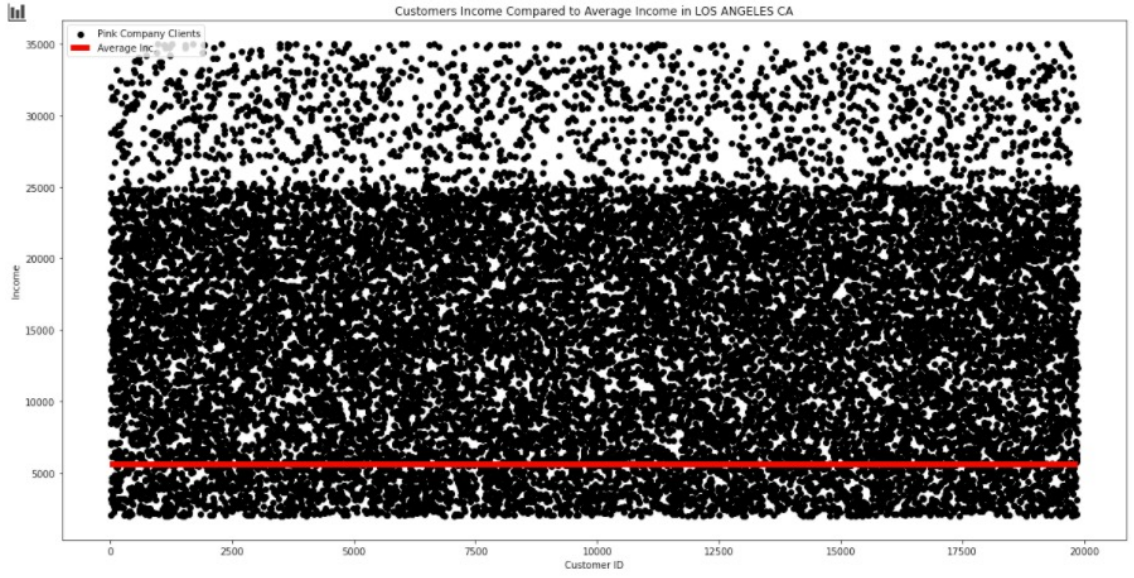
The correlation between KM Travelled and Price Charged for the Yellow Cab company is slightly higher than the Pink Cab company at 0.999411. Meaning that the prices do increase slightly faster per KM travelled when comparing it to the Pink Cab company.

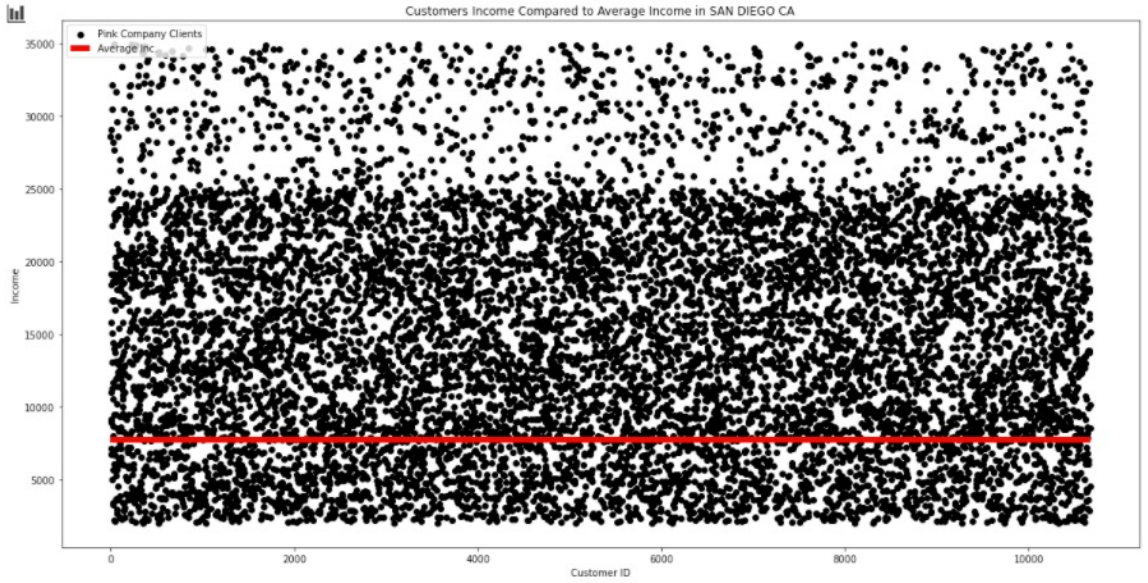
Same goes for the profit made. As the KM travelled increase, the profit made increases faster for the Yellow Cab company than the Pink Cab company.

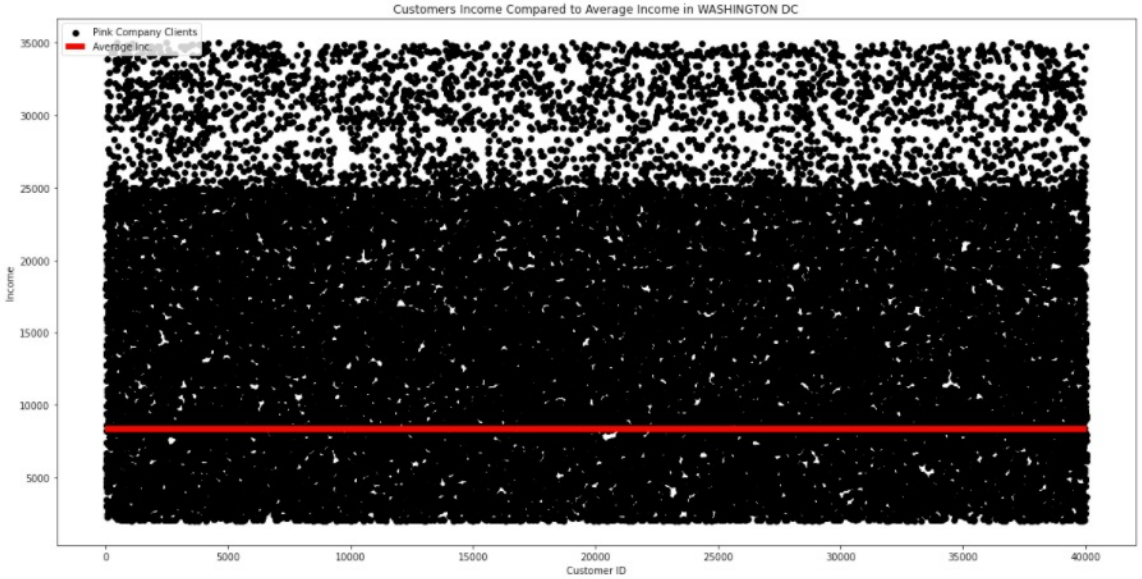
5

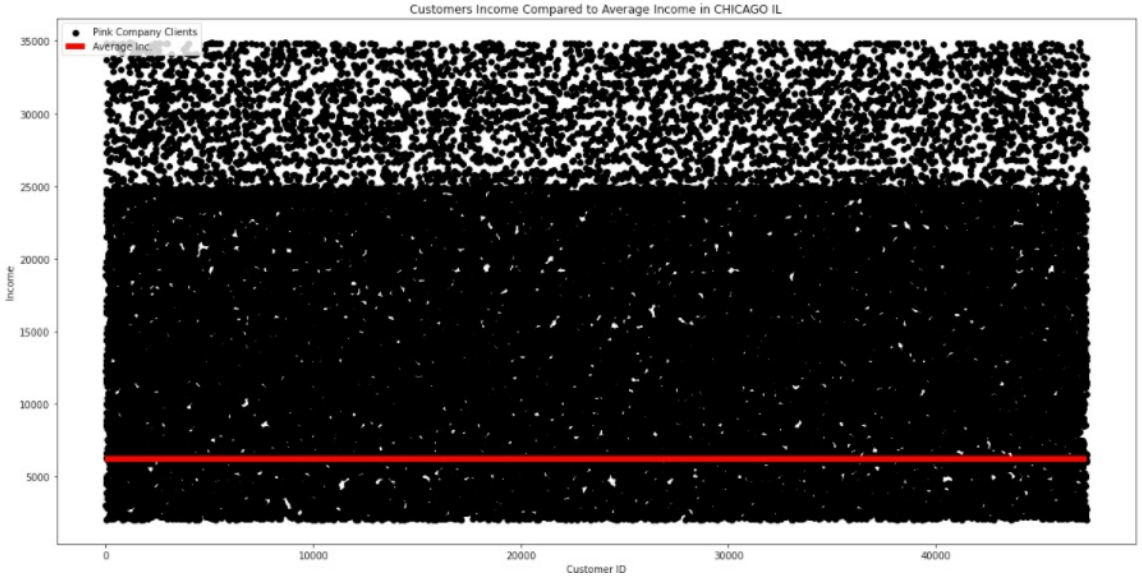
Pink Cab Company



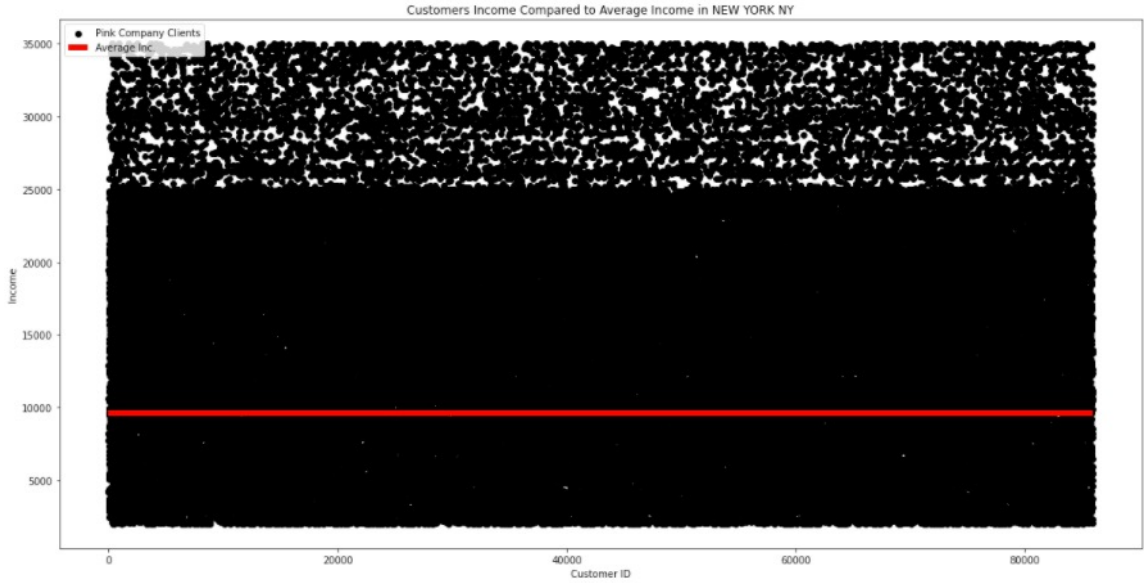





Customers Income Compared to Average Income in WASHINGTON DC
Pink Company Clients
Average Inc
Income
Customer ID


Customers Income Compared to Average Income in CHICAGO IL
Pink Company Clients
Average Inc
Income
Customer ID

Yellow Cab

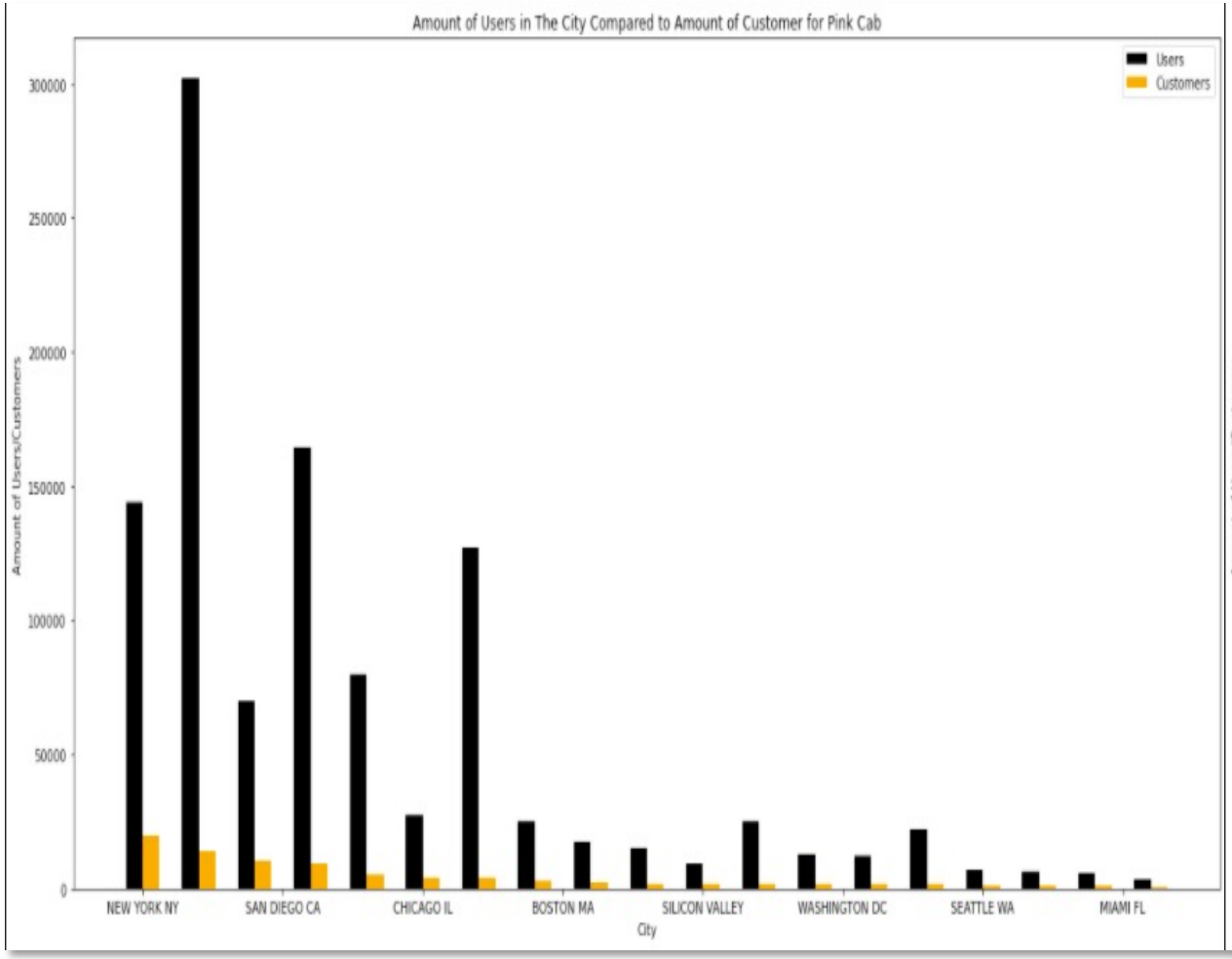

Customers Income Compared to Average Income in NEW YORK NY
Pink Company Clients
Average Inc
Income
Customer ID

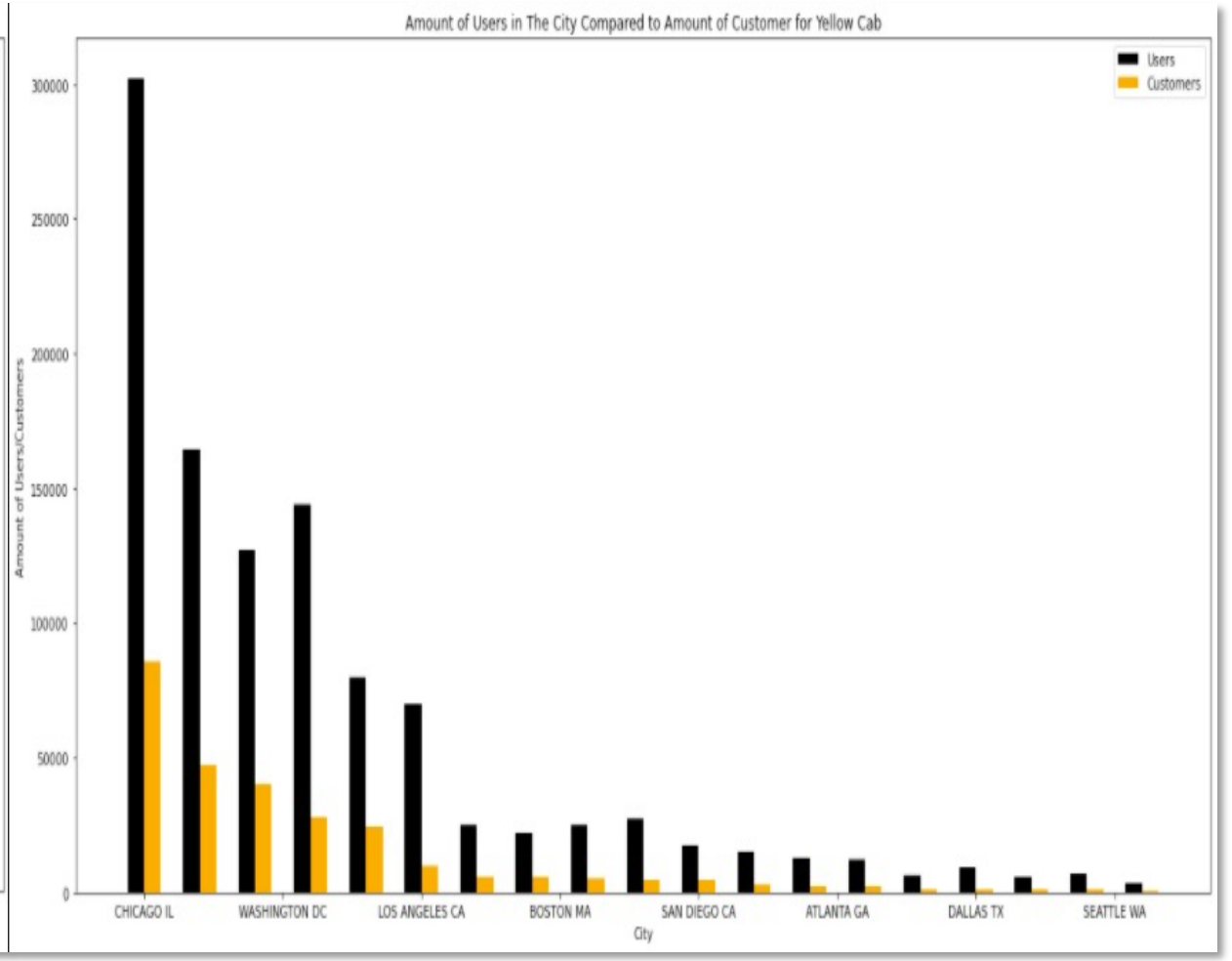
For the previous two slides - inference

The red horizontal line is the average income per month for the city. We can tell that majority falls above it. Which is positive for both companies as it means shortage of funds to use cabs won't fall short any time soon.

6

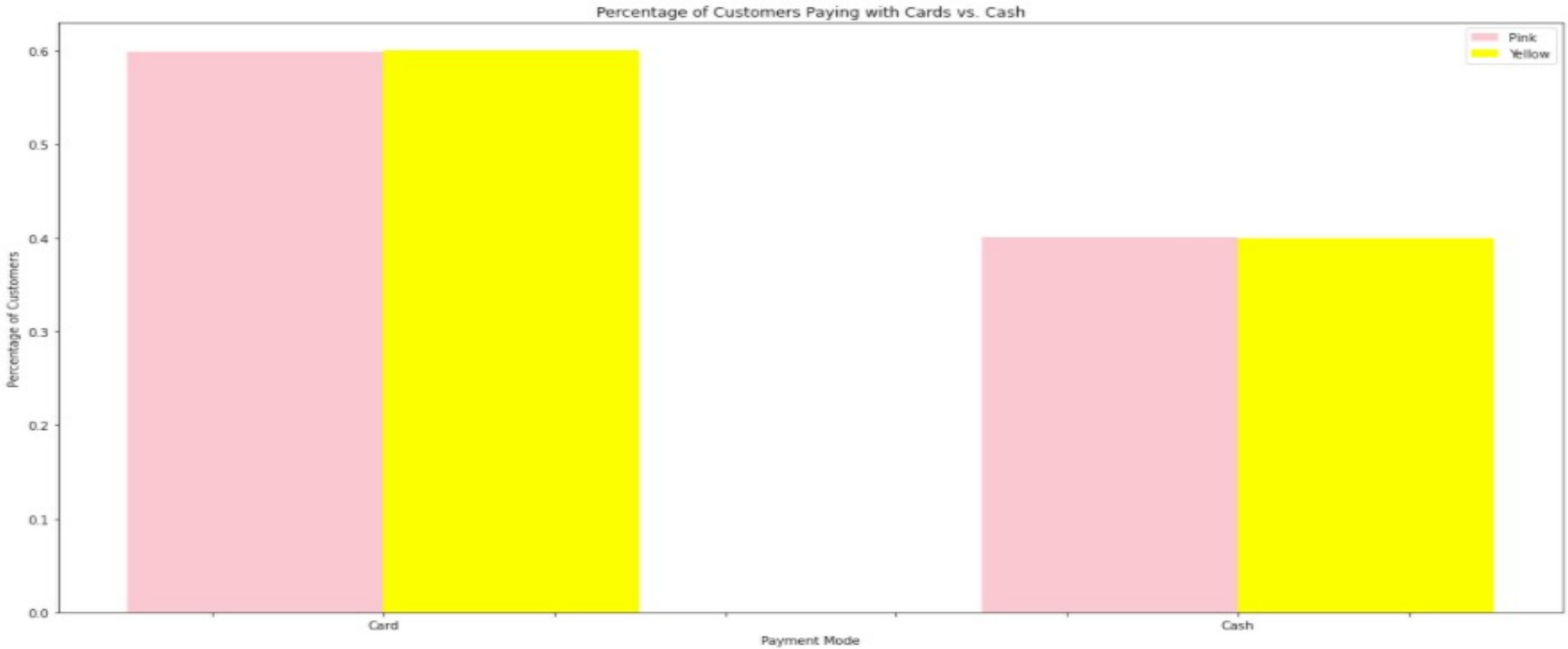
The number of transactions the Yellow Cab company has in their most popular city is more than double the number of transactions the Pink Cab company has in their most popular city.





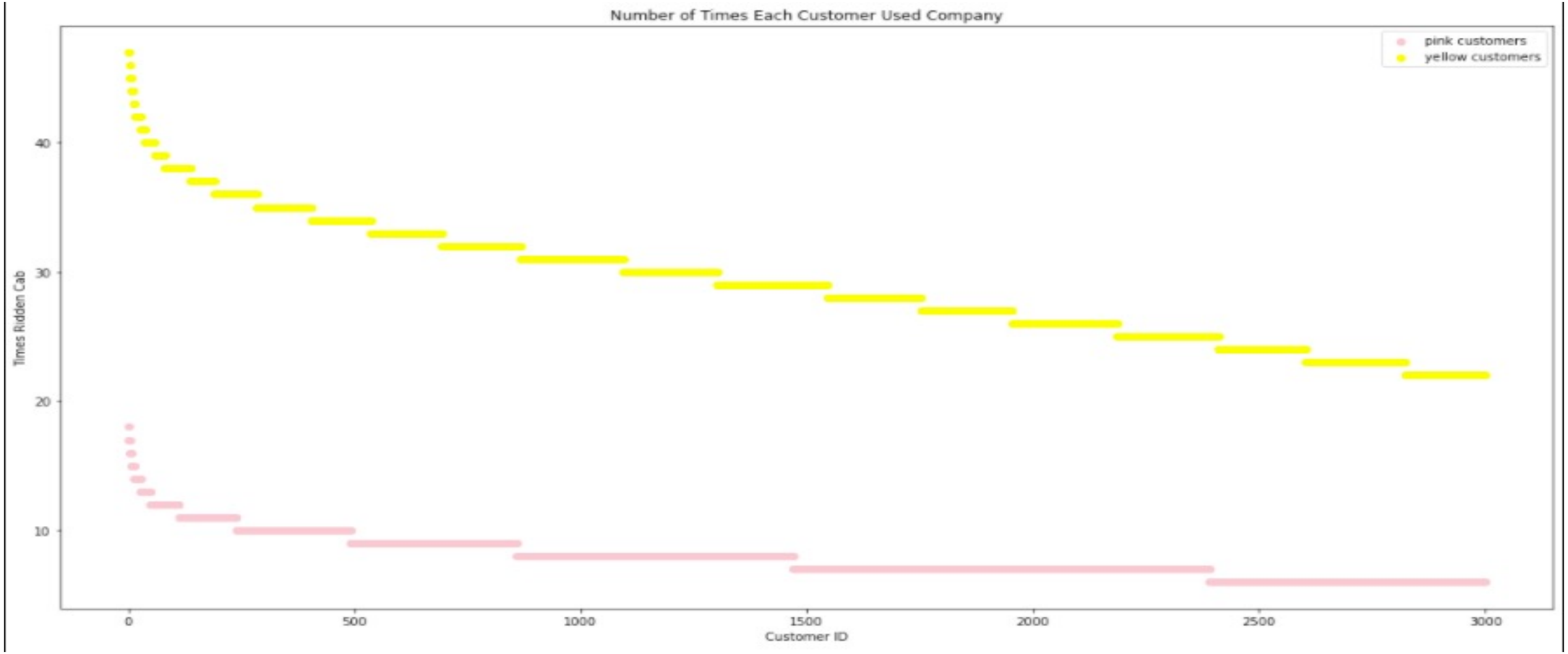
7

Here we look at the percent of transactions that were completed using card and cash. It is equal for both companies. It is important to point out that more customers use card as we are continuing to have more innovative and convenient ways to use our card rather than cash.



8

This is important to see because both companies have between 3000 and 4000 DISTINCT customers. Yet, for the Yellow Cab company, we see that it has more customers coming back. Customers will come back almost double the amount of times the Pink Cab customers would go back.



Recommendation

After analyzing the data, we will recommend the Yellow Cab company.

Thank You