G2M Case Study

virtual Internship

10-Nov-2024

Agenda

Agenda: 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-Market(G2M) strategy they want to understand the market before taking final decision.

Objective: Provide Data Analysis so,Firm XYZ can easily understand which company is better for Investment

Analysis:

1. Data Understanding/Merging
2. Stat info Through Hist. Plot
3. Find Relation through Box plot
4. Profit Analysis

Data Glacier
Your Deep Learning Partner

Data Merging

There are 4 types of Data:

1. **Cab_Data.csv –** this file includes details of transaction for 2 cab companies
2. **Customer_ID.csv** – this is a mapping table that contains a unique identifier which links the customer's demographic details
3. **Transaction_ID.csv –** this is a mapping table that contains transaction to customer mapping and payment mode
4. **City.csv –** this file contains list of US cities, their population and number of cab users

For, Data Analysis we need to do Data Merging-combine all data.So,we can easily Analyze data and understand it.

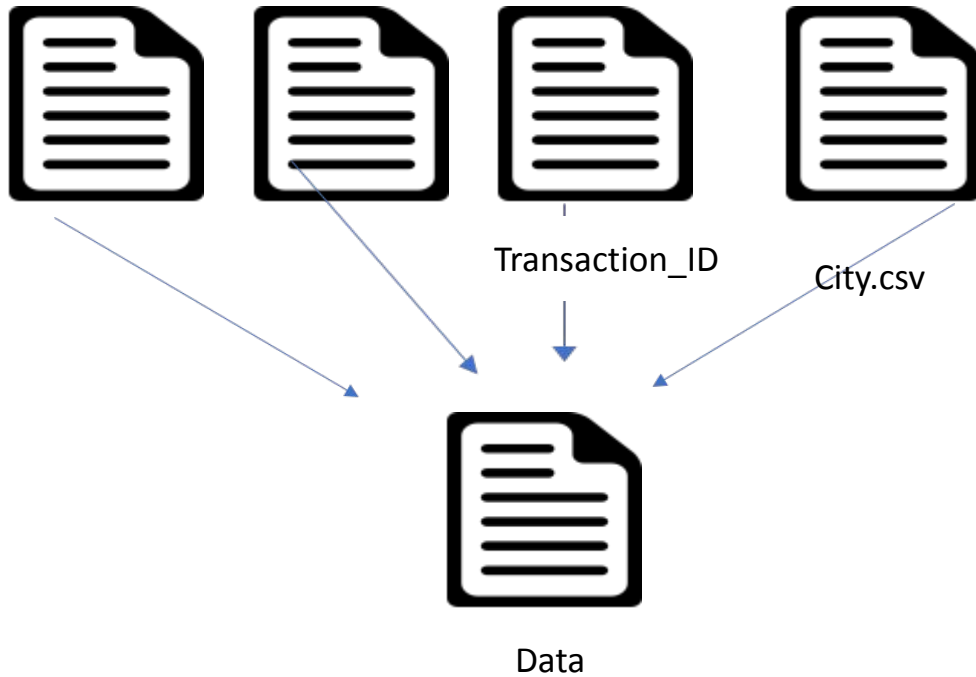**After Merging:**

New Dataframe : Data

Total Data:359392

Features:13(With Profit Column)

Timeframe of the data: 31/01/2016 to 31/12/2018.

Data Glacier
Your Deep Learning Partner

Data Merging

Cab_Data.csv   Customer_ID.csv

Transaction_ID   City.csv

Data

Assumption/Analysis:

1. Two Cab Companies Yellow and Pink.

1. If we need to compare these both companies as a Business Analysis then we have to focus on Profit. And Profit column is not available in data. so, we need to find Profit per ride using 'Price Charged' and 'Cost of Trip'

   Data['Profit'] = Data['Price Charged'] - Data['Cost of Trip']

1. Users column treated as a cab user per cities. Except these all datas are easily understandable.
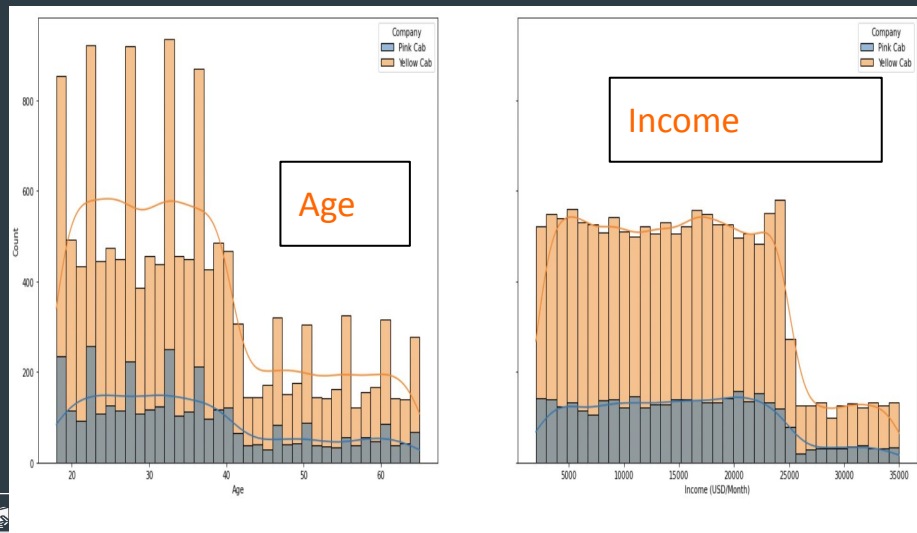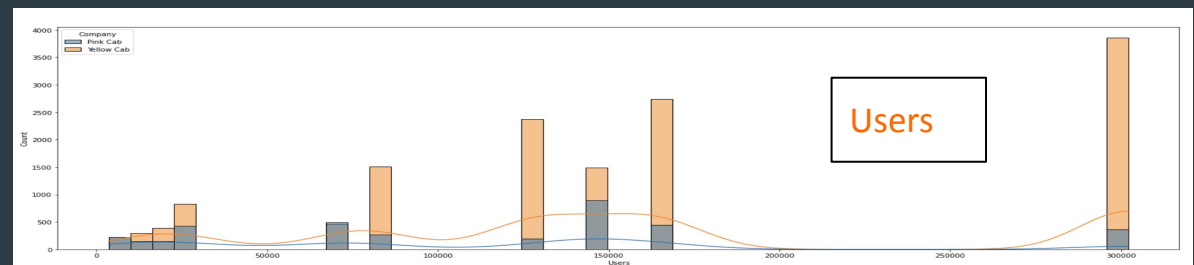
Hist Plot(7 Features)
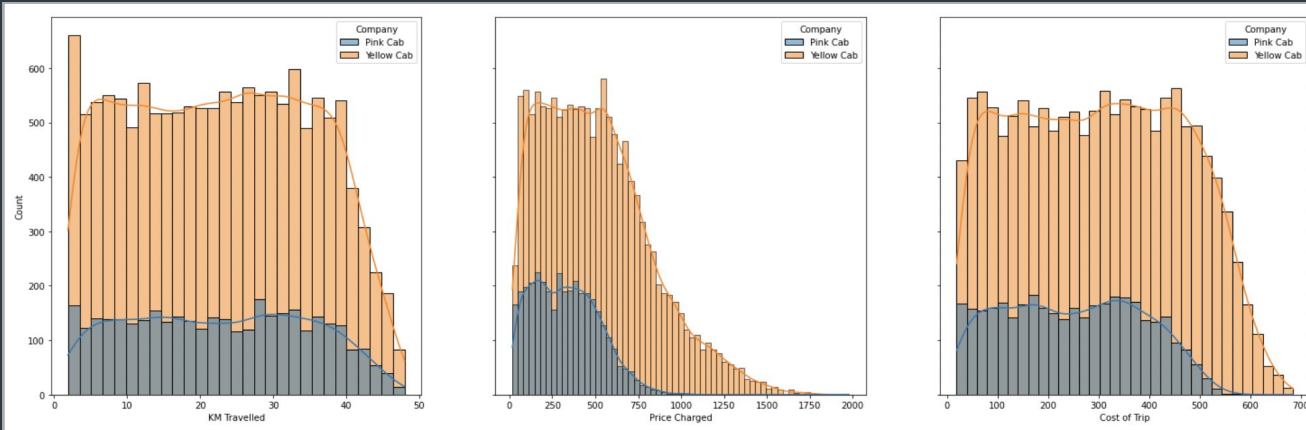
KM_Travelled

Price Charged

Cost of Trip

Hist Plot or Distribution Chart for both Companies 7 Features which are Most Important for Data Analysis.So, we can get stat_info of Both Companies These 3 Features.
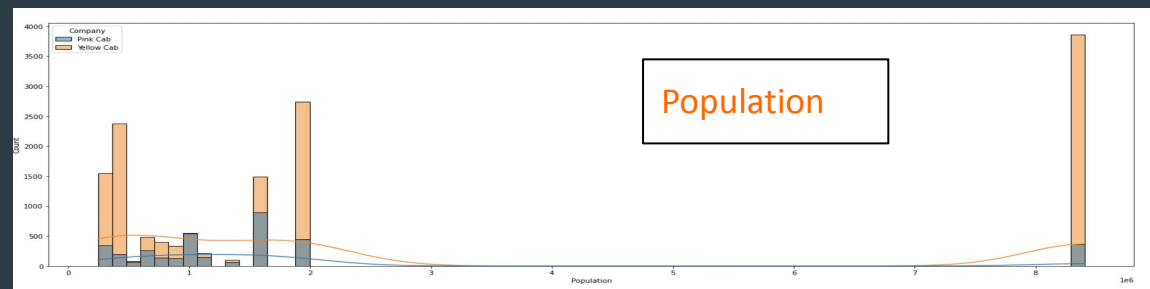
Age

Income

Users

Population

Stat_info of Pink Cab

Index	KM Travelled	Price Charged	Cost of Trip	Age	ome (USD/Mor	Population	Users
count	84711.00	84711.00	84711.00	84711.00	84711.00	84711.00	84711.00
mean	22.56	310.80	248.15	35.32	15059.05	2350641.51	125590.81
std	12.23	182.00	135.40	12.64	7991.08	2734890.26	94593.43
min	1.90	15.60	19.00	18.00	2000.00	248968.00	3643.00
25%	12.00	159.97	131.87	25.00	8371.00	814885.00	27247.00
50%	22.44	298.06	246.33	33.00	14713.00	1595037.00	144132.00
75%	32.96	441.50	360.18	42.00	21055.00	1955130.00	164468.00
max	48.00	1623.48	576.00	65.00	35000.00	8405837.00	302149.00

Stat_info of Yellow Cab

count	274681.00	274681.00	274681.00	274681.00	274681.00	274681.00	274681.00
mean	22.57	458.18	297.92	35.34	15045.67	3373228.31	168473.25
std	12.23	288.39	162.55	12.58	7962.73	3439014.40	100570.56
min	1.90	20.73	22.80	18.00	2000.00	248968.00	3643.00
25%	11.99	226.68	158.40	25.00	8439.00	671238.00	80021.00
50%	22.44	425.06	295.60	33.00	14676.00	1595037.00	144132.00
75%	32.96	633.88	432.43	42.00	21023.00	8405837.00	302149.00
max	48.00	2048.03	691.20	65.00	34996.00	8405837.00	302149.00
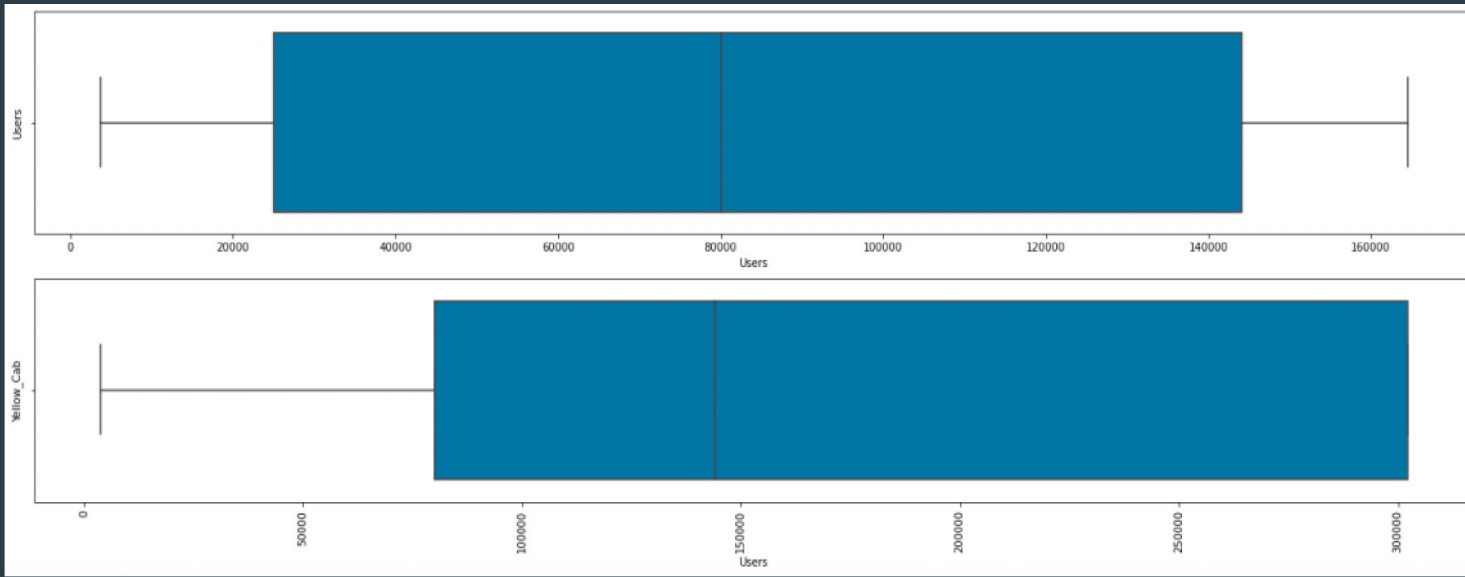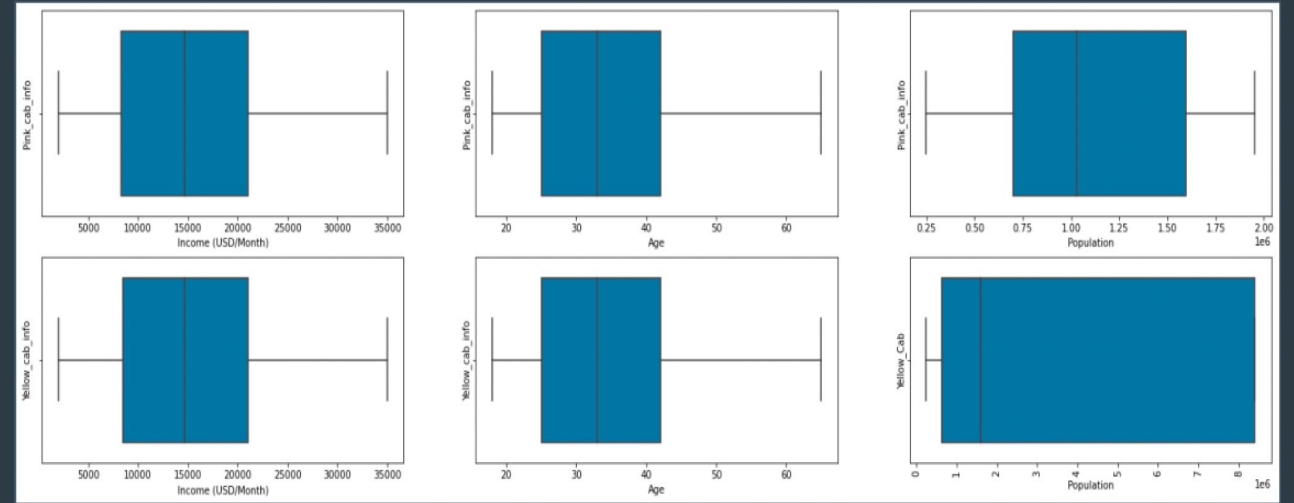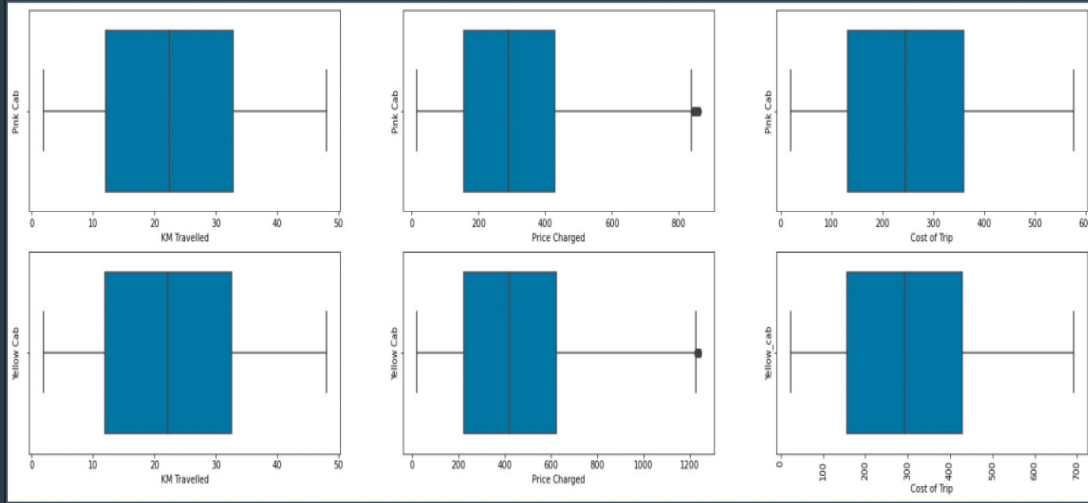
From Box Plot we can find relationship Between Features

Data Glacier
Your Deep Learning Partner

BOX PLOT

Box Plots are used to show Distributions of Numeric data values,especially when you want to compare them between groups……

So From, Box Plots we derived that:
 There is Positive Relations between
1.   Km Travelled and Price Charged
2.   Price Charged and Cost of Trip
3.   Cost of Trip and Km Travelled
4.   Population and Users

Profit Analysis

Now, It's time for Profit Analysis of Both Companies with different Features.

1. Statinfo of Profit in Both cab companies
2. Profit Analysis per ride
3. year wise Profit Analysis
4. Gender wise Profit Analysis
5. Profit Analysis Age wise and Incomewise

Hist Plot(Profit)

Distribution for Profit
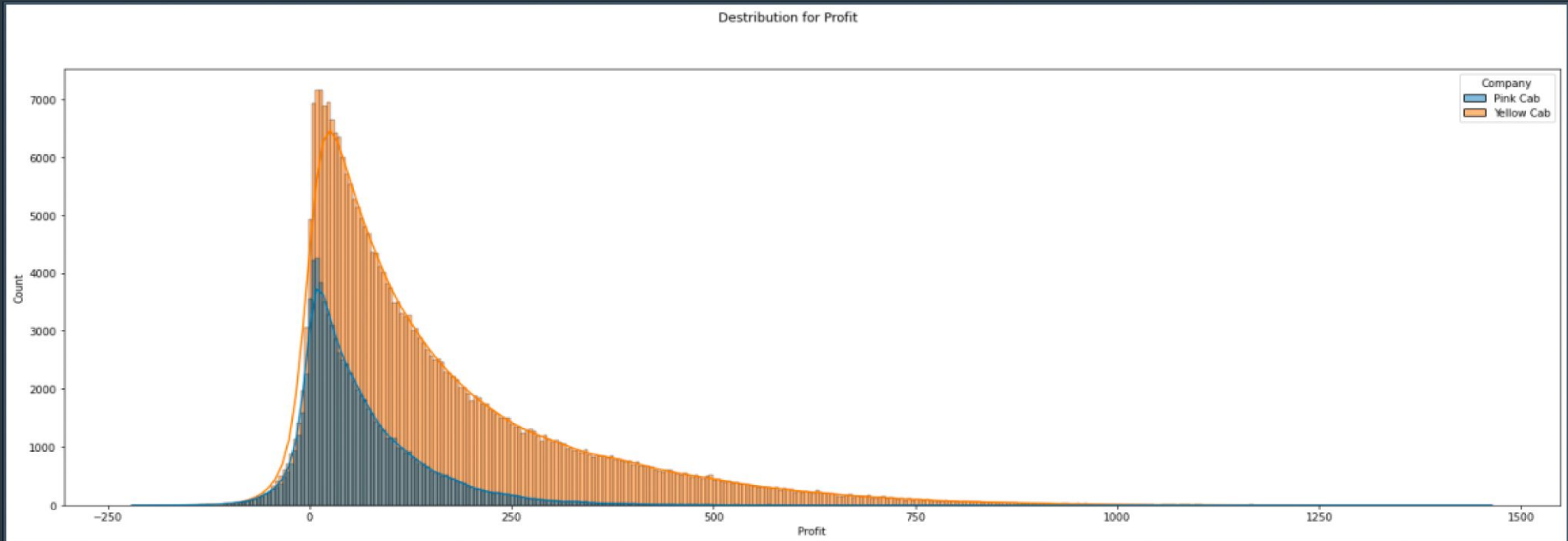
This Is Profit Distribution Plot for Both companies From this we can get max,min and Average values of Profit.

Data Glacier
Your Deep Learning Partner

Pink_Cab statinfo for Profit

Index	Profit
count	84711.00
mean	62.65
std	77.59
min	−220.06
25%	11.21
50%	41.50
75%	93.97
max	1119.48

Yellow_Cab statinfo for Profit

Index	Profit
count	274681.00
mean	160.26
std	171.82
min	−176.93
25%	37.18
50%	102.00
75%	228.67
max	1463.97

Data Glacier
Your Deep Learning Partner

Profit Analysis per ride

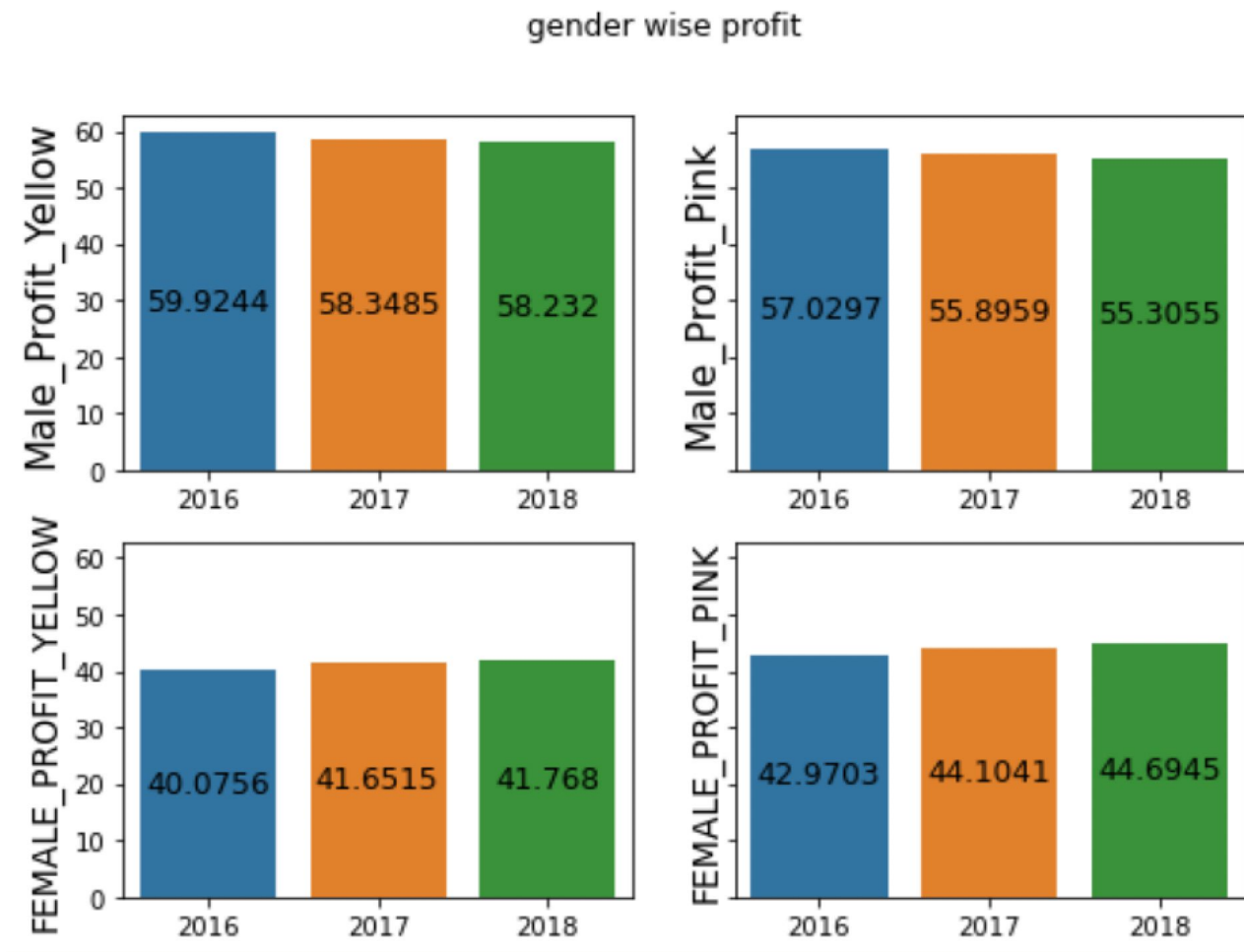| Company	Profit	Total Rides	Profit per ride
Pink Cab	5307328.321	84711	62.65217410961977
Yellow Cab	44020373.17080002	274681	160.25998584103021
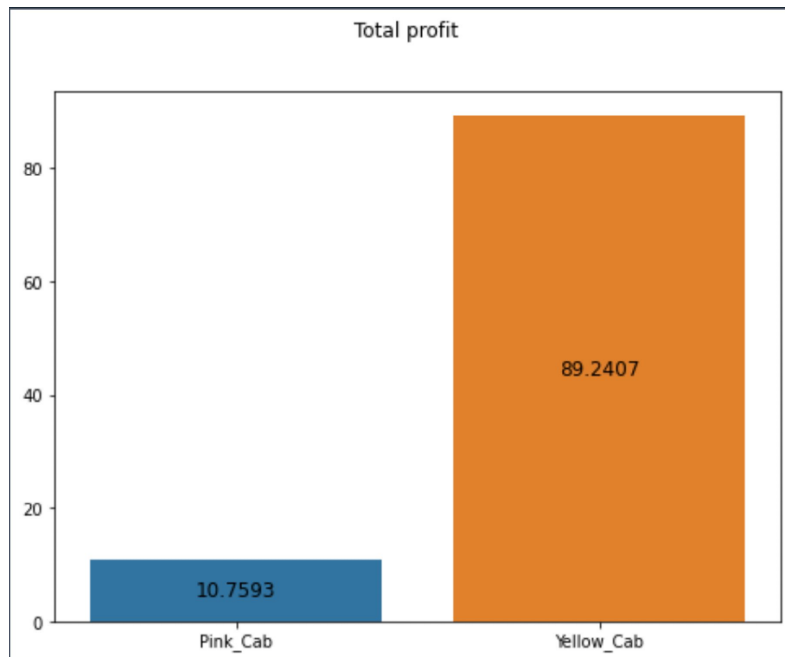
If we compare Both companies with total profit then we can derived that Yellow cab total profit is 9 time more than pink cab
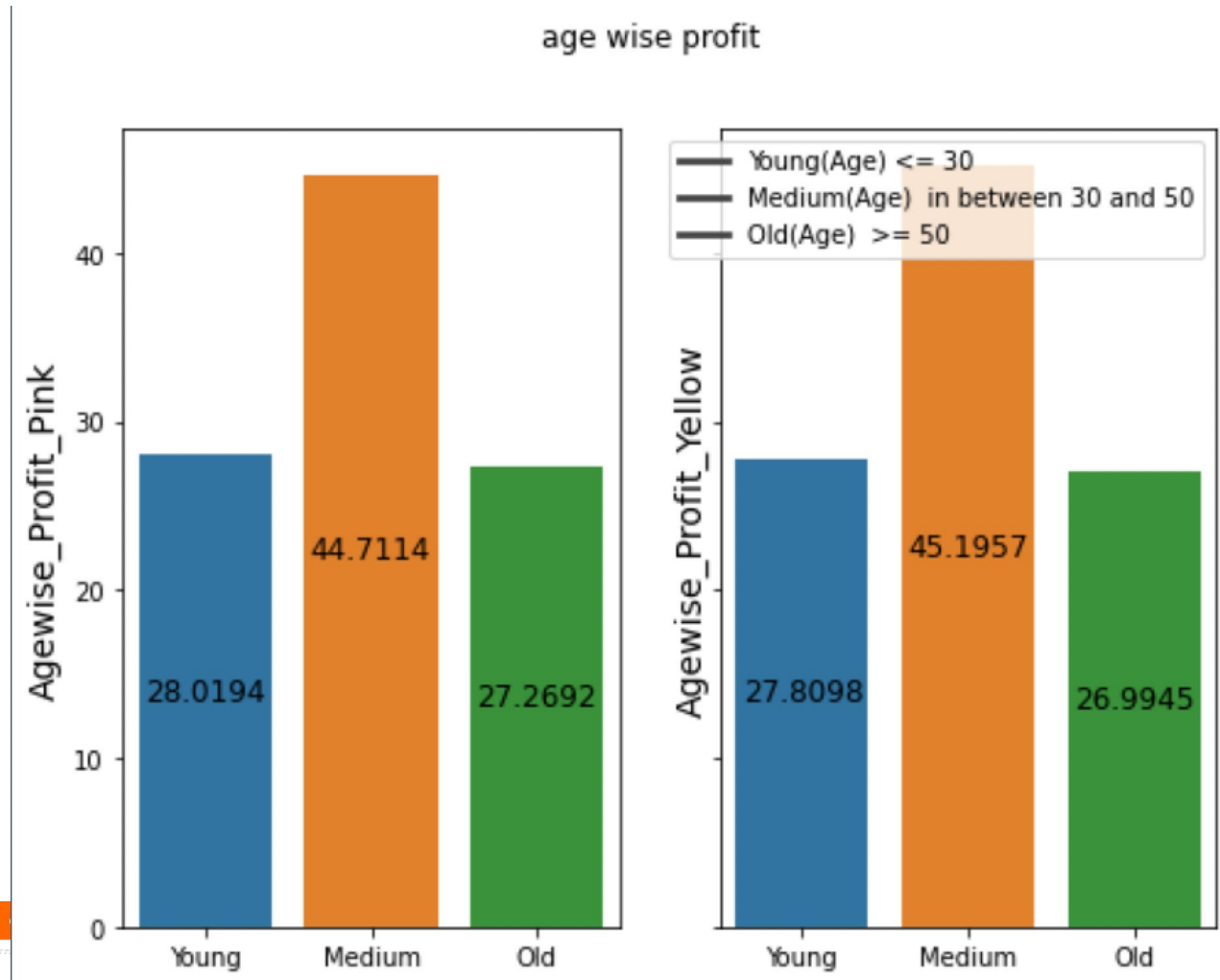
Year wise Profit analysis

Data Glacier
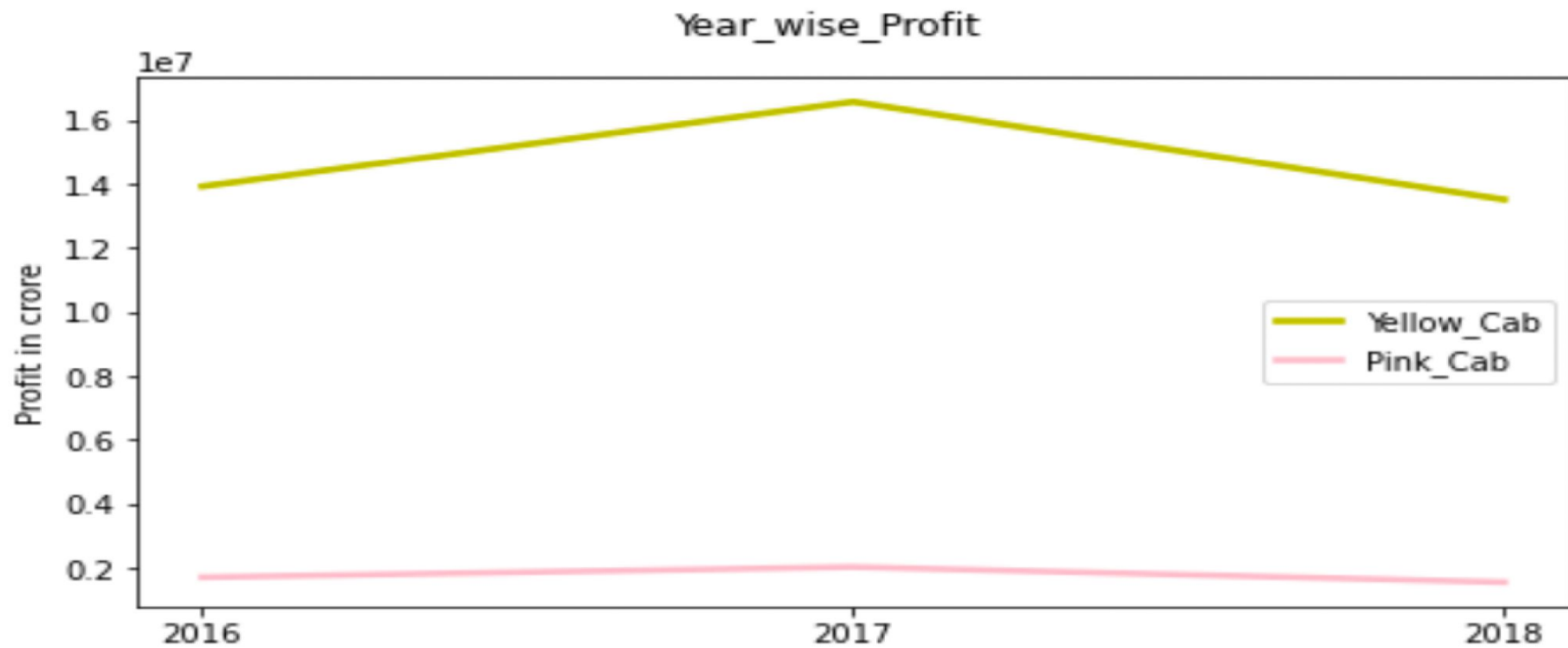
Profit Analysis Gender wise

gender wise profit

There is not more difference in Profit percentage of both companies. That mean equal distribution of Profit based on gender.

In Both companies Male Users are more Than Female Users.

Profit Analysis based on Age

age wise profit

First I divide users in three categories
1. Young (Age <= 30)
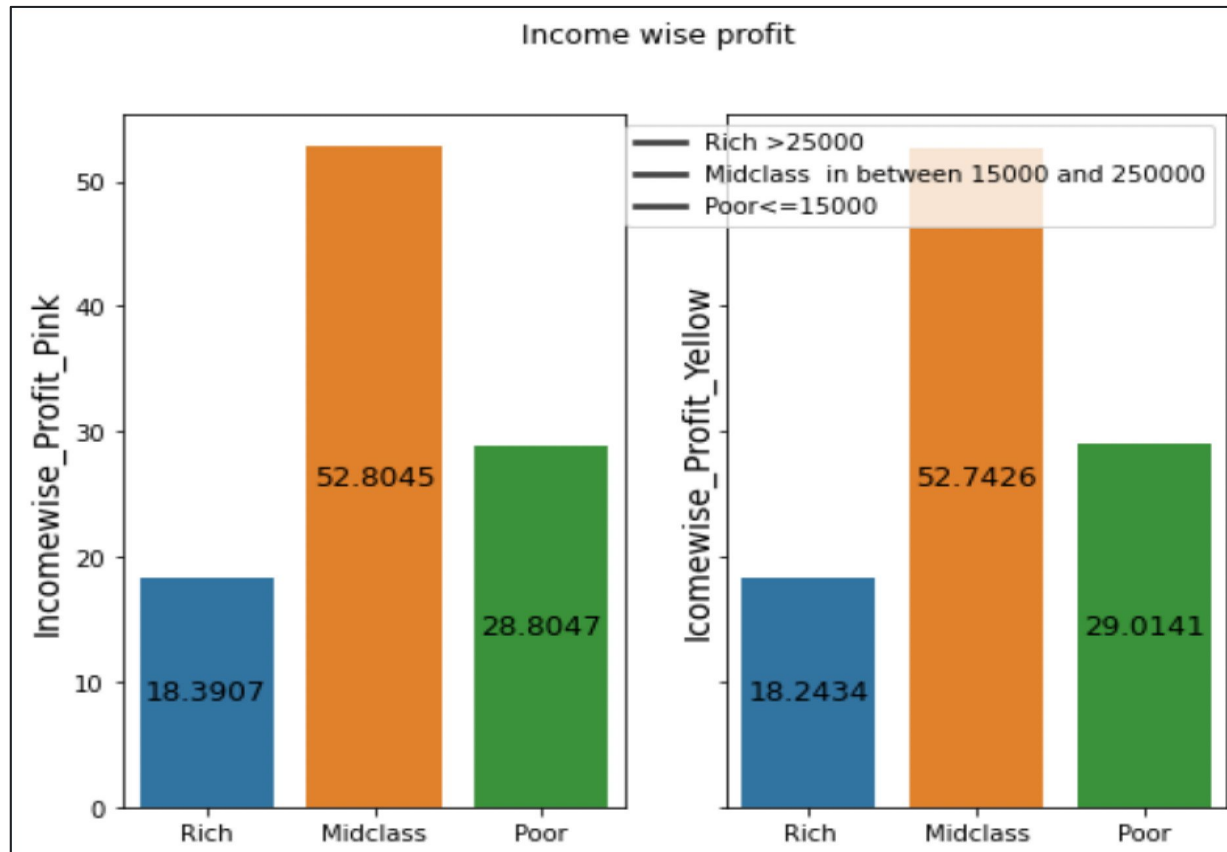2. Mid(Age in between 30 and 50)
3. Old(Age>= 50)

Then,did Profit analysis.
Mid age customers are more as compared to other in both companies.

This Bar graph shows Profit percentage based on generation.

Profit Analysis based on Income

Income wise profit

First I divide users in three categories
1. Rich (Income >= 25000)
2. Middle Class(Income in between 15000 and 25000)
3. Poor(Income<= 15000)

If we did analysis between two companies then there is almost same distribution based on Income..
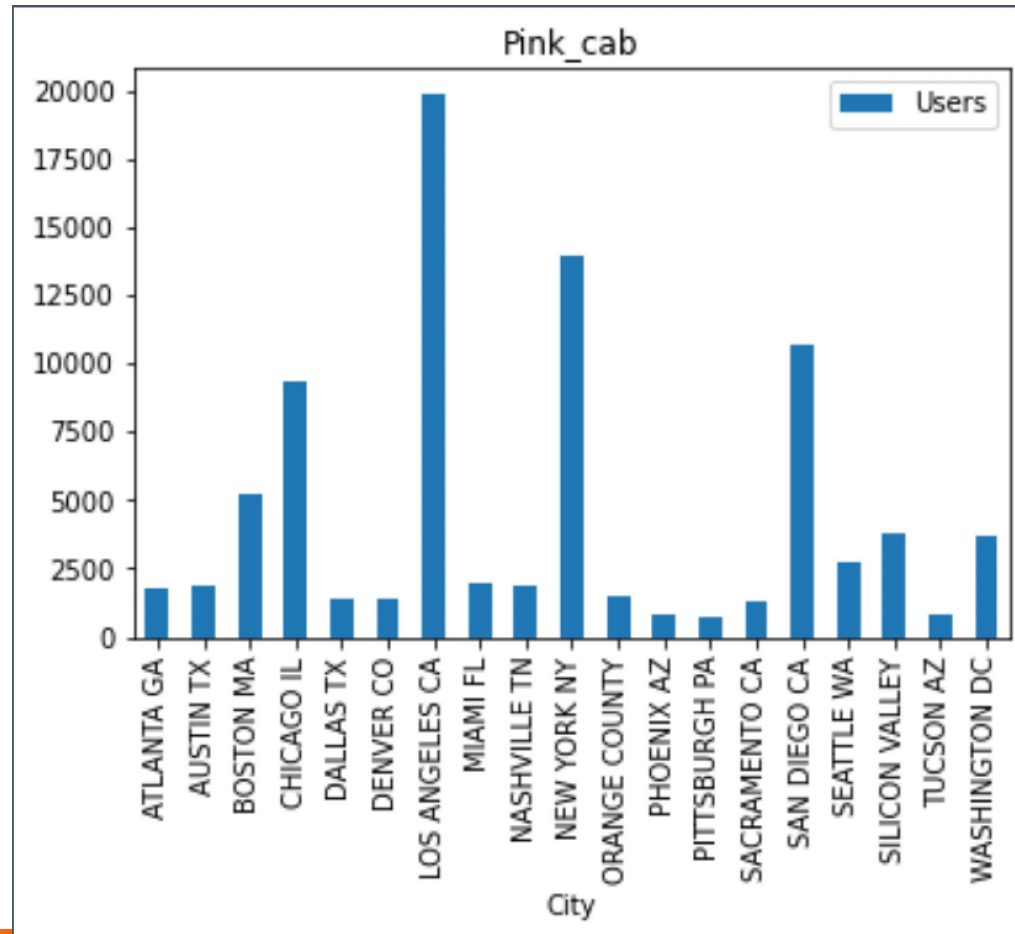from this bar chart we can say Middle class customers most use cabs either Yellow or Pink..

City wise cab Users covered by company
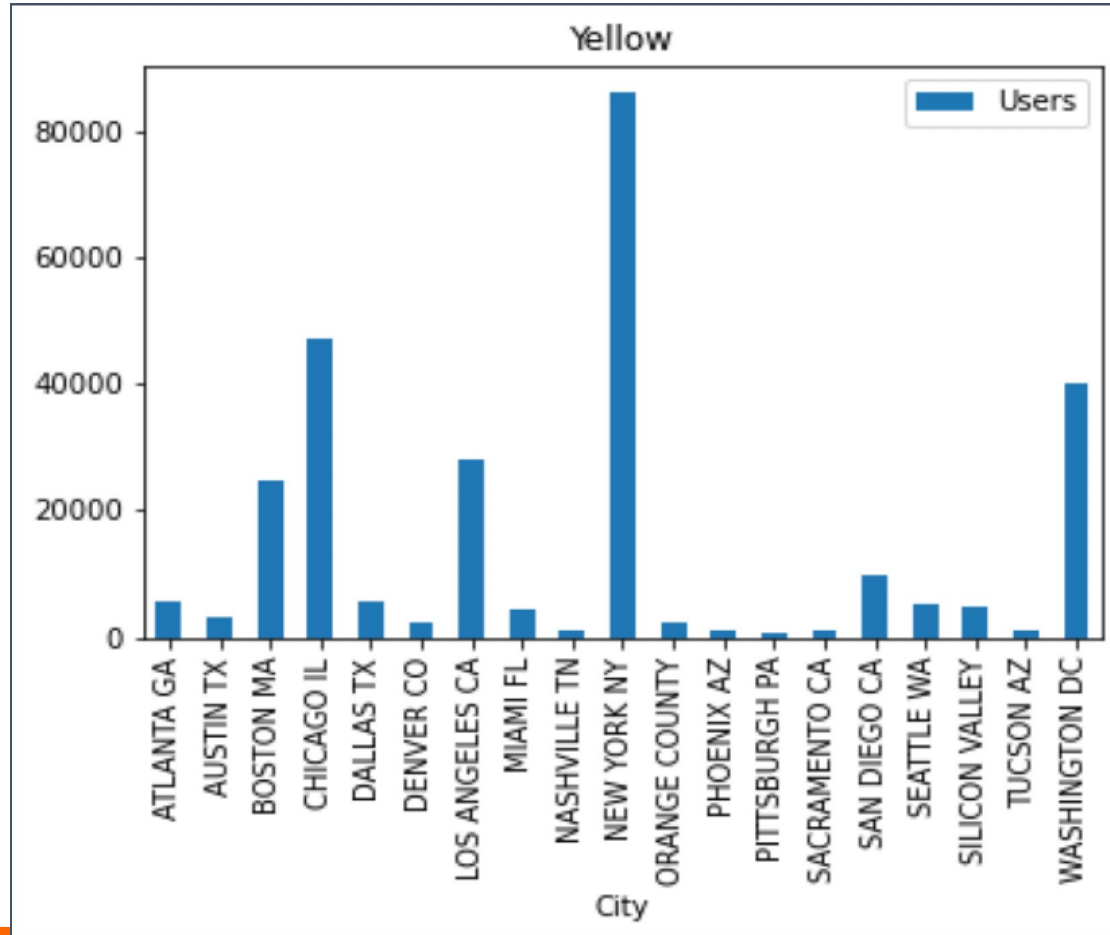
Yellow cab's most Users are from New York NY,Chicago IL,Washington DC,Boston MA and Los Angeles CA.

City wise cab Users covered by company

Pink cab's most Users are from Los Angeles CA,New York NY,San Diego CA and Chicago IL.

Recommendation

We evaluated both the companies on following points and found Yellow cab company is better than pink cab company:

1. Customer Reach: Yellow cab has higher customer reach in almost 25 cities while Pink has only in 4 Cities.
2. Age-wise Reach: Yellow cab has customer in all age group and it's been observed that it's even popular in 60+ age group as equally as its in 18-25 age group.
3. Profit per Ride: Yellow cab has almost 3 time more profit than Pink cab per ride
4. Income wise Reach :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)

From, all the points we can recommend Yellow cab is far better then Pink cab for investment..

Thank You