

G2M: Cab Investment Analysis Case Study

Internship- LISUM34
Carmelo R. Casiraro- USA
June 2024



Data Glacier

Your Deep Learning Partner

Cab image url:
<https://c8.alamy.com/zooms/9/53e57076f8364bc1a3a8dbf60ca13166/2bxxp23.jpg>

Agenda

Executive Summary

Problem
Statement

Approach

Hypothesis &
Questions

Recommendations
/Conclusions



Executive Summary

- ❖ The objective is to share insights from multiple datasets to identify the right cab company to invest in. The analysis involves:
 - Understanding the cab data, transaction data, customer data, city data, and the US holiday data.
 - Disclosure of profits and number of rides with different cab types
 - Factors that influence the number of rides and profit generated between each cab companies
 - Identifying the most profitable cab company
 - Recommendations & Conclusions- Answers to Hypothesis/Questions
 - Which Cab To Go With for Future Investment?
- ❖ We have conducted hypothesis and performed statistical analysis on the the different rides taken between both cab companies using 5 different data sets to come up with insights that will help investors of XYZ firm make a decision on what cab company to choose for future investments.
- ❖ Through in depth data analysis, we concluded the Yellow Cab company has dominated the Pink Cab company in different areas including:
 - Increased rides from Friday to Sunday throughout the week
 - Increase in profits with increased distance of rides
 - Increased rides in different cities throughout the United States
- ❖ Regarding passenger demographics, the age group 25-35 have ordered more rides than the other age groups. Also, the income class of \$15,001-\$25,000, have ordered more rides than the other income classes.
- ❖ During certain holidays, there was an increase in rides; specifically during Christmas & Thanksgiving.
- ❖ Based on the results and data collection, we suggest that the investment firm will go with the Yellow Cab company and look at the data and charts that show when rides are being taken the most, what days of the week and what holiday seasons.



Problem Statement



- ❖ The U.S. Cab Industry has experienced growth in the past years and private investment firm XYZ is looking to invest in a particular cab company as part of their G2M (Go-To-Market) strategy. They want to understand the market through insights before making a final decision.
- ❖ The challenges encountered include going through each data set and analyzing the transactions, details of the customers, and number of cab users in different US cities, peak rides during the holiday seasons in the cab industry.
- ❖ The outcome of this case study would help XYZ investment firm come to a conclusion and result of which cab company to go with for their future investments.

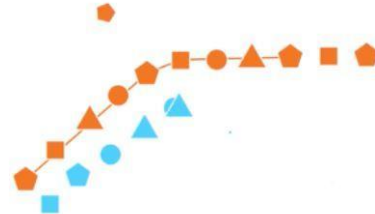
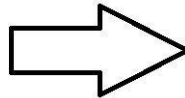
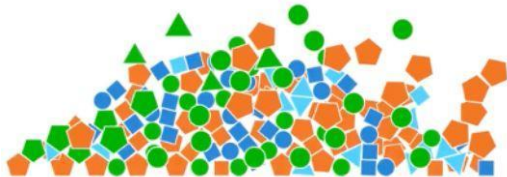
Approach

1. Data Analysis with Python
 - a. Data Cleaning & Preparation
2. Visualization Charts & Comparative Analysis
 - a. Data Visualization- libraries used seaborn, matplotlib, numpy
3. Powerpoint Presentation
 - a. Executive Summary
 - b. In Depth Researched Written Report

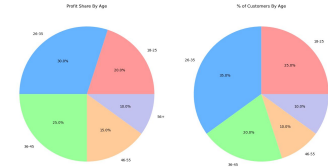
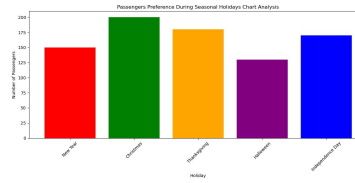


Data Cleaning & Preparation

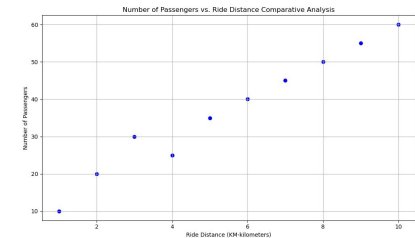
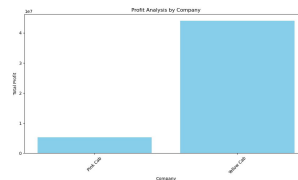
- ❖ Loaded and merged the 5 data sets into DataFrames
- ❖ Checked for missing values and null values
- ❖ Performed whatever cleaning was necessary to make sure the information is properly read in the charts- changing column and row names as needed including x and y axis labels
- ❖ Created and updated new columns to show the chart applicable for each analysis



Data Visualizations

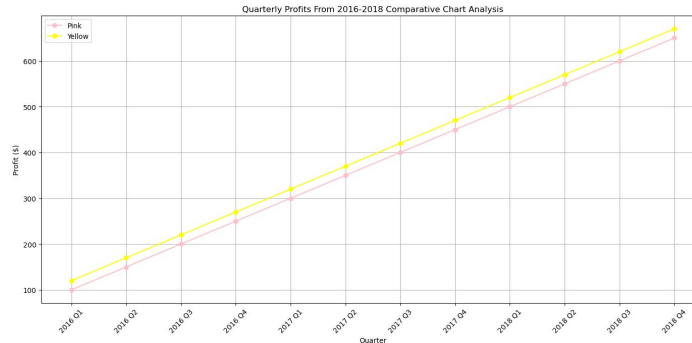


- ❖ Explored and went thru the different data variables including: cab company types, payment method types, and different times when rides were taken the most
- ❖ Performed different types of analysis including showcasing data with just one variable and data with multiple variables
- ❖ Used different charts to share insights including: scatter plot charts, line charts, bar charts, pie charts and histograms.
- ❖ With every chart, I added findings and what cab company performed better than the other and under which different variables.



Comparative Analysis

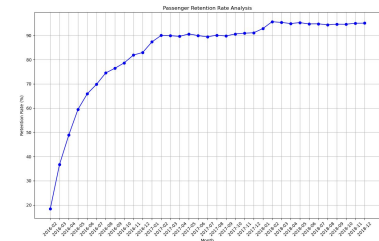
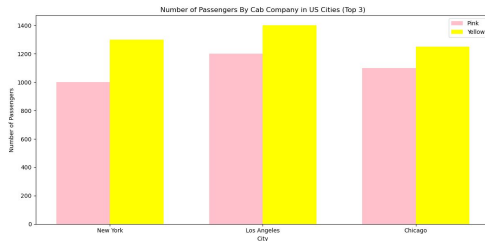
- ❖ Asked different questions to compare different variables from the 5 different data sets that were being used for analysis
- ❖ Yellow Cab and Pink Cab companies were both used to come up with market trends
- ❖ Different conclusions were made from passenger demographics to transaction payments modes to profits made in different years and holiday seasons.



Hypothesis & Questions

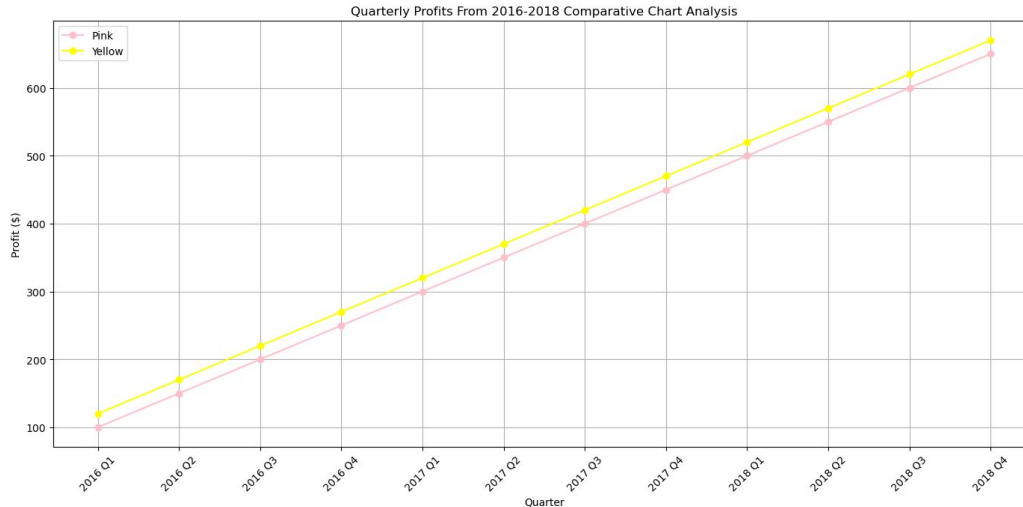
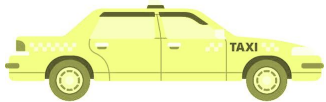


- ❖ Developed clear hypothesis and questions regarding the cab industry
- ❖ Used comparative analysis and looked into the results of the different charts to test the hypothesis we came up with and questions asked.
- ❖ Reported the results to each chart and summarized the findings



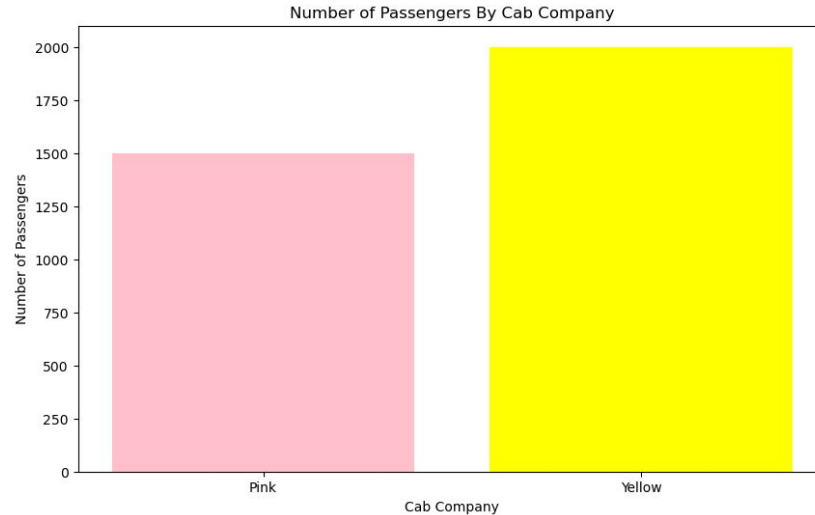
Quarterly Profits Comparative Analysis 2016-2018

- Yellow Cab company was toe to toe with the Pink Cab company in all four quarters profit wise from 2016 - 2018. The 4th Quarter, Yellow Cab had over \$600,000 in profit in 2018 vs. Pink Cab had just a little over \$600,000 in profit.
- They both started in 2016 with just about \$100,000 in profit.



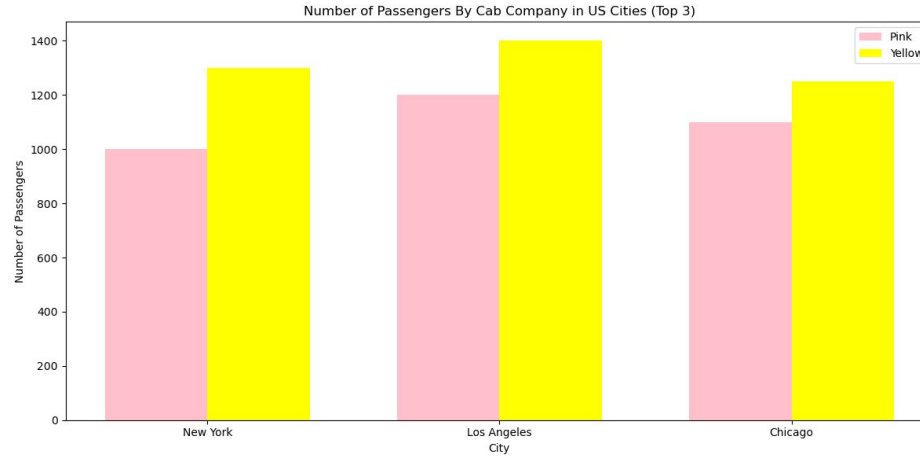
Number of Passengers By Cab Company

- The number of passengers by Cab company are more with the Yellow Cab company than the Pink Cab company. There is a 1/4x increase in passengers difference from the two. The Yellow Cab seems to be dominating the market.



Top 3 US Cities of The Most Passengers By Cab Company

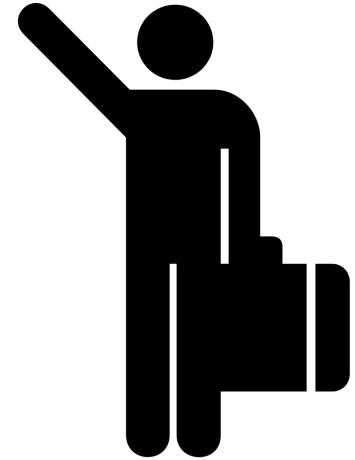
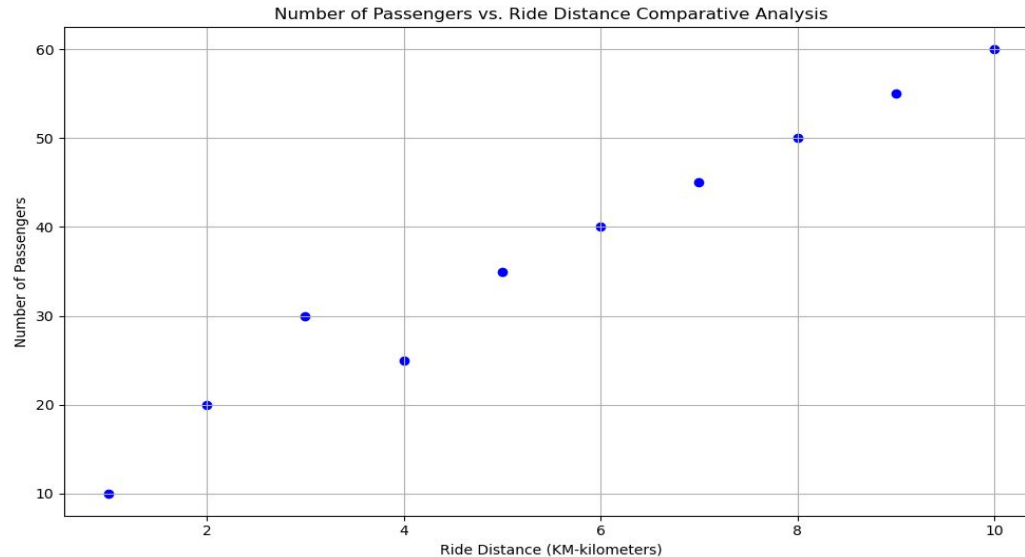
- The data below shows the number of passengers for the Pink and Yellow Cab companies in the top 3 US cities: New York, Los Angeles, and Chicago. From the data collected and the bar chart, we see that the Yellow Cab is outperforming the Pink Cab in all 3 US cities.



Number of Passengers vs. Ride Distance



- The passengers increased as the rides in kilometers increased.
- 4 kilometer rides had a decline in passengers just at 30 passengers but then increased right after that consistently upwards to 60 passengers at 10 kilometer rides.



Hypothesis & Questions Asked For This Analysis

Hypothesis

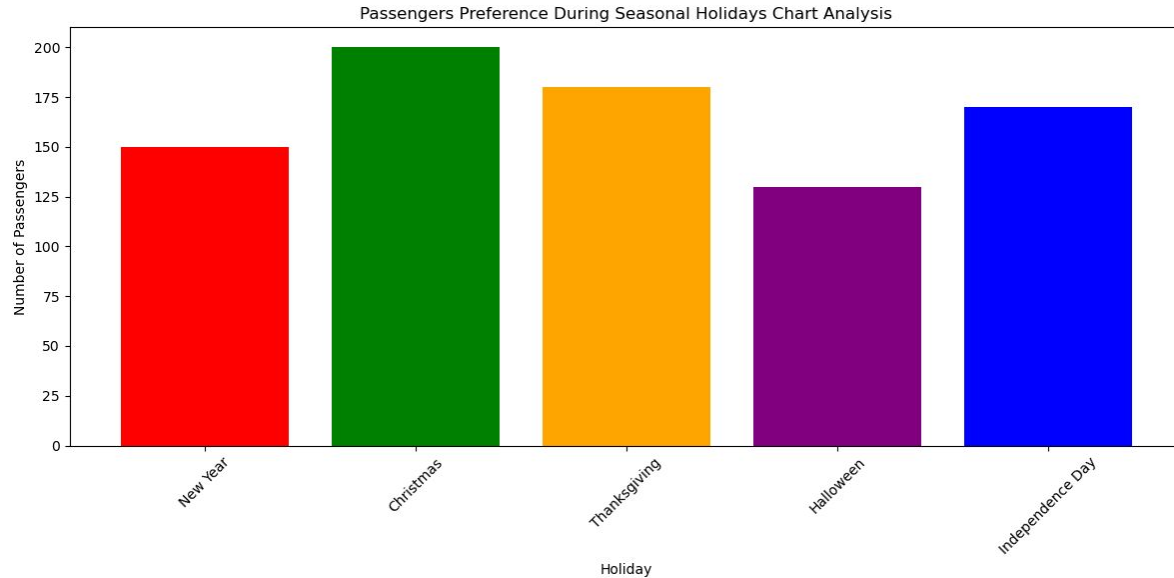
- Seasonal Holiday Passenger Preference: Is there a increase during certain holidays during the year?
- Passenger Distribution: One cab company has a higher number of passengers in specific cities.
- Passenger Age Groups & Income: Certain age groups and income range groups will take more rides and spend more money than the others.
- Profit Trends: One cab company will outperform the other in profits.
- Future Ride & Profit Forecast: The number of rides and profit will increase in 2019 for both companies.



Questions

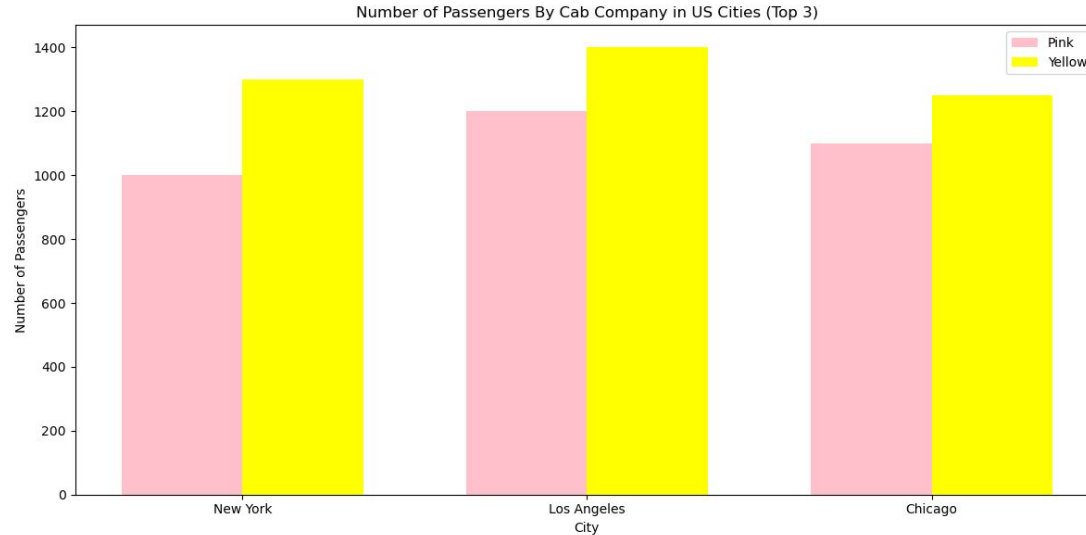
- What is the profit according to distance, year and city with each cab company?
- What is the yearly profit by each cab company?
- How does gender play a role in the profit gained per cab company?
- How does passenger income levels play a difference in the profit?
- What is the profit share, if we look at the age of each passenger?
- What is the retention rate of passengers by cab company?
- What are the number of passengers taken by each cab company?
- What are the quarterly profits by each cab company from 2016-2018?
- What are the top 3 cities that have the highest numbers of passengers in the US?
- How does the season affect the number of passengers using the cab service?
- Which company has the highest amount of passengers at one time?
- Does profit increase with the number of passengers?
- Which cab company is recommended for future investments?

Test #1: Is there a increase during certain holidays during the year?



- From reading the data and the chart depicted above, we can come to the conclusion that the most passengers riding with the cab companies are on Christmas being at the top of the list and Thanksgiving holiday coming in 2nd place. The lowest amount of passengers is during Halloween.

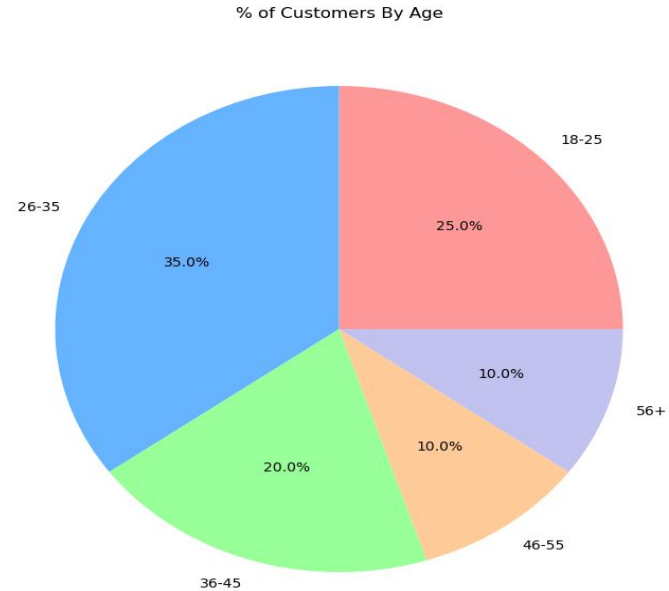
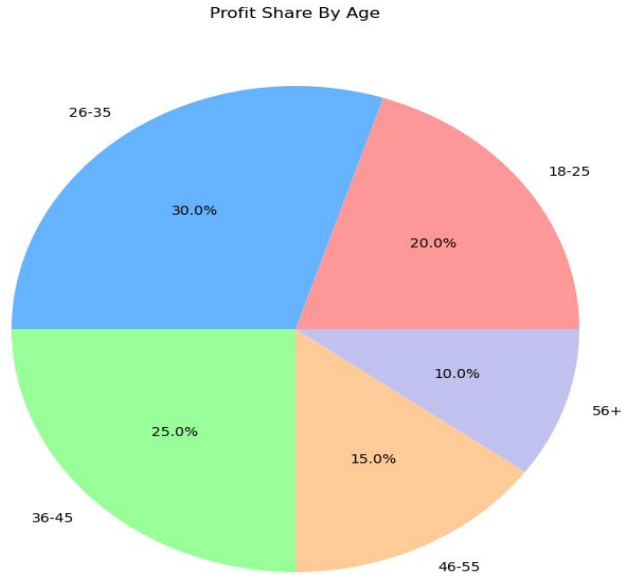
Test #2: One cab company has a higher number of passengers in specific cities in the United States.



- The data above shows that the Yellow Cab is outperforming the Pink Cab in all 3 US cities. Los Angeles being the number one city in the United States.

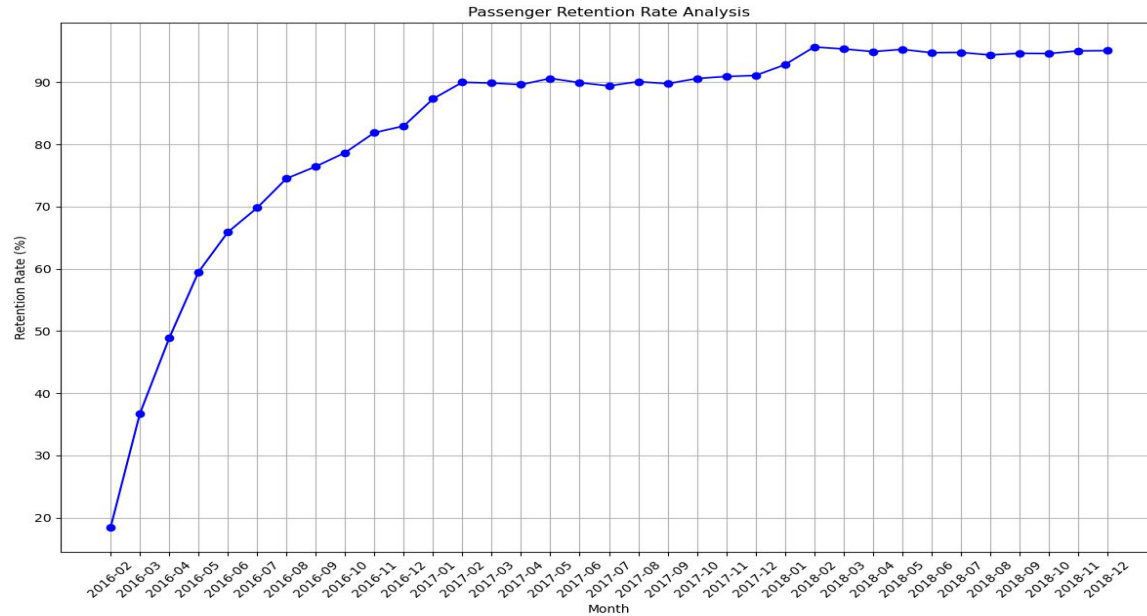
Test #3:

Certain age groups and income range groups will take more rides and spend more money than the others.



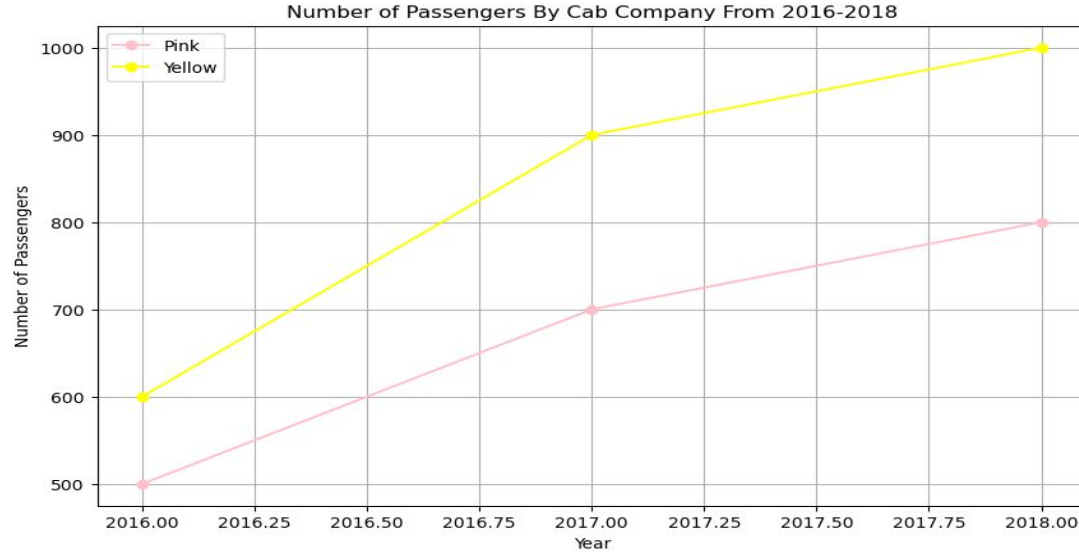
- By analyzing the data in the pie charts, we see that: there is a higher profit share in percentage from the age group: 26-35. We also see that there is a 35% customers in that age group that represent the profit share. The lowest profit share comes from the age group of 56+.

Test #4: What is the retention rate of passengers by cab company?



- By analyzing the data in the line chart above, we see that: the amount of passengers from the years 2016-2018. From the line chart depicted below, it shows that in 2016 it was less than 20%, the rate of retention. In 2018, it shot up to over 90%.

Test #5: What is the comparison in passengers between the Yellow and Pink Companies throughout 2016-2018?



- By analyzing the data in the line chart above, we see that: the amount of passengers have increased for the Yellow Cab company from the years 2016-2018.

Summary/Recommendations/Conclusions

Summary

- 1) The method of payment used does not matter whether it is card or cash.
- 2) Demand is higher during the weekends.
- 3) There is an impact of the income class and age group range vs the amount of rides being taken.
- 4) There are (2) top holidays that outperform in rides and profits versus the other holidays throughout the year.
- 5) The data analysis and information collected can help XYZ investment firm make a better collective decision in what company to go with for their future investment.

Conclusions

- 1) Passenger Reach : The Yellow Cab has generated more passengers and rides than the Pink cab company.
- 2) Transactions in Forms of Payments- there are more card payments than cash payments for rides.
- 3) Passenger Retention Conclusion: passenger retention rate is retained from 20% to 90% from 2016 to 2018.
- 4) Passenger Age Group Observation: the age group 25-35 seemed to order more rides than the other age groups.
- 5) Passenger Income Class: there is a increased profit by 2x from passengers from the income class of earning between 15,000-25000 USD per month.
- 6) Distance Relationship to Amount of Passengers: the higher the kilometers the increased number of passengers but at 4 kilometers there was a slight decline of passengers.
- 7) Top 3 US Cities in Rides: NYC, Los Angeles and Chicago have the most rides in the US. Los Angeles being at #1, NYC at #2, and Chicago as #3.
- 8) Quarterly Profits- Which company performed better?: the Yellow Cab company outperformed the Pink Cab company in all quarters from 2016-2018.
- 9) Most Rides Taken During Which Day of The Week?: The most rides taken were Friday through Sunday for both cab companies. The Yellow Cab company has taken more rides than the Pink Cab company.
- 10) Passenger Preference During the Holidays: the passengers are riding the most during Christmas and Thanksgiving. Halloween has performed the lowest amount of rides in the US.

Recommendations

- 1) Based off the data collected and the data analysis conducted, investors from the XYZ firm should look into investing into the Yellow Cab company.
- 2) The Yellow Cab company has increased profits and rides in all four quarters of every year and has superseded and dominated the cab industry compared to of the Pink Cab company.
- 3) Investors should also consider that profits increase with the distance driven per cab, so to look into these insights and capitalize on the demand of longer trips.
- 4) Based off of certain passenger demographics, certain age groups and income classes are more likely to spend more money on rides versus the other groups.
- 5) It can be a lucrative and profitable investment, but market research and data collection should be analyzed and compared before diving into this investment.