



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

G2M Insight for Investment in Cab Firm

August 2023

Siobhan Hwang

Agenda

Objective Statement

Candidate Summary

Approach

Exploratory Data Analysis (EDA)

EDA Summary

Recommendation

Objective Statement

Investor: XYZ

Purpose: G2M strategy for investment

Candidates: Yellow Cab, Pink Cab

Intended Result: Recommend
investment in one of the candidates
based on historical data exploration

Candidate Summary 2016-2018

	Yellow Cab	Pink Cab
Average Yearly Profit	14,673,460.00	1,769,109.33
Total Customers	39,896	32,330
Total Rides	27,4681	84,711
Top Cities (Total Rides)	1. New York, NY 2. Chicago, IL	1. Los Angeles, CA 2. New York, NY
Top Cities (% Total Customers)	1. Washington DC 2. New York, NY	1. Nashville, TN 2. Sacramento, CA

Approach

Statistically manipulate the available data to provide visual insight into the candidates' history from 2016-2018

Available Data 359,352 Data Points

- Transaction IDs
- Cities of Service
 - Population
 - Total Cab Users
- Customer Information
 - Customer ID
 - Gender
 - Age
 - Monthly Income
- Ride Information
 - Date of Travel
 - Kilometers Travelled
 - Price Charged
 - Cost of Trip
 - Payment Method
 - Profit (=Price-Cost)

Ride Analysis

Total Rides by Company

Total Rides by Company per Year

Percent of Total Rides by Year

Total Rides by Company per Month

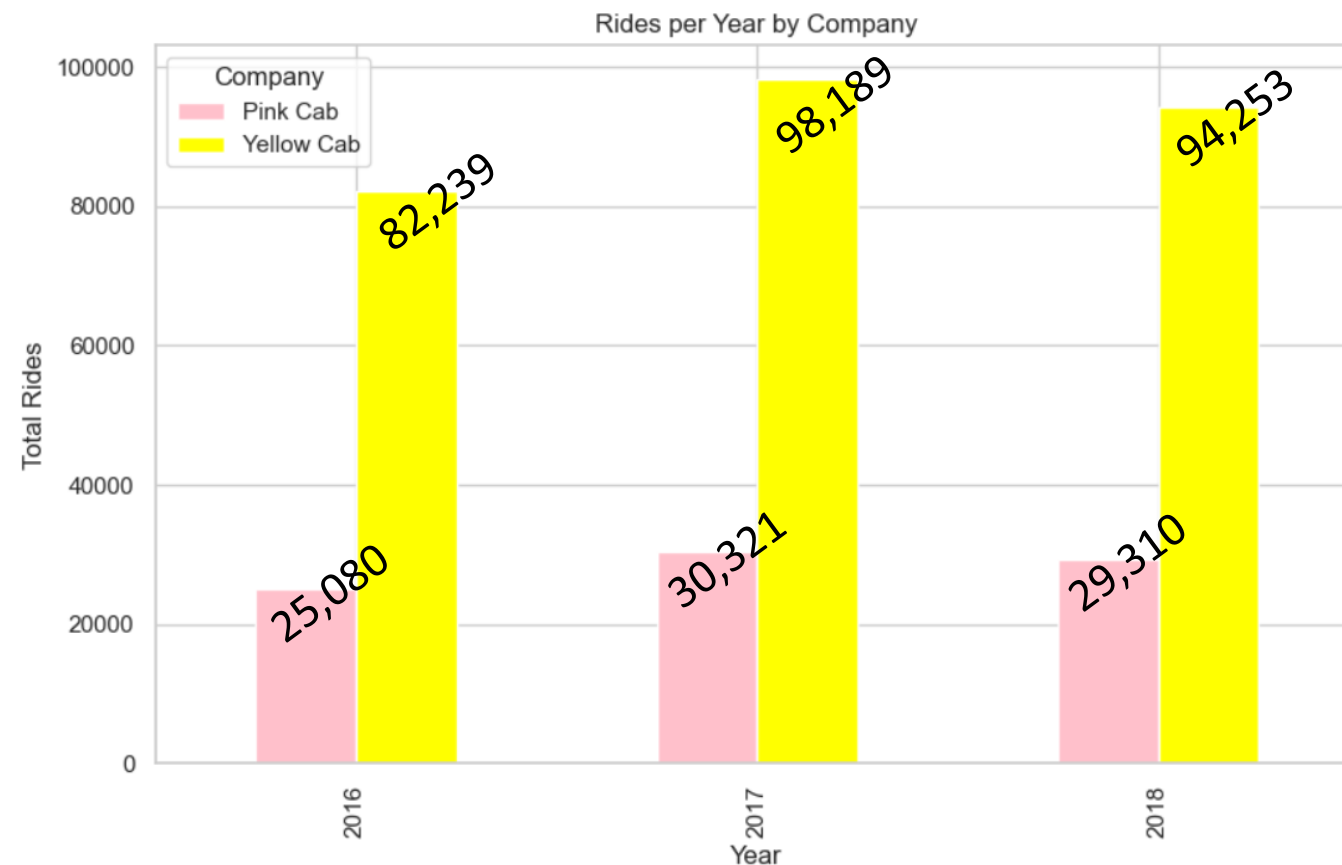
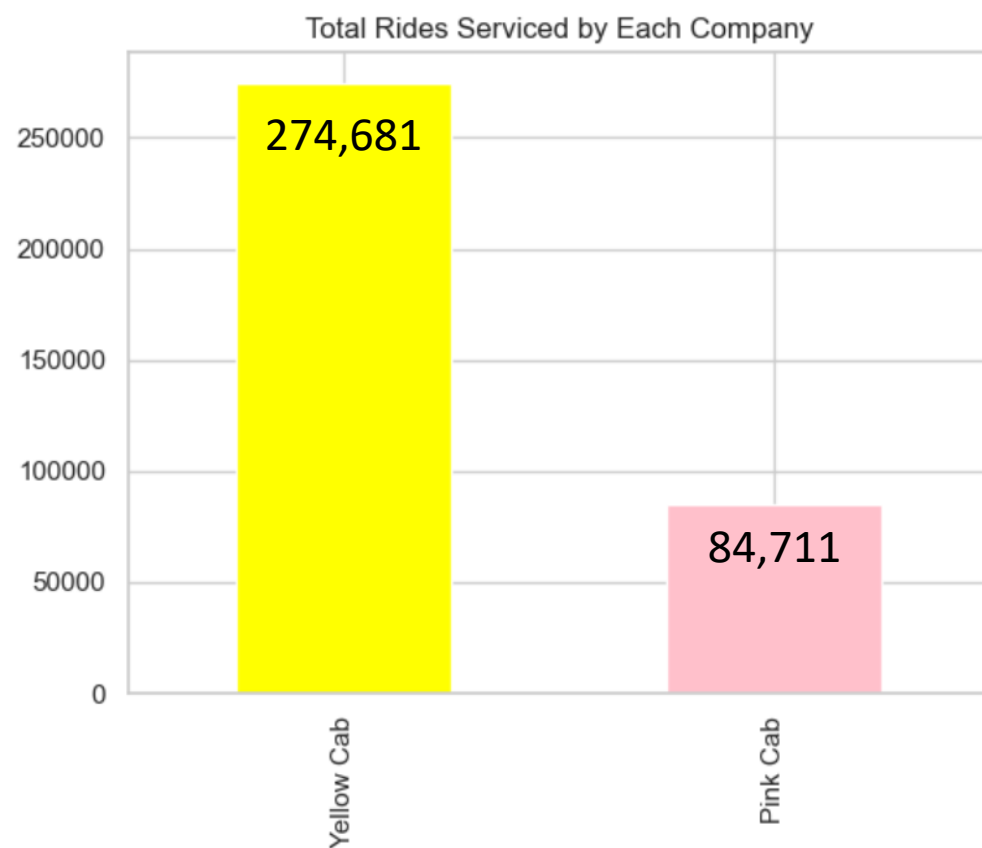
Percent Change in Total by Year

Total Rides by City

Percent of Total Rides by City

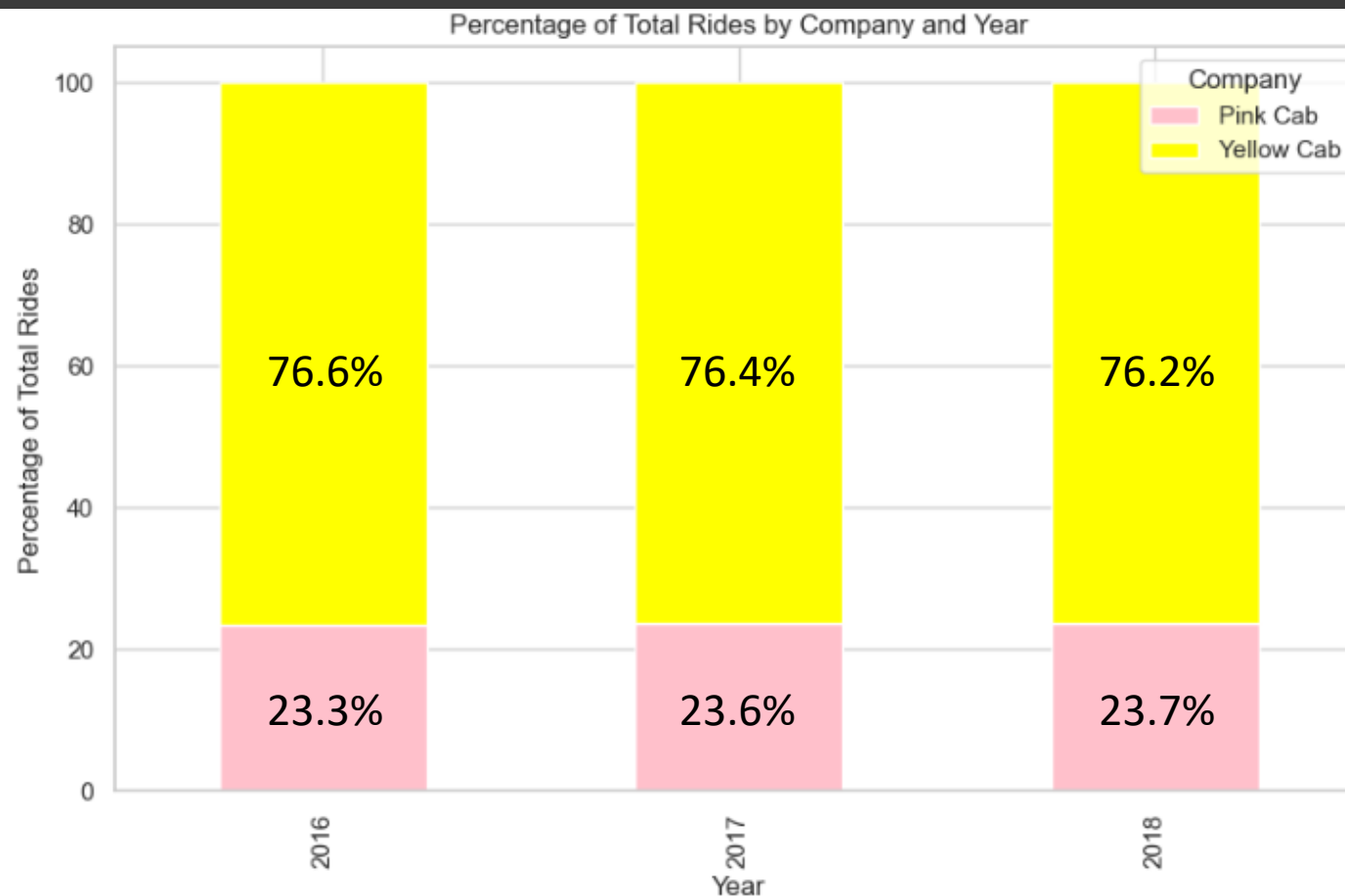
Ride Analysis

Total & Year Total



Ride Analysis

Percentage by Year

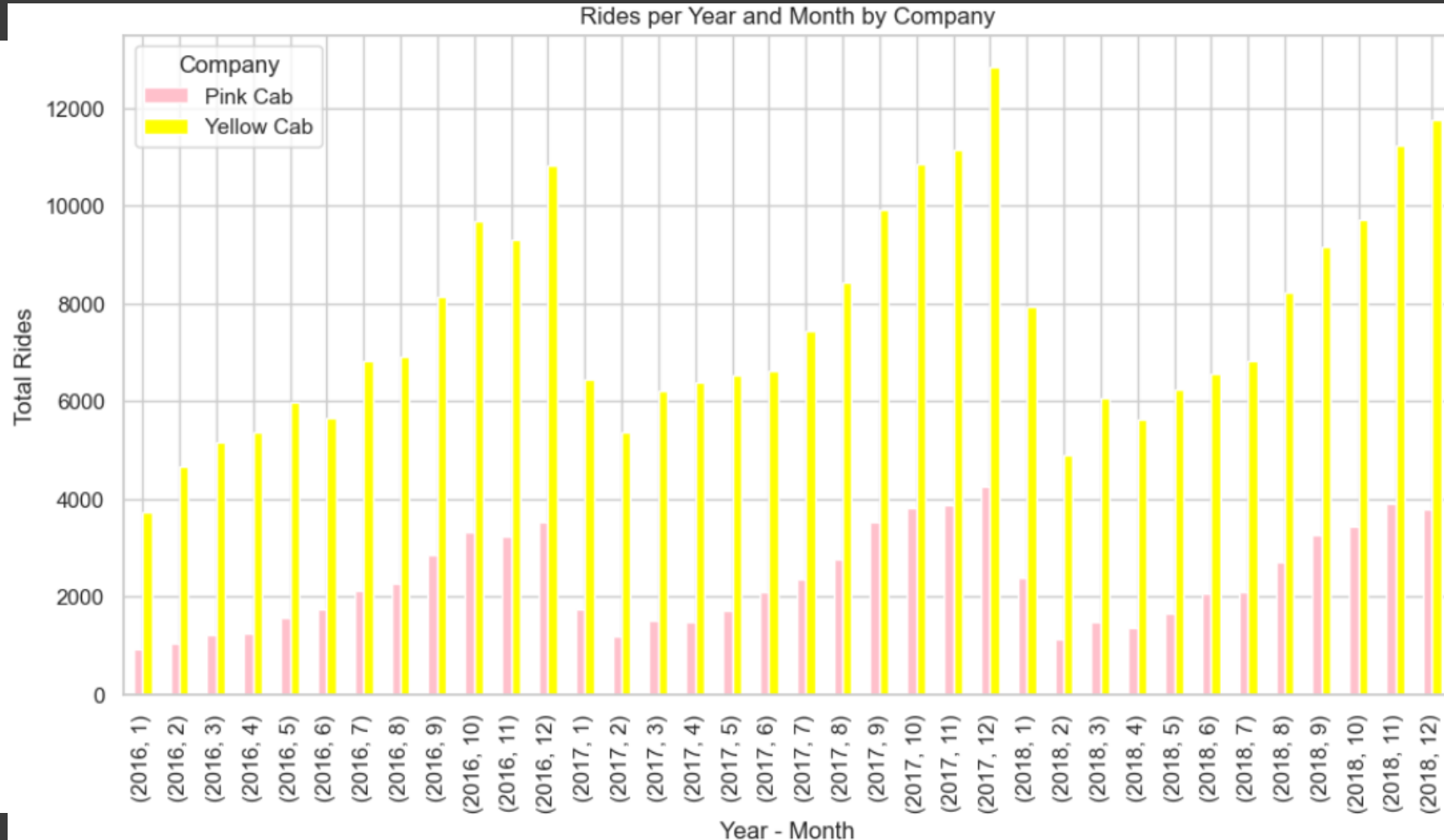


Ride Analysis

Total per Month

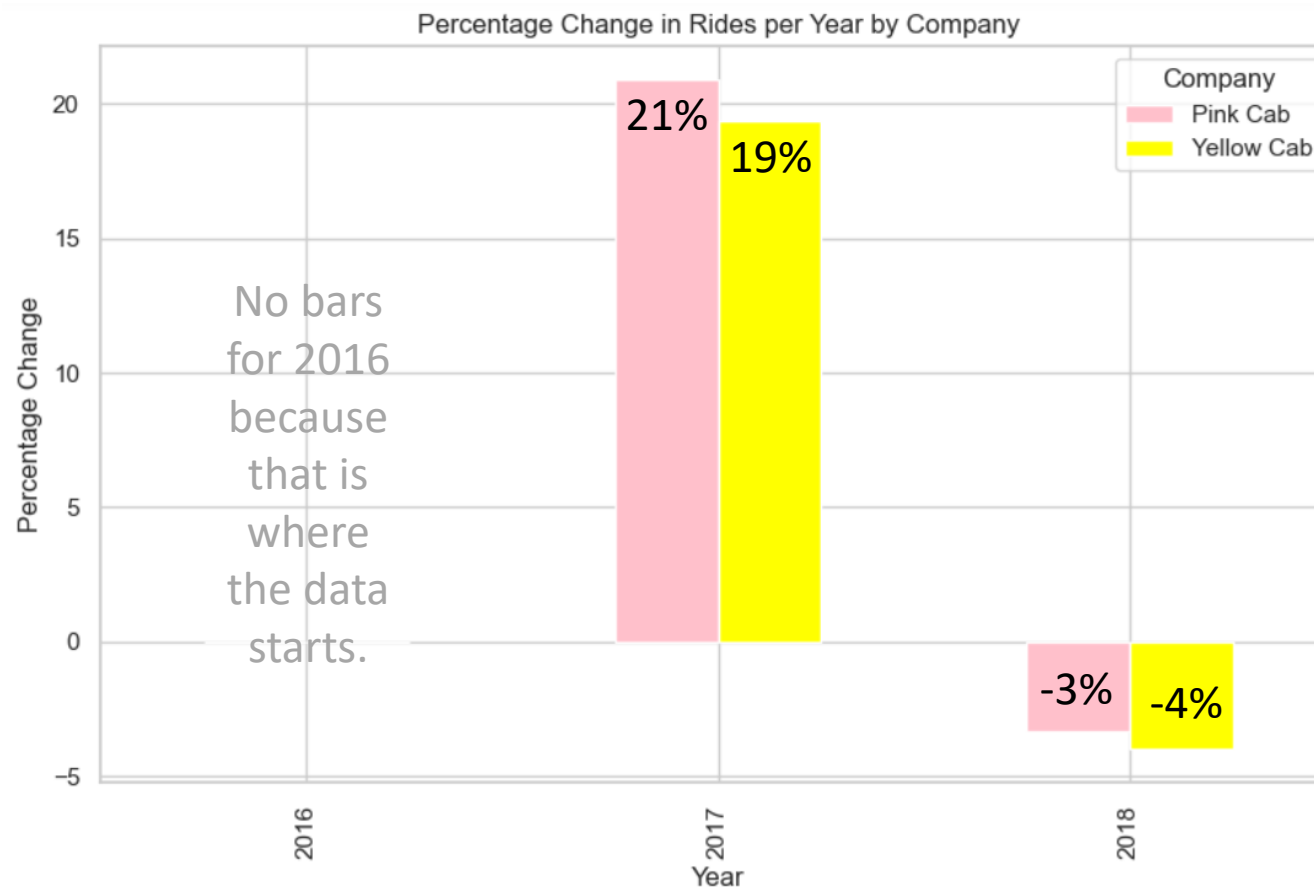
Trends

- Changes in number of rides by month follow a similar pattern irrespective of company
- December consistently sees the most rides serviced for Yellow Cab over all years and for Pink Cab in 2016 and 2017.



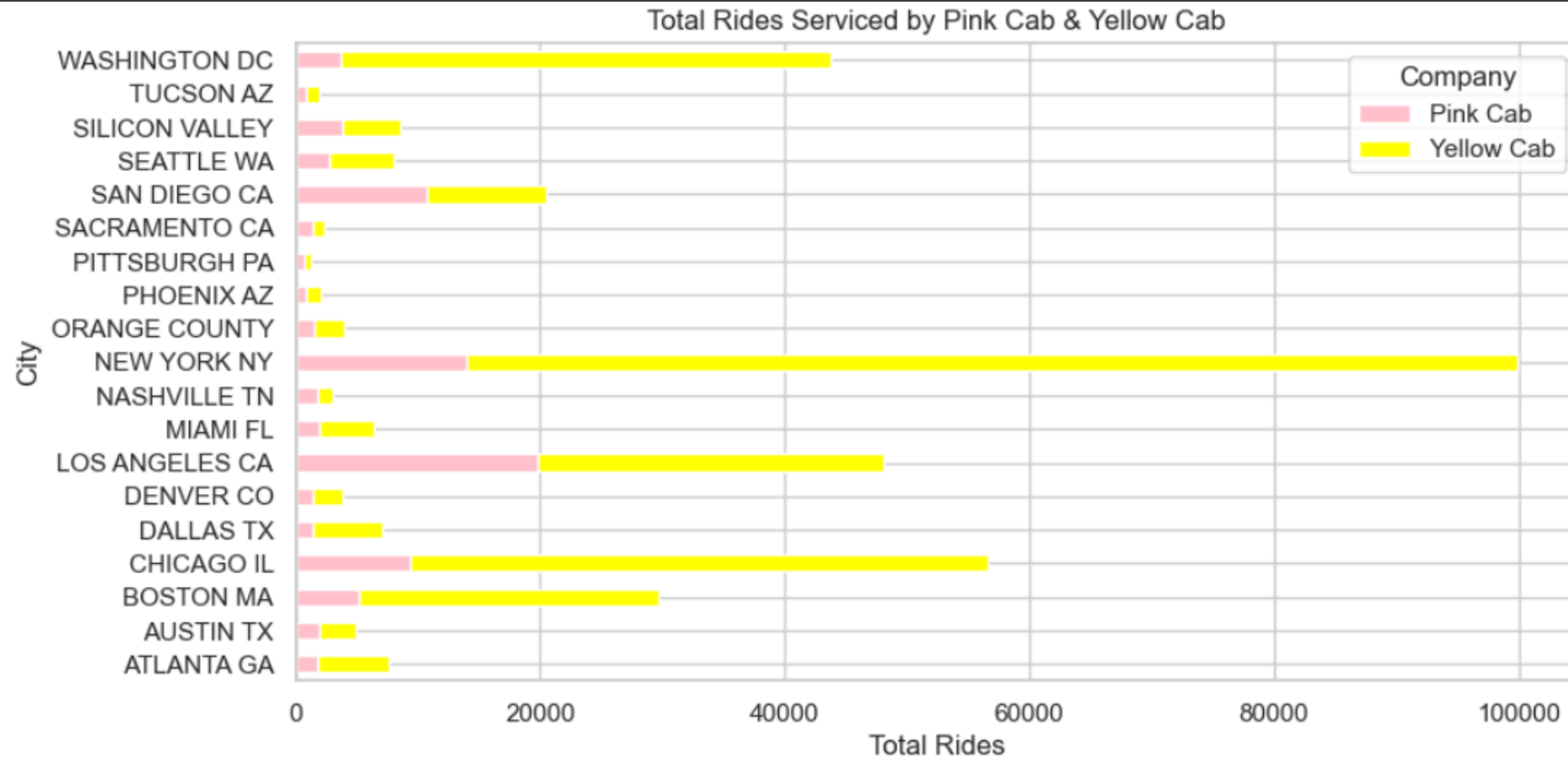
Ride Analysis

Percent Change



Ride Analysis

City Totals



Top Yellow Cab Cities

1. New York, NY
2. Chicago, IL
3. Los Angeles, CA
4. Washington DC
5. Boston, MA

Top Pink Cab Cities

1. Los Angeles, CA
2. New York, NY
3. San Diego, CA
4. Chicago, IL
5. Boston, MA

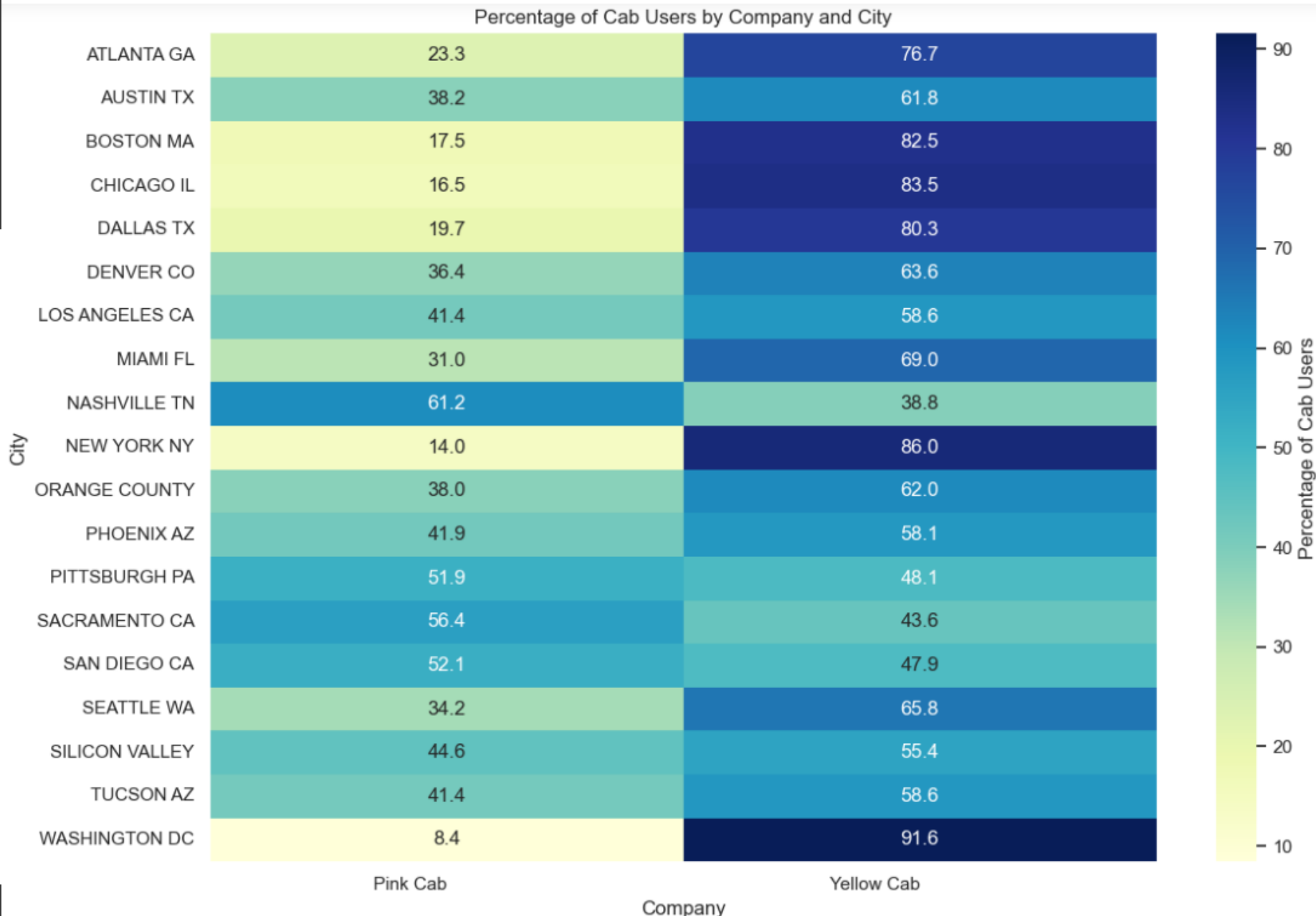
Ride Analysis

Yellow Cab Majority Cities

1. Washington DC
2. New York, NY
3. Chicago, IL
4. Boston, MA
5. Dallas, TX
6. Atlanta, GA
7. Miami, FL
8. Seattle, WA
9. Denver, CO
10. Orange County
11. Austin, TX
12. Tucson, AZ
13. Los Angeles, CA
14. Phoenix, AZ
15. Silicon Valley

Pink Cab Majority Cities

1. Nashville, TN
2. Sacramento, CA
3. San Diego, CA
4. Pittsburgh, PA



Customer Analysis

Total Customers by Company

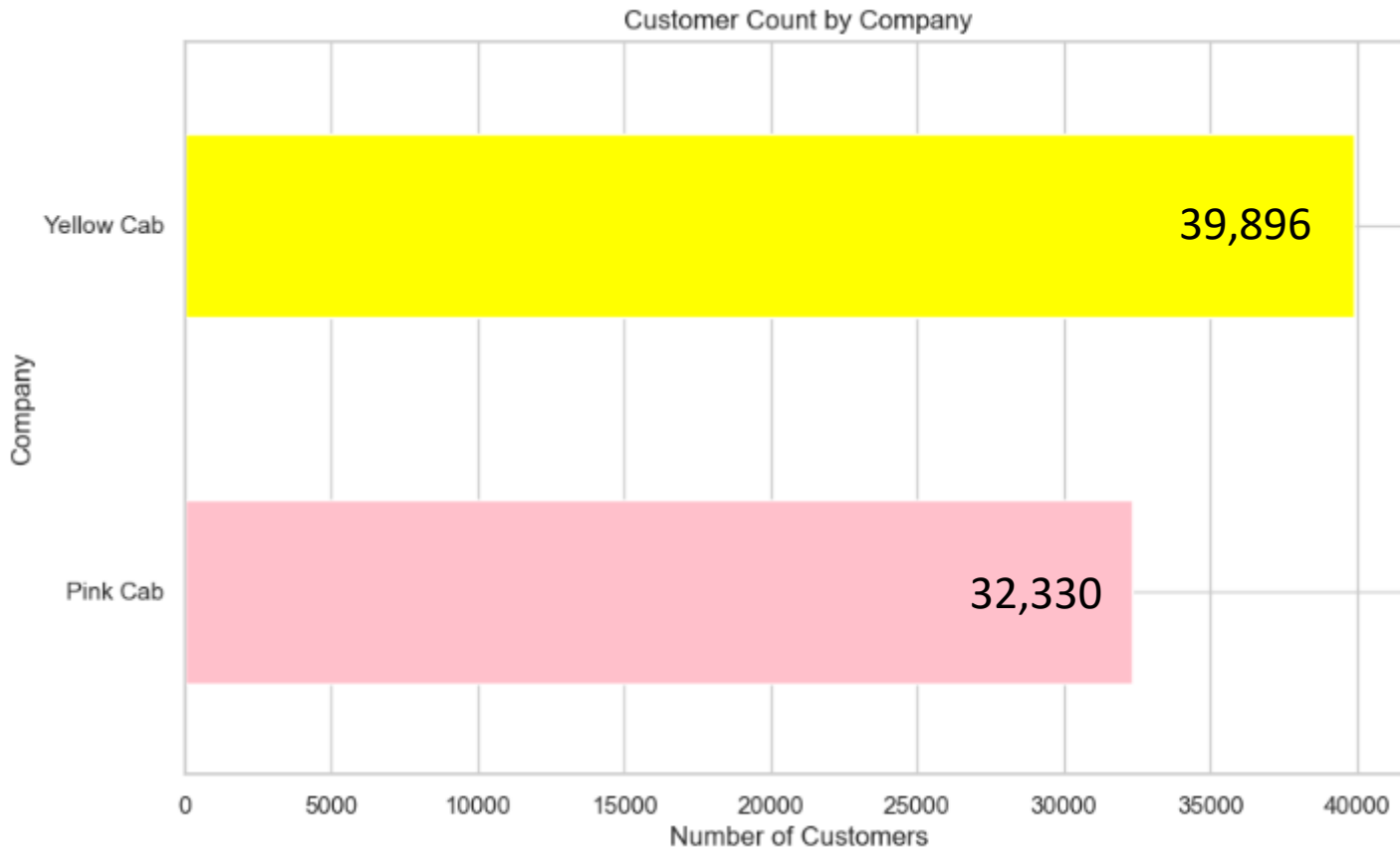
Distribution of Age by Gender

Distribution of Monthly Income by Age Group

Percentage of Multiple-Ride Customers

Customer Analysis

Total



Noteworthy Points

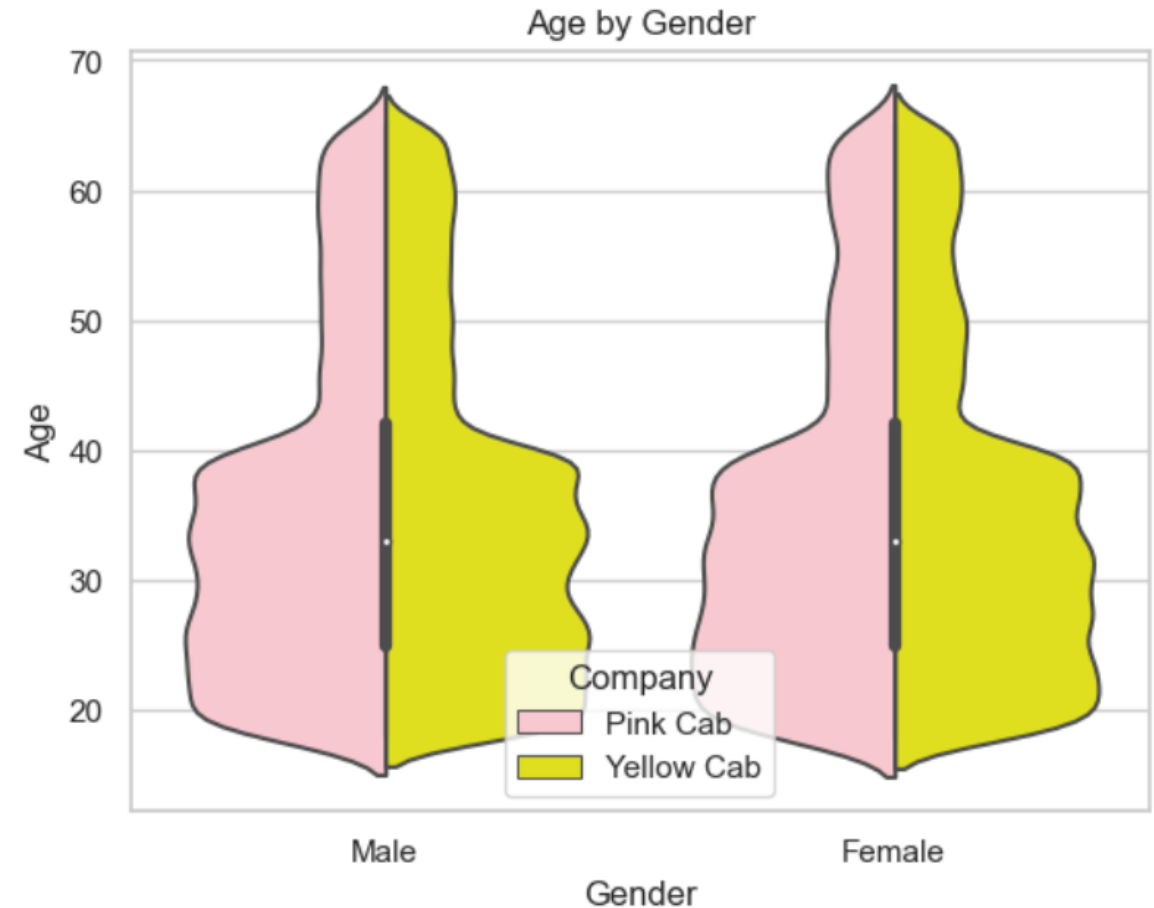
- Yellow Cab has about 23.4% more customers than Pink Cab.
- According to previously included data*, Yellow Cab also consistently services more than 76% of total rides per year.

* [Percentage by Year](#)

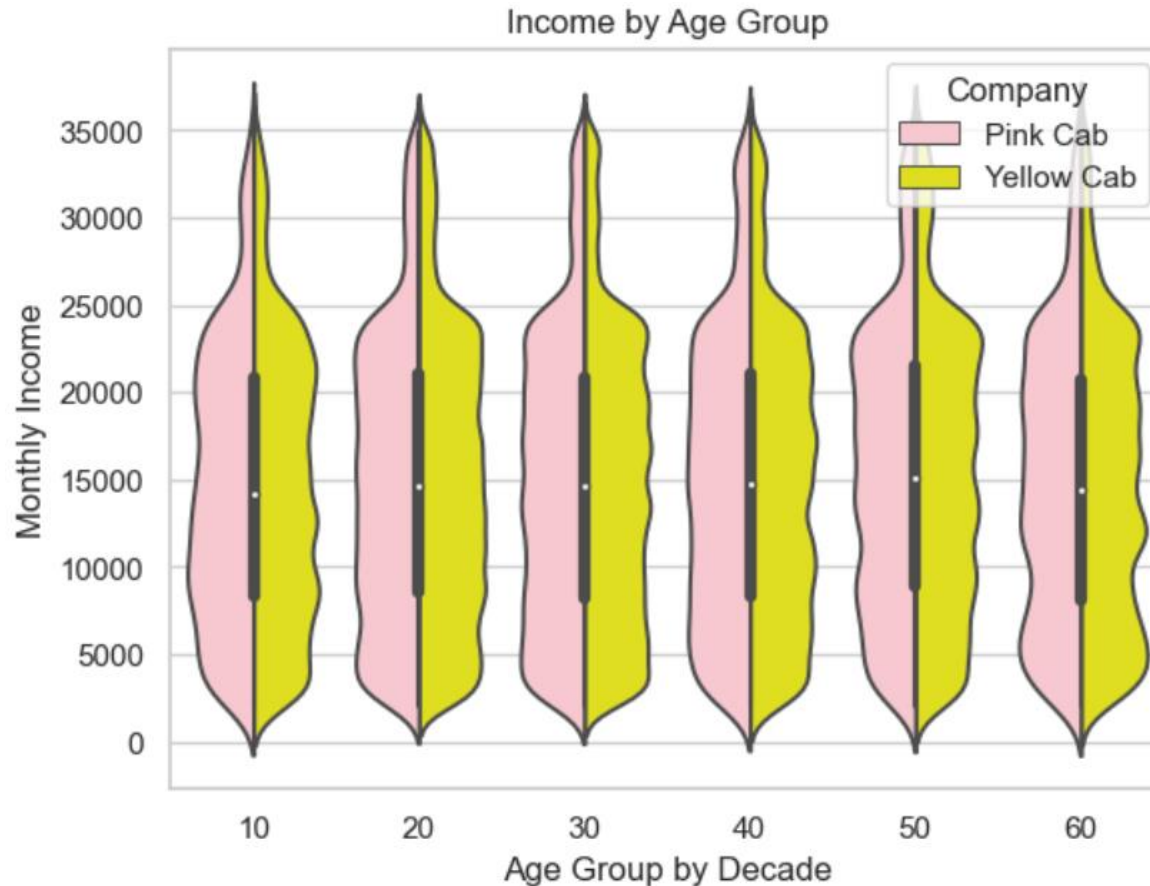
Customer Analysis Age-Gender

Interpretation

- Both Yellow Cab and Pink Cab have very similar demographics in their customers when looking at age and gender.
- The mean age of customer for both companies is 35.



Customer Analysis Age-Income



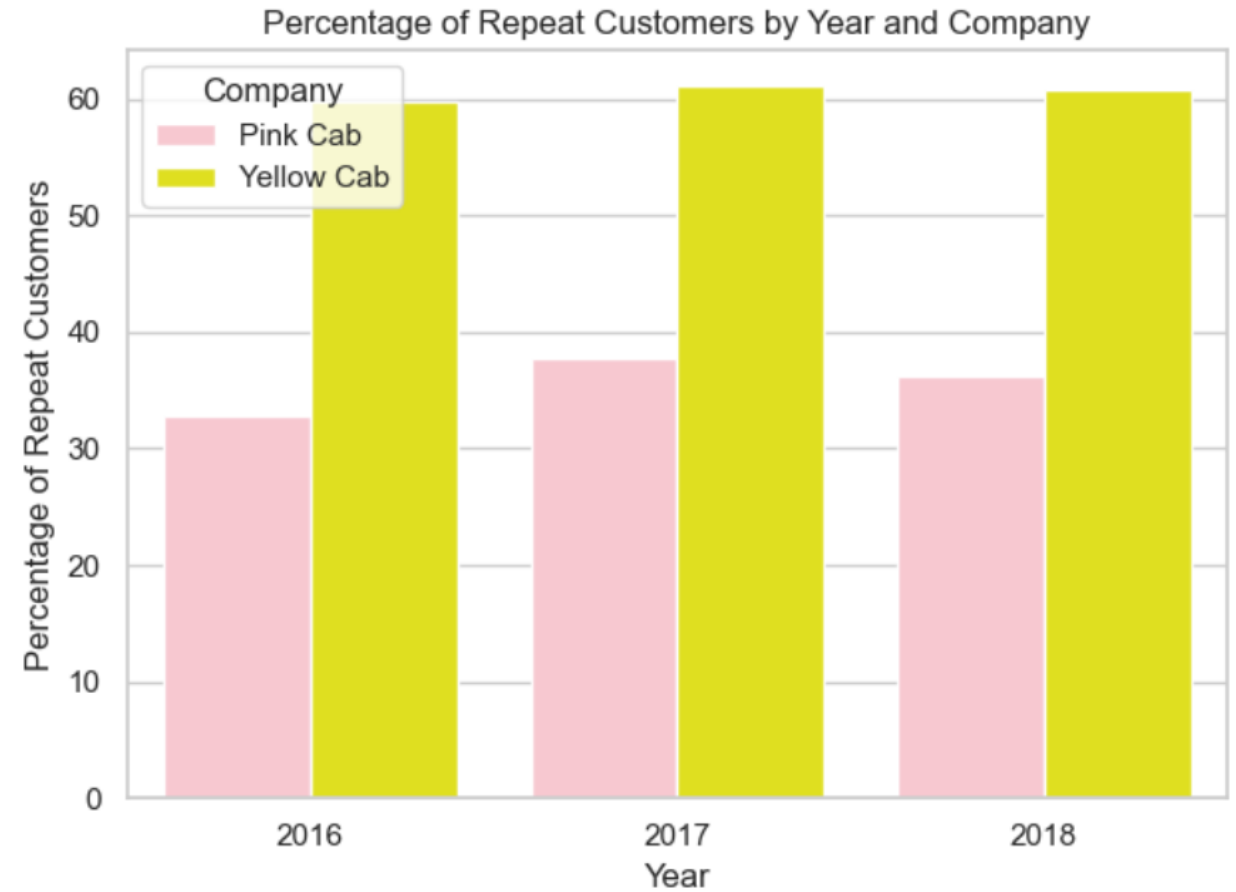
Interpretation

- Both Yellow Cab and Pink Cab have very similar distribution of wealth across their customer base.
- The highest earning group is 50-year-olds with a monthly average of 15,400-15,600 USD.
- The lowest earning group is teenagers with a monthly average of 14,700-14,800 USD.

Customer Analysis Multiple Rides

Trends

- Yellow Cab had multiple transactions from roughly 60% of their customers per year.
- Pink Cab had multiple transactions from less than 40% of their customers per year.



Profit Analysis

All figures in USD

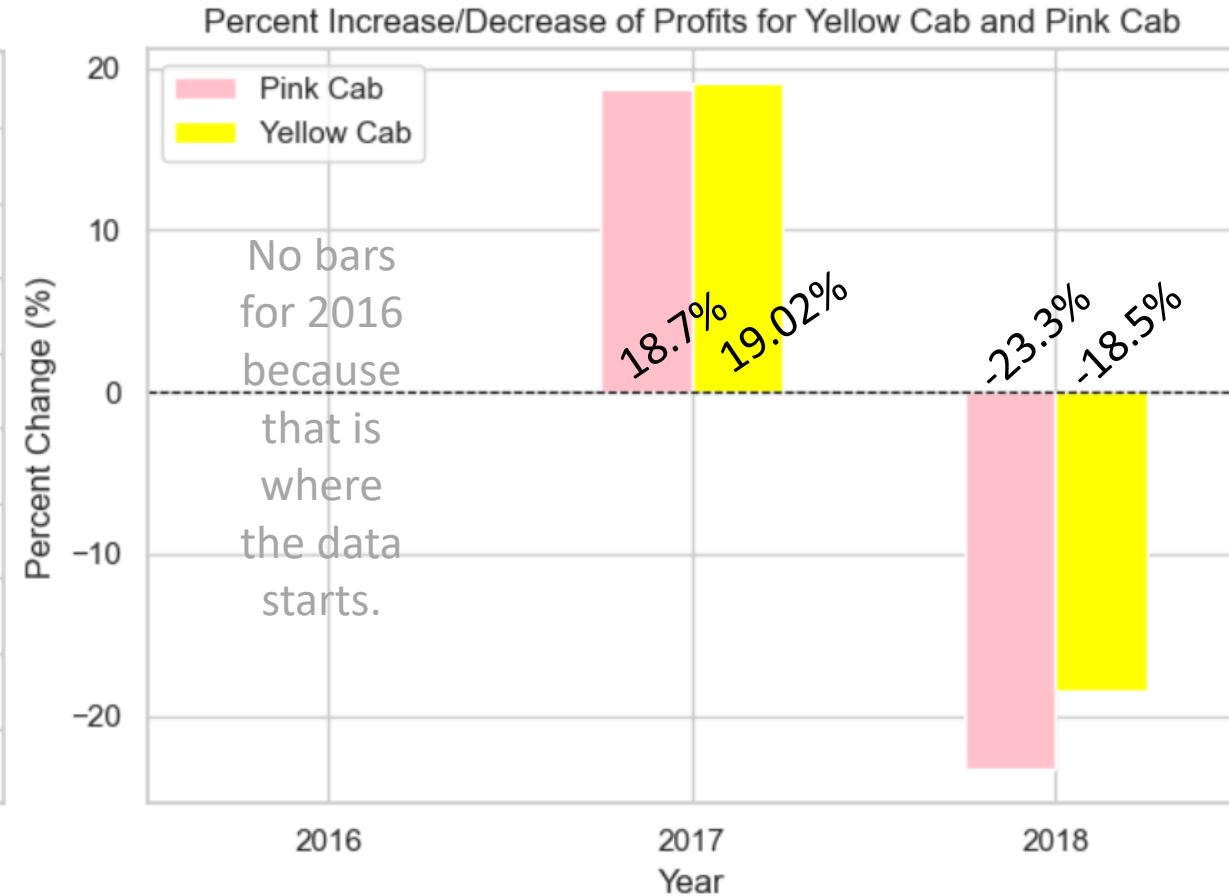
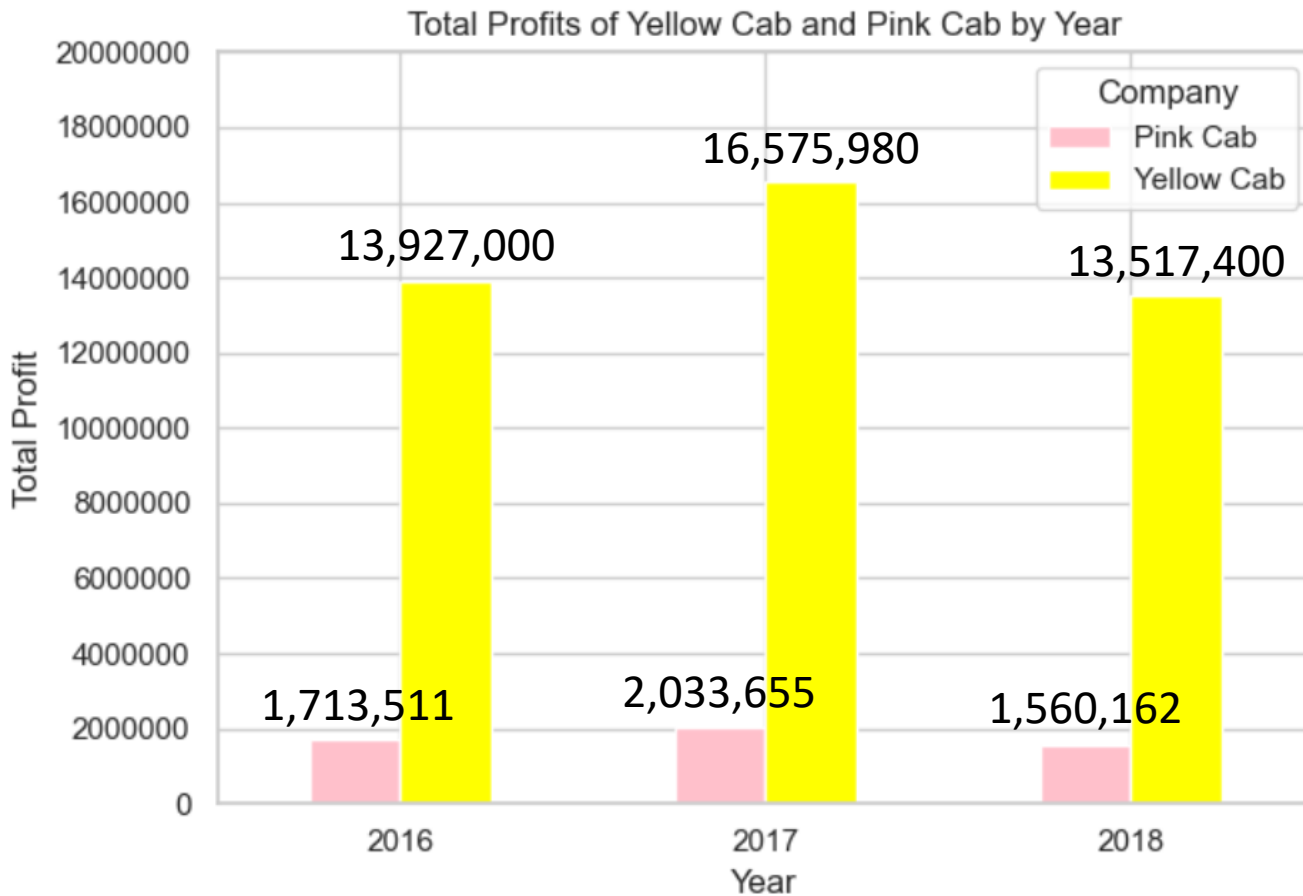
Total by Year

Average Cost per Ride

Average Profit per Ride

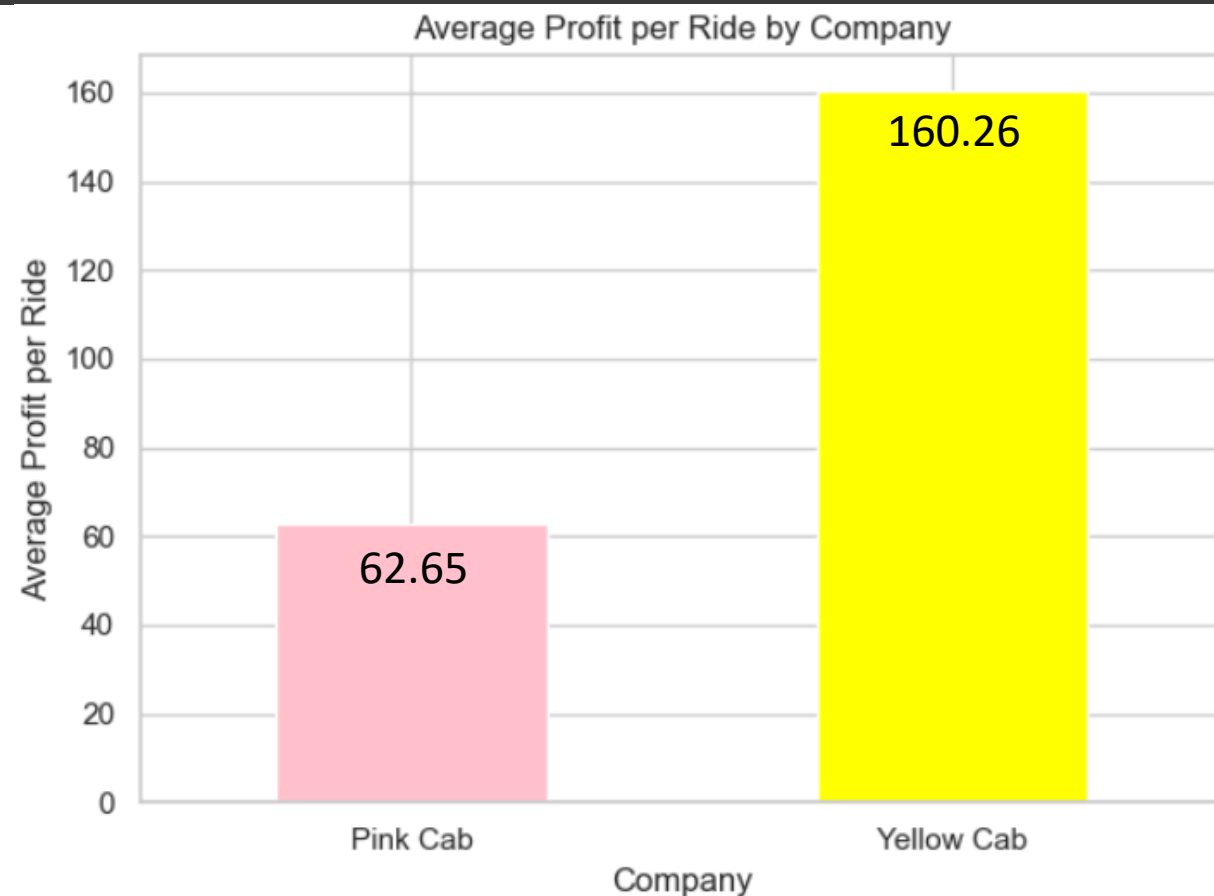
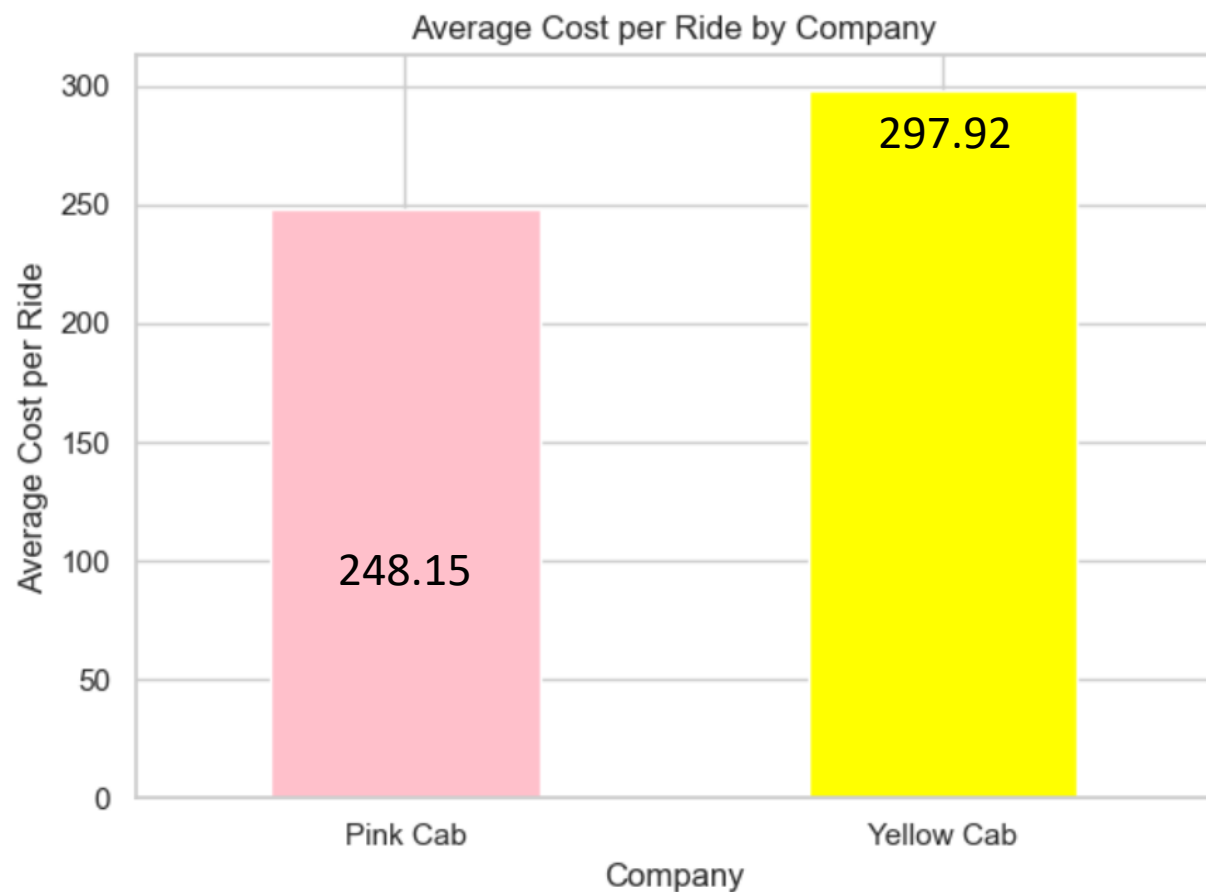
Profit Analysis

Total by Year Percent Change



Profit Analysis

Average per Ride



EDA Summary

Rides

- Yellow Cab serviced the most rides in all categories

Cities

- Yellow Cab dominated the market in over 70% of cities

Customers

- Yellow Cab had roughly 7.5k more customers, 23% more than Pink Cab.
- Yellow Cab serviced about 76% more rides per year than Pink Cab.
- Age, gender, and monthly income distributions are comparable for both companies.
- About 60% of Yellow Cab customers have multiple rides each year, whereas less than 40% of Pink Cab customers have multiple rides a year.

Profit

- Yellow Cab had more total profits by far per year.
- Yellow Cab's average cost per ride is about \$50 more than Pink Cab's.
- Yellow Cab's average profit per ride is about \$100 more than Pink Cab's.
- Pink Cab's profits decreased by 23.3% in 2018, whereas Yellow Cab's profits decreased by 18.5%.

Recommendation

Yellow Cab

Thank You