CYCLISTIC

A case study

By Yuvraj

Email: yuvipahari@gmail.com
LinkedIn: Yuvraj.

INTRODUCTION

- This case study is a portfolio project by Yuvraj, to showcase the skills of various BI tools and his analytical prowess. This case study was the capstone project of the Google Data Analytics Professional Certificate.
- This project was made using R programming and SQL for data cleaning and manipulation. Tableau was used for visualization.
- A detailed report of the cleaning process and data manipulation with step-by-step action and all the relevant codes is given in the attached R document. (Link)
- The Tableau workbook link is also provided here. (Link)

About Cyclistic

In 2016, Cyclistic successfully launched a bike-share program in Chicago, which has since expanded to include 5,824 geotracked bicycles across 692 stations. The innovative system allows users to unlock bikes from one station and return them to any other station within the network. Cyclistic's initial marketing strategy focused on creating widespread awareness and appealing to diverse consumer segments. The key to this approach lay in the flexible pricing plans, including :

- single-ride passes - Includes the casual users who subscribe for only one ride.
- full-day passes - These people are also categorized as casual users who subscribe for a single day.
- annual memberships - Annual members subscribe to the yearly pass of the company.

This flexibility in pricing has been instrumental in accommodating various customer preferences and promoting the program's growth.

Characters

- **Cyclistic**: A bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike. The majority of riders opt for traditional bikes; about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day.
- **Lily Moreno**: The director of marketing and your manager. Moreno is responsible for the development of campaigns and initiatives to promote the bike-share program. These may include email, social media, and other channels.
- **Cyclistic marketing analytics team**: A team of data analysts who are responsible for collecting, analyzing, and reporting data that helps guide Cyclistic marketing strategy. You joined this team six months ago and have been busy learning about Cyclistic's mission and business goals — as well as how you, as a junior data analyst, can help Cyclistic achieve them.
- **Cyclistic executive team**: The notoriously detail-oriented executive team will decide whether to approve the recommended marketing program.

Problem Statement

- We are a junior data analyst working in the marketing analyst team at Cyclistic.
- The director of marketing believes the company's future success depends on maximizing the number of annual memberships.
- Therefore, the team wants to understand how casual riders and annual members use Cyclistic bikes differently.
- From these insights, the team will design a new marketing strategy to convert casual riders into annual members.
- But first, Cyclistic executives must approve our recommendations, so they must be backed up with compelling data insights and professional data visualizations.
- Three questions will guide the future marketing program:
    1. How do annual members and casual riders use Cyclistic bikes differently?
    2. Why would casual riders buy Cyclistic annual memberships?
    3. How can Cyclistic use digital media to influence casual riders to become members?
- Moreno has assigned the first question to answer:

  How do annual members and casual riders use Cyclistic bikes differently?

Link to Problem Statement

Deliverables

We had to produce a report with the following deliverables:

1. A clear statement of the business task
2. A description of all data sources used
3. Documentation of any cleaning or manipulation of data
4. A summary of your analysis
5. Supporting visualizations and key 3 findings
6. Your top three recommendations based on your analysis

Business Goal

To analyze the historical data of 1 year and find out how the Members and casuals use the bikes differently.

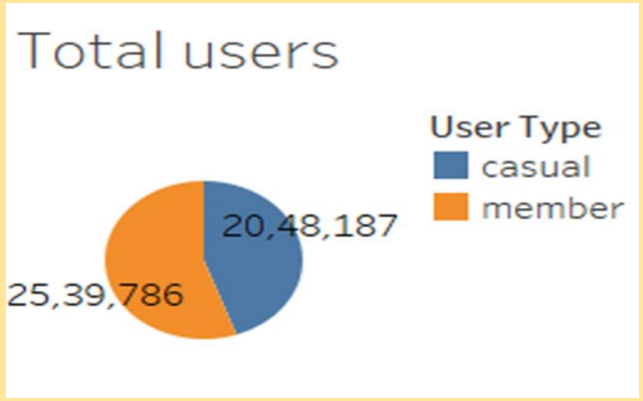
Hypothesis

- Members are the group of people who use the bikes for work-related travel.
- Casuals include the group who uses the bikes majorly for leisure purposes.
- People may be taking the bikes out of the station for trips.

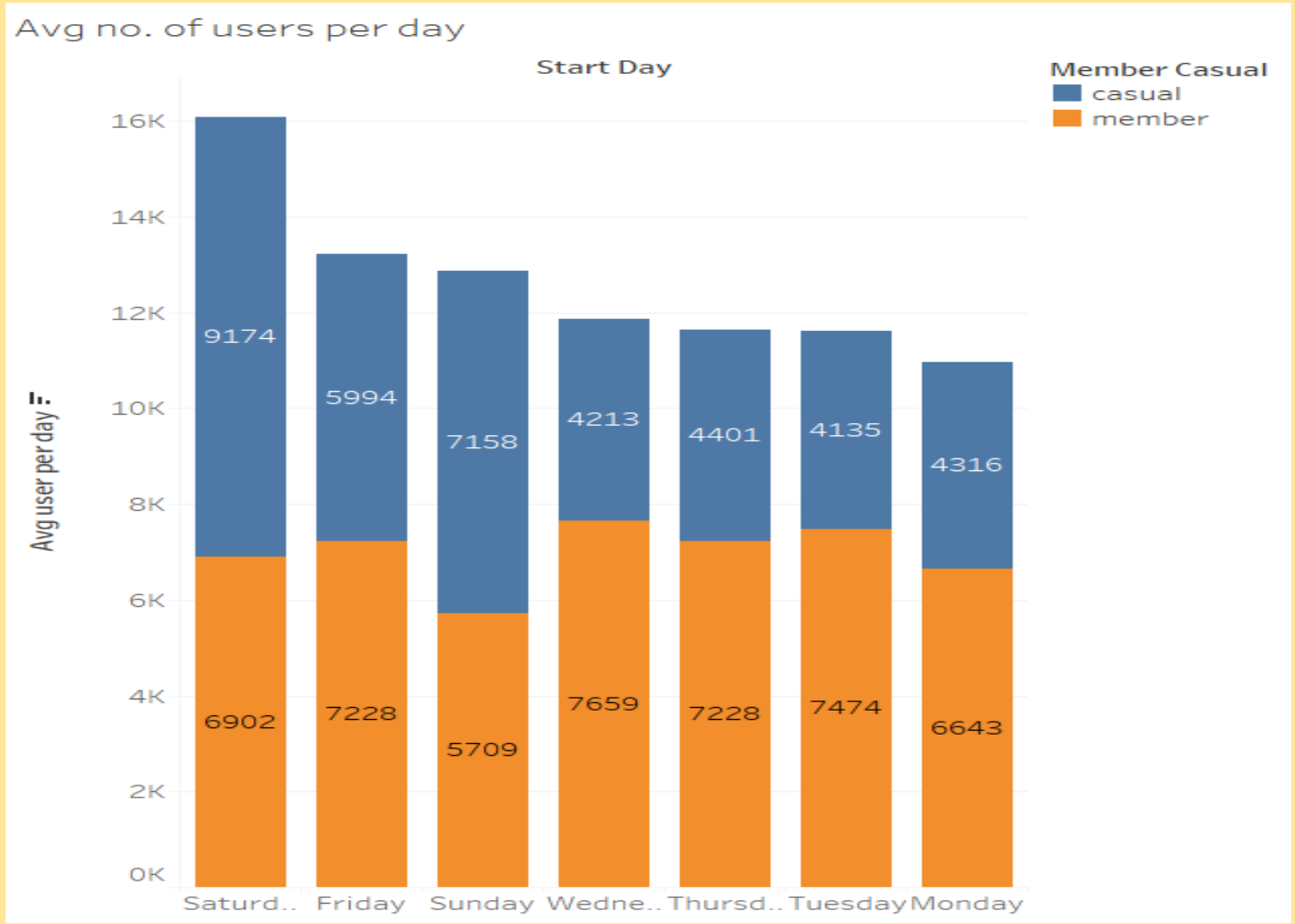
Tasks

1. Find out the total number of members and casual users.
2. Find out the average ride time of both members and casuals.
3. Find out the average distance traveled by members and casuals respectively.
4. Find out the number of members and casual users who are there daily.
5. Find out the average distance traveled by both types of users daily.
6. Find out the top 50 boarding stations of members and casuals respectively.
7. Find out the top 50 ending stations of both types respectively.
8. Find out the number of people who return bikes after one day.
9. Find the average distance traveled by people who return the bike after one day.
10. Find the average distance traveled by members and casuals every month.
11. Find the number of members and casuals traveling monthly.
12. Find the number of members and casuals using it at different hours of the day.

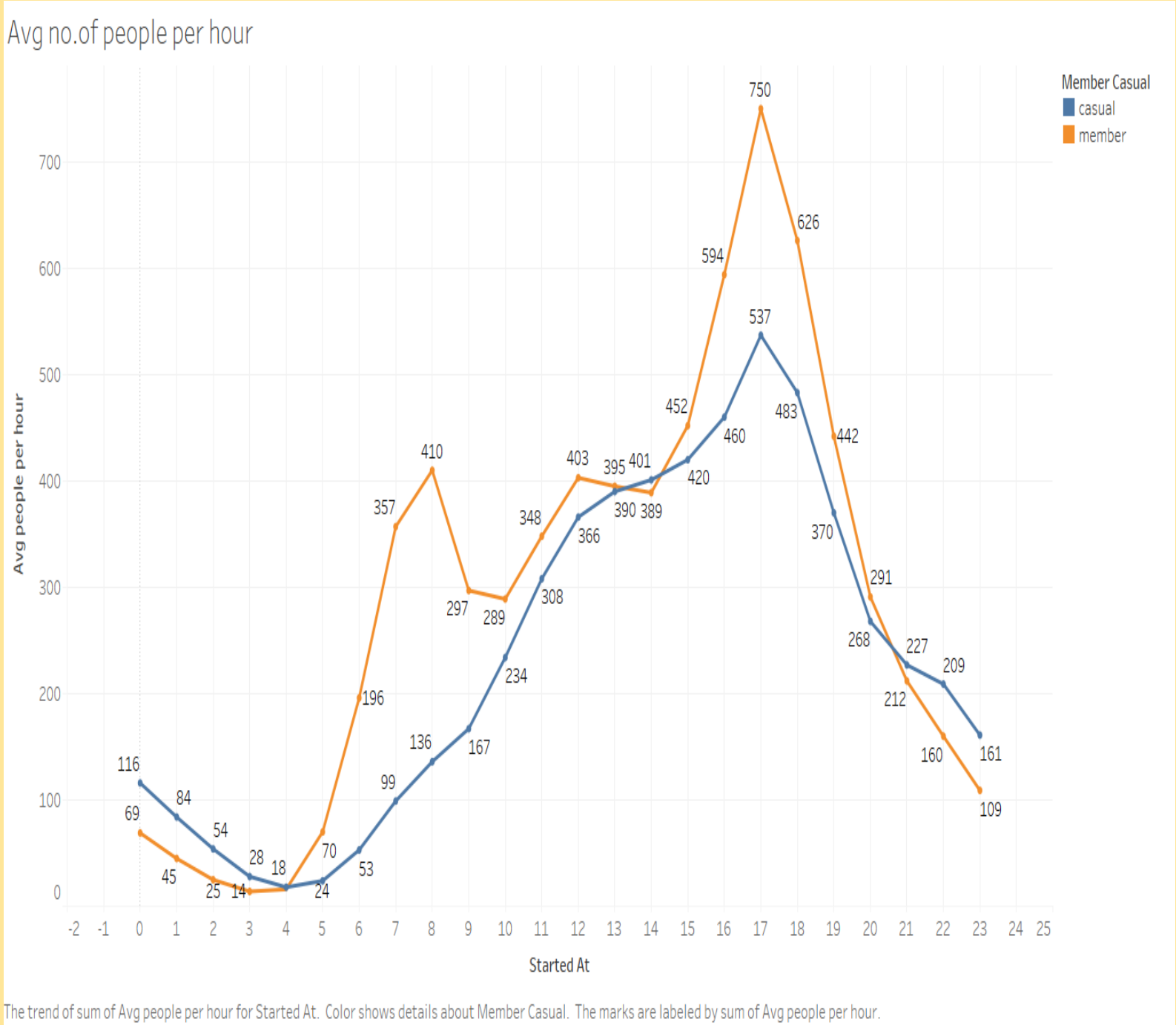
Note: The tasks have been executed and the report has been provided in the R document. This presentation is only for analysis of the visualizations made.



Analysis of Total Number of Users:

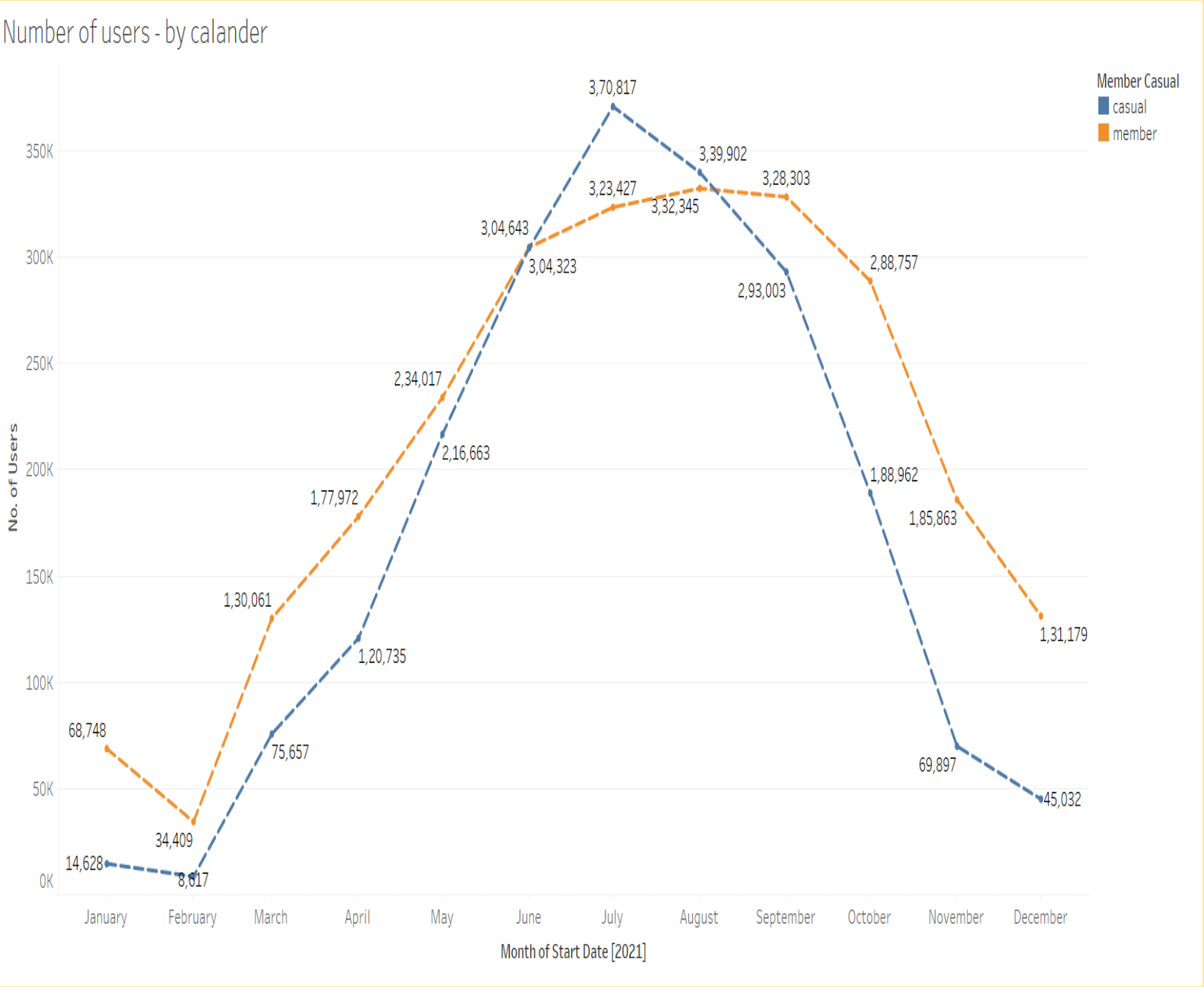


- We can see that the number of Members is greater than the casuals, yet, there are a good number of casual riders who should be tapped for the future growth of the company.
- Looking at the average number of users per day it can be seen that the casuals bar is highest during the weekend and the behaviour pattern of members is exactly opposite to that of the casuals.
- This is expected as according to the hypothesis the casuals are people using for leisure purposes, which people do during the weekends, while members use it extensively for their daily work commute.
- Another notable point is that in the casuals group also, a good number of people are using it for work-related purposes during the weekdays.
- This means that casuals use Cyclist bikes majorly for leisure purposes whereas for work travel they may be choosing different modes of transportation.



Hourly Number of users analysis

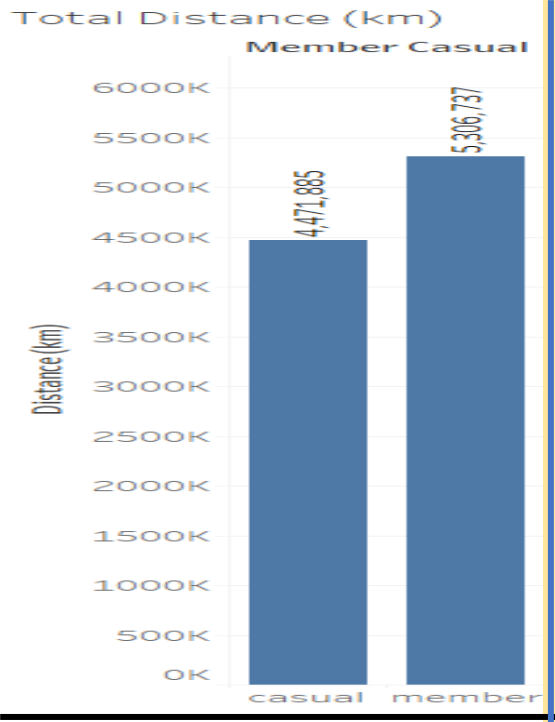
- The number of members as expected increases during the office going, lunch, and office ending hours. A similar trend can be observed with casuals also.
- Looking at the casual trends and comparing them with those of the members, we can say that there are also a good number of office-going people in this group.
- Since there is a spike in members in the early morning, we can say that even school/college-going people use bikes for their commute.
- This morning spike is not seen in the casuals group, rather there is a gradual growth. However, during the returning hours, the casual trend mimics that of the members. We can safely say that college-going people use bikes for their commutes.
- Also at night the number of casuals is greater than that of members. It may be said that, night-shift workers may be using the bikes for their commutes.

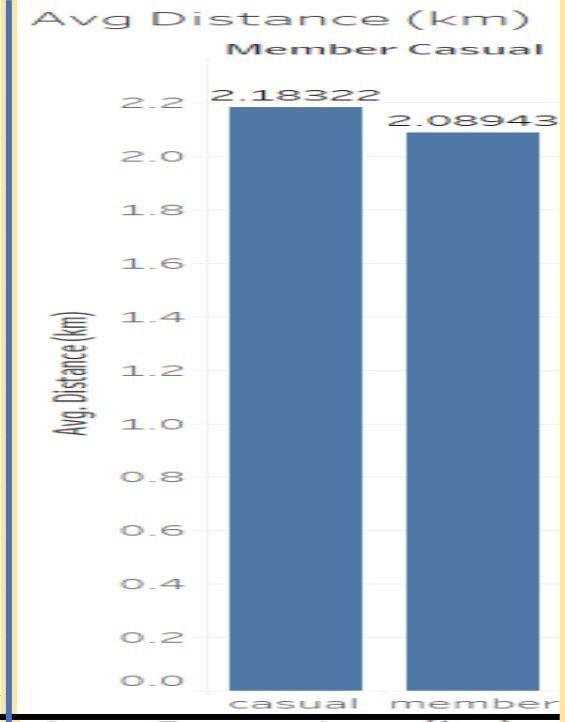


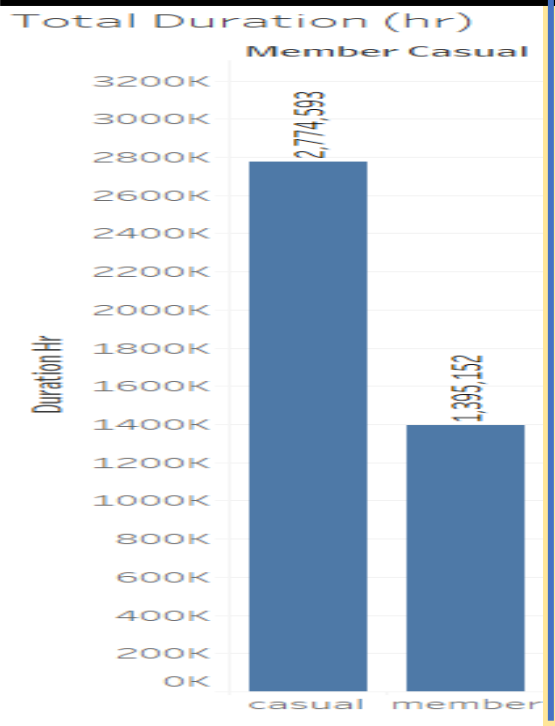
Monthly number of users analysis

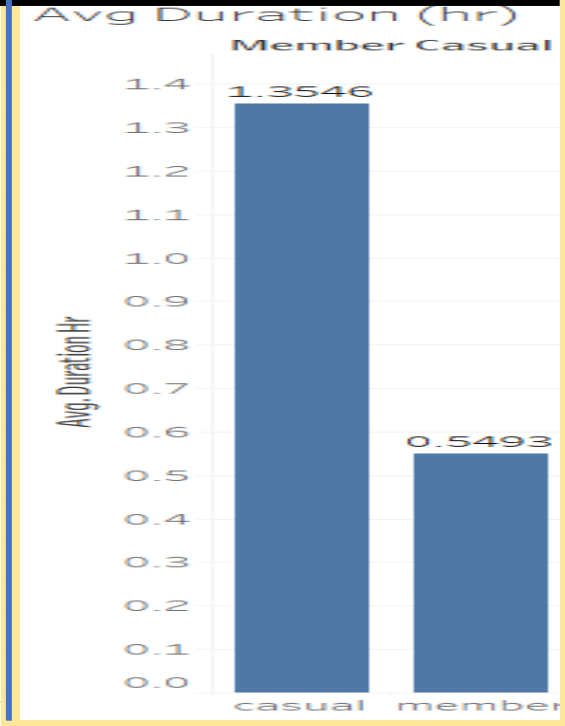
- It can be seen that the graph of Casuals climbs with a steep rise in the number of users from February to July and then falls with the same intensity.
- Members also represent the same trend but it is a gentle rise and also is nearly constant from May to September.
- It is a notable fact that the month in which casuals peak which is July is the time when Chicago gets summer vacations.
- This is in line with our hypothesis, as casuals are using bikes more for leisure during holidays.
- Although members are also using it that way, but the number is lesser than the casuals and a fixed number of members do so.

Comparison of Distance and Duration of the ride

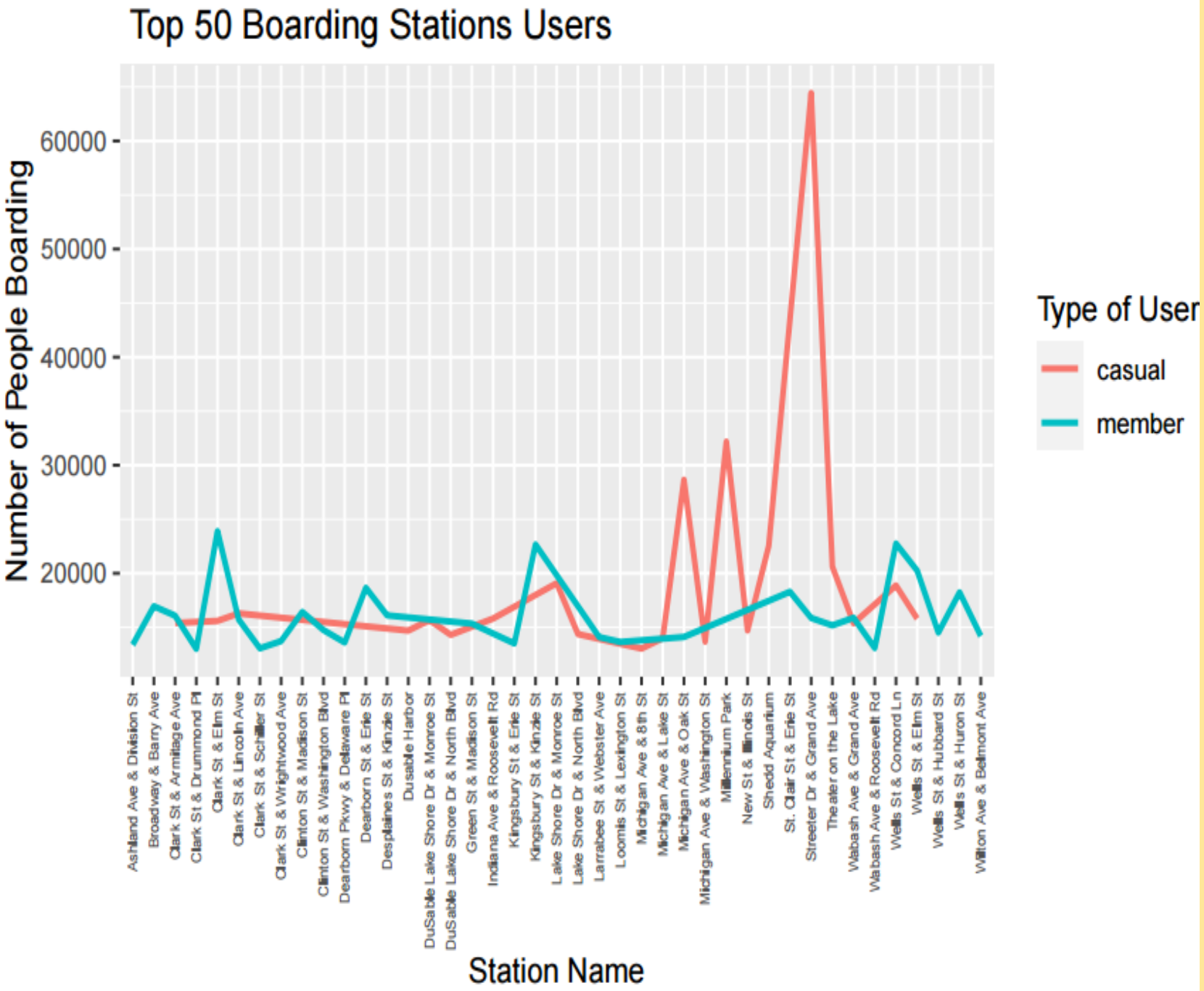








- It can be seen that although the total distance covered by members is greater than that of the casuals when it comes to the average distance covered by individuals of both groups, the casual exceeds here. It is mainly because even though members are riding more than the casuals are, their greater population reduces their average distance. Nonetheless, the differences are minimal.
- But, when we see the duration of trips covered by members and casuals, there is a striking difference. Here the members are riding notably lesser amount of time than what the casuals are, even though they ride more distance.
  - This implies that the larger distance of members is not because they are riding more distance per ride, it is because they ride regularly at a fixed one.
  - Whereas the casuals even if they ride at less frequencies, their distance per ride is greater than that of members.
- The above analysis confirms our hypothesis and tells that members are office goers whereas casuals can be using it for anything, be it going for work or leisure purposes.



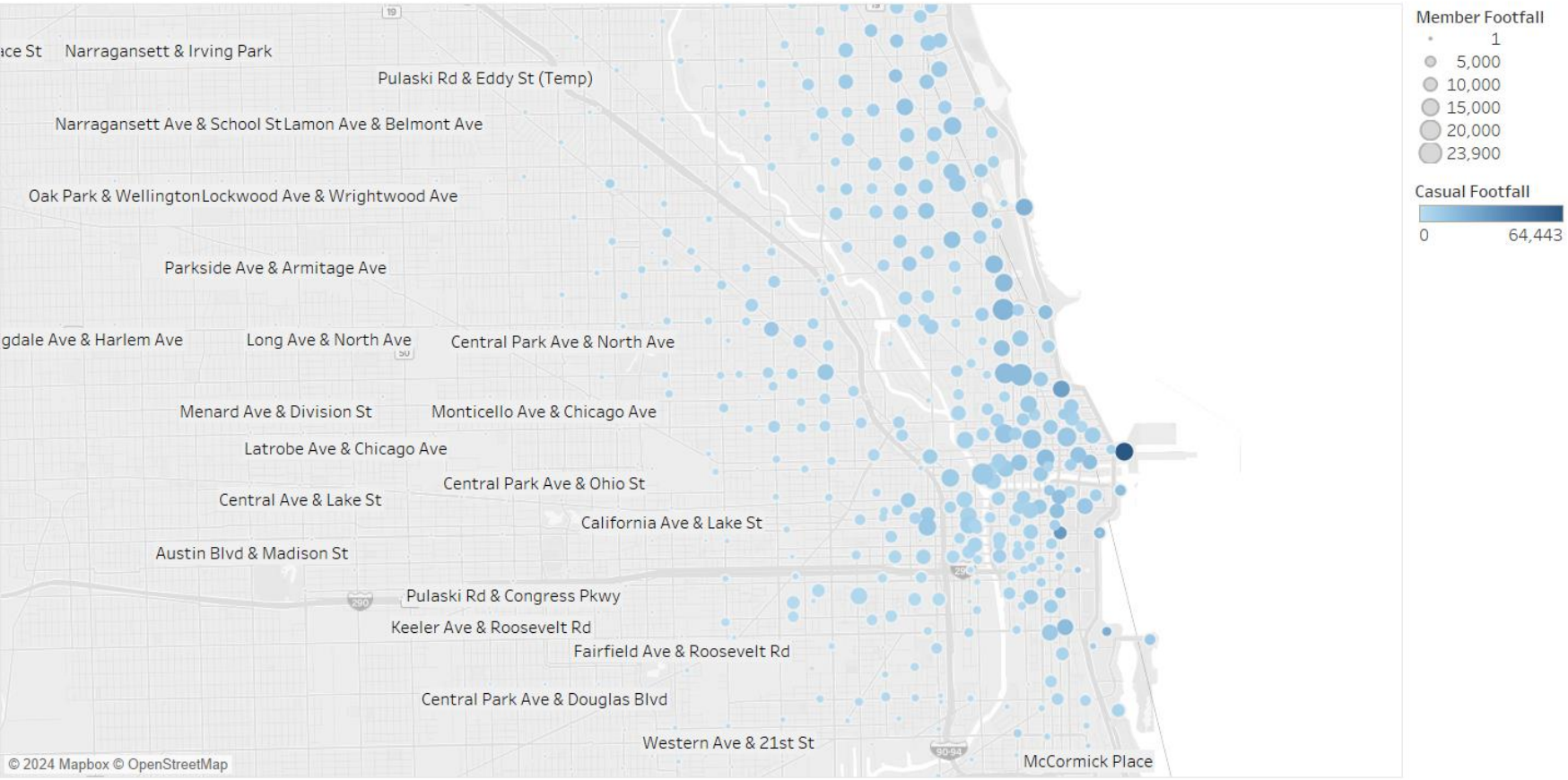
Analysing the footfall of people at different stations

- We can see the top 50 stations where the number of both casuals and members is highest. This information would be crucial to understanding at which stations advertising would be more effective.
- Streeter Dr & Grand Ave is the station with the most casuals who are the target users.

Top_10_stations <chr>
Streeter Dr & Grand Ave
Millennium Park
Michigan Ave & Oak St
Theater on the Lake
Shedd Aquarium
Wells St & Concord Ln
Lake Shore Dr & Monroe St
Lake Shore Dr & North Blvd
DuSable Lake Shore Dr & North Blvd
Wabash Ave & Grand Ave

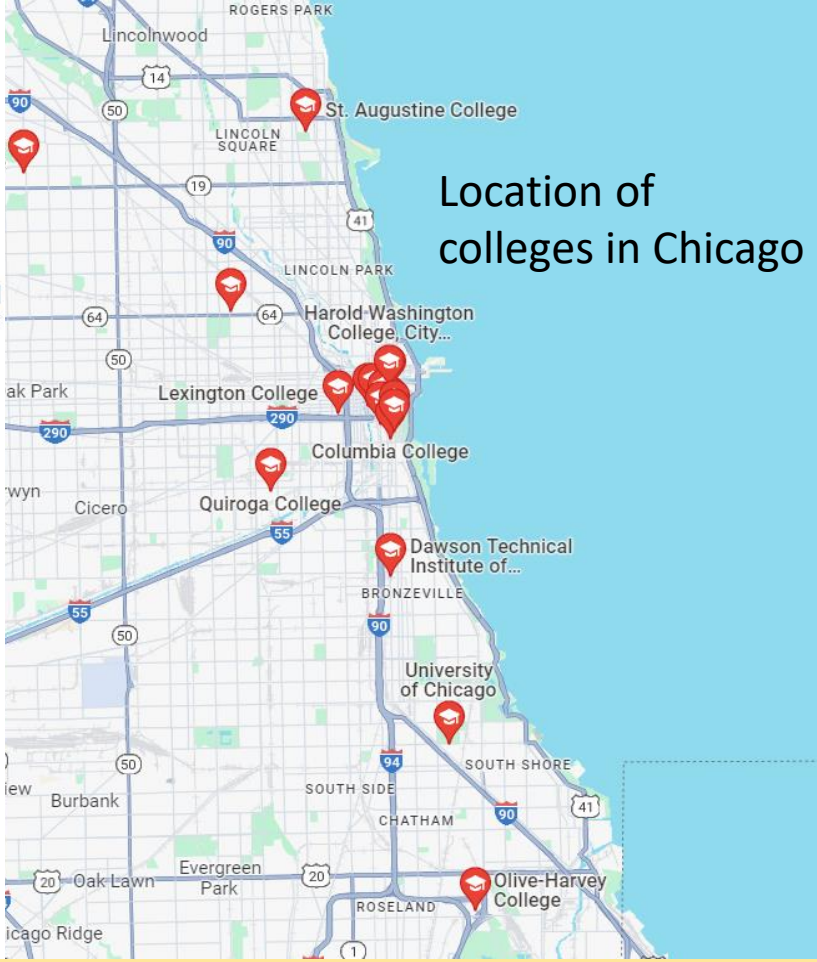
- Apart from this, the next slide uses the map to analyse a crucial aspect of a major segment of the users in the casuals group.

Footfall of users at different stations



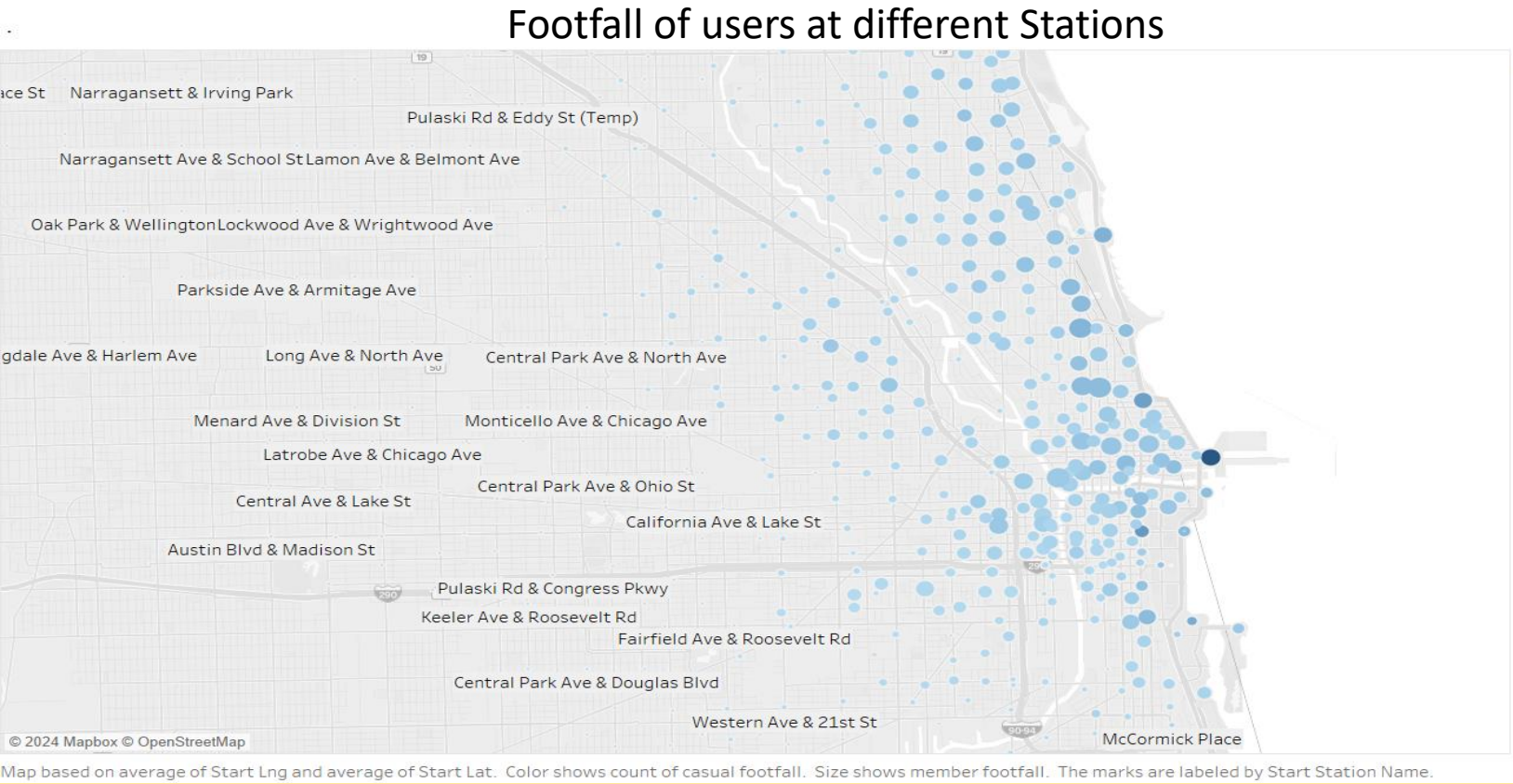
Map based on average of Start Lng and average of Start Lat. Color shows count of casual footfall. Size shows member footfall. The marks are labeled by Start Station Name.

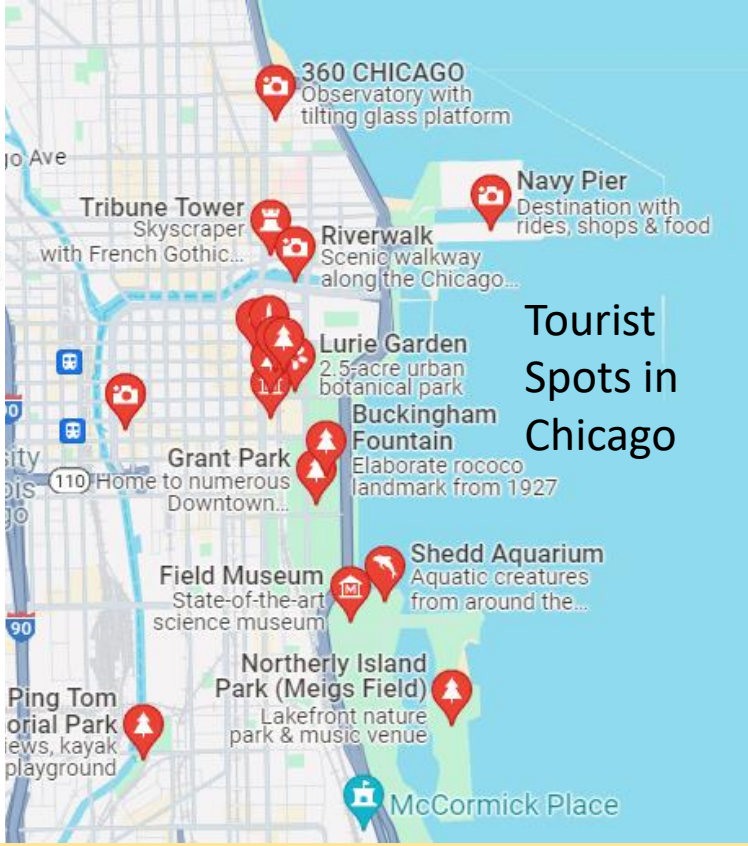
Location of colleges in Chicago



- Above are 2 maps. The left one shows the dominance of users at different stations whereas the right one shows the location of colleges in Chicago. We can see the commonality in both of them.
- The stations where the causal footfalls are the most are at the places where most of the colleges are.
- Hence, now we can confidently say that a significant number of people in the casual group are college-going students.
- Also, these are the places where some major offices in Chicago are located. This also signifies that there are also a good number of working people in the group.

Footfall of users at different Stations



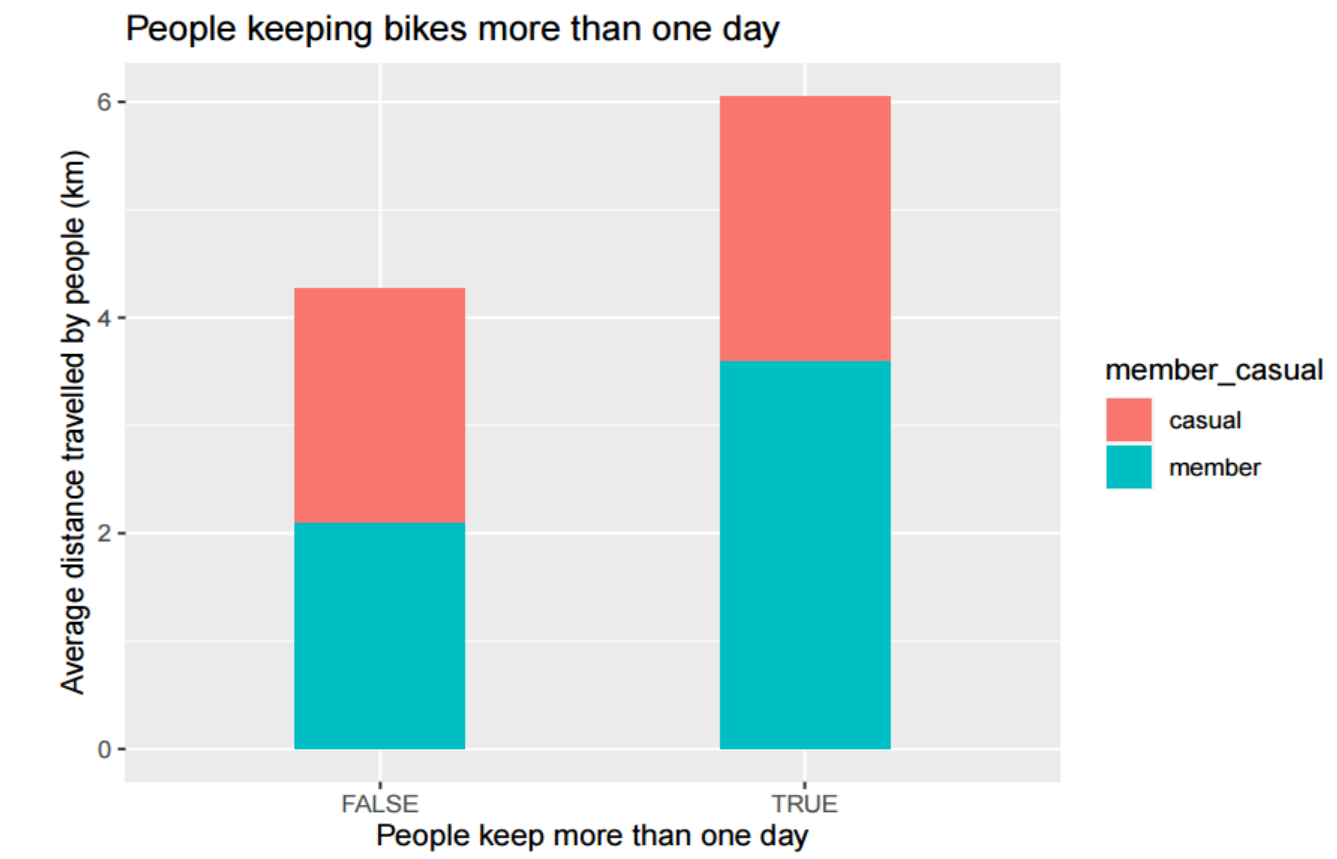


We can also see that the stations with the most number of users, especially casuals are people also places with tourist attraction spots and places from where one can go on trips. There may be people who use the bikes to go on trips and explore the city. The next slide analyses this aspect with a little more detail.

How many people use bikes for trips

People returning bikes the same day vs those who do not

```
# A tibble: 4 x 5
  member_casual not_same_day    Freq total_dist_km avg_dist_km
  <fct>         <fct>          <int>         <dbl>       <dbl>
1 casual        FALSE        2041039      4453857.        2.18
2 casual        TRUE            7263        17962.        2.47
3 member        FALSE        2536066      5293338.        2.09
4 member        TRUE            3736        13410.        3.59
```



- The table shows the number of people who return the bikes on the same day and that of those who do not.
- It can be seen that the number of casuals who do not return the bikes is approximately double that of the members.
- The reason behind seeing this number is that, if people are not returning the bikes on the same day then they may be using them to go on trips and explore the city.
- We can see that the average distance travelled by people not returning their bike on the same day is fairly higher than those who do.
- Also, there are a fairly good number of casuals who do so.
- Hence, we can say that casuals are also using bikes for going on trips, and exploring the city and tourist spots.

SUMMARY OF ALL ANALYSIS

- Our hypothesis that members are majorly the working population is true. Although members also use bikes for leisure activities, the percentage of such members is less.
- Whereas it would not be correct to say that casuals are entirely the people using the bikes for leisure activities. According to the analysis, people also use bikes for work commutes, and other than that, those people who use it entirely for leisure activities may also be in the working population, but use other modes of travel for work commutes.
- A good diversity in the usage patterns of casuals can be seen:
    1. A good number of casuals are college students, who use the bikes for their commute to colleges.
    2. People going on trips and sightseeing are also the ones in this group.
    3. Night shift workers are also using the bikes for their commute.
    4. People who use the bikes for other leisure activities in their free time and on holidays form the majority of this group.

Recommendations

1. Tapping the majority of people using bikes for leisure activities:
   - There can be a discount for members during holidays and weekends when the number of casuals is in the majority. This would persuade them to take memberships.
   - A dynamic pricing system must be made based on an hourly basis.

     Ex- At 1 PM both the members and casuals are high and equal in number. That may act as a sweet spot to provide special pricing for members. It would persuade the casuals to take membership as it is the hour when they also need bikes. If we provide it at 5 PM when the casuals peak, then the number of Members is even higher and we would have to give discounts to even existing members, draining the profit.
   - Advertisement signifying how using Cyclistic bikes enables the person to explore the city and go on trips more cheaply and easily.
2. Tapping the college students:
   - There can be special discounts for students upon showing their college IDs on membership.
   - Use of hoardings near college areas and at stations with higher casual footfalls.
   - Use of social media and youth-relevant influencer marketing can render good results.
3. Tapping the casual working segment and night time workers:
   - Special advertisements can be made that indicate how using Cyclistic bikes eases the commute of people to their offices, and is cheaper than other modes of travel.
   - Special nighttime discounts and membership facilities can be drawn upon for the nighttime workers.

Conclusion and Sources

All the tasks were accomplished, and due analysis was given. It is hoped that a good marketing strategy considering all the stated points would help the company to accomplish the Business Goal of converting casual users to annual members, hence, boosting the profits.

Source of dataset:

This dataset was provided along with the problem statement of the capstone project. Source

THANK YOU!