Data Glacier Data Report

blaisepke@gmail.com

Problem Statement

- 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.

DATA INTAKE

- With an initial look at the data we have 7 columns and 359,392 records.

- The 7 columns are:

1. **Transaction ID:** This is a unique identification for every transaction by a taxi.

2. **Date of Travel:** The date of travel is meant to indicate the date of the transaction/ trip

3. **Company:** This column specifies the cab company that the transaction belongs to, there are two columns the Pink and Yellow Cab company

4. **City:** This column gives information of the location of the transaction

5. **KM Travelled:** This captures the number of KM the trip covered

6. **Price Charged:** This is the total price that was due by the cutomer who took the taxi

7. **Cost of Trip:** The cost of the trip is the amount the tip cost the cab company.

ASSUMPTIONS

- Assuming the start date of collecting the data is 1-1-2016 and the end was 30-12-2018.
- The outliers are to kept as they help understand cab trips/transactions
- There are no other cab companies in the mentioned states
- The difference between the cost of the trip and the price charged is the profit; a negative profit is considered a loss
- In the day_week 0 is Monday through to 6 which is Sunday
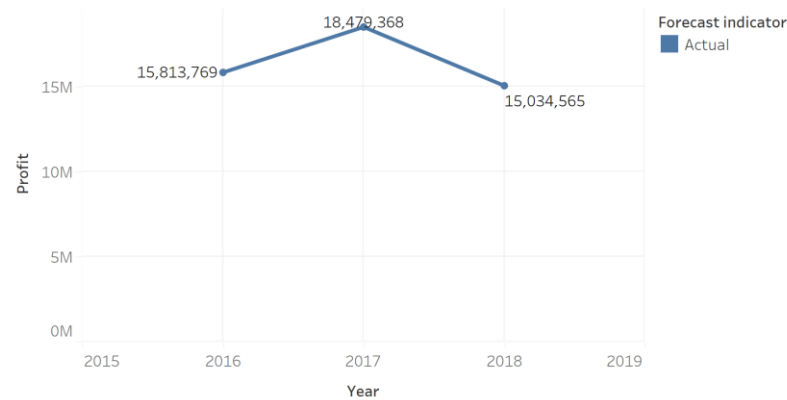- Each user uses only a single account (User ID)

Transactions per year
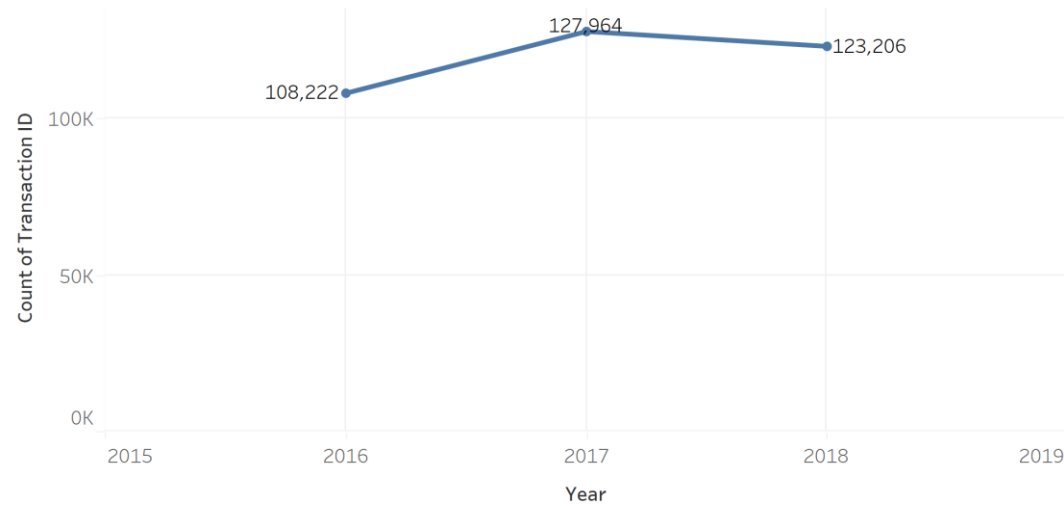
The trend of count of Transaction ID for Year. The marks are labeled by count of Transaction ID.

The cab industry saw an increase in number of rides from the year 2016 to 2017. This however, slightly   dropped in the year 2018.

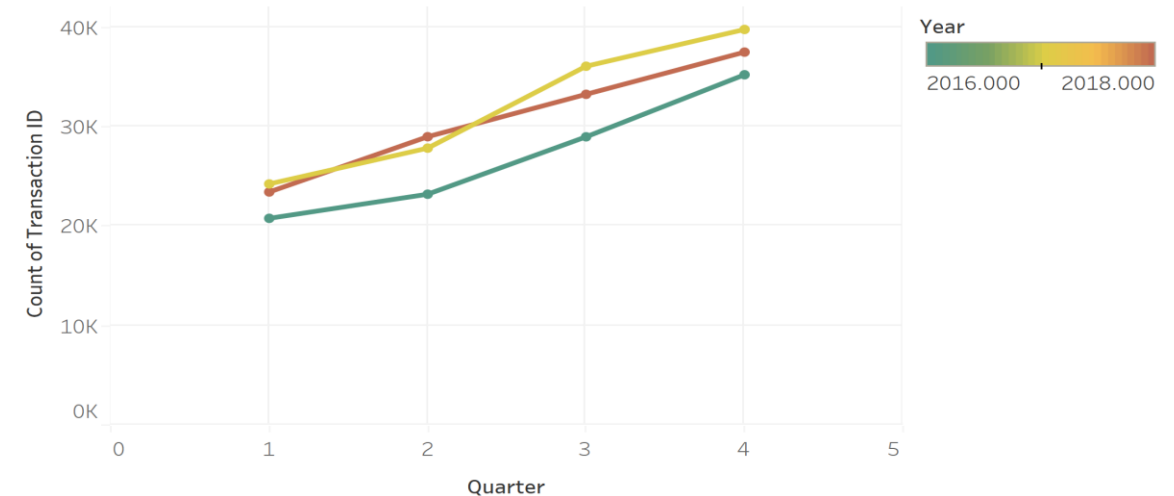# CAB Industry Analysis

Yearly profit analysis per company

The year 2018 saw the lowest profit gain in the cab industry in the past three years with 2017 being the most profitable with a margin of over $2 million from 2018.

This could probably be to the rise of cab applications such as uber in the year 2018. Which had a huge impact on other cab companies
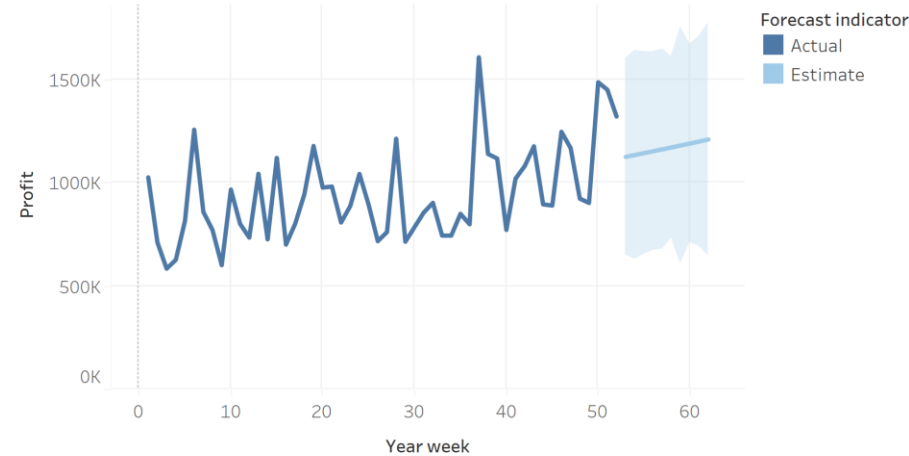
The trend of sum of Profit (actual & forecast) for Year. Color shows details about Forecast indicator. The marks are labeled by sum of Profit (actual & forecast). The data is filtered on Company, which keeps Pink Cab and Yellow Cab.

The cab industry experience a gradual increase in users as the quarters progress. The fourth quarter sees the highest usage of the cabs in every year

Quarter vs Transaction/Traffic

The trend of count of Transaction ID for Quarter. Color shows details about Year.
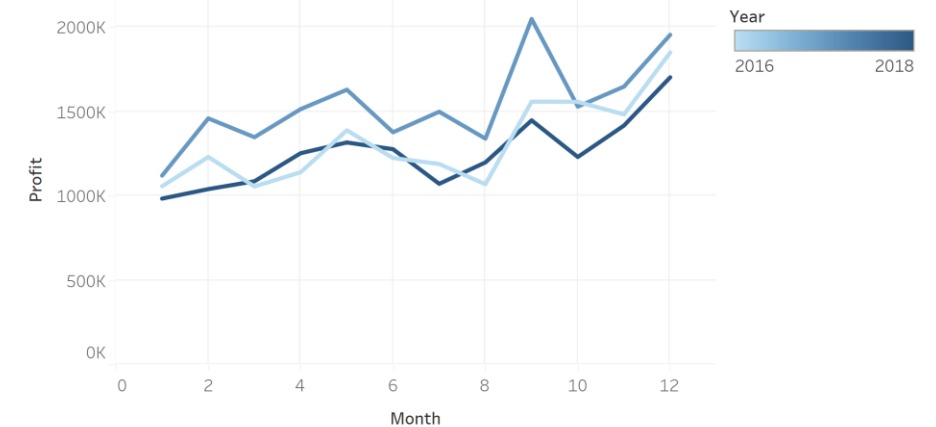
CAB Industry Analysis

Profit per week

The trend of sum of Profit (actual & forecast) for Year week. Color shows details about Forecast indicator.

Distribution of profits over month over years

The trend of sum of Profit for Month. Color shows details about Year. Details are shown for Year.

The industry sees peak profit in week 37 of the year.
Week 37 is early September in the years.
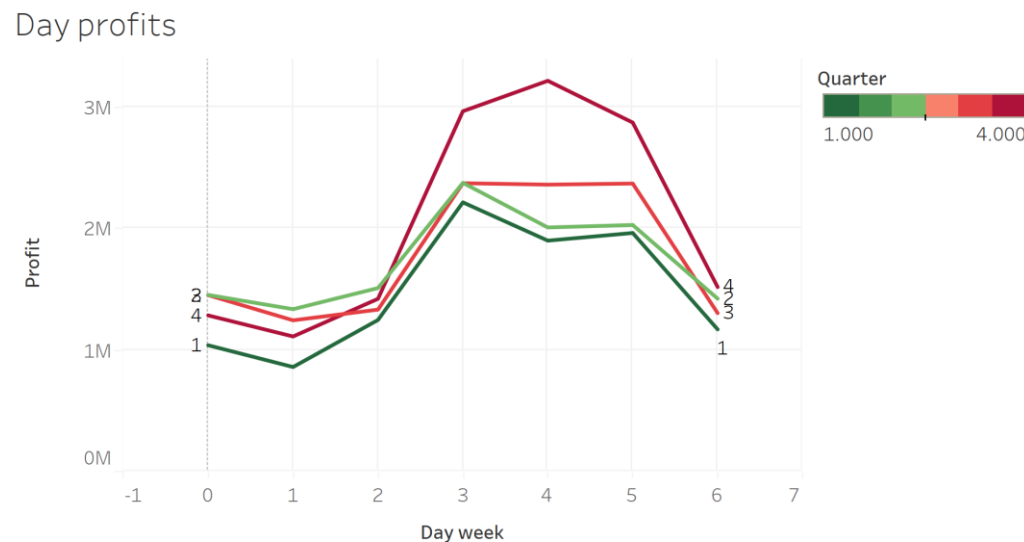This can be due to it being the last few weeks of summer and lots of travelling happens then

With the same analysis we do see that September is the most profitable month, accredited to the summer holidays which generate profit for the cab company

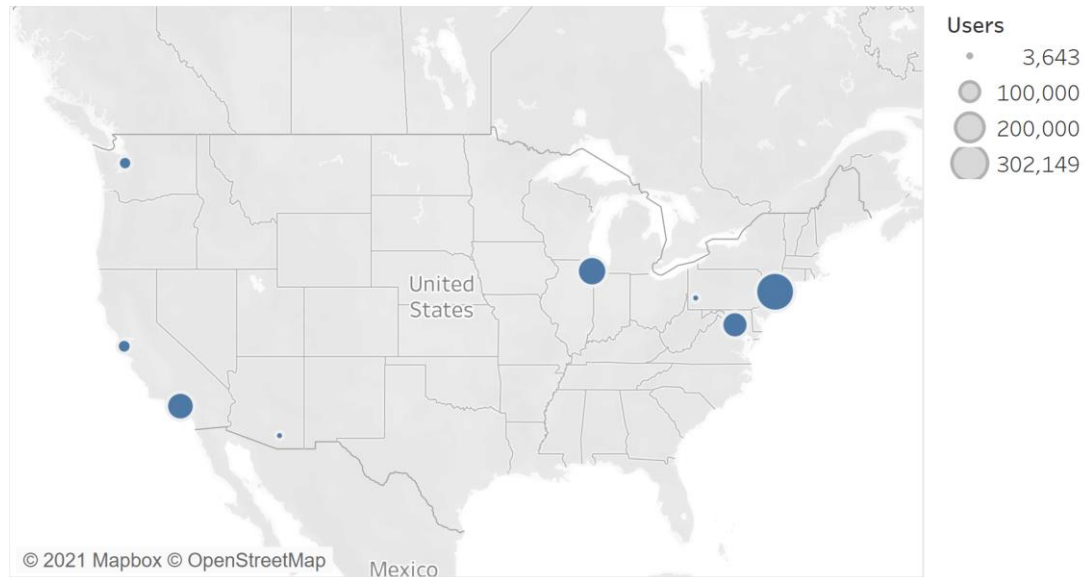# CAB Data Analysis

Day profits

The trend of sum of Profit for Day week. Color shows details about Quarter. The marks are labeled by Quarter.

The cab industry generates most profit mid week with spikes witnessed in the fourth quarter of the year.
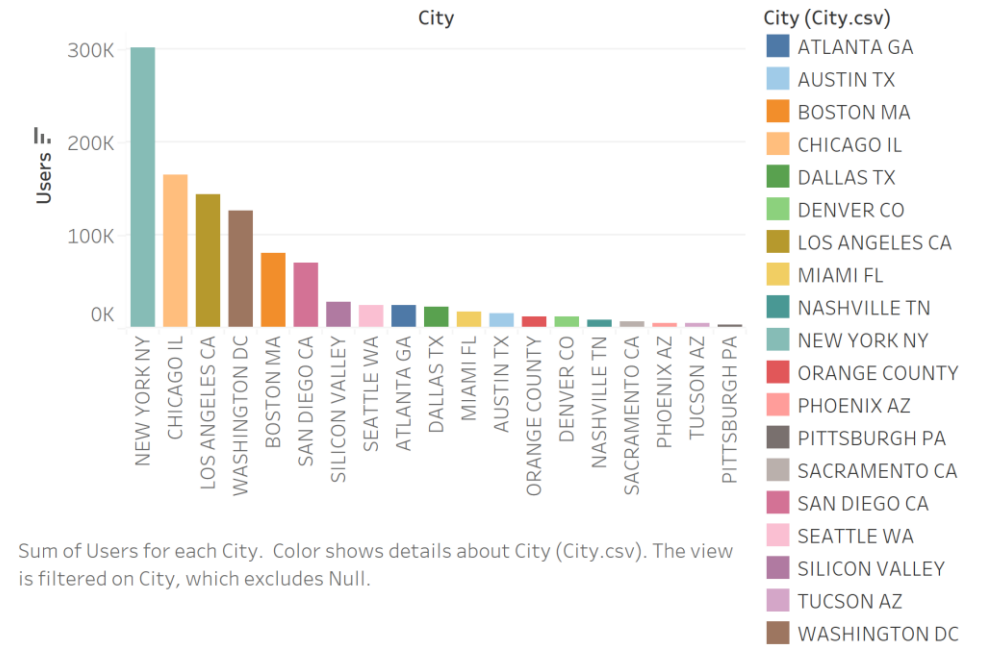
User Analysis

City vs Users

Users
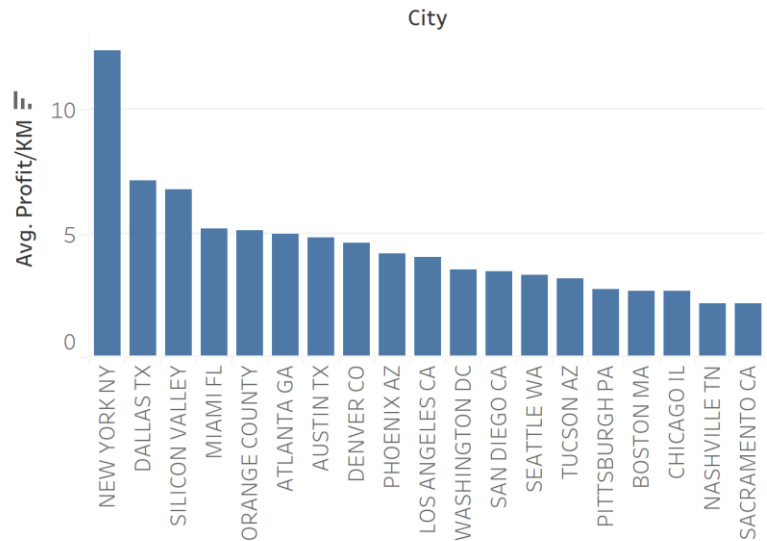- · 3,643
- ○ 100,000
- ○ 200,000
- ○ 302,149

Map based on Longitude (generated) and Latitude (generated). Size shows sum of Users. Details are shown for City. The data is filtered on City (City.csv), which keeps multiple members.
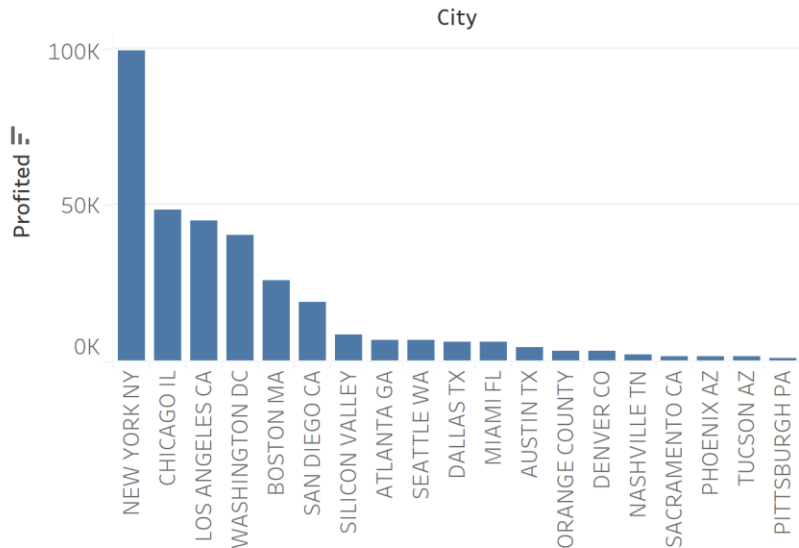
City vs Population

City (City.csv)
- ■ ATLANTA GA
- ■ AUSTIN TX
- ■ BOSTON MA
- ■ CHICAGO IL
- ■ DALLAS TX
- ■ DENVER CO
- ■ LOS ANGELES CA
- ■ MIAMI FL
- ■ NASHVILLE TN
- ■ NEW YORK NY
- ■ ORANGE COUNTY
- ■ PHOENIX AZ
- ■ PITTSBURGH PA
- ■ SACRAMENTO CA
- ■ SAN DIEGO CA
- ■ SEATTLE WA
- ■ SILICON VALLEY
- ■ TUCSON AZ
- ■ WASHINGTON DC

Sum of Users for each City. Color shows details about City (City.csv). The view is filtered on City, which excludes Null.

Avg Profit per City

City

Avg. Profit/KM

NEW YORK NY, DALLAS TX, SILICON VALLEY, MIAMI FL, ORANGE COUNTY, ATLANTA GA, AUSTIN TX, DENVER CO, PHOENIX AZ, LOS ANGELES CA, WASHINGTON DC, SAN DIEGO CA, SEATTLE WA, TUCSON AZ, PITTSBURGH PA, BOSTON MA, CHICAGO IL, NASHVILLE TN, SACRAMENTO CA

Average of Profit/KM for each City.

Most profitable city

City

Profited

NEW YORK NY, CHICAGO IL, LOS ANGELES CA, WASHINGTON DC, BOSTON MA, SAN DIEGO CA, SILICON VALLEY, ATLANTA GA, SEATTLE WA, DALLAS TX, MIAMI FL, AUSTIN TX, ORANGE COUNTY, DENVER CO, NASHVILLE TN, SACRAMENTO CA, PHOENIX AZ, TUCSON AZ, PITTSBURGH PA
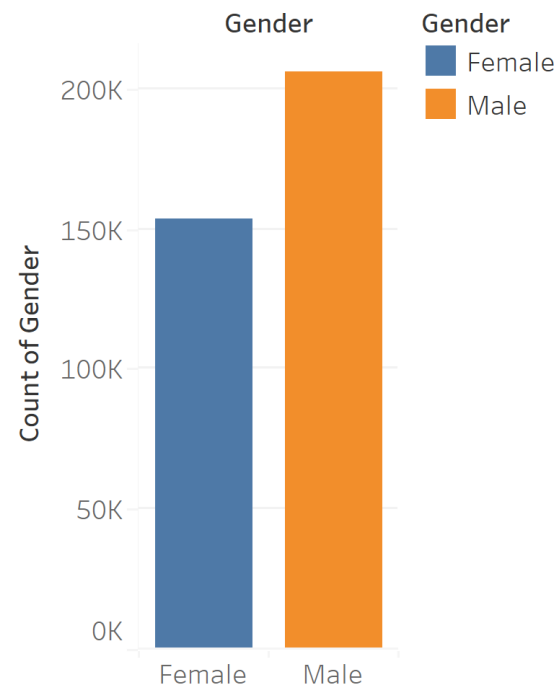
Sum of Profited for each City.

New York city tops the list with it being the most profitable city with the cab companies, silicon valley is also an area of interest that brings great revenue to the cab companies
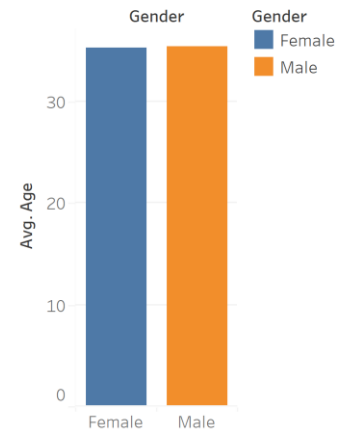
Cab Industry Analysis

Gender Count

Count of Gender for each Gender. Color shows details about Gender.

The industry saw more males use cab services than females in the past three years with a deficit of roughly 50,000 more men using the service than women.
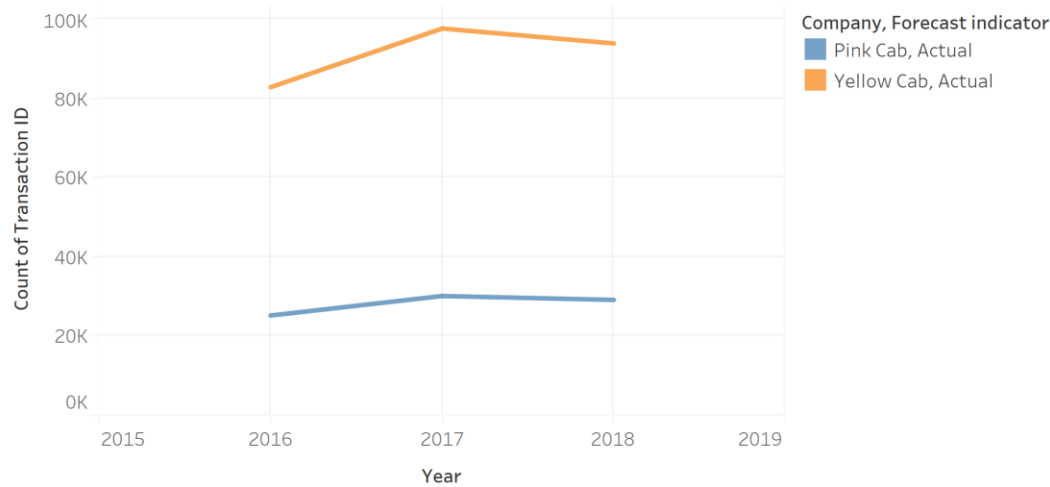

Mean Age of Genders

Average of Age for each Gender. Color shows details about Gender.

Company Analysis

Growth of Transactions

Company, Forecast indicator
- Pink Cab, Actual
- Yellow Cab, Actual

The trend of count of Transaction ID (actual & forecast) for Year. Color shows details about Company and Forecast indicator.

Just like the industry the two cab companies experienced similar effect on the number of transactions/trips they had over the past three years.

Profit of industry per year

Company
- Pink Cab
- Yellow Cab
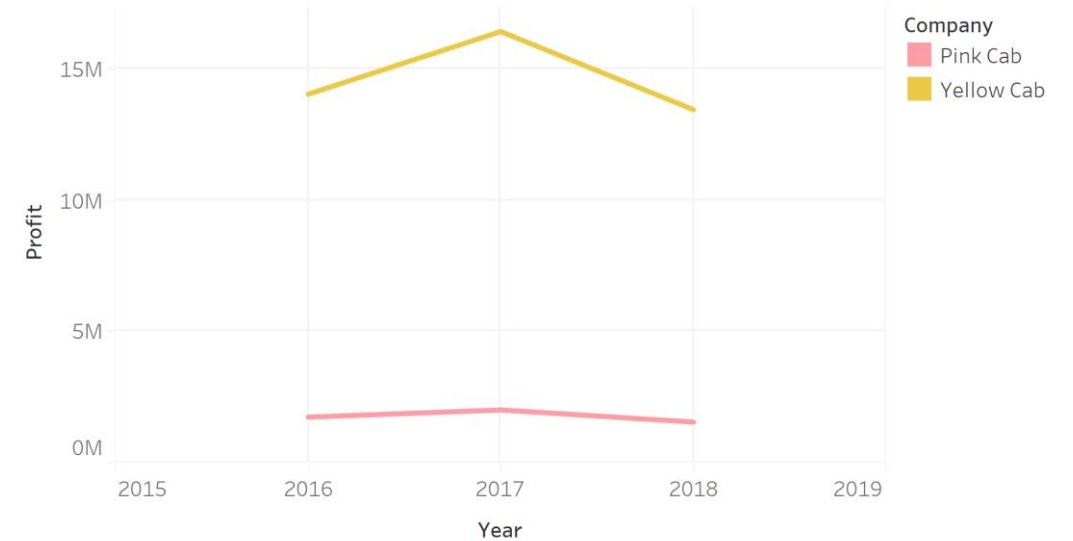
The trend of sum of Profit (actual & forecast) for Year. Color shows details about Company.

The pink cab company has had a rather steady two years with having steady profit changes. The yellow cab company was however heavily impacted in the year 2017-2018 and had a major drop in profit
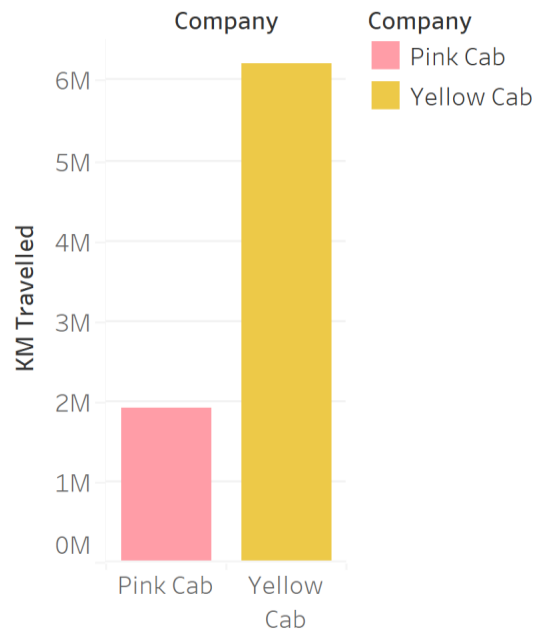
Company vs Users

Company

Company
- Pink Cab
- Yellow Cab
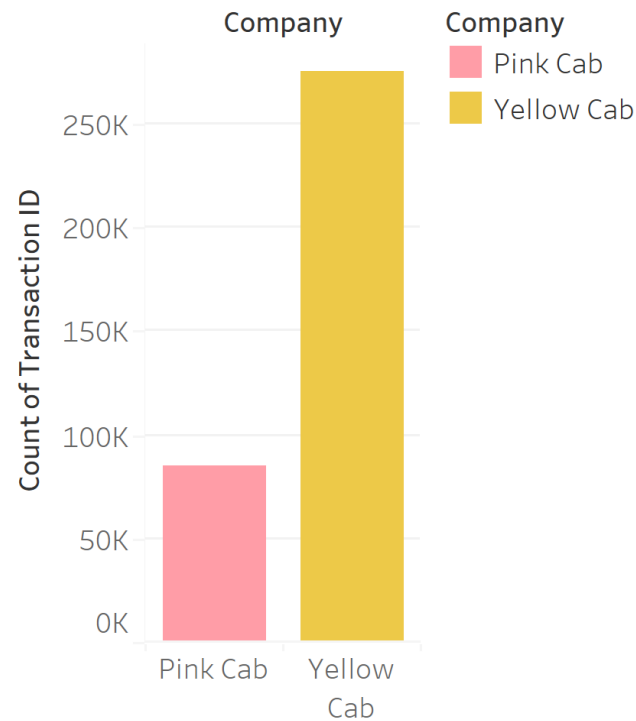
Count of Transaction ID for each Company. Color shows details about Company.

Gender Distribution

Company	Gender
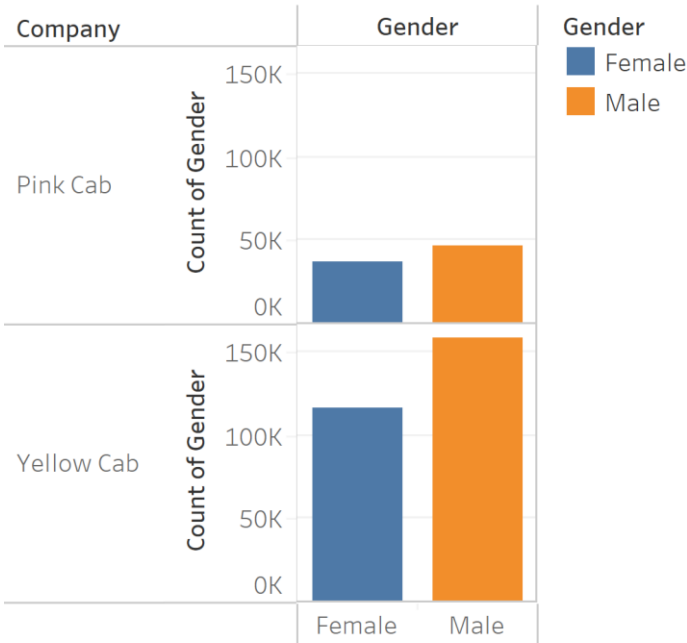
Gender
- Female
- Male

Count of Gender for each Gender broken down by Company. Color shows details about Gender.

Company vs KM Travelled

Company

Company
- Pink Cab
- Yellow Cab
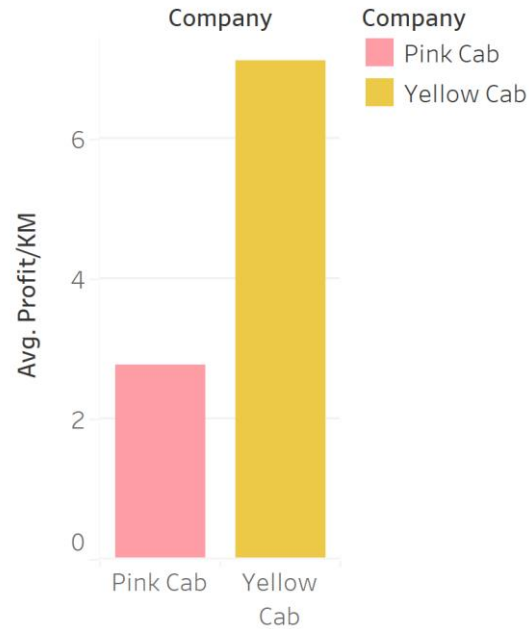
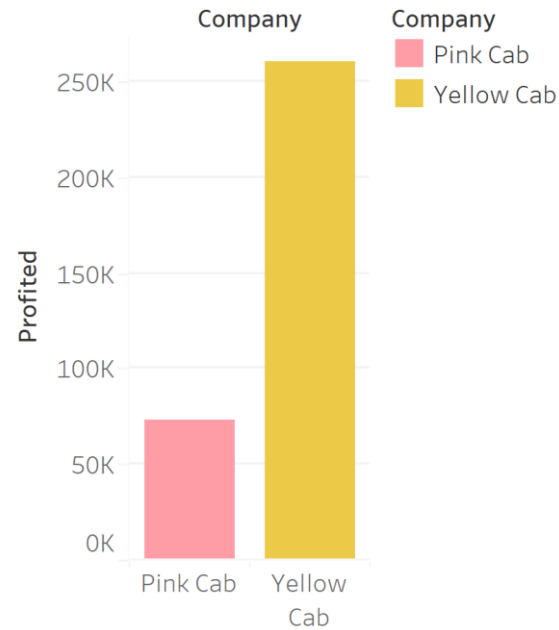Sum of KM Travelled for each Company. Color shows details about Company.

This shows the number of active users per company.
The yellow company had a significantly more userbase as compared to the pink company

Profit/KM per Company

Average of Profit/KM for each Company. Color shows details about Company.

Profitable Rides

Sum of Profited for each Company. Color shows details about Company.

Average Profitable rides

Average of Profited for each Company. Color shows details about Company.

This is calculated by calculating how many rides on average result in a profit for the company.
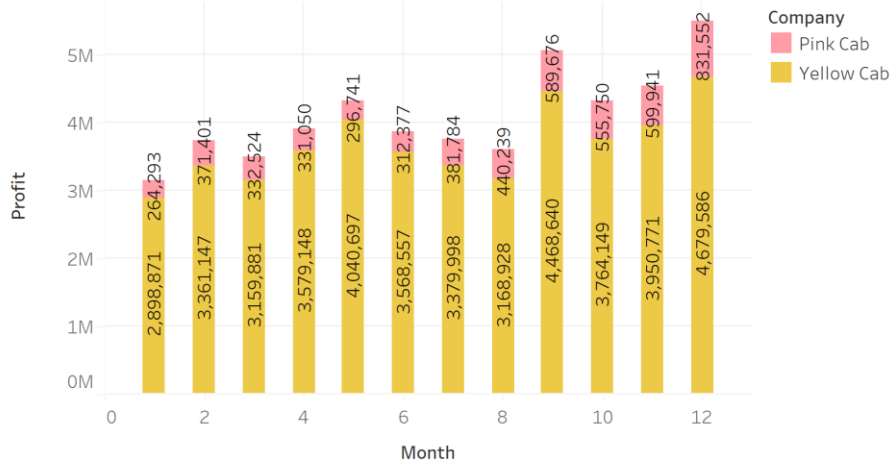
The yellow cab company appears to have dominated the industry in various sectors including profit/km , average profitable rides and number of profitable rides when compared to the pink yellow cab company
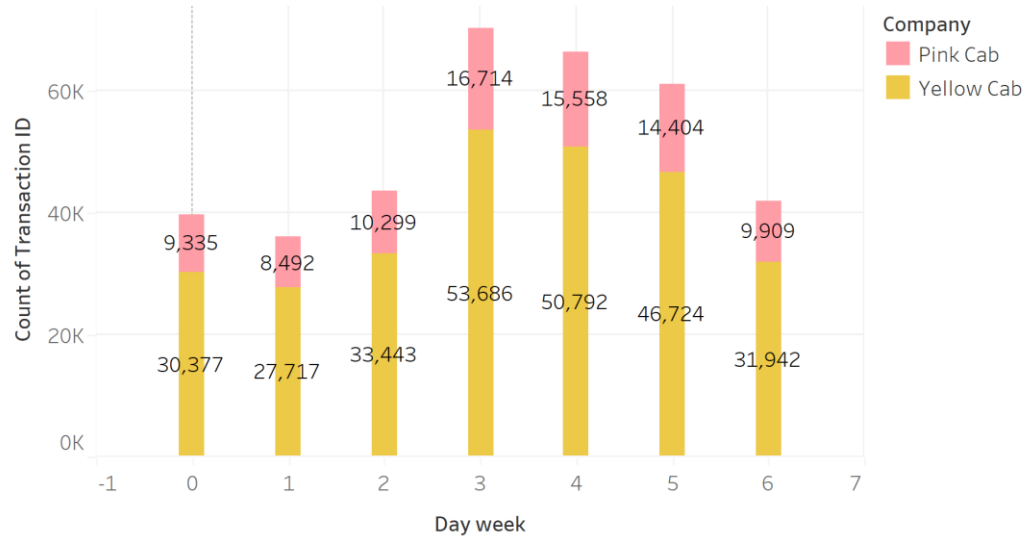
The yellow company yet again edges the pink cab company with profits acquired per month distribution

KM vs Price Charged
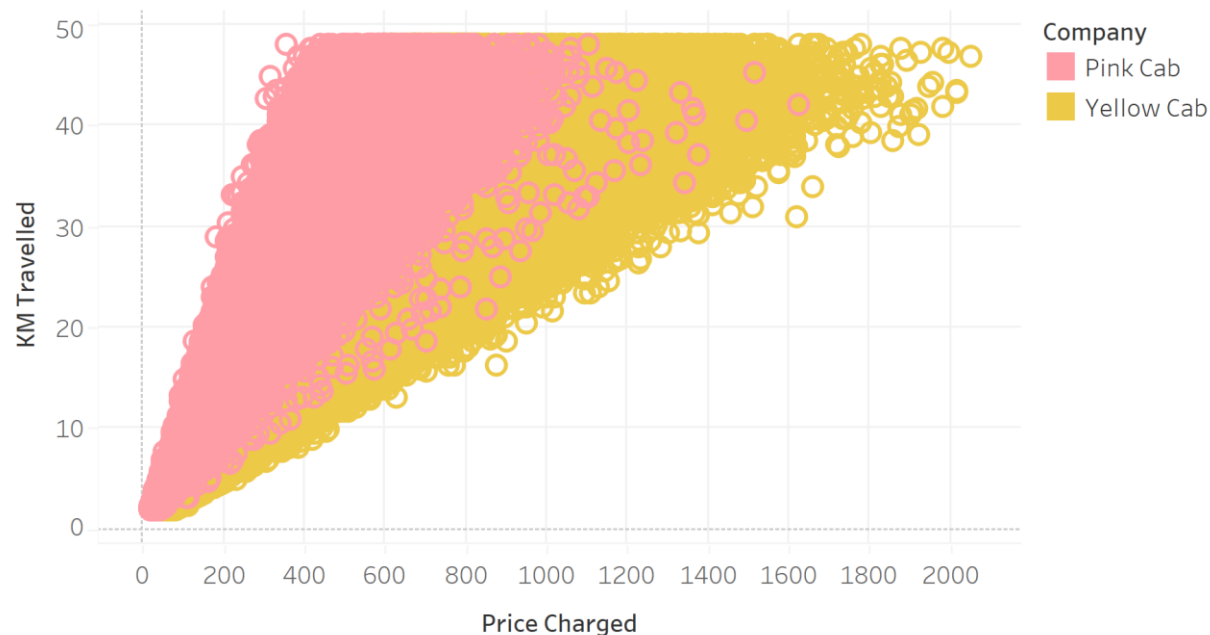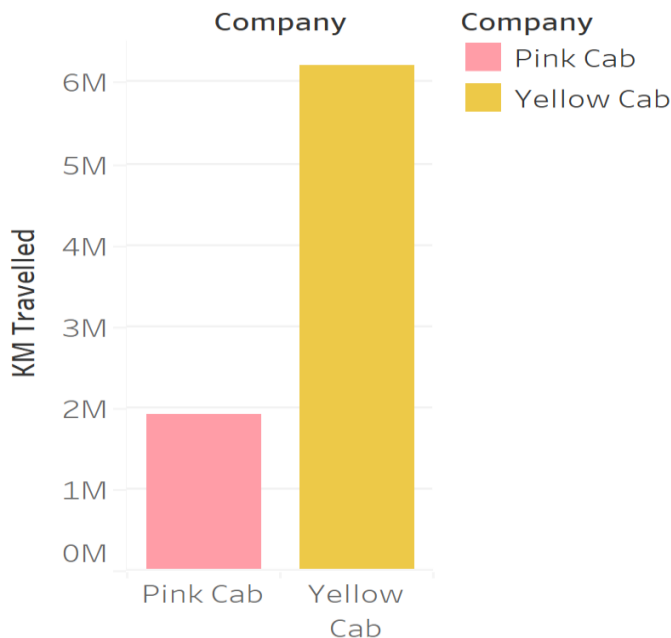

Price Charged vs. KM Travelled. Color shows details about Company.

The yellow cab company charges more per kilometer than the pink cab company. This also translates to more profit per kilometer.
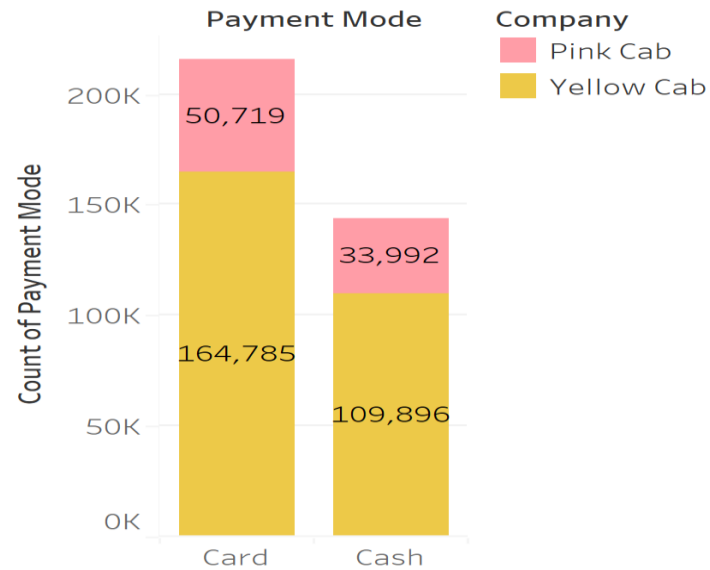
Company vs KM Travelled

Sum of KM Travelled for each Company. Color shows details about Company.

Payment Options

Sheet 10

Payment Mode

Count of Payment Mode

- 200K
- 150K
- 100K
- 50K
- 0K

Card: 50,719 (Pink Cab), 164,785 (Yellow Cab)

Cash: 33,992 (Pink Cab), 109,896 (Yellow Cab)
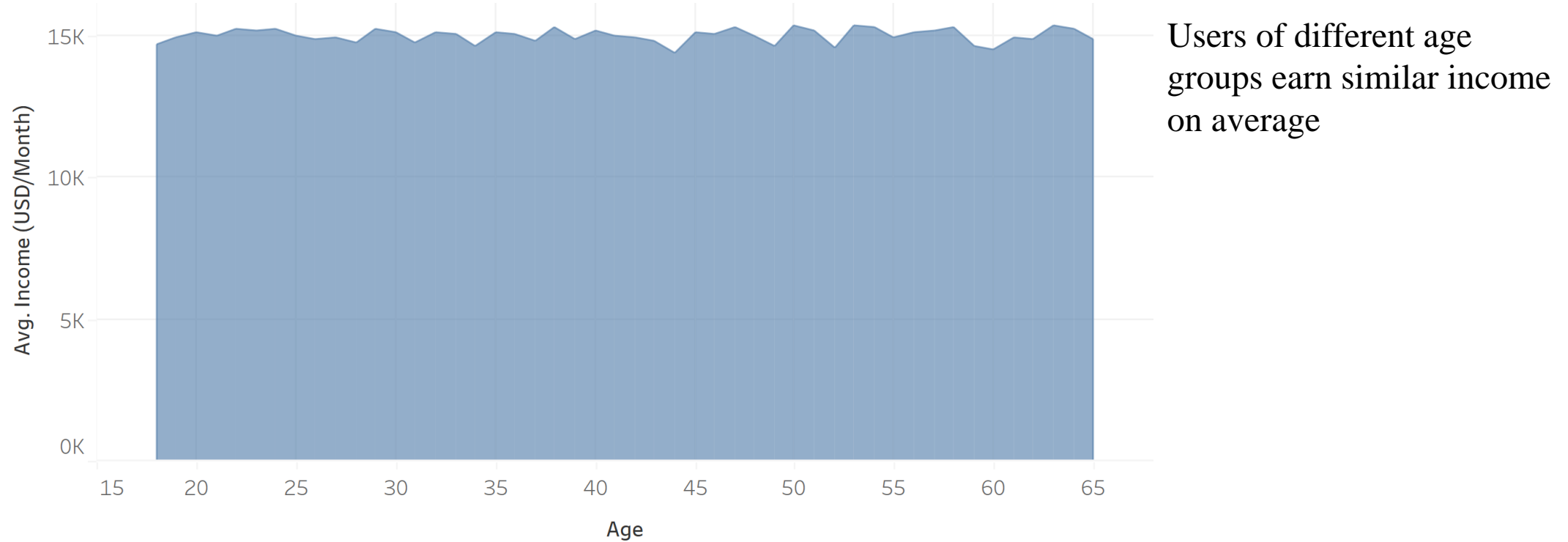
Company
- Pink Cab
- Yellow Cab

Count of Payment Mode for each Payment Mode. Color shows details about Company. The view is filtered on Company and Exclusions (Company,Payment Mode). The Company filter keeps multiple members. The Exclusions (Company,Payment Mode) filter keeps 4 members.

Users of the cab services used two payment options with card being the preferred option between the two.

The chart also shows the distribution of payment methods per company

AVERGAE INCOME IN USERS

Avg income vs User

Users of different age groups earn similar income on average

The plot of average of Income (USD/Month) for Age.

RECOMMENDATIONS

- **Which company is more profitable**

  The yellow Cab Company is more profitable and hence the better option if Profitability is used as a KPI for investment

- **If they are targeting a certain city**

  If the company is aiming at a certain city for them to expand / invest .. The yellow cab company is the best option as it has a higher reach in most of the major cities in the industry

- **Target higher income areas so as to get more revenue**

  If they are targeting higher income areas they should invest in the yellow cab company

- **Customer Churn**

  Yellow Cab has a lower churn(Has a larger loyalty base when compared to the pink company and hence should invest in yellow cab if this is used as a KPI

- **Targeting a particular quarter**

  If they intend to invest during a specific quarter to get higher returns the company should invest in the Yellow cab company during the last two quarters/ second half of the year