# Cyclistic Bike-Sharing Report

*This case study is part of the capstone project under Google-Coursera's Data Analytics course*

Mark Hasegan

STEP 1 - ASK

**Key Stakeholders:** Cyclistic, Cyclistic marketing analytics team, Lily Moreno (director of marketing), Cyclistic executive team.

**Business Task:** Convert Cyclistic casual riders to annual members. How do casual riders and annual members use Cyclistic bikes differently?

STEP 2 - PREPARE

**Data Sources:** Cyclistic trip public data (Jan. 2022 – Dec. 2022)

**How the data is organized:** Data is kept in csv files separated by month. All the files contain usage of Cyclistic bikes including ride type, start and end stations (with station IDs and coordinates), and type of riders (casual or member). Some files are missing their data values for start and end station names and IDs, however, they provide the starting and ending coordinates for all observations.

**Credibility and ROCCC:**

- **Reliable –** Although starting and ending station names and IDs have not been provided, the starting and ending coordinates were provided and can be matched to with other files to identify the existing station names and IDs.
- **Original –** Data is proprietary to Cyclistic.
- **Comprehensive –** Although we are observing all data from 2022, the records date back to 2013. In 2022 alone, there were 5.66M transactions. This abundance of data makes it possible to provide accurate insights and results.
- **Current –** The data remains current and valid to come up with plausible recommendations.
- **Cited –** This is first-party data generated from Cyclistic's bike-sharing technology.

**Privacy, security, accessibility, licensing:** The data downloaded will be stored in my local drive and backed up on Google Drive as I go through the process. Raw data will not be shared by any means and processed data will only be displayed as tables, tibbles, and visualization (keeping the raw data inaccessible to the public. The data has been made available by Motivate International Inc. under the Data Licensing Agreement.

**How did you verify the data's integrity:** Despite the identification of the missing values (start/end station names and their corresponding IDs), the existence of lat/long coordinates ensures that each transaction is geospatially identifiable.

**How does it help you answer your question:** It is necessary to know this because the missing values account for 30% of the transactions. If it would be determined that accounting for this 30% is necessary to come up with a valid analysis, there is secondary data in the form of the coordinates that we can use.

**Are there any problems with the data:** Although about 30% of transactions have missing station names and IDs, data integrity remains intact with the presence of coordinates across all transactions. The challenge for me (member of the analytics team) is to generate the station names and IDs for the missing values if it would be required.

STEP 3 - PROCESS

**Following are the steps in cleaning and manipulating the data:**

**Excel:**

**Columns:**

1. Added “ride_length” (end_time – start_time)
2. Added “day_of_week” (what day the ride was used)
3. Added “most_occurances_day” and “most_occurances_count” (which days had the most rides used)
4. Added “day_of_week” (day of week that ride was used)
5. Added “ride_time” and “avg_time” (total sum and average of ride lengths for members and casuals)
6. Added “Hours” (what hour of the day the ride was used)
7. Added “daily_count” and “daily_avg” (daily ride count and daily averages members and casuals)
8. Added “num_ride”, “classic_bikes”, “docked_bikes”, and “electric_bikes” (total count for members and casuals)
9. Added percentages for “ride_time”, “num_ride”, “classic_bikes”, “docked_bikes”, and “electric_bikes ”

**Tables:**

10. Added a Hours table (what hour of the day the ride was used for members and casuals)
11. Added a AvgTimes Table (daily average times for members and casuals)
12. Added a Bike Table (total and daily counts for classic, docked, electric bikes for members and casuals)

STEP 4 – ANALYZE

**What days and hours do they use the service the most:**

- Members: Usage hours had a large surge at 8am and peaked at 5pm. Thursdays were the most used day of the week (532,261) and the lowest were Sundays (387,223) throughout the year. Members used the service most on the weekdays.
- Casual: Usage hours peaked at 5pm. Saturdays were the most used day of the week (473,190) and the lowest were Tuesdays (263,746) throughout the year. Casual riders used the service most on the weekends.

**How are they using the service:**

- Members: Total time spent on rides was approx. 37,480 minutes (624h & 40m) and had a total of 3,345,685 rides which made up approx. 59% of all rides (5,667,717) in the year. The average time spent using a ride was a little over 12 minutes and the highest day average were Saturdays (13m & 13s). Classic bikes were the most popular bike type used at about 51.1% (1,709,755) with the rest going to electric bikes because no member used any docked bikes this year. Classic bikes were used on Tuesdays the most (268,366) and electric bikes were used on Thursdays the most (264,806).
- Casual: Total time spent on rides was approx. 44,665 minutes (744h & 25m) which is approx. 54.4% of all time spent on rides this year and had a total of 2,322,032 rides. The average time spent using a ride was a over 22 minutes (shows they were using it for longer periods which equates to more money spent) and the highest day average were Saturdays. Electric bikes were the most popular bike type used at about 53.95% (1,253,099), casual riders were the only ones to use docked bikes which was about 7.65% of their rides (more electric and docked bikes shows that casual riders like to go for cost-efficient methods). All bikes were used on Saturdays the most.

**Conclusion:**

- Both members and casual riders have peak usage hours at around 5pm, showing that both use the service the heaviest during the end of work day rush hour. However, members seem to use the service most during the weekdays meanwhile the casual riders use the service the most on the weekends. Members had 3,345,685 rides, approx. 59% of the total 5,667,717 rides during this year and also made up 45.6% (624h & 40m) of the time that was spent on all rides. Casual riders only had 41% (2,322,032) of the total rides and made up 54.4% (744h & 25m) of the total time spent on rides which was approx. 1,369h & 5m. The average time spent on a ride for members was right over 12min while the casual riders was over 22min. This shows that casual riders use the service for longer periods of time than members which costs more for them, also, more electric and docked bikes shows that casual riders like to go for cost-efficient methods (this highlights an opportunity that is present for converting casual riders into members). Members used the classic bikes majority (51.1%) of the time with the rest going with electric bikes, there were no accounts for members using docked bikes. Casual members used the electric bikes the most (53.95), docked bikes were used about 7.65% of the time with the rest using classic bikes. Members used classic bikes on Tuesdays the most and electric bikes on Thursdays the most, and for casual riders all bikes were used the most on Saturdays.

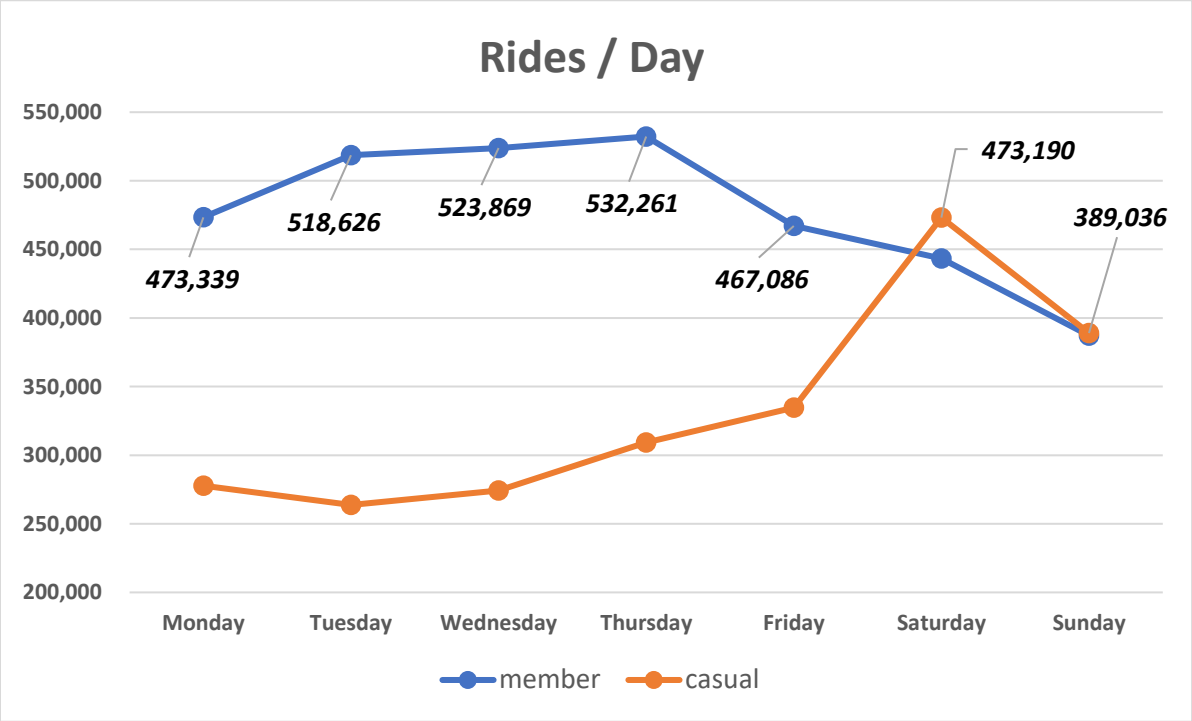
STEP 5 – SHARE

**Numbers:**

- TOTAL number of rides in 2022: 5,667,717
- Number of rides by members: 3,345,685
- Number of rides by casuals: 2,322,032

59% of rides were members, 41% of rides were casual riders.

**Figure 1.** Shows the number rides per day through the week for both members and casual riders with members peaking on Thursdays and casuals peaking on Saturdays. Members ride the most on weekdays and casual riders ride the most on the weekends.

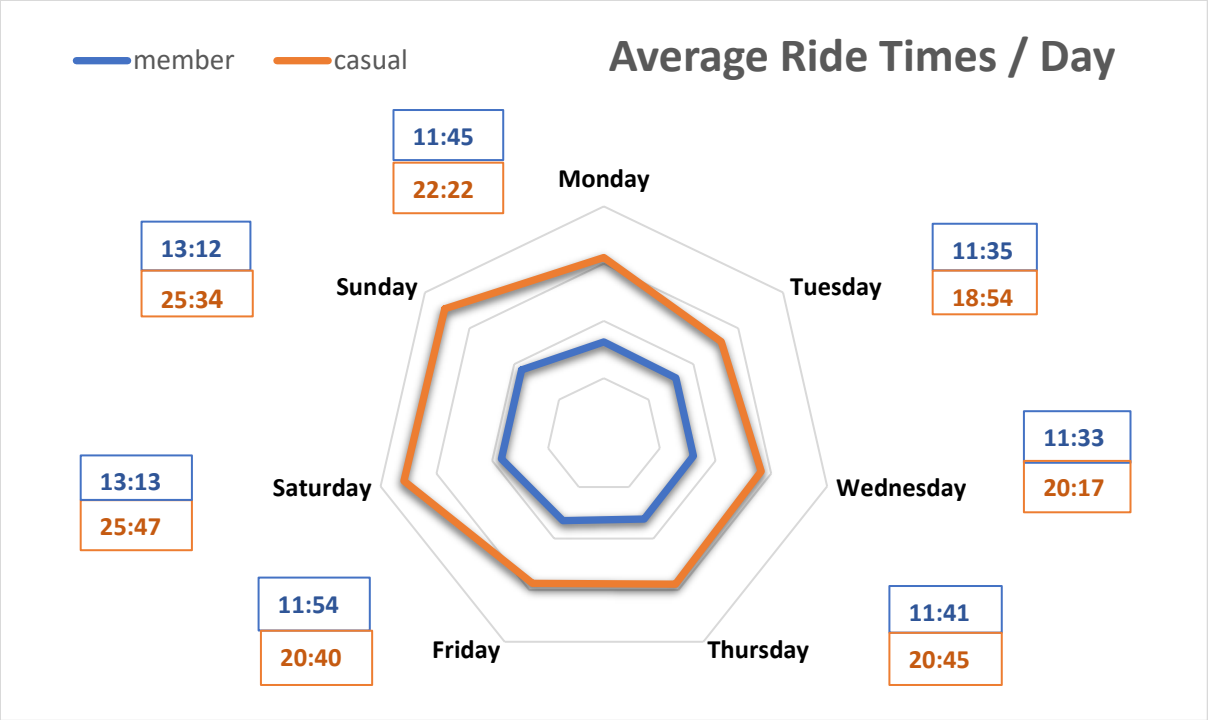



**Numbers:**

- TOTAL time used on rides in 2022: 1,369 hours & 5min
- Time used on rides by members: 624 hours & 40min
- Average ride length for members: 12min & 5sec
- Time used on rides by casuals: 744 hours & 25min
- Average ride length for casuals: 22min & 39sec

45.6% of rides were used by members, 54.4% of rides were used by casual riders.

**Figure 1**. Shows the average ride lengths for the days through the weeks in the year for both members and casual riders, Saturdays being the highest for both members and casual riders. Members have consistent ride lengths throughout the days while casual riders tend to have higher averages on the weekends.

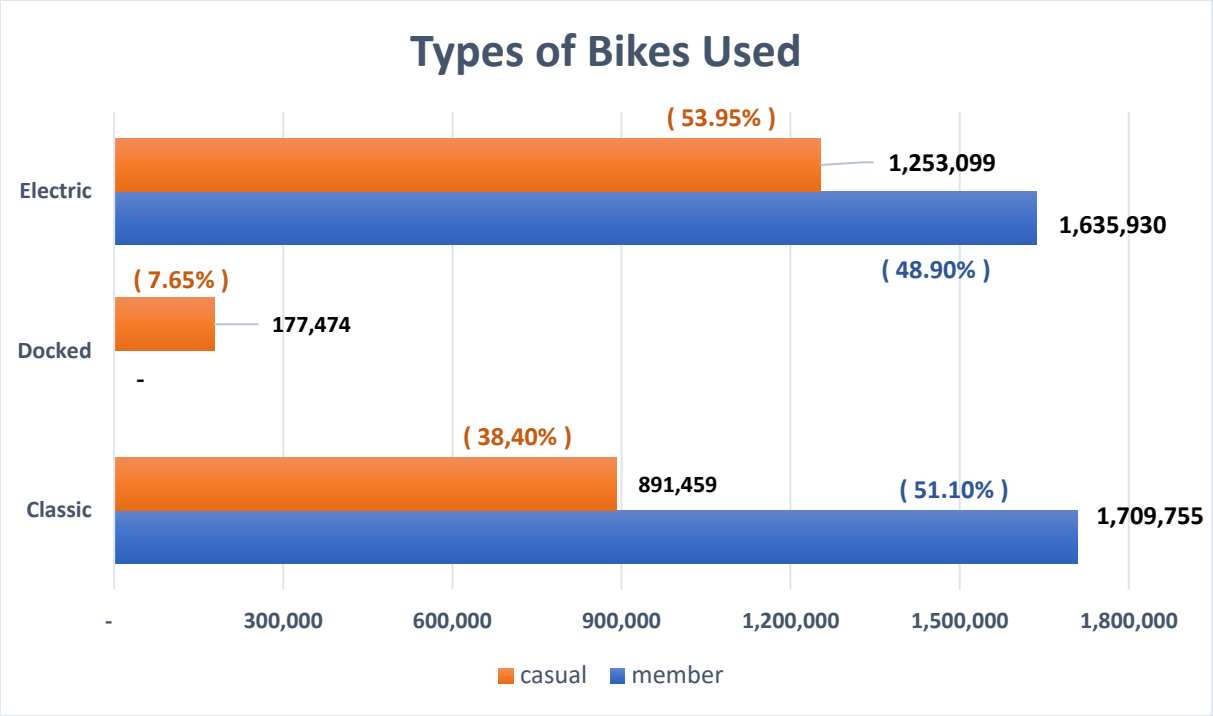



**Numbers:**

- TOTAL number of classic bikes used in 2022: 2,601,214
- Number of classic bikes used by members: 1,709,755
- Number of classic bikes used by casuals: 891,459
- TOTAL number of docked bikes used in 2022: 177,474
- Number of docked bikes used by members: 0
- Number of docked bikes used by casuals: 177,474
- TOTAL number of electric bikes used in 2022: 2,889,029
- Number of electric bikes used by members: 1,635,930
- Number of electric bikes used by casuals: 1,253,099

51.1% of all member rides were on classic bikes and 48.9% were on electric bikes. Meanwhile, 38.4% of all casual riders rides were on classic bikes, 53.95% were on electric bikes, and 7.65% were on docked bikes.

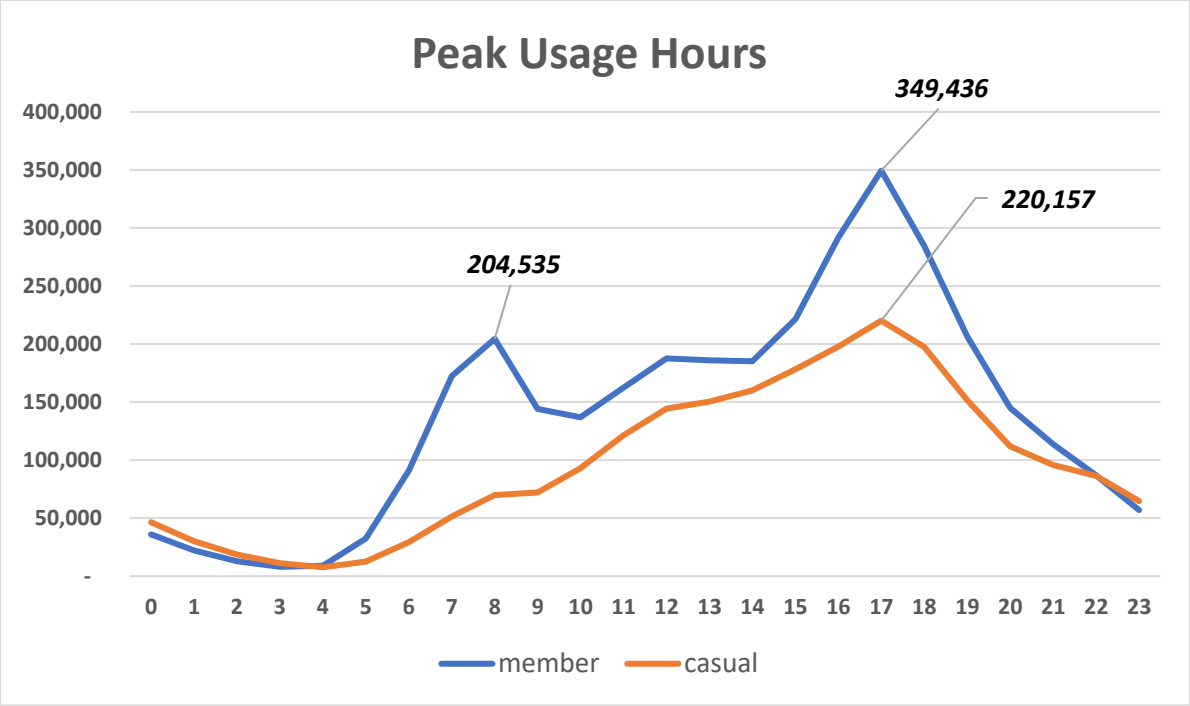
**Figure 3**. Shows the types of bikes used by members and casual riders in correlation with each other.




**Numbers:**

- TOTAL number of rides at peak hour in 2022: 569,593
- Number of rides at peak hour by members: 349,436
- Number of rides at peak hour by causals: 220,157

**Figure 4.** Shows the number of rides for each hour in the day throughout the year for members and casual riders.




**Observations on how casual riders and members use Cyclistic bikes differently:**

1. Although members had a higher number of rides, casual riders made up 54.4% of time spent on rides with higher daily averages through the whole week.
2. Members tend to ride more on the weekdays and casual riders tend to ride more on the weekends.
3. Casual members are the only ones who used docked bikes.
4. Majority of member's rides were on classic bikes, majority of casual rider's bikes were on electric and docked (presumably cost-efficient bikes).

**Potential to convert casual riders in members:**

1. Casual riders might have tighter budgets which would be explained by the higher use of more cost-efficient bikes.
2. Casual riders use the service for much longer, indicating that this might be their main form of transportation.

STEP 6 – ACT

**Recommendation:**

1. Create a quick survey for casual riders to come up with different customer profiles based on consumer behavior.
2. Use customer profiles to create promotions for casual riders to convert to members.
    a. Promotions can be over email, app, or messaging.
3. Offer weekday discounts or membership free trials.

    a. Discounts can be offered more on the weekends (casual riders use the service on the weekends the most)
4. Offer rewards program for using cost-efficient bikes more.
5. Offer seasonal promotions that incentivize longer rides or rides on specific days and times.