G2M Cab Insight Case

virtual internship
-Yitong Lu

Background

- Our client XYZ is plaining for an investment in Cab industry. They want to understand the market better before they make final decision.

- We have the information of two cab companies: one is Pink cab, the other one is Yellow cab. XYZ will choose the better company which can bringthem more benefits to invest.

- focus: we will analysis the data of two companies to see which company is better.

Data Exploration

- total 5 cvs files: Cabdata, ID of customers, ID of transaction, City, US holidays.

- total 20 variables: date, city, population, age, payment mode, etc.

- number of observations: 359392
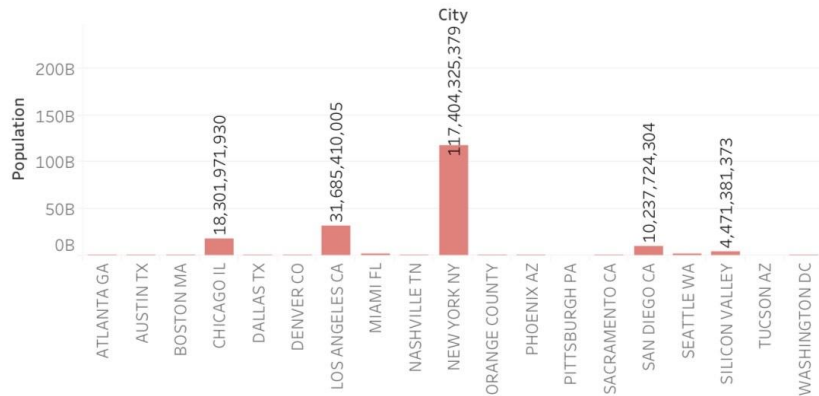
Year vs. profit in each company
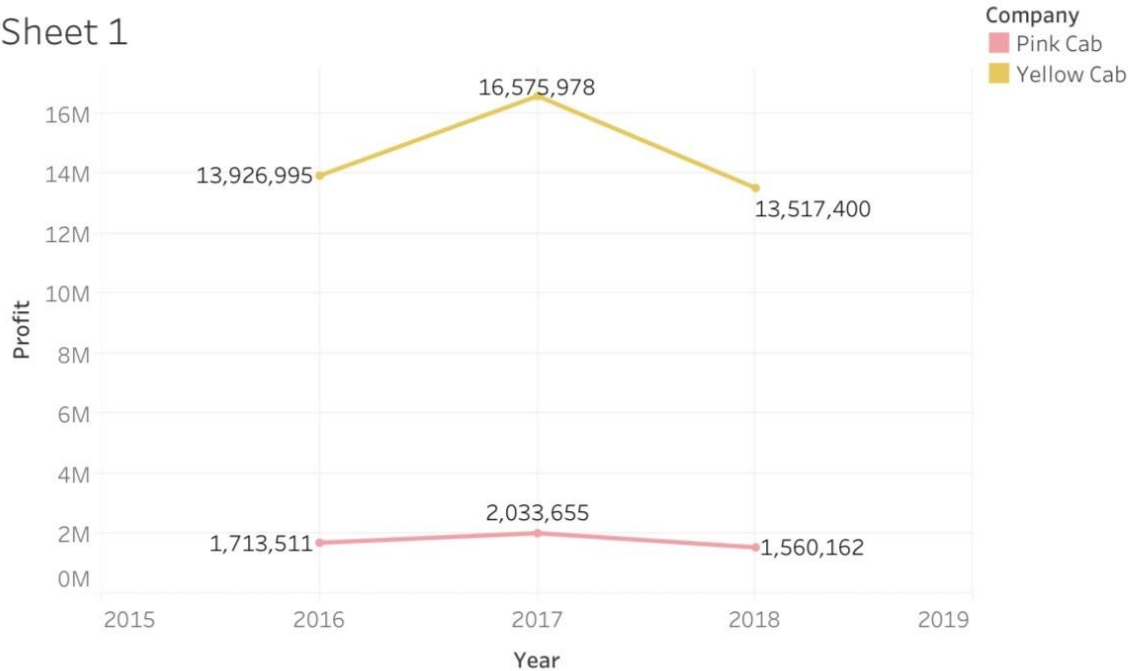
- **both companies have the best profits in 2017**

- **Compare the two companies' profits in 2016,2017,2018:**
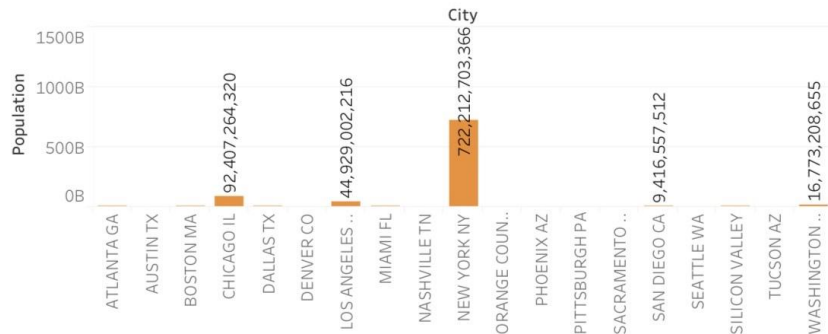- **Yellow Cab has profits 8 to 10 times bigger than Pink Cab's profits.**

City vs. Population

city vs population (pink cab)

yellow cab

- This plot is showing each city activated customers base on each company.

- from the pink cab's plot, we can see the top 5 activated cities during 2016-2018 were:Chicago,Los Angeles,New York, San Diego, Silicon Valley
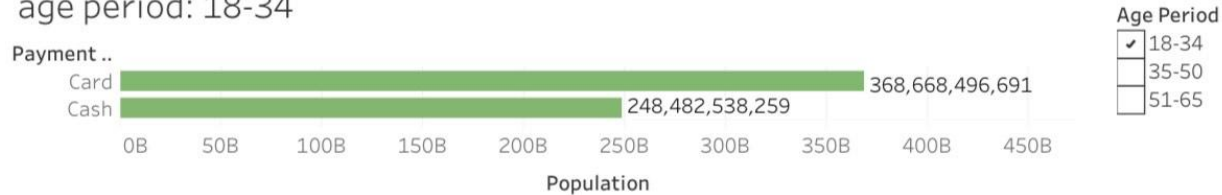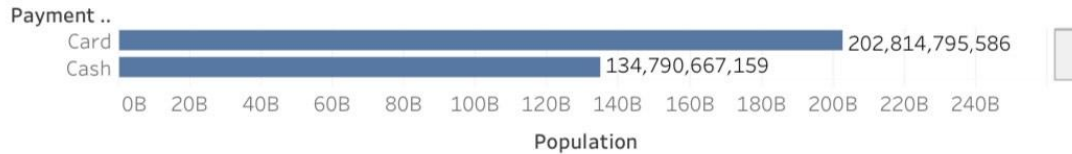
- from the yellow cab's plot, we can see the top 5 activated cities during 2016-2018 were:Chicago,Los Angeles,New York, San Diego, Washington

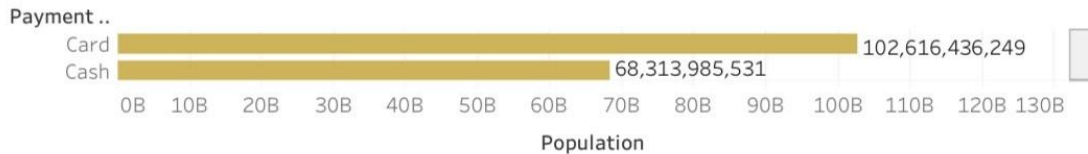# payment method vs population within different age periods

age period: 18-34

Payment ..

Card	368,668,496,691
Cash	248,482,538,259

0B 50B 100B 150B 200B 250B 300B 350B 400B 450B
Population

Age Period
- ✓ 18-34
- ☐ 35-50
- ☐ 51-65

35-50

Payment ..

Card	202,814,795,586
Cash	134,790,667,159

0B 20B 40B 60B 80B 100B 120B 140B 160B 180B 200B 220B 240B
Population

51-65

Payment ..

Card	102,616,436,249
Cash	68,313,985,531

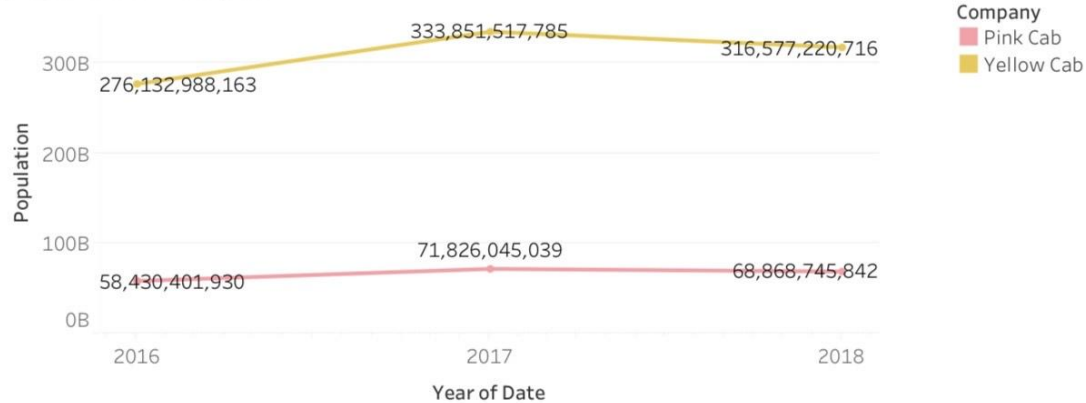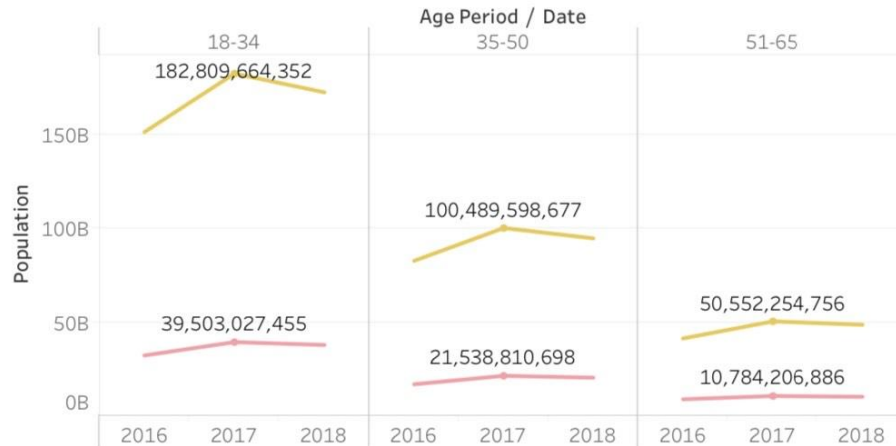0B 10B 20B 30B 40B 50B 60B 70B 80B 90B 100B 110B 120B 130B
Population

- From the plot we can see all 3 age periods all prefer to use card more than cash

- apparently, young people(18-34) use cab relatively often comparing to other 2 age periods

population vs.year
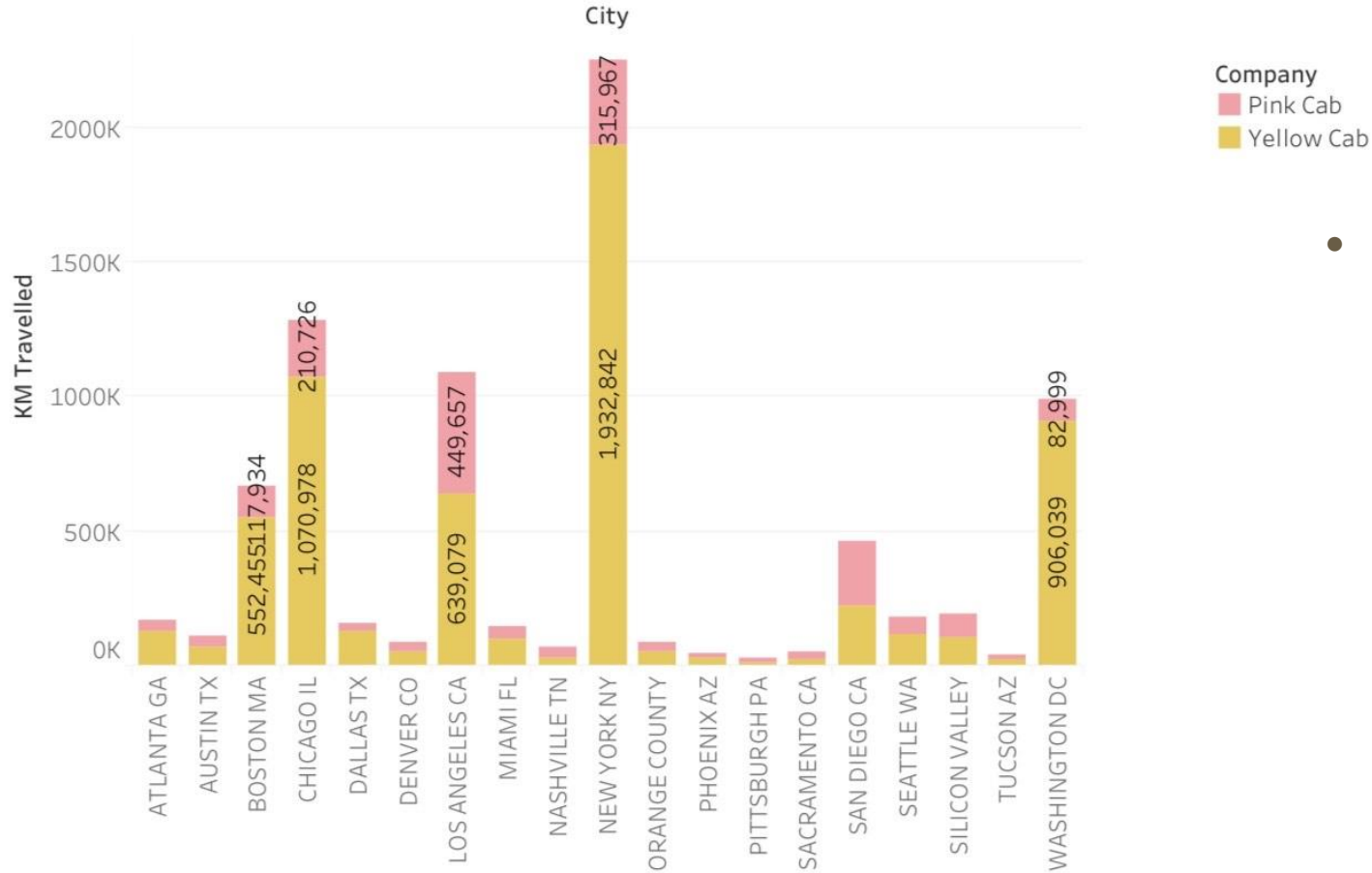

Company
- Pink Cab
- Yellow Cab

population vs. date within different age periods

- for both companies, the total users were increasing from 2016 to 2017 and were decreasing from 2017 to 2018

- the fluctuation of pink cab total customers from 2016 to 2018 is smaller than Yellow cab.

City vs. KM Travelled

Company
- Pink Cab
- Yellow Cab

- New York, Los Angeles, Chicago and Boston have bigger accumulated travelled KM than other cities.

- Company should be aware of the fluctuation of yellow cab total users. Maybe can make a questionnaire survey to see the satisfied level of how customers think about the Yellow cab service.
- the total users of Yellow cab are 8-10 times bigger than Pink cab.
- base on travelled KM, population, Citie with high user activity are: Los Angeles, Washington, New York, Chicago. Company can focus on maintaining the users on these city. (discount, coupon, etc.)
- the most important we should focus on is the profit part, the data shows that Yellow cab has larger profit than Pink cab
- overall, XYZ should invest Yellow cab since it has larger user base and makes better profits.