Annual Members vs Casual Riders

A comparison to understand different behaviors

Business Task:

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

Objective

To design marketing strategies to convert casual riders into annual members

Prepare Step

1) Data were downloaded from the Divvy Bikes website https://www.divvybikes.com/system-data
2) Data were organized by month and stored directly by the company
3) It was requested to start working on January, February and March 2021, but especially on March 2021 data

Process step

1) Some anomalies were present:
   - Ride lengths negative or equal to 0
   - Many NULL cells in the Start or End Station columns
   - Time formatting was not consistent
   - Ride ending time preceding ride starting time for some rows

2) I chose to work with Google Sheets and Big Query

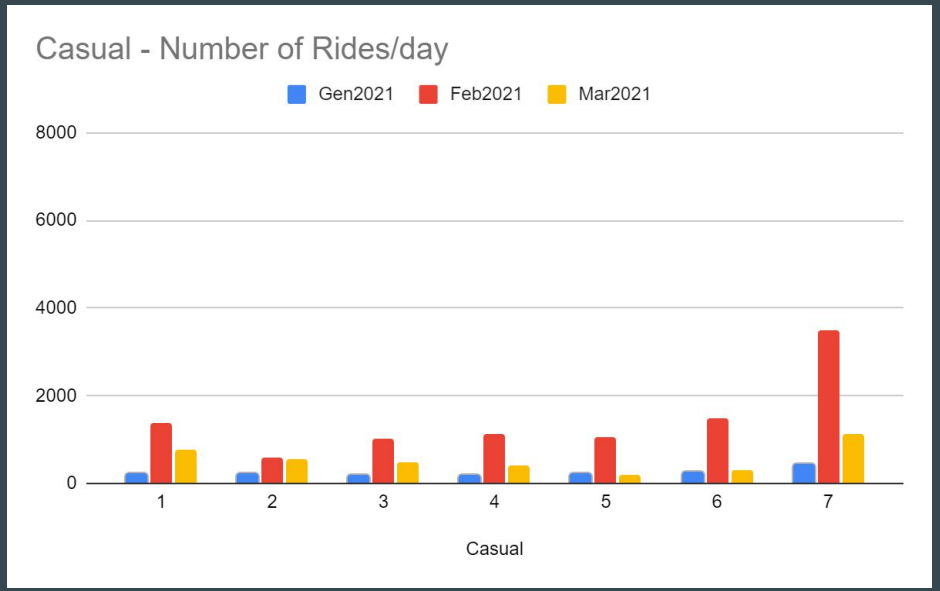
How I transformed data to make them clean

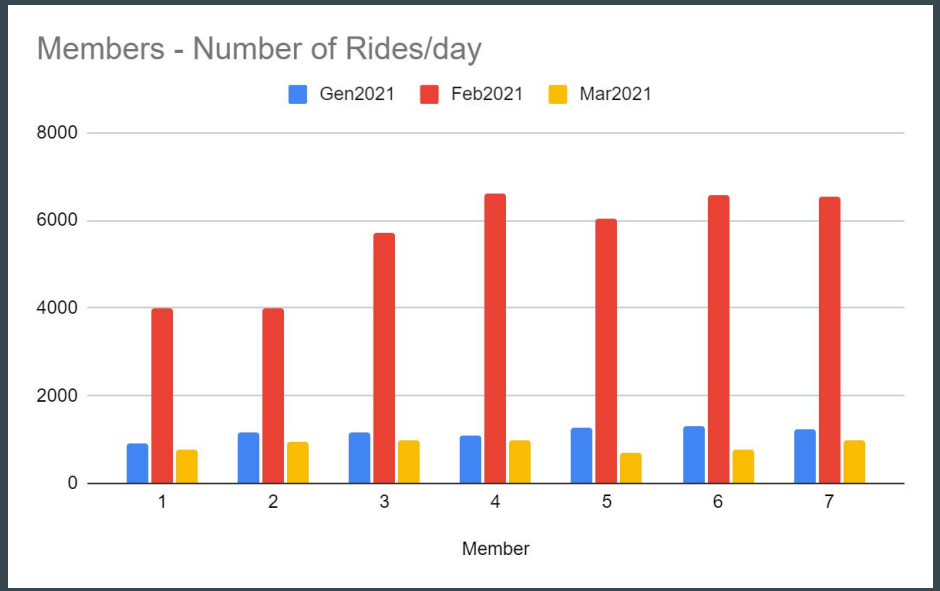
1) Removed duplicates
2) Conditional formatting especially for date and time cells and duration time
3) Removed inconsistent outliers
4) Opted for saving in another sheet rides with duration equal to 0 (to check what was the reason)
5) Checked again through aggregate functions (example: minimum function to check if any negative ride length was still present)

Analysis

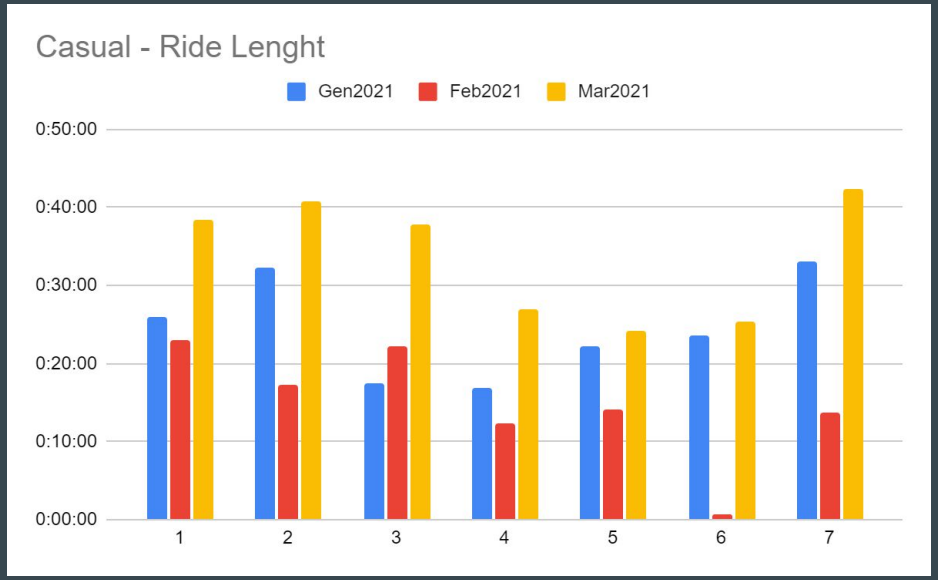
1) I calculated ride lengths and days of the week when rides were made along the months
2) I calculated the maximum, the average and the minimum ride length
3) I pivoted data to explore trends regarding members and casual riders

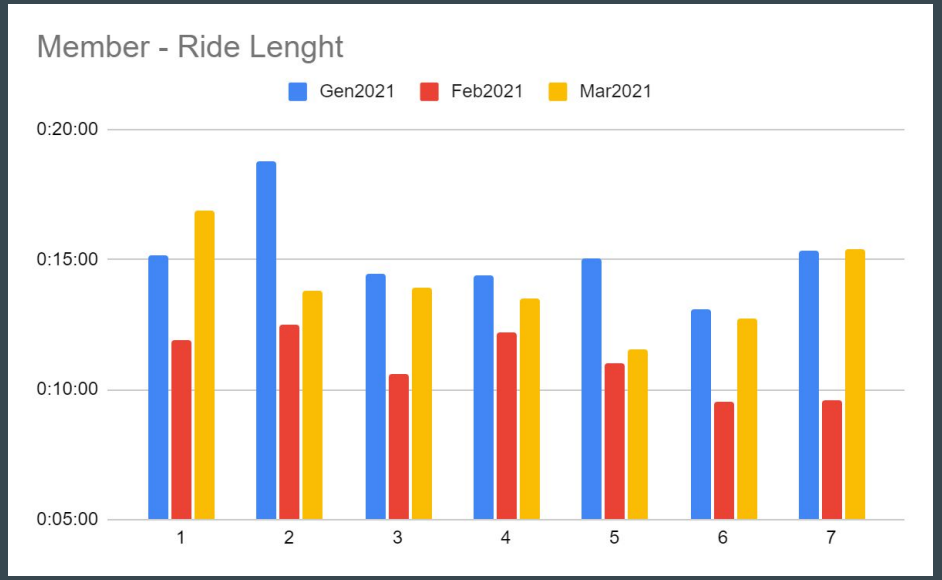
Comparison: Number of rides





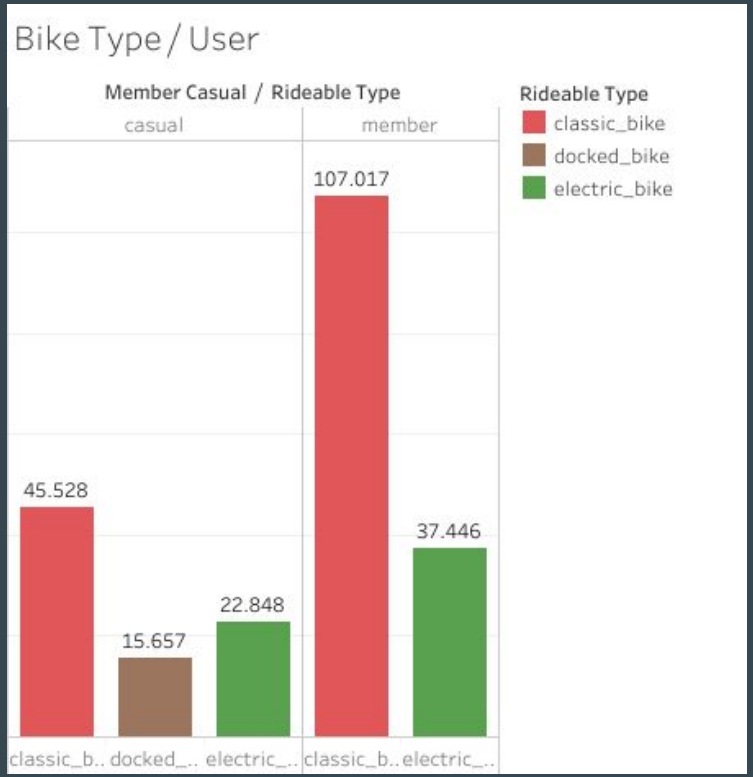
Comparison : Ride Lenght



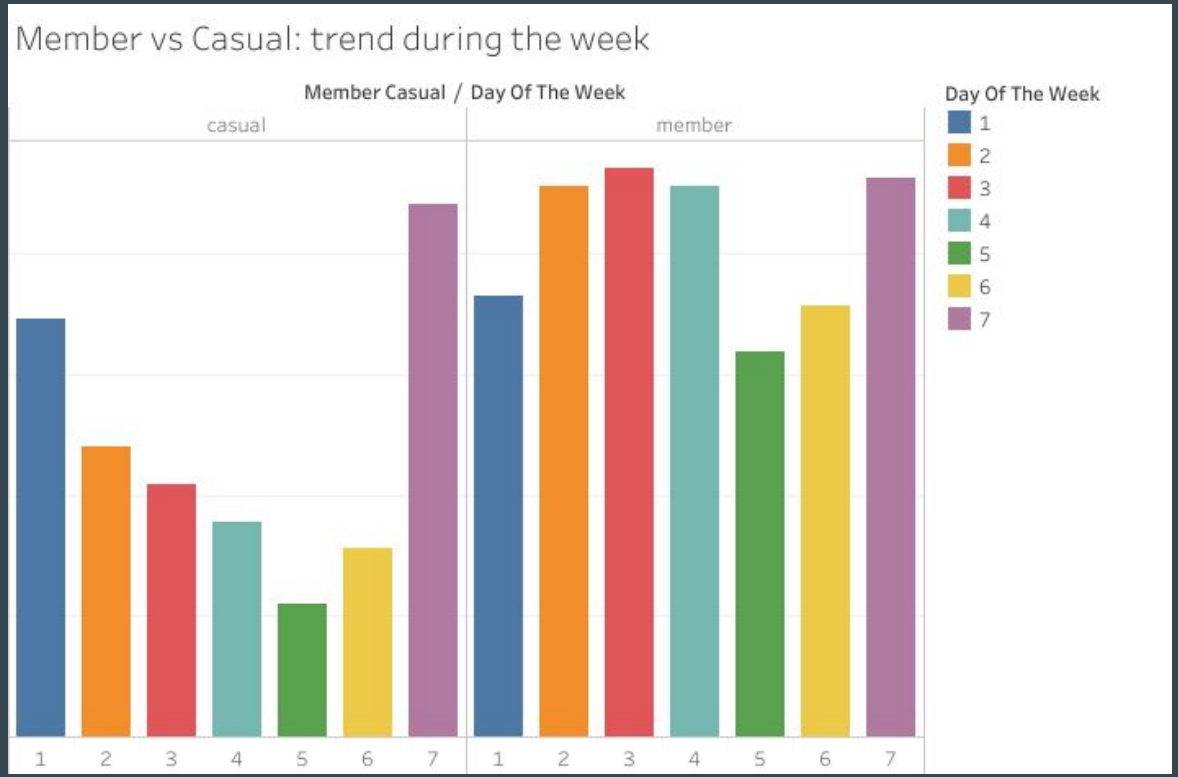


Now let’s focus on March 2021...

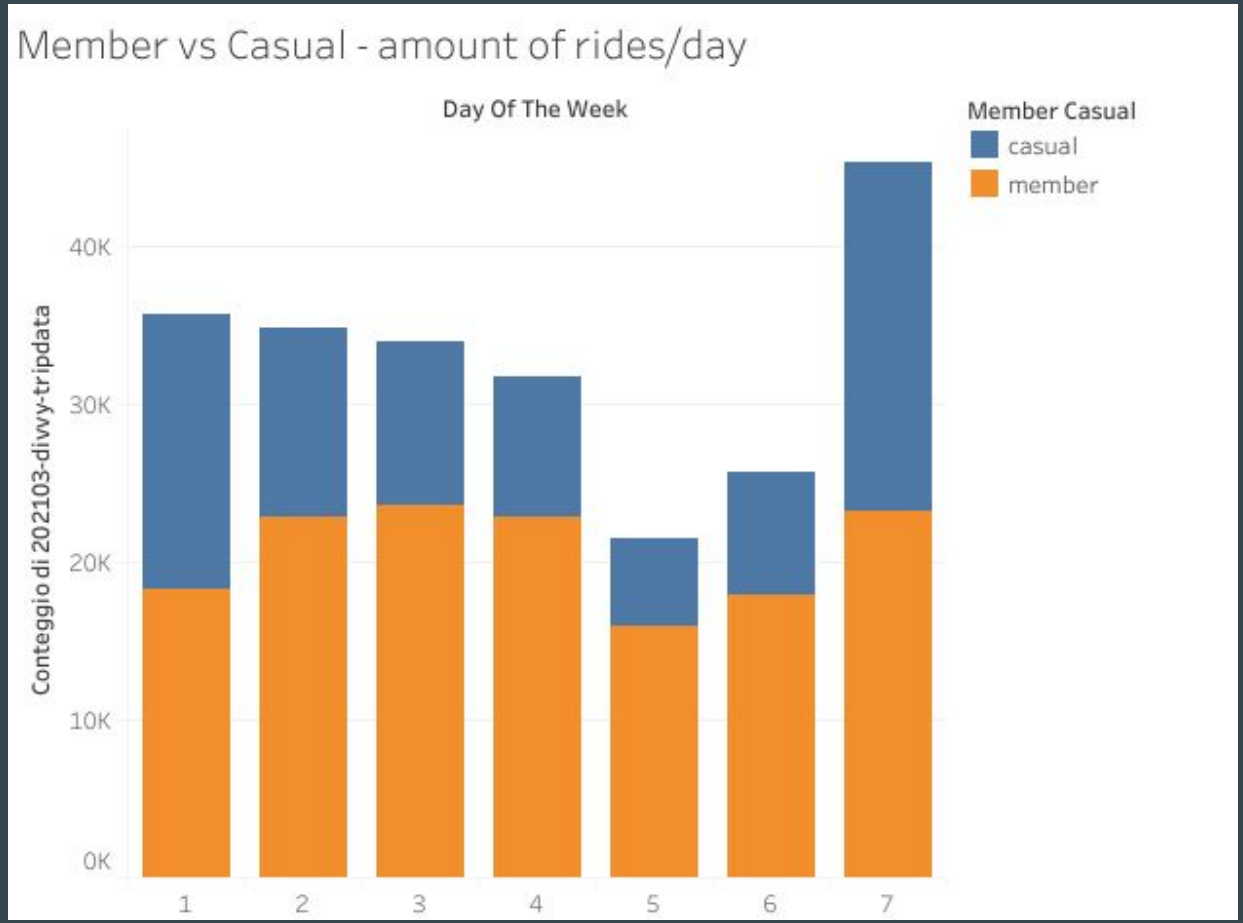
Bike Type / User



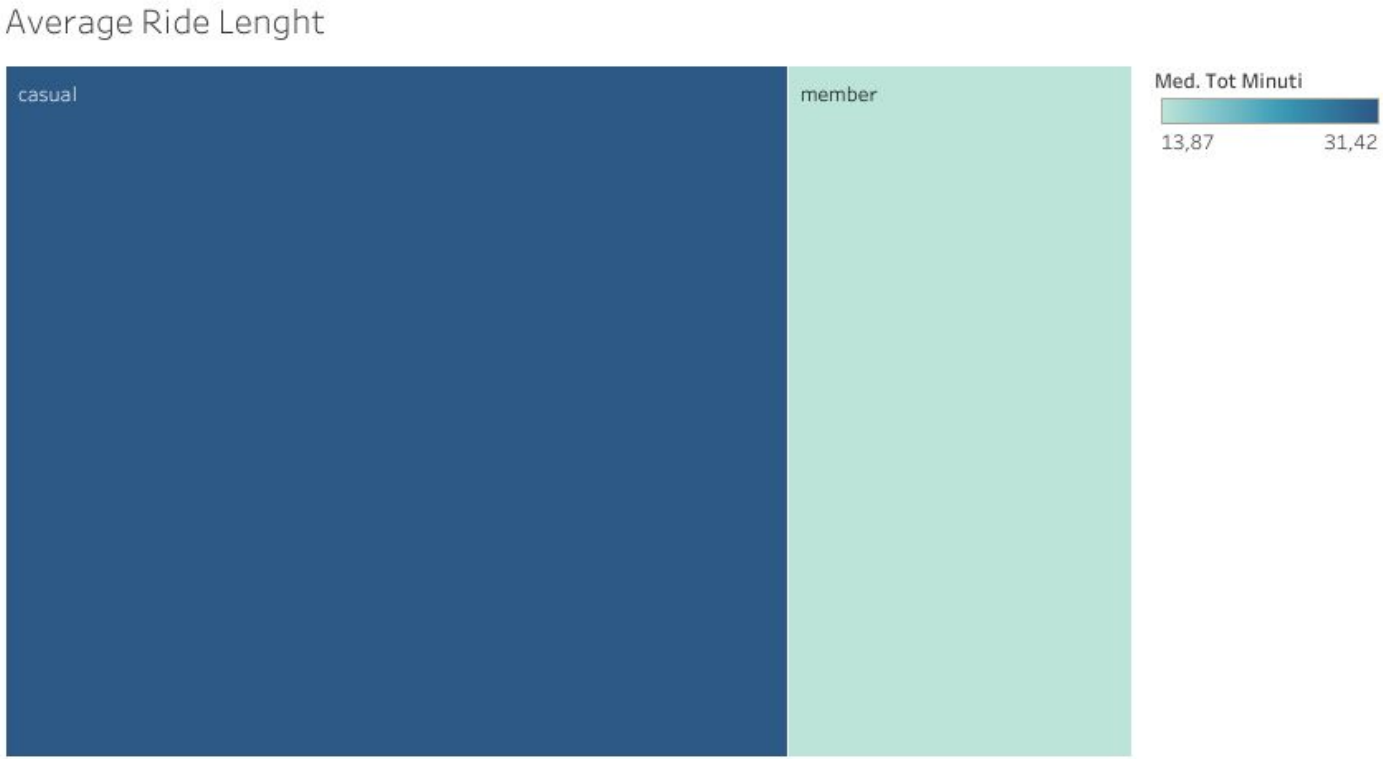
Number of rides during weekdays



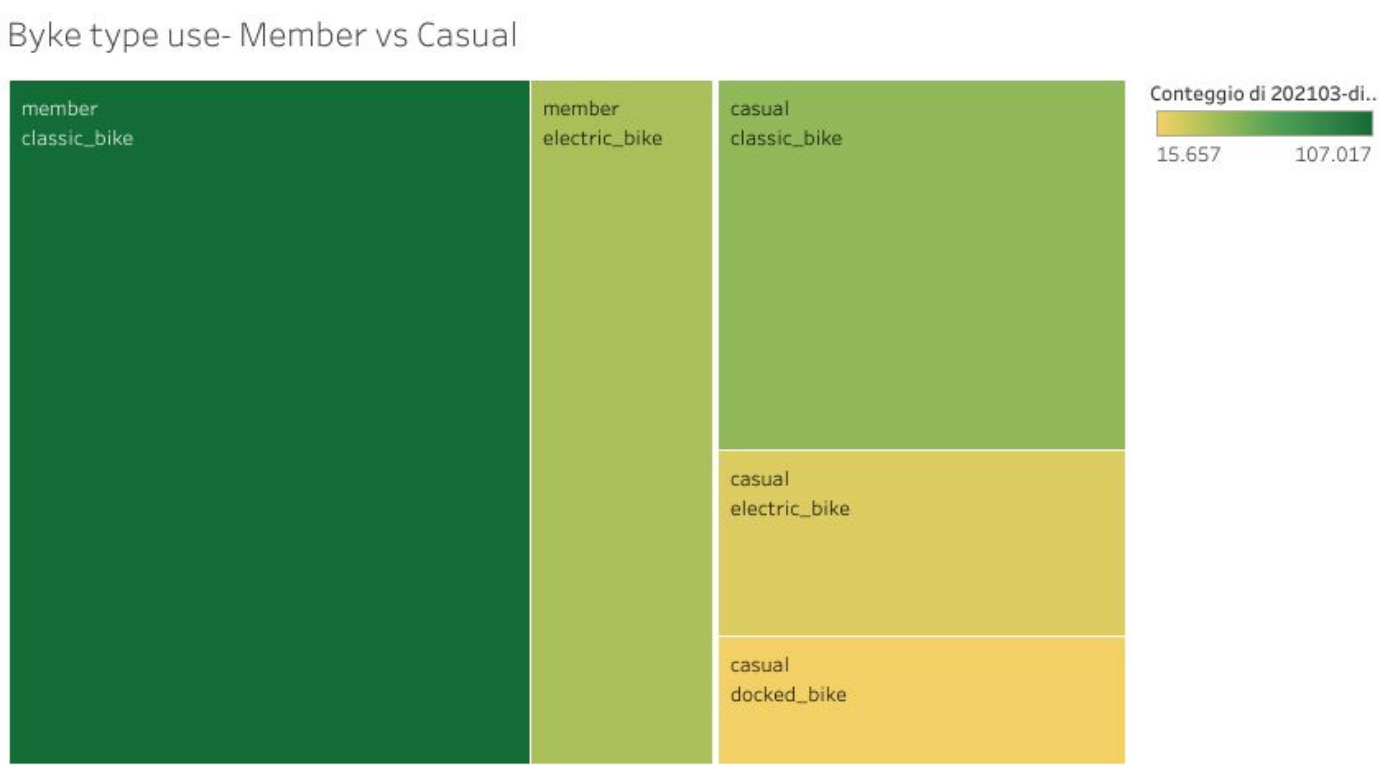
Member vs Casual - amount of rides/day



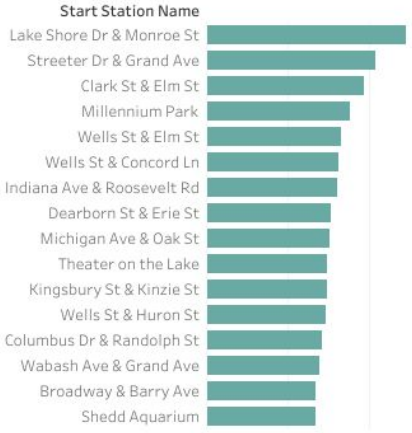
Average Ride lenght



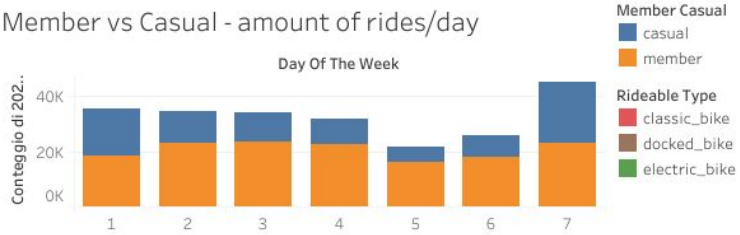
Bike Type / User



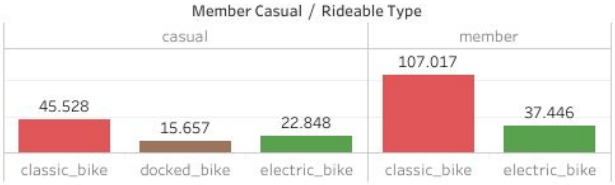
Start Station Hits



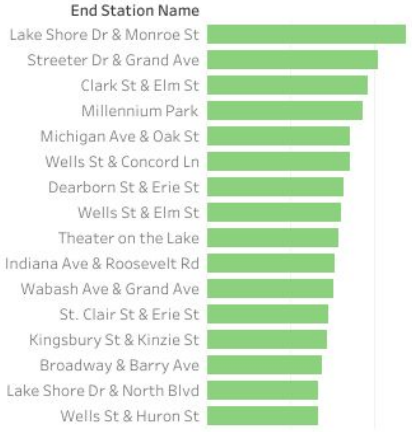
Member vs Casual - amount of rides/day



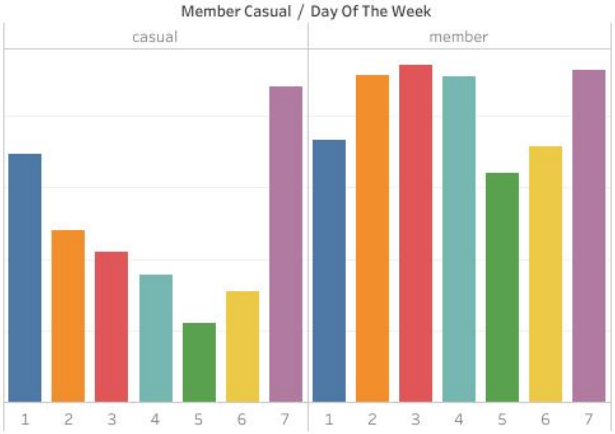
Bike Type / User



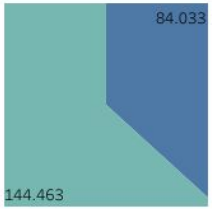
End Station Hits



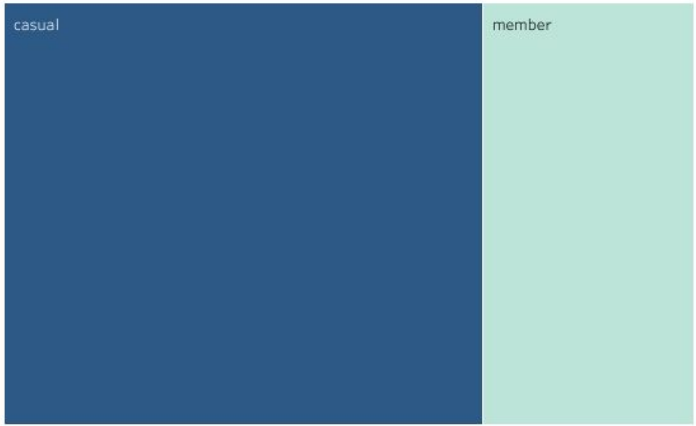
Member vs Casual: trend during the week



Member vs Casual - Number of rides



Average Ride Lenght



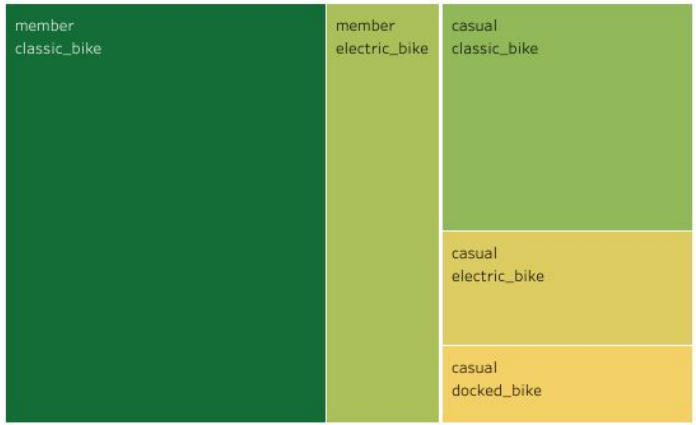
Med. Tot Minuti
13,87 31,42

Member Casual
casual
member

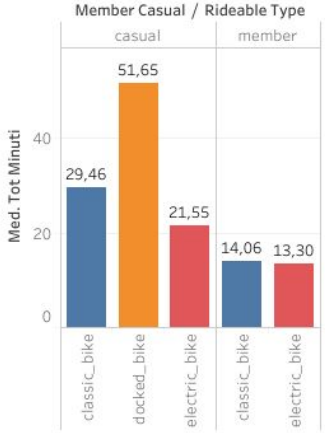
Rideable Type
classic_bike
docked_bike
electric_bike

Conteggio di 202103-divv..
15.657 107.017

Byke type use- Member vs Casual



Ride Lenght /Bike Type



Takeaways

1) - Casuals use Cyclistic service especially in the weekend (peak ->Sunday) -> LEISURE ACTIVITY

   - Members use the service all the week (quite uniform trend in the month) -> this lead to think of a WORK COMMUTING behaviour

2) Casuals ride for MORE TIME than members do (almost the double of the duration).

   Members: ride length are shorter BUT number of rides are much more!

3) - CLASSIC bikes are the most used in both categories

   - Docked bikes compared only among casuals than members (in March 2021)

4) Knowing the more hit Stations for departing and arriving may be important to know where allocate more available resources.

5) Among casuals DOCKED BIKE is the least used (in number of rides) but the kind of bike ridden for much more time

SO, WHAT NEXT?

1) Collect more data concerning Cyclistic Users (Survey?) to know :
   - AGE
   - WORKER/STUDENT
   - HOME LOCATION
   - ...

2) Collect and work on data covering 1 OR 2 YEARS to highlight annual trends

3) Match Cyclistic service Data with Weather Data to better understand behaviours

4) Check the origin of many NULL cells in the database

THANK YOU!

...questions?