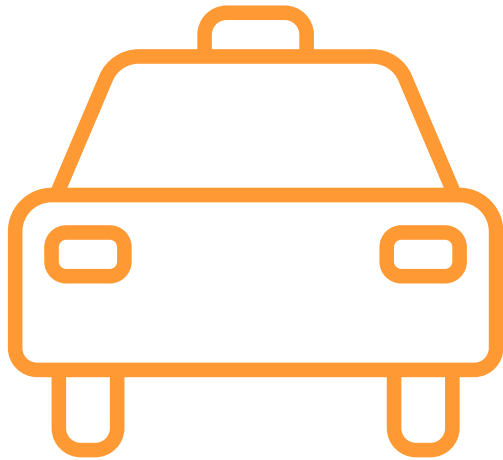**Data Glacier**
Your Deep Learning Partner

# G2M Insights of Cab Companies
## Virtual Internship

**17/04/2023**

*Presented by Patryk Potocki*

# Agenda

- 1.Introduction

- 2.Used Data-Sets (Explanation)

- 3.Visualization and Statistical Analysis

- 4.Summary of findings

- 5.Recommendation

### Introduction.

XYZ is a private equity 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.

### Objective.

Run necessary analysis which will help our client to make a right decision, on which company they should make an investment.

## Analysis and order included in our presentation.

Part. 1:
Explanation of datasets which were used in our research.

Part. 2:
Visualization of datasets provided to see which company is more popular.

Part. 3:
Summary of our findings and recommendation for our client.

**Data Glacier**

Details of Datasets.

1. Cab_data.csv
2. City.csv
3. Customer_ID
4. Transaction_ID

Details of Master dataset.

We have merged all four datasets to get
as accurate advice as possible for you.
In our first visualization, User's data is
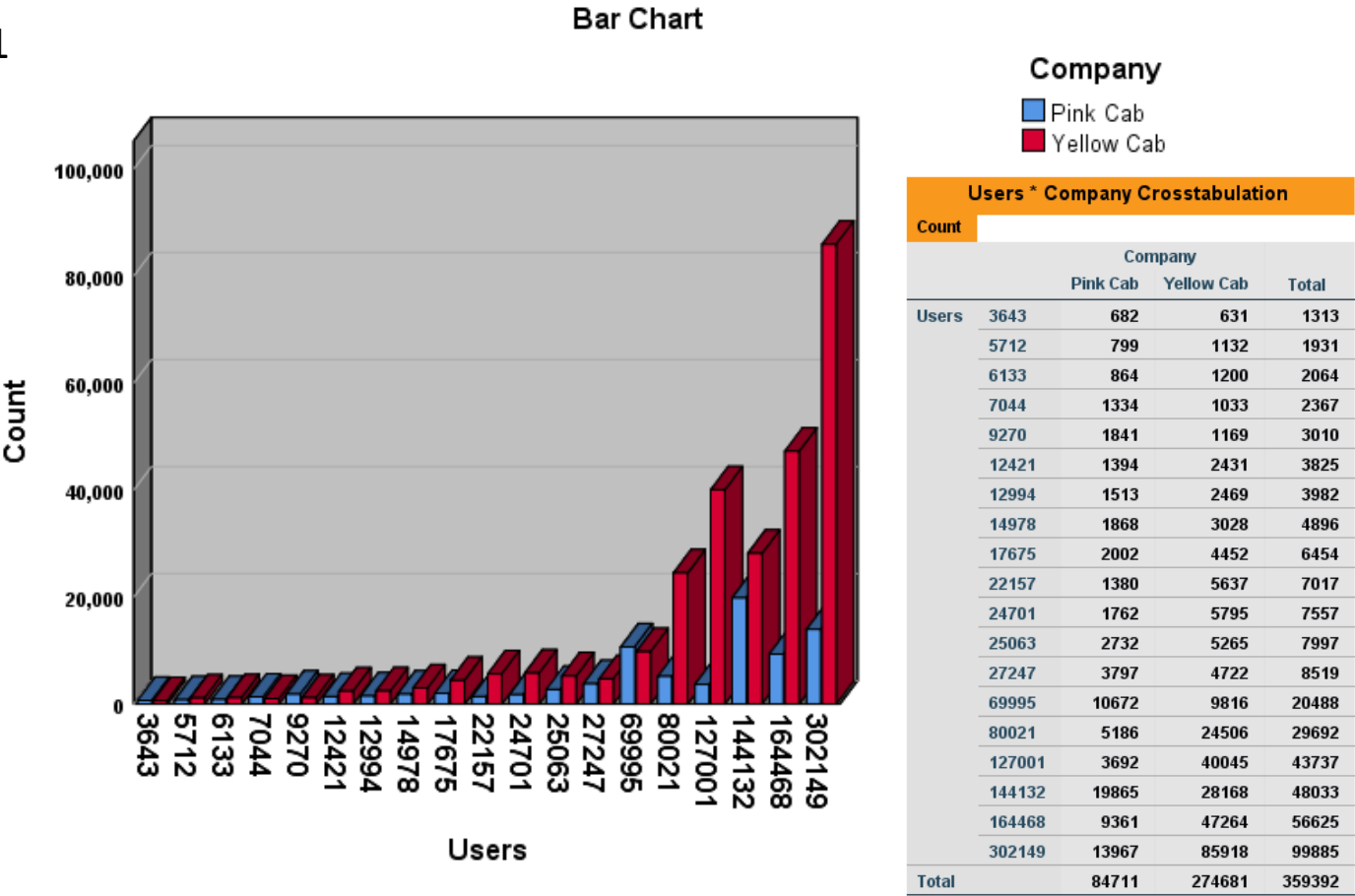anticipated by us as group of people from
different Cities.

**Total Number:  359393 Rows**

**14 Columns**

PART 1

## Users per City.(Survey)

Fig. 1

### Bar Chart



**Company**
- Pink Cab
- Yellow Cab

**Users * Company Crosstabulation**

Count

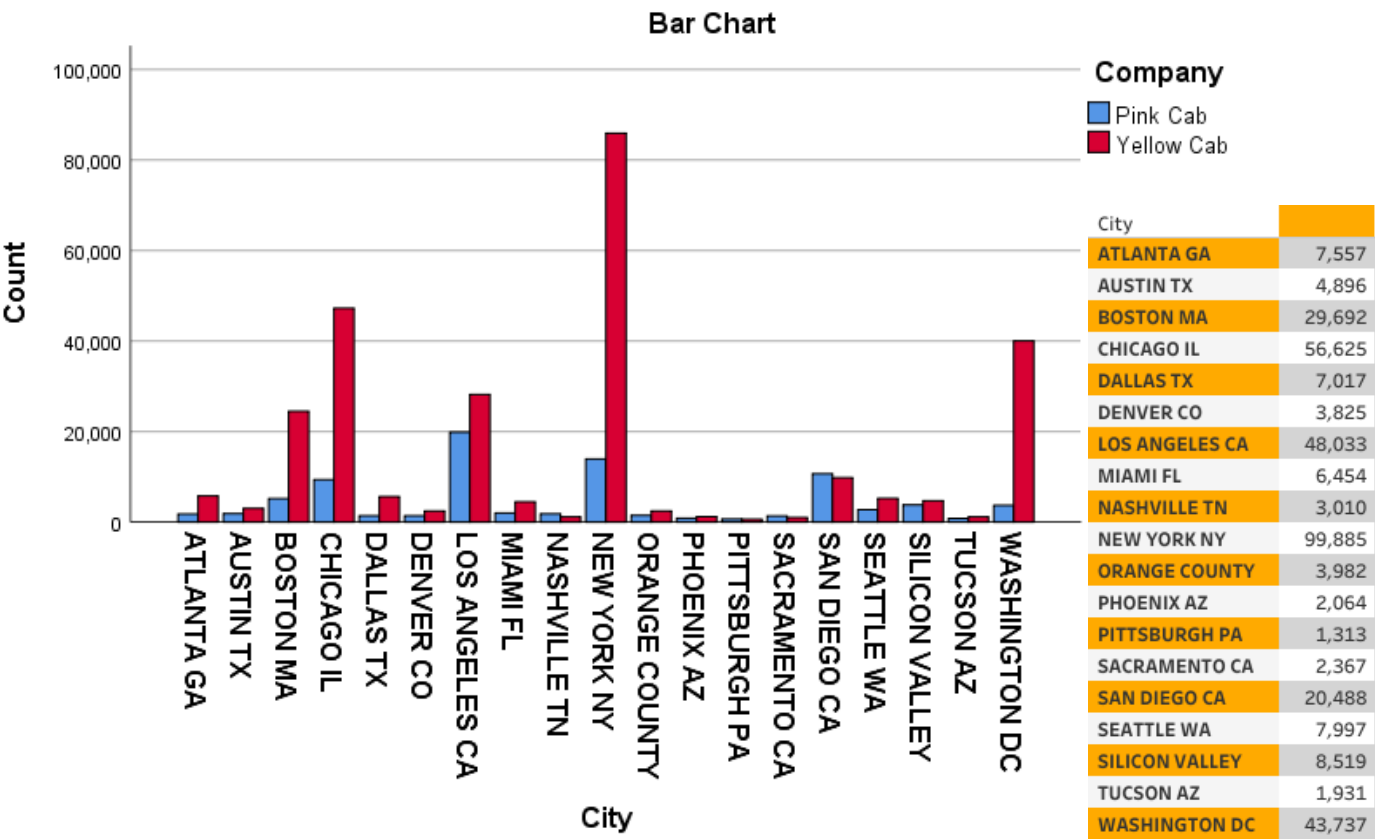| Users | | Company | | Total |
|---|---|---|---|---|
| | | Pink Cab | Yellow Cab | |
| Users | 3643 | 682 | 631 | 1313 |
| | 5712 | 799 | 1132 | 1931 |
| | 6133 | 864 | 1200 | 2064 |
| | 7044 | 1334 | 1033 | 2367 |
| | 9270 | 1841 | 1169 | 3010 |
| | 12421 | 1394 | 2431 | 3825 |
| | 12994 | 1513 | 2469 | 3982 |
| | 14978 | 1868 | 3028 | 4896 |
| | 17675 | 2002 | 4452 | 6454 |
| | 22157 | 1380 | 5637 | 7017 |
| | 24701 | 1762 | 5795 | 7557 |
| | 25063 | 2732 | 5265 | 7997 |
| | 27247 | 3797 | 4722 | 8519 |
| | 69995 | 10672 | 9816 | 20488 |
| | 80021 | 5186 | 24506 | 29692 |
| | 127001 | 3692 | 40045 | 43737 |
| | 144132 | 19865 | 28168 | 48033 |
| | 164468 | 9361 | 47264 | 56625 |
| | 302149 | 13967 | 85918 | 99885 |
| Total | | 84711 | 274681 | 359392 |

Visualization of comparation users for both Cabs. In this case we can see that of our Users group data, people are divided equally in smaller Cities, but as we go up Yellow Cab it's a lot frequent choice for Cab users.
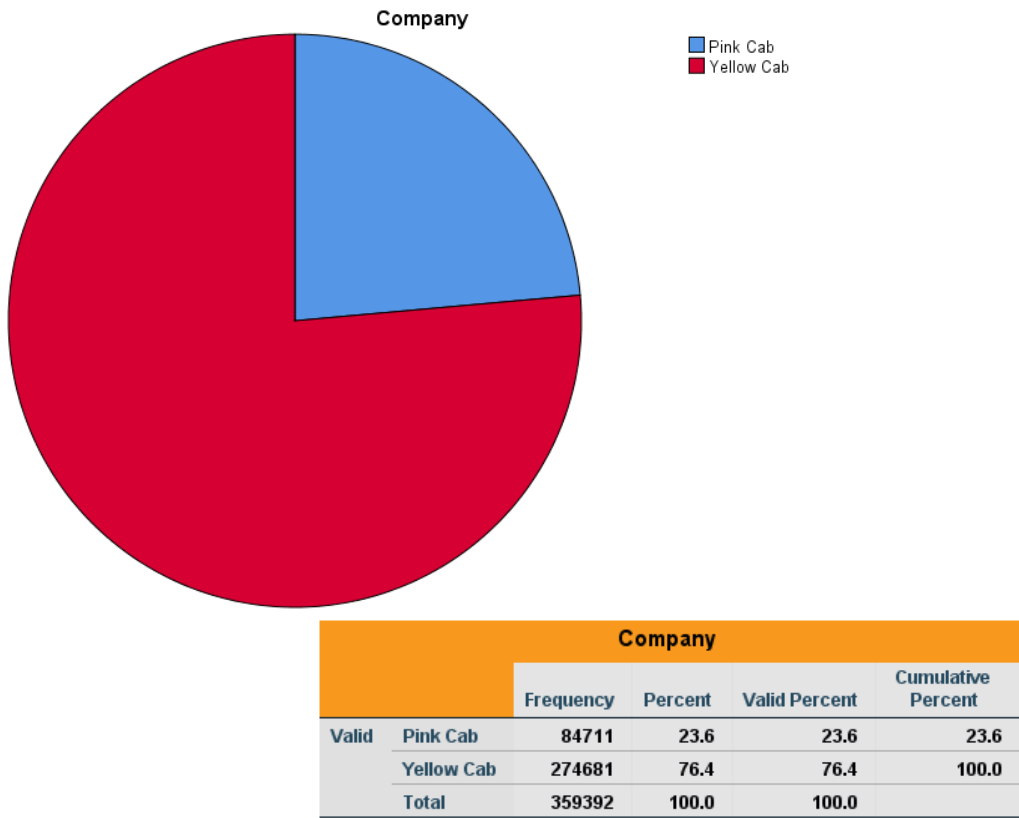
Data Glacier

PART 2

Frequency of both Cab company services use Per City.

Visualization of most popular Cab.

Fig. 3

Fig. 2



Bar Chart

Company
- Pink Cab
- Yellow Cab

| City | |
|------|------|
| ATLANTA GA | 7,557 |
| AUSTIN TX | 4,896 |
| BOSTON MA | 29,692 |
| CHICAGO IL | 56,625 |
| DALLAS TX | 7,017 |
| DENVER CO | 3,825 |
| LOS ANGELES CA | 48,033 |
| MIAMI FL | 6,454 |
| NASHVILLE TN | 3,010 |
| NEW YORK NY | 99,885 |
| ORANGE COUNTY | 3,982 |
| PHOENIX AZ | 2,064 |
| PITTSBURGH PA | 1,313 |
| SACRAMENTO CA | 2,367 |
| SAN DIEGO CA | 20,488 |
| SEATTLE WA | 7,997 |
| SILICON VALLEY | 8,519 |
| TUCSON AZ | 1,931 |
| WASHINGTON DC | 43,737 |



Company
- Pink Cab
- Yellow Cab

| | | Company | | | |
|------|------|------|------|------|------|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | Pink Cab | 84711 | 23.6 | 23.6 | 23.6 |
| | Yellow Cab | 274681 | 76.4 | 76.4 | 100.0 |
| | Total | 359392 | 100.0 | 100.0 | |

Data Glacier

Profit per Company in each City.

Fig.4

| Company | Total Profit |
|---|---|
| Yellow Cab | 43,251,775 |
| Pink Cab | 5,237,540 |



| City | Pink Cab | Yellow Cab |
|---|---|---|
| NEW YORK NY | 3,475,780 | 25,515,727 |
| CHICAGO IL | 2,317,464 | 14,136,541 |
| WASHINGTON DC | 913,740 | 11,960,094 |
| LOS ANGELES CA | 4,945,494 | 8,438,422 |
| BOSTON MA | 1,296,658 | 7,294,378 |
| SAN DIEGO CA | 2,648,427 | 2,900,790 |
| SILICON VALLEY | 952,161 | 1,415,236 |
| SEATTLE WA | 682,842 | 1,566,789 |
| ATLANTA GA | 436,306 | 1,701,503 |
| DALLAS TX | 334,105 | 1,679,591 |
| MIAMI FL | 490,363 | 1,331,471 |
| AUSTIN TX | 451,127 | 899,314 |
| ORANGE COUNTY | 370,101 | 727,771 |
| DENVER CO | 344,441 | 720,505 |
| NASHVILLE TN | 454,993 | 354,583 |
| SACRAMENTO CA | 331,684 | 312,602 |
| PHOENIX AZ | 206,601 | 357,140 |
| TUCSON AZ | 198,958 | 333,362 |
| PITTSBURGH PA | 169,678 | 187,695 |

Data Glacier

Fig.5

Monthly income per Company



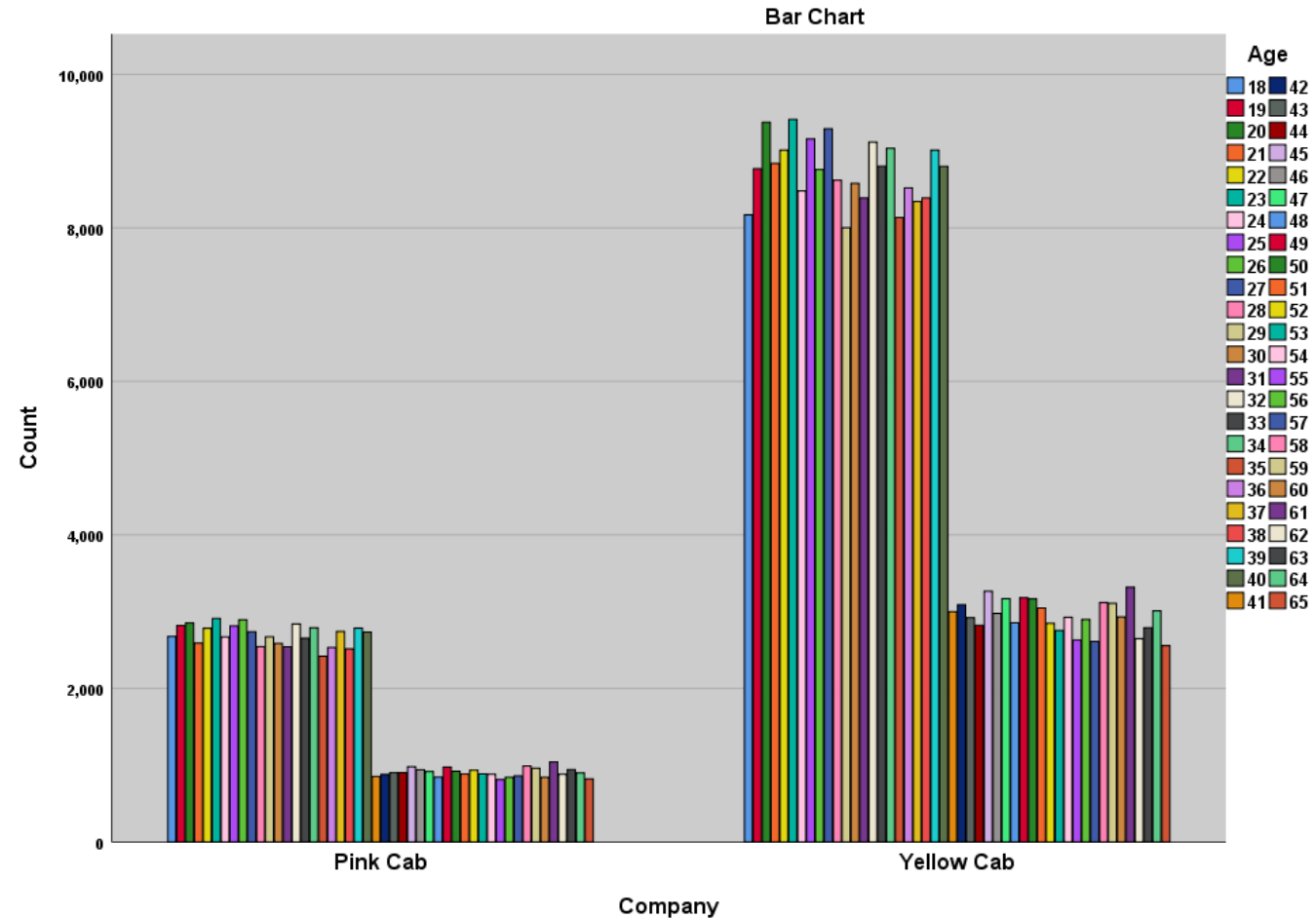| Company | |
|---------|-------|
| Pink Cab | 84,711 |
| Yellow Cab | 274,681 |
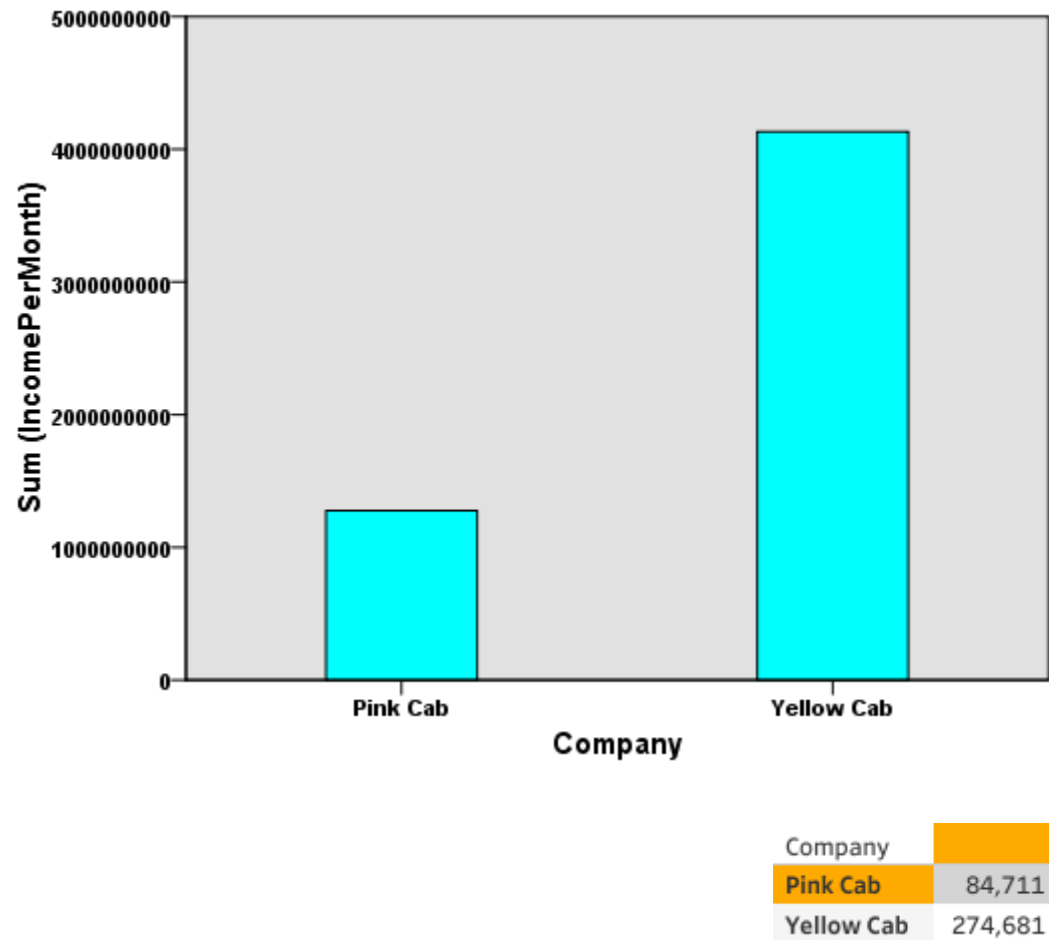
After looking at total profit per year, we want to check also monthly income per company. As we can see income per month for Yellow Cab its 274,681 and Pink Cab income is 84,711.
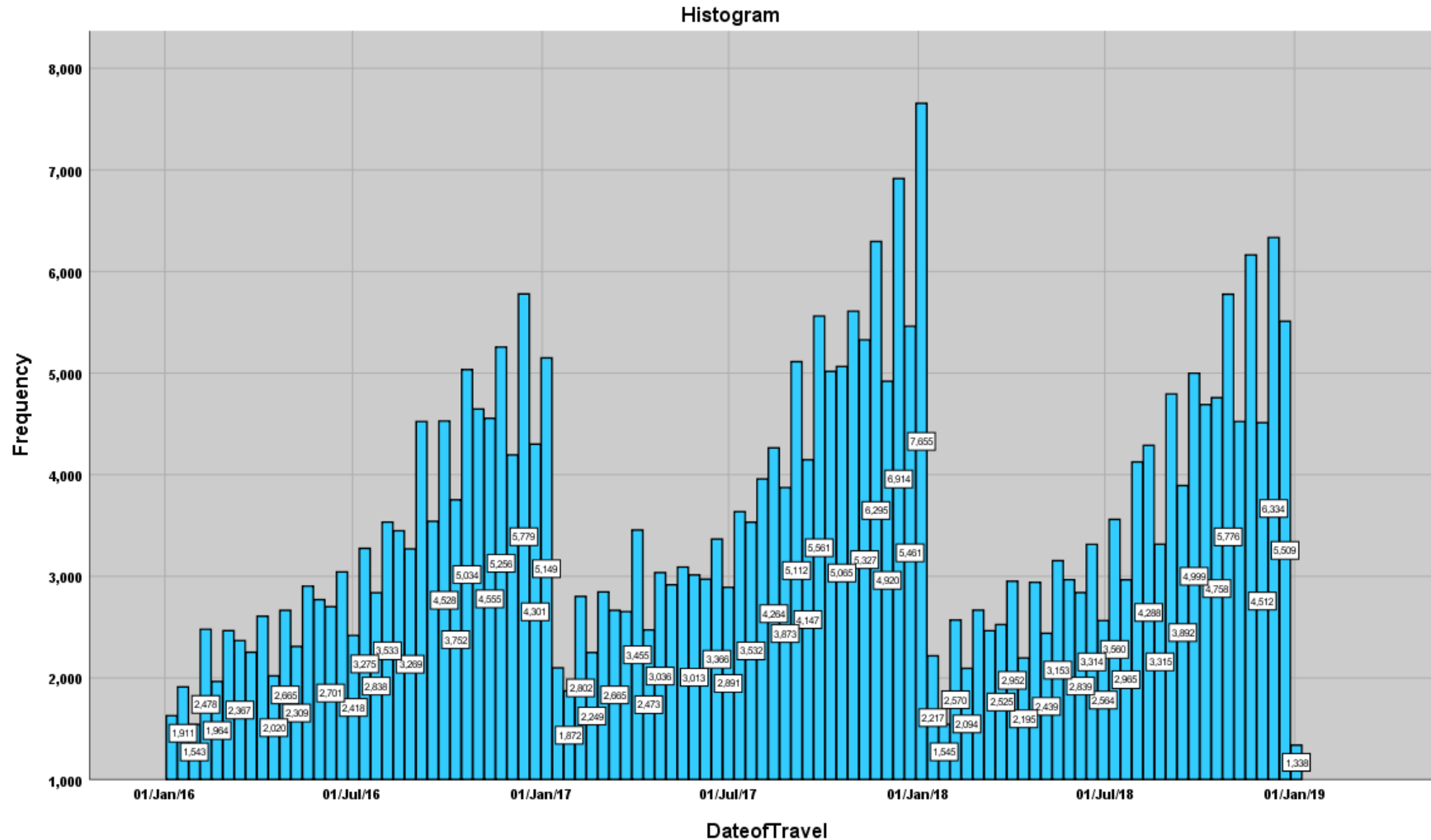
Data Glacier

Fig.6
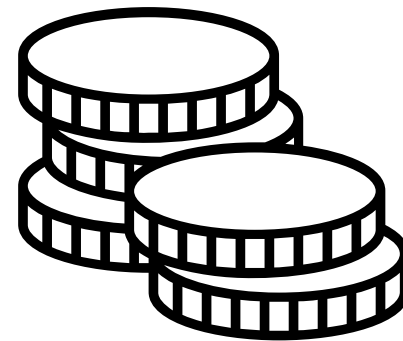
Here we can see once again domination of Yellow Cab, which we could expected as already in our previous analysis Pink Cab didn't stand a chance in every aspect. But from this Fig. we can also see that people in age between 18-40 are in the highest demand for Cab services in general.



Users of both companies per age.

Fig.7



In our next visualization we are looking at dates from our dataset to identify period of time when Cab services are in highest demand. As we can see patterns are different but very similar, also that in most cases time between July and Jan are the times for highest demand of Cab services.

Data Glacier

Fig.8        Preferable payment method per City.
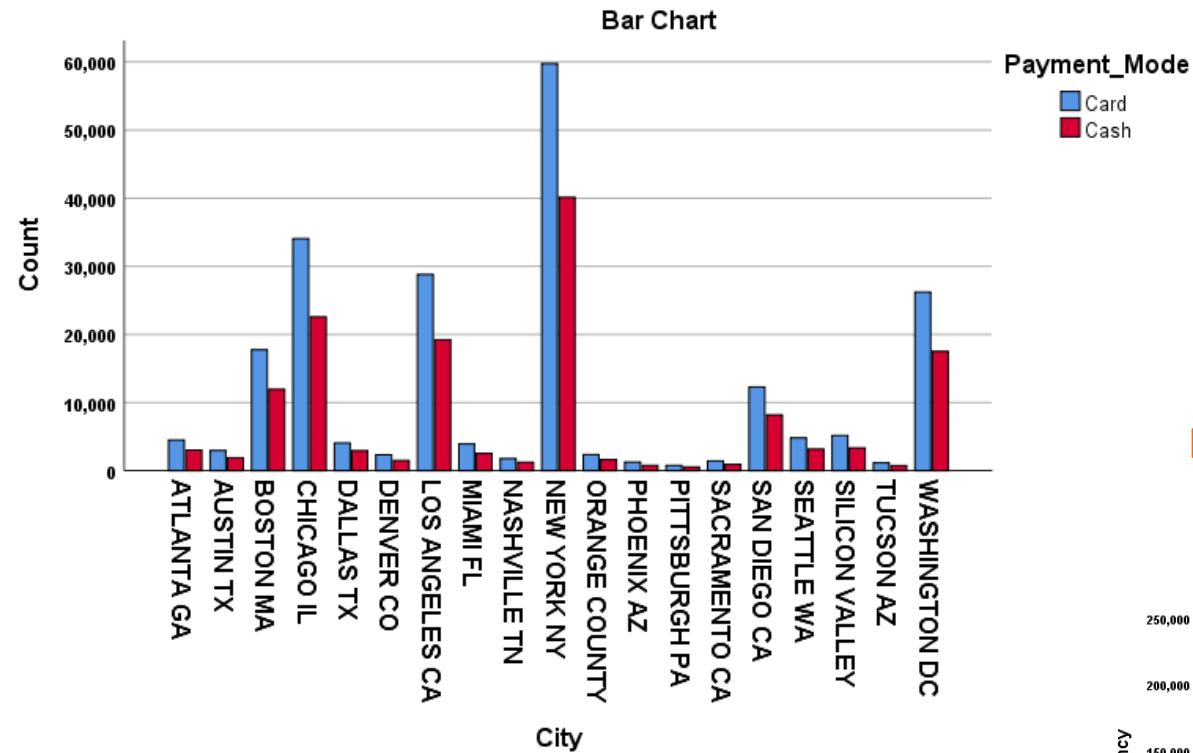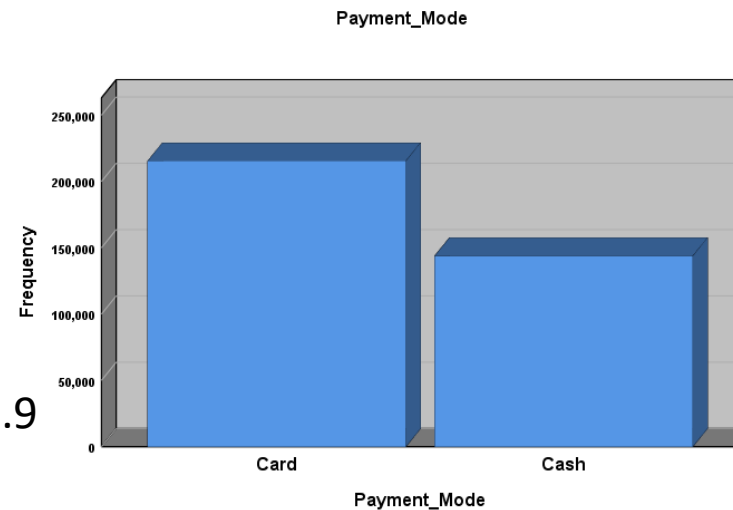


Preferable payment method.



Fig.9

For our last insights from analyzed dataset we checked the most popular payment method and as we can clearly see Card payment it's the most common choice in every city but also in general.

Data Glacier

- Participant from groups of people in each City has higher liken towards Yellow Cab then Pink. Only in smaller Cities Pink Cab overtakes Yellow Cab.

- Yellow Cab services availability overwhelm Pink Cab back far, with Yellow Cab count as 274681, compared to Pink Cab count of 84,711.

- While checking the profit per year for each company once again Yellow Cab results were a lot higher looking at Pink Cab. The total income for both companies recorded were:

a) Yellow Cab - 43,251,775 per year
b) Pink Cab - 5,237,540 per year

- Next up we checked income per month, Yellow Cab had over 3 times more income than Pink Cab. Income for both companies:

a) Yellow Cab – 274,681 per month
b) Pink Cab – 84,711 per year

- In additional analysis we find out groups per age who uses the Cab services most frequently. The age group was between 18-40 years of age, also next we checked the time of the year when Cab services had highest demand which was between July and January of each year. On the last page of our analysis, we could see that most preferable payment method is Card Payment.

PART 3

From the beginning of presentation to the end of it we could clearly see that Yellow Cab it's a lot more popular and a lot more used Cab company in US. From our survey of preferable Company per City to service availability ending up at income per month/year, in every aspect Yellow Cab overtakes Pink Cab by quite a lot.

After all our findings we can safely say that we recommend  Yellow Cab company to make your investment in. Would like to also point out as in the analysis on page 5, as mentioned in 3 Cities which are Sacramento CA, San Diego CA and Nashville TA the use of Yellow Cab its lower then Pink. If we come back to our last page and check Fig.8 we could see that Card Payments are the most common way chosen by customers in every City, we assume the lack of Card Payment method availablity in those Cities for the Yellow Cab is the cause of less used company compare to Pink Cab.

Data Glacier