

Bike Case Rapport:

Background information:

I am working for Cyclistic bike-sharing company.

Bikes can be unlocked from one station and returned to any other station in the system anytime. Cyclistic has flexible pricing plans: single-ride passes, full-day passes, and annual memberships.

Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual

memberships are Cyclistic members.

Director of marketing believes the company's future success depends on maximizing the number of annual memberships as

finance analysts have concluded that annual memberships are much more profitable than casual riders.

She also believes that there is a good chance of converting casual riders to members as they are already aware

of Cyclistic program and have chosen it for their mobility needs.

Business task:

To analyze user behaviors on how annual members and casual riders use Cyclistic bikes differently to make recommendations on how to convert casual riders into annual members

Dataset used:

I used the 12 month dataset from the year 2021 taken from <https://divvy-tripdata.s3.amazonaws.com/index.html>, made publicly available by Motivate International Inc. Due to data privacy issues, I won't be able to connect pass purchases to credit card numbers to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes.

Data cleaning:

I used rstudio for cleaning data

I started by combining the sheets for trip data for 12 months from january 2021 to december 2021

The data set has 5595063 rows and 13 variables:

ride_id: Unique ID assigned to each individual ride

rideable_type: Type of bicycle (classic, docked, electric)

started_at: Date and time at the start of the trip

ended_at: Date and time at the end of the trip

start_station_name: Name of the station the trip started at

start_station_id: ID number of the station the trip started at

end_station_name: Name of the station the trip ended at

end_station_id: ID number of the station the trip ended at

start_lat: Latitude of starting location

start_lng: Longitude of the start location

end_lat: Latitude of ending location

end_lng: Longitude of ending location

member_casual: Type of membership (Casual or Cyclistic Member)

I noticed that the data set contains some missing values in the variables of (start_station and end_station) .

i started cleaning by formatting column names to be consistent

*i removed duplicate records and dropped rows with missing values

*i added a column for weekday the rides start in letters and a column for month in plain letters

* in order to calculate trip duration i converted started_at and ended_at to string to characters then i added a column for trip duration

* i filtered the dataset to remove rows that contain negative durations

* i added columns in which i extracted start and end hours

*i summarized mean /max trip durations by weekdays , then i sorted them . i noticed that the longest rides are made in weekends

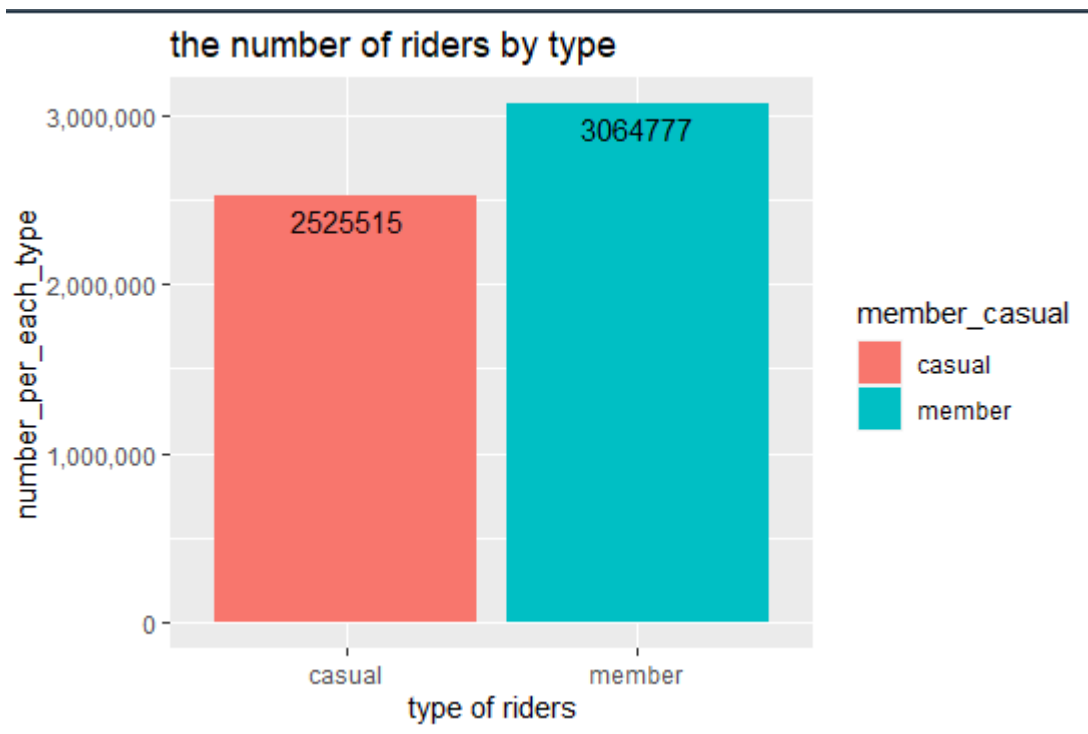
	week_day	mean Ride Time
1	Sunday	26.51248
2	Saturday	24.96771
3	Friday	20.03652
4	Monday	20.00690
5	Tuesday	17.81256
6	Thursday	17.76857
7	Wednesday	17.64118

* i also calculated the number of members and casuals :

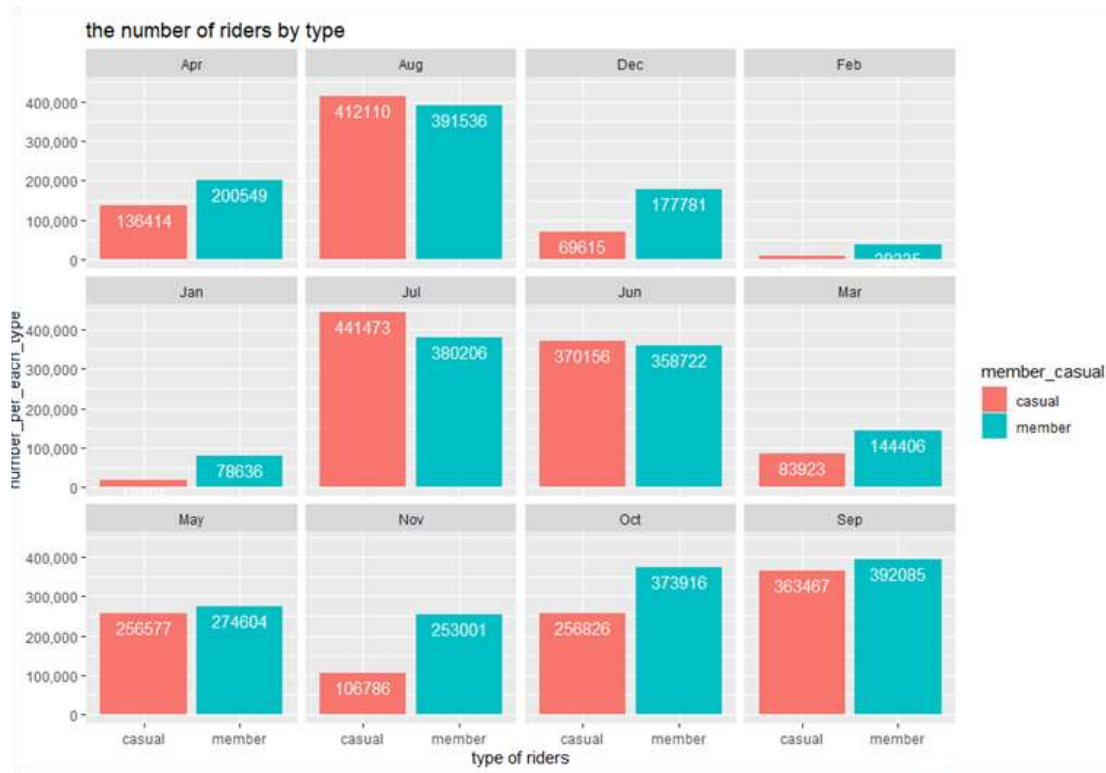
	member_casual	n
1	casual	2525515
2	member	3064777

Data visualization :

*I plotted the number of riders and their type in a bar chart



*then i plotted the number of riders and their type all year round in the differant months



i've come with this initial remarks:

all riders:

Riders in general are most likely to ride in hot weather (may-juin-july-august -sep-oct)

Then the number of riders decreases as the weather gets colder (nov-dec-jan-fev-mar-april)

casual riders:

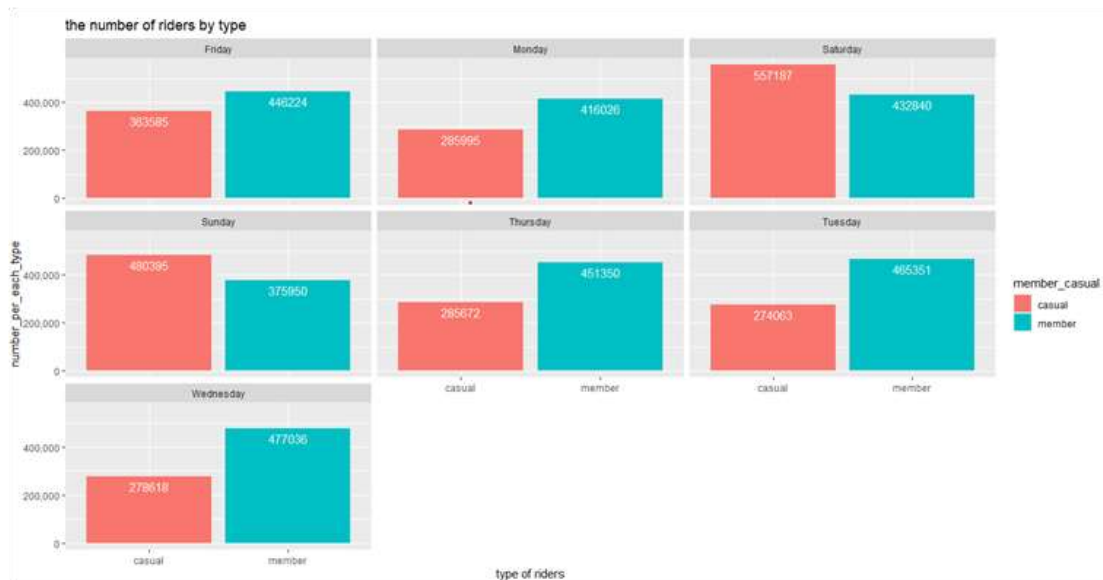
Summer season(Juin -july-august) the number of casual riders surpasses members

conclusion:

we can conclude from these bar charts that Casual rides for pleasure meanwhile

Members ride for commuting to work/study ect...

* then i plotted the same plot as earlier but with weekdays this time



Members:

Member riders use bikes for Monday to Friday more than casual ones

Casuals:

casual riders surpasses members on weekends

Conclusion:

this confirms that members use rides to work /school/to commute

and casual riders use bikes for leisure purposes.

*The plot of number of rides in different year months in one single plot also confirms that Members tend to use bikes more all year round especially in work seasons meaning from sept to may

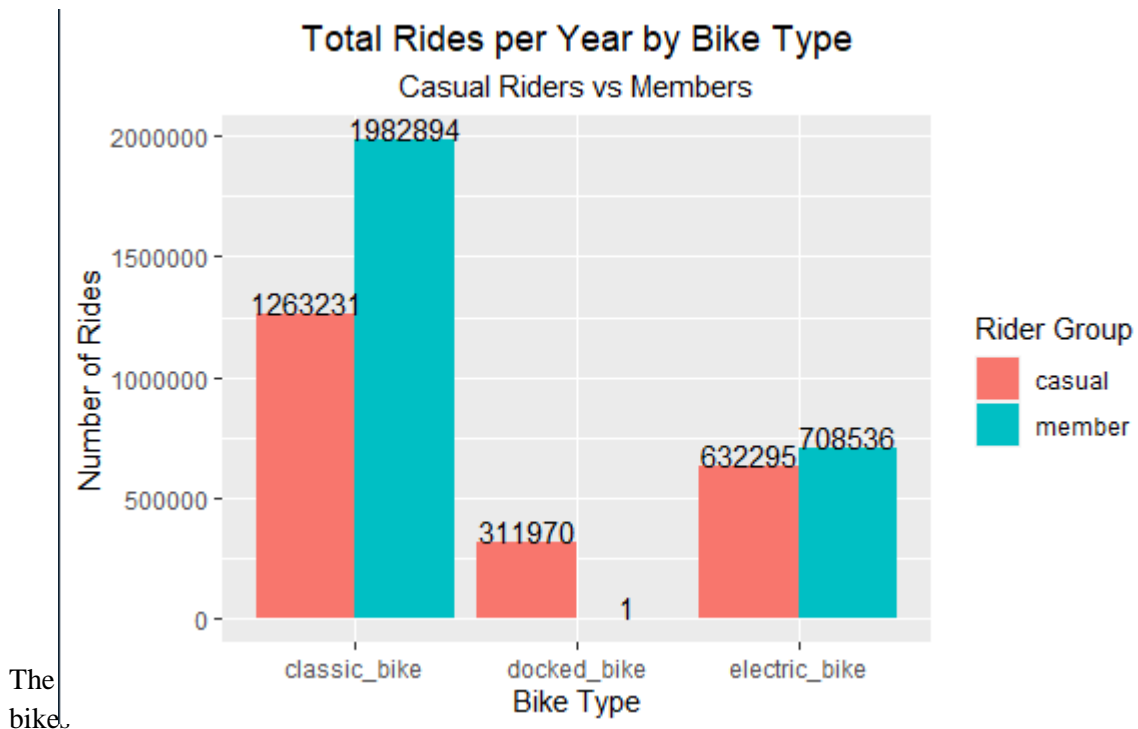
Casual riders surpasses members on summer season only.

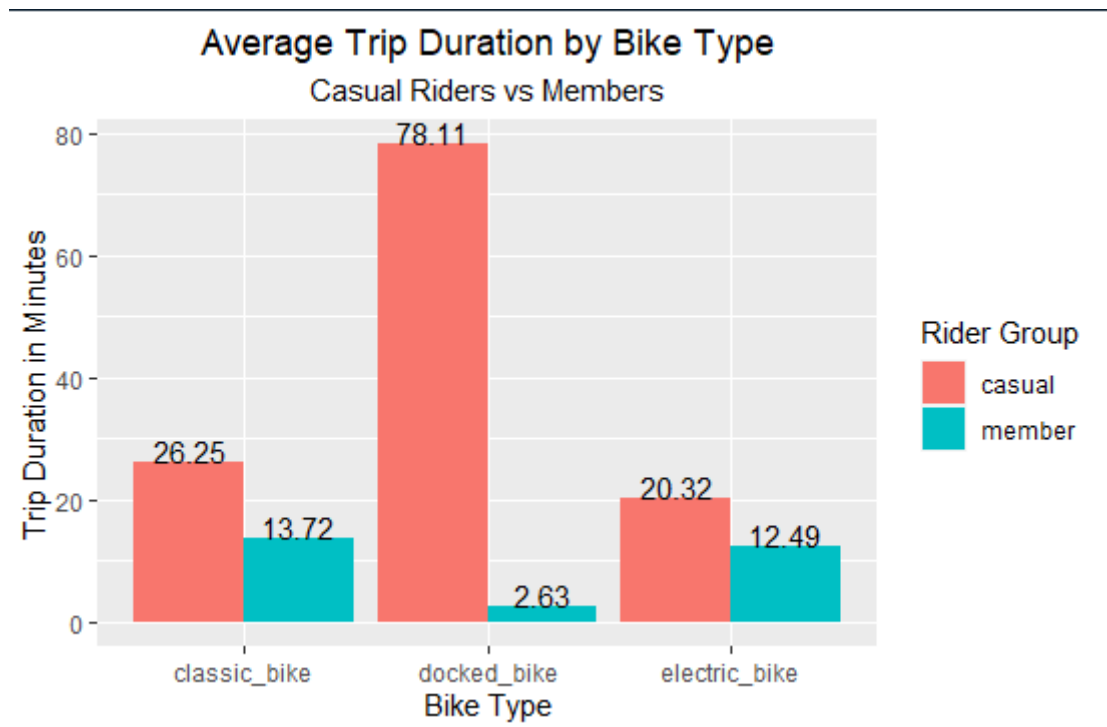
As we move on to the next metric to compare which is trip duration, I plotted the average trip duration of both type of riders:

I noticed that Casuals tend to use bikes twice longer than members because apparently casuals use rides for trip purposes while members use it for short destined rides to commute.



As I compared the number of rides with each type of bicycle we can see that the most preferred type of bike is classic bikes by both rider groups





the plot of average trip durations by bike type in weekdays shows that docked bikes are mainly used by casuals all around the week and for longer ride times especially in weekends.



Conclusion:

I was presented with the task of analyzing historical data of the year 2021 to determine how casual and member riders use the service differently in order to offer recommendations on how to convert casual riders into members and I've come with these conclusions:

*Casual riders tend to use the service in nicer weather and they ride for longer periods also they tend to ride more on weekends which is an indicator for riding for fun purposes but not a concrete explanation.

*Members tend to be more consistent in riding all around the year and not only as the weather becomes more tolerable, also they take shorter rides which is an indicator for purpose-oriented courses such as commuting to work or study but again there is no concrete explanation for this pattern.

*There appears to be a cross population preference of classic bikes over electric bikes, moreover docked bikes seem to be mostly casual riders' preferences.

Some suggestions to convert casuals to members is for example offering discounts or bonuses for subscribing for all year round and present this form of discounted membership in the summer season since it's the most time in which the service is used by casuals and that way not only they will subscribe but also this will increase their use of the service around also colder weather since they have the advantage of the discount.

Also gathering riders' feedbacks on the service would be insightful to determine the cause of hesitation into committing to a membership.