# Cyclistic

**A fictional Bike-share company in Chicago**

**Case study: How does a Bike-share navigate speedy success?**

**August 31, 2024**

## Purpose of Case Study

This case study was the final capstone assignment in the Google Data Analytics Professional Certificate course offered by Coursera.

The objective of the case study was to demonstrate knowledge and skills in performing typical tasks of a Junior Data Analyst, and to follow the steps of the data analysis process: **Ask**, **Prepare**, **Process**, **Analyze**, **Share**, and **Act**.

The case study was also included in a portfolio to showcase data analysis competence of the preparer to potential employers.

## Step 1 – Ask

### Objective of understanding the task

To clearly understand the problem, it's context, and to ask relevant questions early on, so that the assigned task can solve the problem in an efficient manner.

### Background

Cyclistic is a fictional operating Bike-share company in Chicago with two categories of customers; casual riders and members. Casual riders purchase single-ride and full-day passes, while customers who purchase annual memberships are considered members. The Bike-share offering was launched in **2016** with **5,824** geo-locked bicycles in a network of **692** stations across Chicago.

### Problem Statement

Cyclistic's Finance Analysts has identified that annual members are more profitable than casual riders.

Consequently, Cyclistic's Director of Marketing, Lily Moreno, wants to maximize the number of annual memberships by converting casual riders into annual members. The Marketing Analytics team, therefore, is required to;

1. Analyze Cyclistic's historical Bike trip data to identify key insights
2. Understand how casual riders and annual members use Cyclistic's Bikes differently, and to
3. Design a new marketing strategy to convert casual riders into annual members.

This case study will address question 2 above; **to understand how casual riders and annual members use Cyclistic's Bikes differently**.

# 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 the analysis performed
5. Supporting visualizations and key findings
6. Top three recommendations based on the analysis

## Key Stakeholders

The key stakeholders include:

1. **Cyclistic's Director of Marketing, Lily Moreno**: Responsible for developing and implementing marketing strategies.
2. **Cyclistic Executive Team**: Responsible for the overall growth and profitability of the company. This is the team that will approve the recommended marketing program.
3. **Cyclistic Marketing Analytics Team:** Responsible for analyzing data to provide insights that will inform marketing strategies. This case study is assumed to be conducted by a Junior Data Analyst who joined the team six months ago.

# Step 2 – Prepare

## Objective of Data Preparation

To determine the data and its source, how it is organized, and to ensure it is adequate to address the problem statement. This includes verification of the data's integrity, bias, and credibility.

## Key Tasks

1. Download Cyclistic's historical 2023 trip data to analyze and identify trends – provided by Motivate International Inc. under license, with restricted use that prohibits use or generation of personally identifiable information. **Data**: **divvy_tripdata**, public data source.

2. Use the ROCCC framework to ensure data credibility.
   a. R – Reliable: Data is sourced from Divvy Bikes website and it contains attributes that can describe Cyclistic's customer ride share patterns. In general, each trip includes the following basic trip information;
      i. Ride_id as the unique trip identifier
      ii. Trip start day and time
      iii. Trip end day and time
      iv. Trip start station or geographical location
      v. Trip end station or geographical location
      vi. Rider type; Member and Casual (Single Ride and Day Pass)

   b. O – Original: Although Cyclistic is a fake company, the data is from a real Bike-share company called Divvy Bikes. The data can also be accessed from Divvy Bikes website; therefore, it can be assumed to be original for the purpose of this case study.
   c. C – Comprehensive: Data for 2023 has more than 5 million records - it is therefore comprehensive.
   d. C – Current: Current year is 2024 while the data is for last year, 2023. The year 2023 is the period of interest because it is the most recent whole year that covers all the months from January to December. Hence, the data is current.
   e. C – Cited: Data is referenced from a licensed source; Motivate International Inc. Also, the data is available from Divvy Bikes official website; https://divvyBikes.com/system-data.

3. Data was downloaded and stored securely in a single folder with restricted access. Each file has consistent naming; cyclistic_divvy_tripdata_yyyy_mm.csv for ease of retrieval and sorting.

4. Data was originally organized in a structured format within CSV files for each month of the year 2023. One file for January 2023 was uploaded to MS Excel for an initial inspection of the type of data contained in the files. The file had more than 190,000 rows and 13 columns. Since 12 such files were to be joined to form data for the whole year, the resultant data size was deemed too large for analysis using MS Excel. Consequently, data consolidation, cleaning and analysis was done using SQL and R.

a. Following is a breakdown of the 13 columns (attributes) and examples of the possible information each attribute may provide in addressing the problem statement.
   1. **ride_id**: Unique identifier for each ride (useful for tracking and referencing specific rides, and combining tables)
   2. **rideable_type**: Type of Bike used (e.g., electric Bike, classic Bike) - to see if there's a preference between members and casual users
   3. **started_at**: Start date and time of the ride to analyze peak usage times
   4. **ended_at**: End date and time of the ride to calculate trip duration
   5. **start_station_name**: Name of the station where the ride started to identify popular starting points
   6. **start_station_id**: Identifier for the start station (useful for precise location tracking) – redundancy relationship with **start_station_name**
   7. **end_station_name**: Name of the station where the ride ended to identify popular ending points
   8. **end_station_id**: Identifier for the end station (useful for precise location tracking) – redundancy relationship with **end_station_name**
   9. **start_lat**: Latitude of the start location for geographical analysis
   10. **start_lng**: Longitude of the start location for geographical analysis
   11. **end_lat**: Latitude of the end location for geographical analysis
   12. **end_lng**: Longitude of the end location for geographical analysis
   13. **member_casual**: Type of user (e.g., member, casual) to differentiate between user types

Below is the list of columns and type of data stored in each column.

| Field name | Type | Mode |
|---|---|---|
| ride_id | STRING | NULLABLE |
| rideable_type | STRING | NULLABLE |
| started_at | TIMESTAMP | NULLABLE |
| ended_at | TIMESTAMP | NULLABLE |
| start_station_name | STRING | NULLABLE |
| start_station_id | STRING | NULLABLE |
| end_station_name | STRING | NULLABLE |
| end_station_id | STRING | NULLABLE |
| start_lat | FLOAT | NULLABLE |
| start_lng | FLOAT | NULLABLE |
| end_lat | FLOAT | NULLABLE |
| end_lng | FLOAT | NULLABLE |
| member_casual | STRING | NULLABLE |

b. The following column (attribute) was deemed unnecessary after the data preparation step was completed;

   i. **ride_id**: Not necessary for analysis because there was no need to track specific rides and to reduce data size

c. Derived columns were added to simplify analysis and provide deeper insights into Bike usage patterns. Additional columns that were considered include the following;

   i. **trip_duration**: Calculate the duration of each trip using the started_at and ended_at columns.

**Formula**: trip_duration = ended_at - started_at

ii. **start_trip_day_of_week**: Extract the day of the week from the started_at column to analyze usage patterns by day.

**Formula**: day_of_week = WEEKDAY(started_at)

iii. **start_trip_hour_of_day**: Extract the hour from the started_at column to analyze peak usage times during the day.

**Formula**: hour_of_day = HOUR(started_at)

iv. **start_trip_month**: Extract the month from the started_at column to analyze seasonal trends.

**Formula**: month = MONTH(started_at)

v. **start_trip_weather_condition**: Weather conditions during each trip would provide insight on how weather affects Bike usage. It was not easy to include 2023 hourly weather data in the data set. This attribute was not included in the analysis, but could be included in follow-up studies if deemed necessary by stakeholders.

**Source**: External weather data API or dataset.

vi. **Is_start_day_holiday**: To indicate if the trip occurred during a bank or public holiday, which might influence usage patterns. 10 holidays were identified as shown below.

**Source**: Calendar of major bank & public holidays in 2023

| | Holiday_2023 | When | Date |
|---|---|---|---|
| 1 | New Year's Day | January 1 | 2023-01-01 |
| 2 | Martin Luther King Jr. Day | The 3rd Monday in January | 2023-01-16 |
| 3 | President's Day | The 3rd Monday in February | 2023-02-20 |
| 4 | Memorial Day | The last Monday in May | 2023-05-29 |
| 5 | Independence Day | July 4 | 2023-07-04 |
| 6 | Labor Day | The 1st Monday in September | 2023-09-04 |
| 7 | Columbus Day | The 2nd Monday in October | 2023-10-09 |
| 8 | Veteran's Day | November 11 | 2023-11-11 |
| 9 | Thanksgiving Day | The 4th Thursday in November | 2023-11-23 |
| 10 | Christmas Day | December 25 | 2023-12-25 |

vii. **miles_to_return_location**: To calculate the distance from the start trip location to end trip location. This value, if non-zero, will show how many Bikes were returned to a different location and the furthest return distance for both customer types.

d. The original column names were renamed to make them more readable and easier to understand. Following is the list of all the original and derived column names;

| current column name | re-named column name |
|---|---|
| rideable_type | bike_type |
| start_station_name | start_trip_station_name |
| start_station_id | start_trip_station_id |
| end_station_name | end_trip_station_name |
| end_station_id | end_trip_station_id |
| member_casual | customer_type |

Original Column Names

| Field name | Type | Mode |
|---|---|---|
| ride_id | STRING | NULLABLE |
| rideable_type | STRING | NULLABLE |
| start_station_name | STRING | NULLABLE |
| start_station_id | STRING | NULLABLE |
| end_station_name | STRING | NULLABLE |
| end_station_id | STRING | NULLABLE |
| member_casual | STRING | NULLABLE |
| start_trip_location | STRING | NULLABLE |
| end_trip_location | STRING | NULLABLE |
| start_lat | FLOAT | NULLABLE |
| start_lng | FLOAT | NULLABLE |
| end_lat | FLOAT | NULLABLE |
| end_lng | FLOAT | NULLABLE |
| miles_to_return_location | FLOAT | NULLABLE |
| started_at | TIMESTAMP | NULLABLE |
| ended_at | TIMESTAMP | NULLABLE |
| trip_duration_hours | FLOAT | NULLABLE |
| start_trip_day_of_week | STRING | NULLABLE |
| start_trip_hour_of_day | INTEGER | NULLABLE |
| start_trip_month | STRING | NULLABLE |
| is_start_day_holiday | STRING | NULLABLE |

Updated List of Column Names

| Field name | Type | Mode |
|---|---|---|
| bike_type | STRING | NULLABLE |
| started_at | TIMESTAMP | NULLABLE |
| ended_at | TIMESTAMP | NULLABLE |
| start_trip_station_name | STRING | NULLABLE |
| start_trip_station_id | STRING | NULLABLE |
| end_trip_station_name | STRING | NULLABLE |
| end_trip_station_id | STRING | NULLABLE |
| start_lat | FLOAT | NULLABLE |
| start_lng | FLOAT | NULLABLE |
| end_lat | FLOAT | NULLABLE |
| end_lng | FLOAT | NULLABLE |
| customer_type | STRING | NULLABLE |
| start_trip_location | STRING | NULLABLE |
| end_trip_location | STRING | NULLABLE |
| miles_to_return_location | FLOAT | NULLABLE |
| trip_duration_hours | FLOAT | NULLABLE |
| start_trip_day_of_week | STRING | NULLABLE |
| start_trip_hour_of_day | INTEGER | NULLABLE |
| start_trip_month | STRING | NULLABLE |
| is_start_day_holiday | STRING | NULLABLE |

e. The relevant column names used for analysis are the following;

| Relevant Columns for Analysis | |
|---|---|
| 1 | bike_type |
| 2 | start_trip_station_name |
| 3 | start-trip_station_id |
| 4 | end_trip_station_name |
| 5 | end_trip_station_id |
| 6 | customer_type |
| 7 | start_trip_location |
| 8 | end_trip_location |
| 9 | miles_to_return_location |
| 10 | trip_duration_hours |
| 11 | start_trip_day_of_the_week |
| 12 | start_trip_hour_of_day |
| 13 | start_trip_month |
| 14 | is_start_day_holiday |

f. **Tools**: MS Excel, SQL, R, Adobe PDF, MS Word, Tableau

| TOOL | PURPOSE |
|---|---|
| | |
| SQL | Data Preparation and Processing |
| R | Data Analysis and Visualization |
| MS EXCEL | Initial Review of Data to determine type, size, attributes, etc |
| MS WORD | Case Study Documentation |
| ADOBE PDF | Case Study Documentation |
| TABLEAU | Visualization |
| | |

g. **Tables**: using SQL, data was loaded into a single dataset and combined into a table created. The dataset name is **cyclistic_case_study_2023** and the table name used for analysis was **cyclistic_2023_r_analysis** (to show **company_year_tool_task**).

h. Data was collected objectively and remotely for both customer and non-customer rides. However, data used for this case study includes only customer rides in the selected period. Data collection was considered to be unbiased.

i. Additionally, during the data processing step, data was checked for missing values, outliers, anomalies, over-and-under representation of the different customer groups so as to further assess bias and to ensure accuracy and fairness of the analysis step.

# Step 3 – Process

## Objective of Data Preparation

To meticulously clean and organize data, ensuring its consistency, completeness, and efficiency in analysis, thereby enhancing the reliability and accuracy of the derived insights.

## Key Tasks

1. The key tool used for processing was SQL within BigQuery's data warehouse and analytics platform. SQL was chosen because of its efficiency in handling large data sets.
2. The data was checked for missing values and rows with missing values in the key columns e.g. ride identification column, ride_id, were removed. As for the other cells with null values or blanks, the work BLANK or 0.00 or 1900:01:00:00 was entered depending on whether the column type is a string, float, or timestamp respectively.
3. The data was checked for errors i.e. syntax, logical, and data type errors. Records with errors were either corrected or deleted as necessary.
4. New tables were created for each quarter of 2023 to ease uploading/downloading and quarterly analysis if required.
5. Counted rows with blanks in the customer type column for the tables of the four quarters of 2023. Since the gist of the case study is based on customer-type behavior, it was vital to eliminate rows without customer type data. There were no blanks in all the four tables.

6. See table below for the list of preparation and processing SQL scripts.

Queries

Shared queries

- 01 deleted blank ride_id because unique trip identifier q1234
- 02 delete duplicates of ride_id q1234 - no dupes in all 4
- 03 create start_location and end_location columns for q1234
- 04 add miles_to_return_location in q1234
- 05 calculates miles to return calculation
- 06 deleted nulls started_at, ended_at, rideable_type, member_casual
- 07 count rows q1234
- 08 count rows with non-essential blanks q1234
- 09 cleaned q1
- 10 cleaned q234 and renamed to qx_cleaned
- 11 confirm no nulls and blanks and trims done
- 12 created 2023 combined table
- 13 confirm all rows added to clean 2023 table
- 14 create duplicate of cleaned table for processing
- 15 derived cols duration day hour month
- 16 change months to JAN to DEC
- 17 added holiday col
- 17 create new col TRUE FALSE holiday
- 18 create new table to rename cols and delete ride id column
- 19 rename cols in analysis table
- 20 create final table for analysis in R with relevant columns
- 22 Analysis R size 1gb rows 5 719 877
- 23 split large R table into 3 for upload to R
- 24 check all rows in new tables equal in R

# Step 4 – Analysis

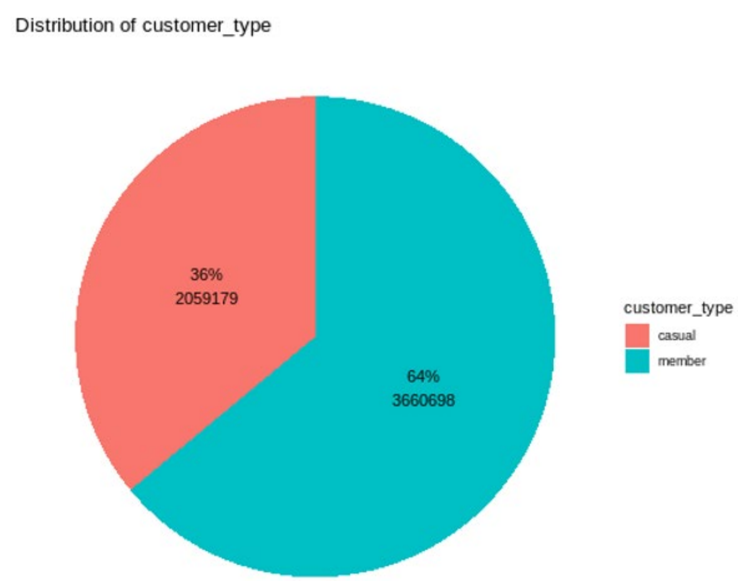
## Objective of Data Analysis

To examine the organized and formatted data to discover trends, relationships, and insights that can answer the initial problem statement. This step collectively aims to turn raw data into actionable information that can guide decision-making by providing possible solutions to the identified problem.

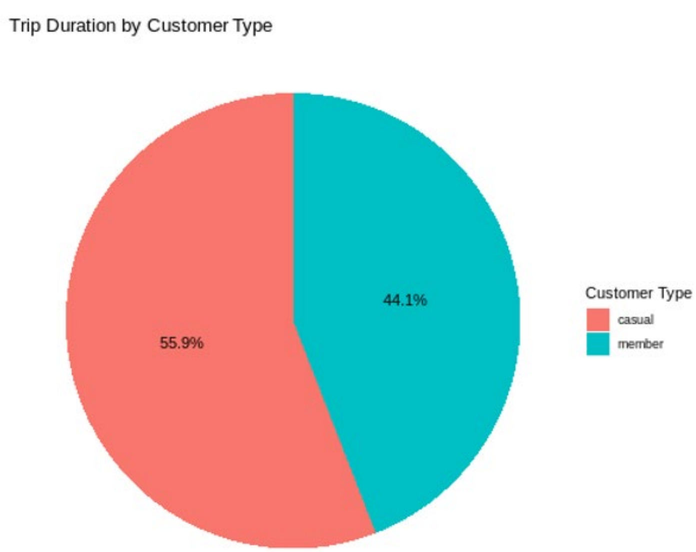
## Key Results

1. Number of rides used in analysis are 5,719,877.

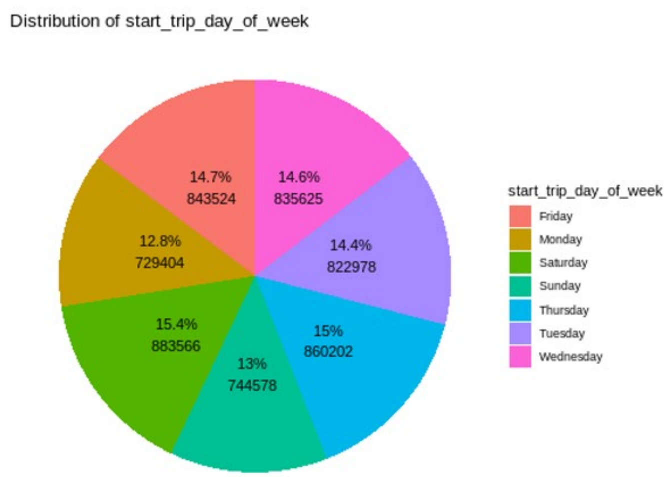
**2.** Distribution of Customer Types – 2 – **Member 64%** and **Casual 36%**

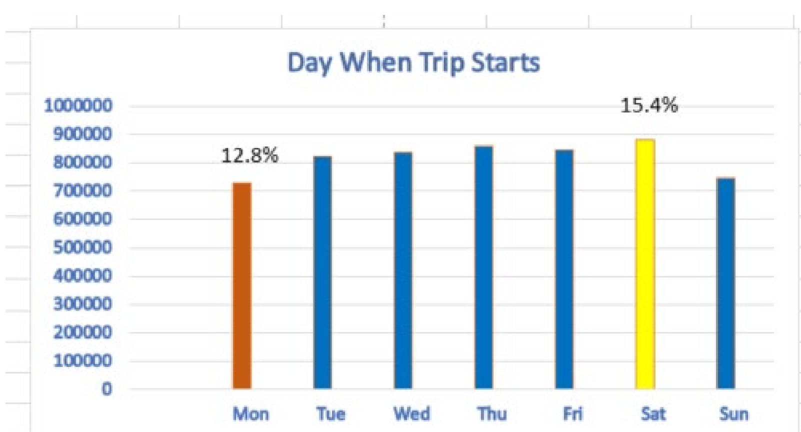



**3.** Distribution of Trip Duration by Customer Types – 2 – **Member 44%** and **Casual 56%**

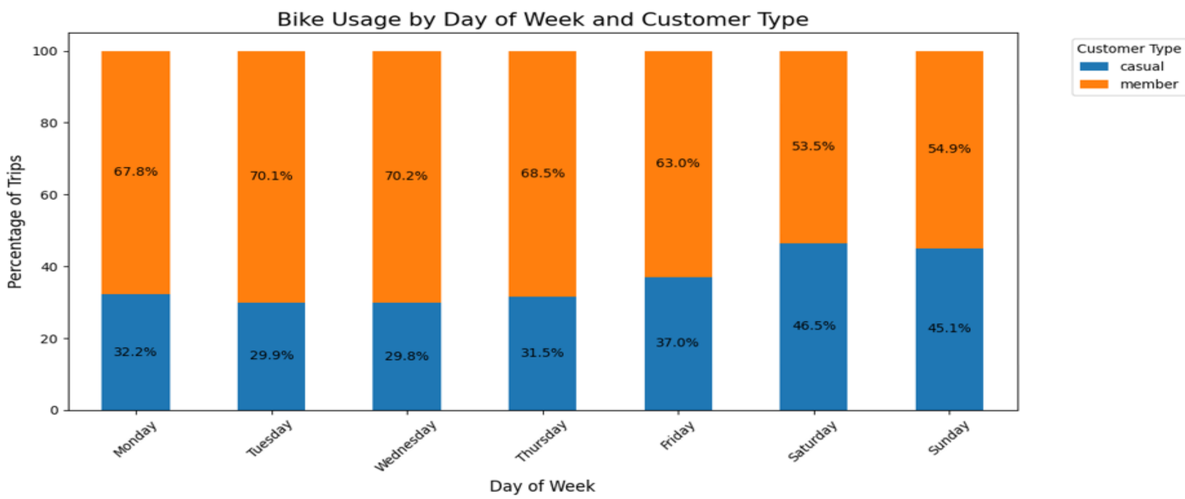


**4.** Distribution of Day of the Week when a trip starts – **Monday is lowest 12.8%** and **Saturday is highest 15.4 %.**

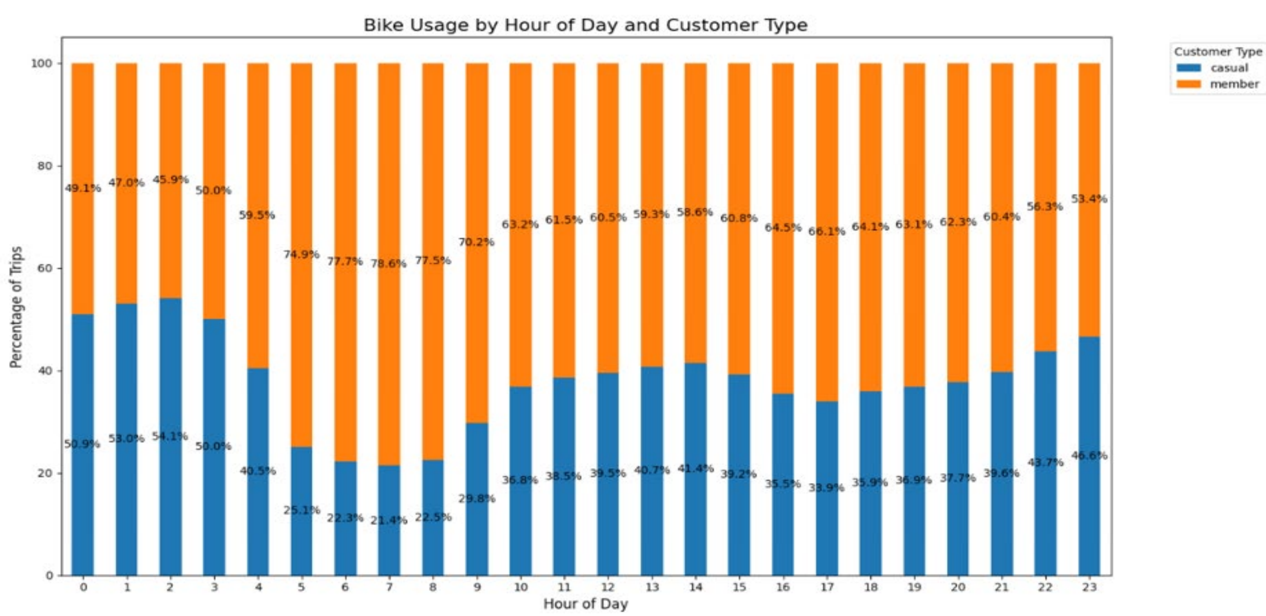


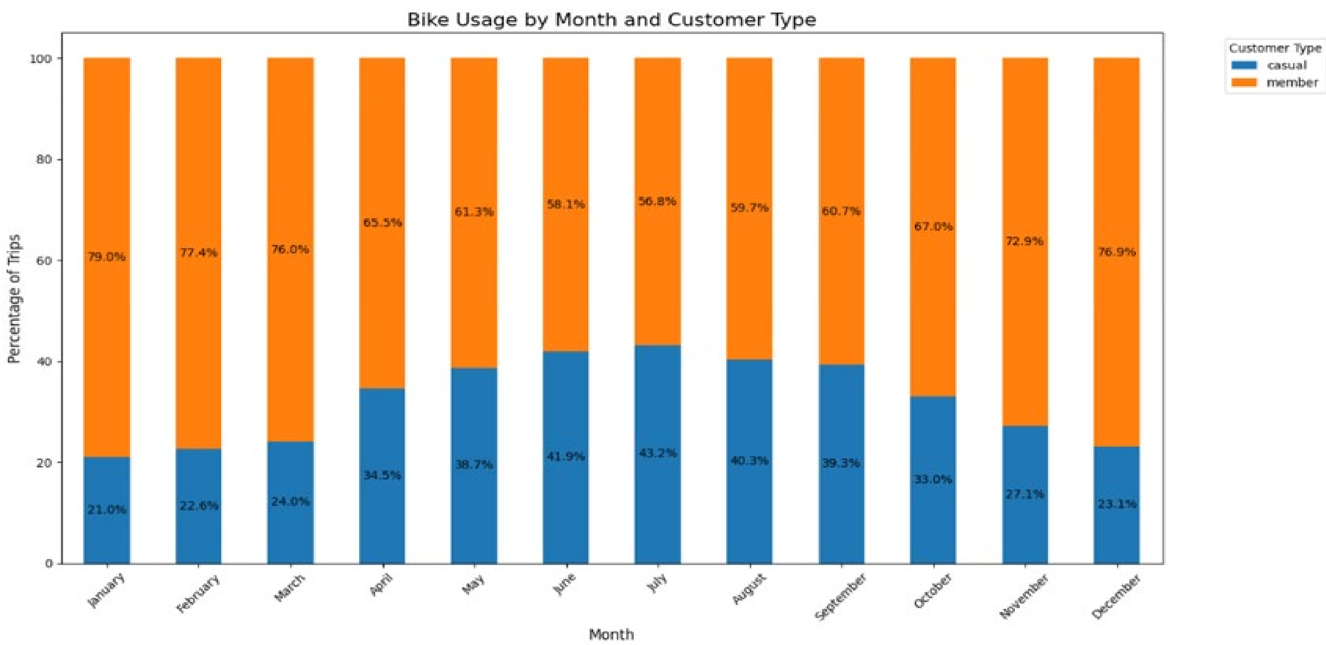
**5.** Bike Usage by Day of the Week – **Average - Member 64%** and **Casual 36%.**



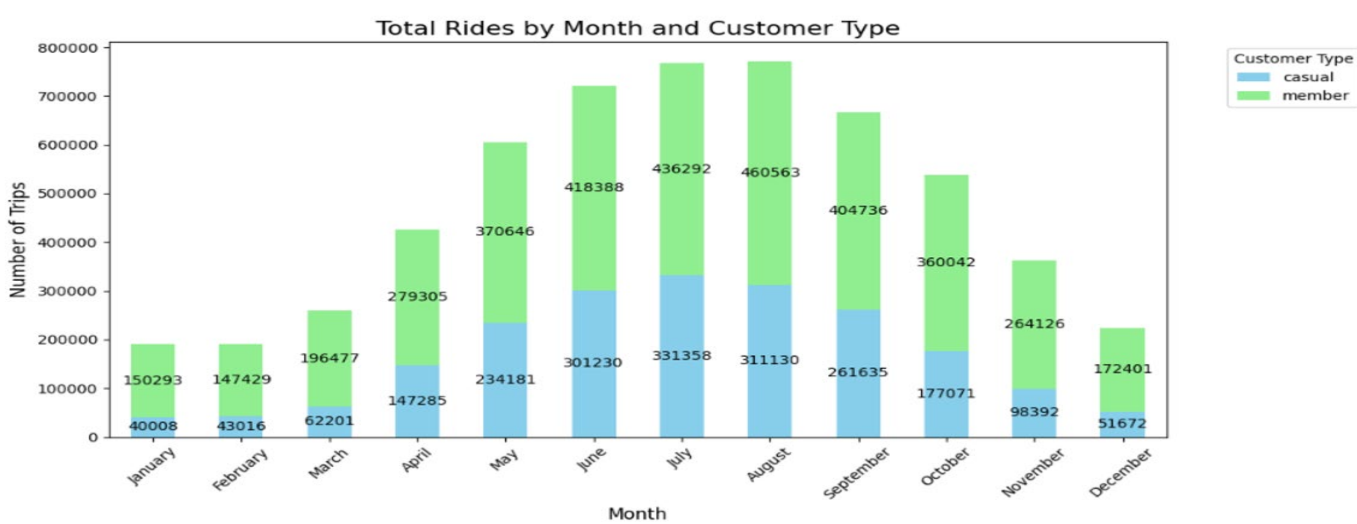

**6.** Bike Usage – **Hourly Average - Member 62%** and **Casual 38%.**
Lowest Bike usage - **Member Mid-night to 4 a.m.** and **Casual 5 a.m. to 8 a.m.**



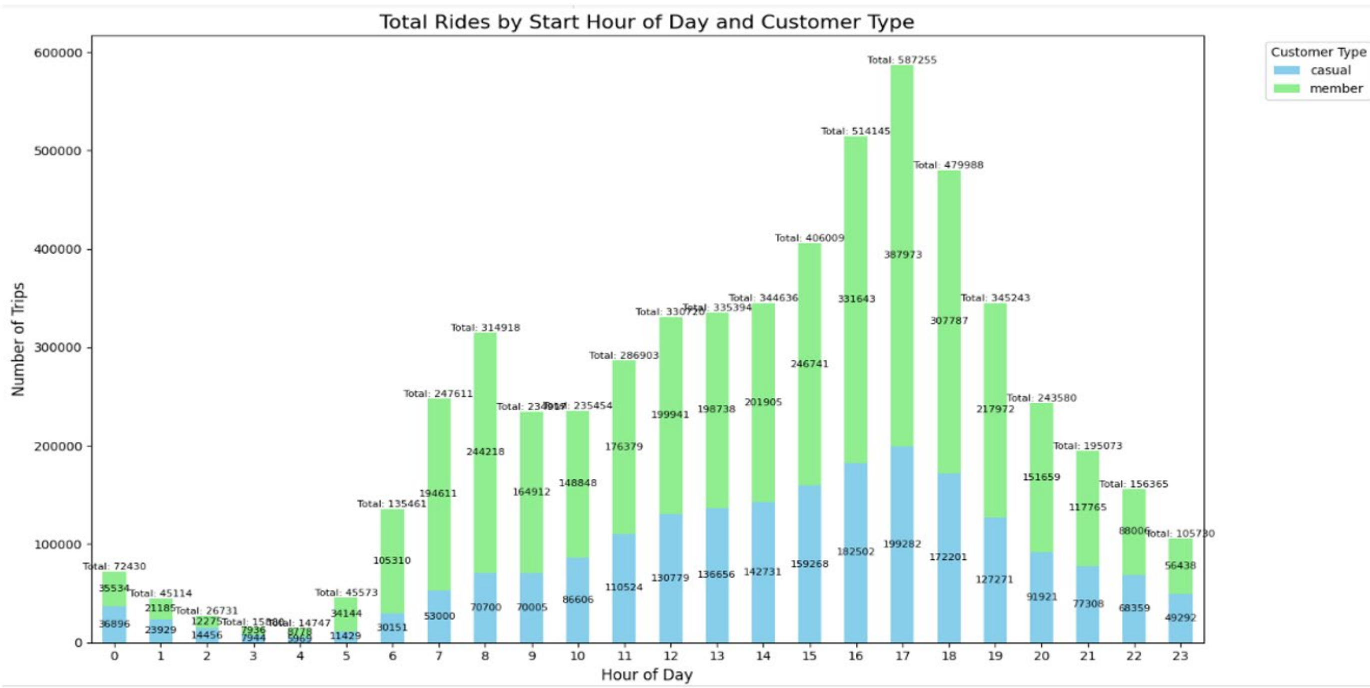
**7.** Bike Usage Distribution – **Monthly Average - Member 68%** and **Casual 32%**
Lowest Bike usage - **Member Jun, Jul, & Aug = Summer Season**
Lowest Bike usage - **Casual Jan, Feb, & Mar = Winter Season**



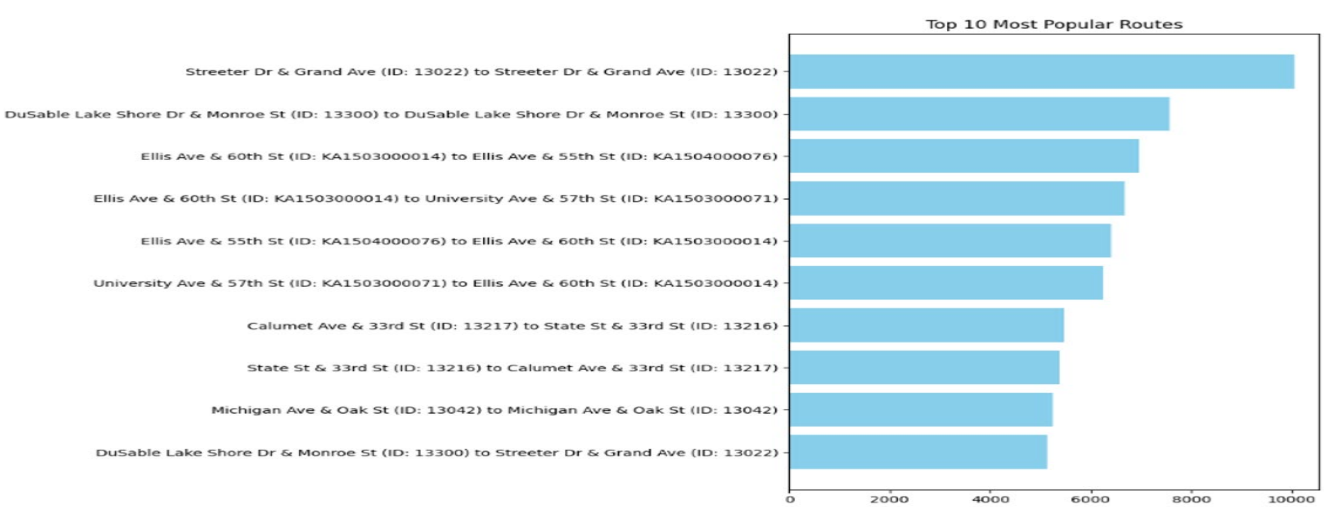

**8.** Trips Distribution – **Quarter 1 11%, Quarter 2 31%, Quarter 3 38%, Quarter 4 20%**
Lowest Trips – **December to March** Highest Trips **June to September**



9. Trips by Start Hour – **Peak Hours for both Member & Casual = 3 p.m. to 6 p.m.**
Lowest Trip Starts – **Mid-night to 5 a.m.**




10. Top 10 most popular Routes ( where start location is different than end location)



11. Top 5 most popular Routes **Casual** customers

| Trips | Trip Start Station Name | Trip End Station Name | Trips |
|---|---|---|---|
| 107,091 | Streeter Dr & Grand Ave | Streeter Dr & Grand Ave | 8,653 |
| 38,752 | DuSable Lake Shore Dr & Monroe St | DuSable Lake Shore Dr & Monroe St | 6,732 |
| 37,972 | DuSable Lake Shore Dr & Monroe St | Streeter Dr & Grand Ave | 4,626 |
| 73,549 | Michigan Ave & Oak Street | Michigan Ave & Oak Street | 4,260 |
| 75,254 | Millenium Park | Millenium Park | 3,427 |
| 332,618 | | | 27,698 |

12. Top 5 most Routes for **Member** customers

| Trips | Trip Start Station Name | Trip End Station Name | Trips |
|---|---|---|---|
| 170,948 | Ellis Ave & 60th St | University Ave & 57th St | 5,208 |
| 138,694 | Calumet Ave & 33rd St | State St & 33rd St | 5,202 |
| 244,023 | State St & 33rd St | Calumet Ave & 33rd St | 5,145 |
| 249,915 | University Ave & 57th St | Ellis Ave & 60th St | 4,935 |
| 170,872 | Ellis Ave & 60th St | Ellis Ave & 55th St | 4,929 |
| 974,452 | | | 25,419 |

13. Top 5 most popular End Trip Locations for **Casual** customers

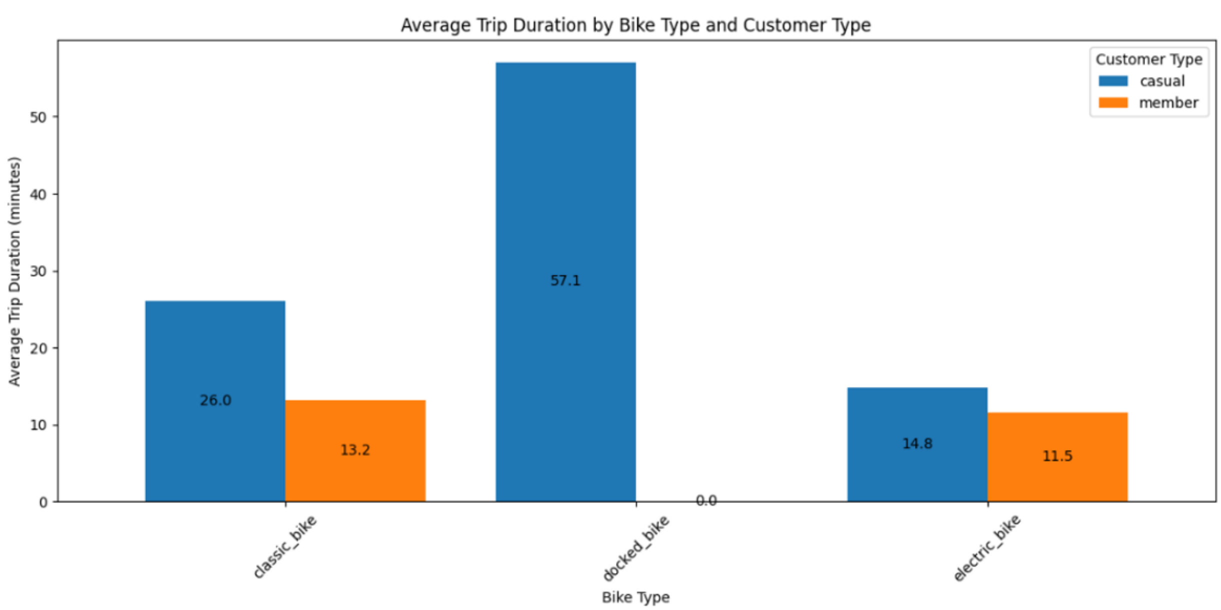
| Trip Start Station Name | Trips |
|---|---|
| Streeter Dr & Grand Ave | 47,718 |
| DuSable Lake Shore Dr & Monroe St | 26,356 |
| Michigan Ave & Oak Street | 22,486 |
| DuSable Lake Shore Dr & North Blvd | 22,338 |
| Millenium Park | 21,320 |
| | 140,218 |

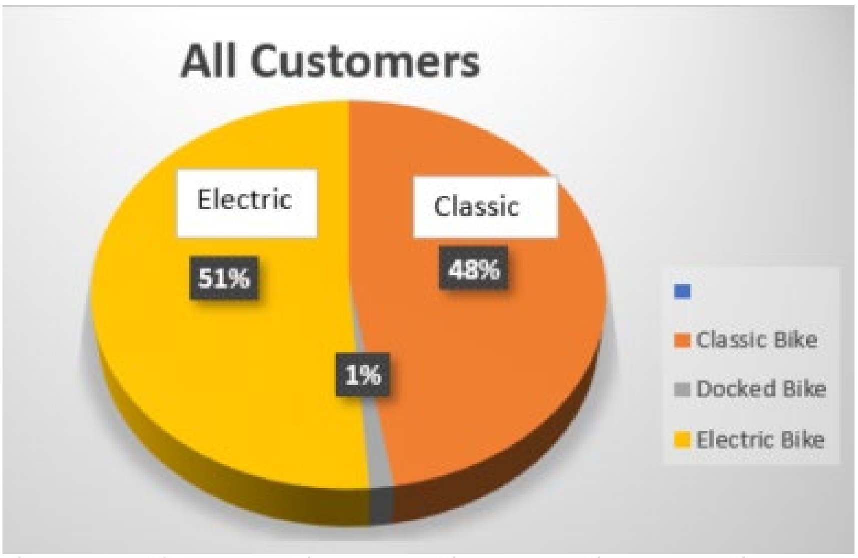
14. Top 5 most popular End Trip Locations for **Member** customers

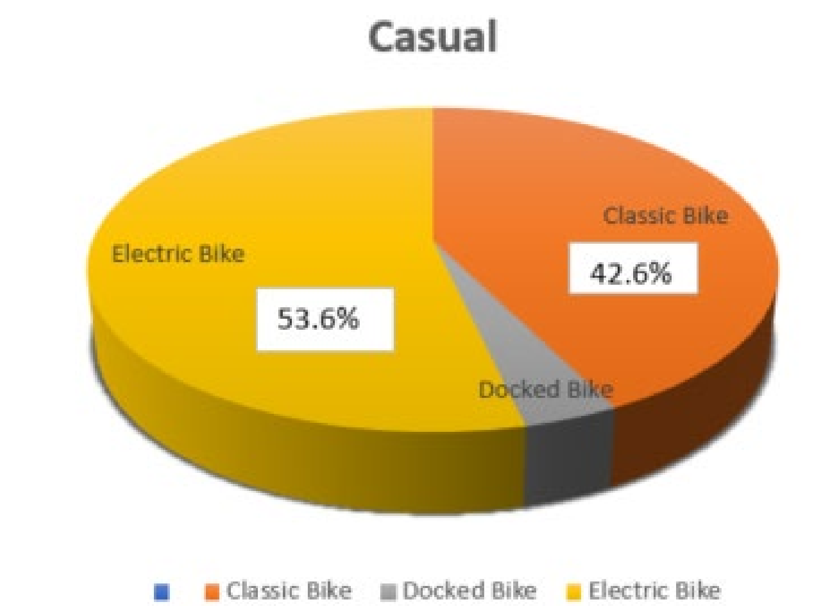
| Trip Start Station Name | Trips |
|---|---|
| | |
| Clinton St & Washington Blvd | **25,497** |
| Kingsbury St & Kinzie St | **24,315** |
| Clark St & Elm St | **22,963** |
| Wells St & Concord Ln | **20,453** |
| Clinton St & Madison St | **20,097** |
| | **113,325** |

15. Bike Usage by Type and Customer

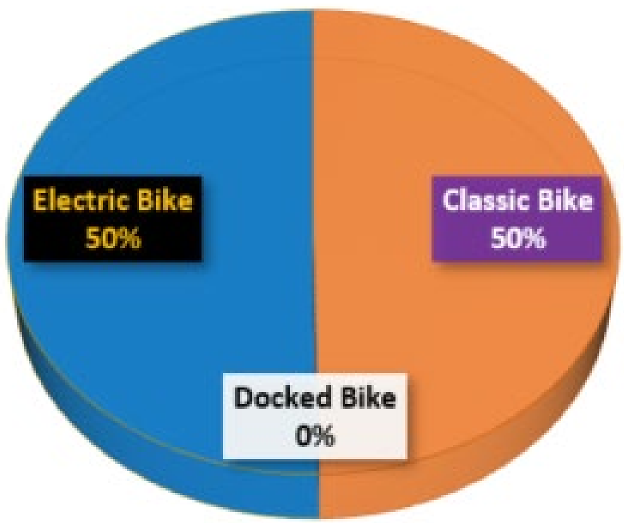
Members did not use the docked Bike and they used Classic & Electric Bikes equally.
Docked Bike had least usage = 1% and longest trips = 2.8 times more than electric Bike
Classic Bike had the second longest trips = 76% more than electric Bike
Electric Bike usage = 58% Classic Bike usage = 48%



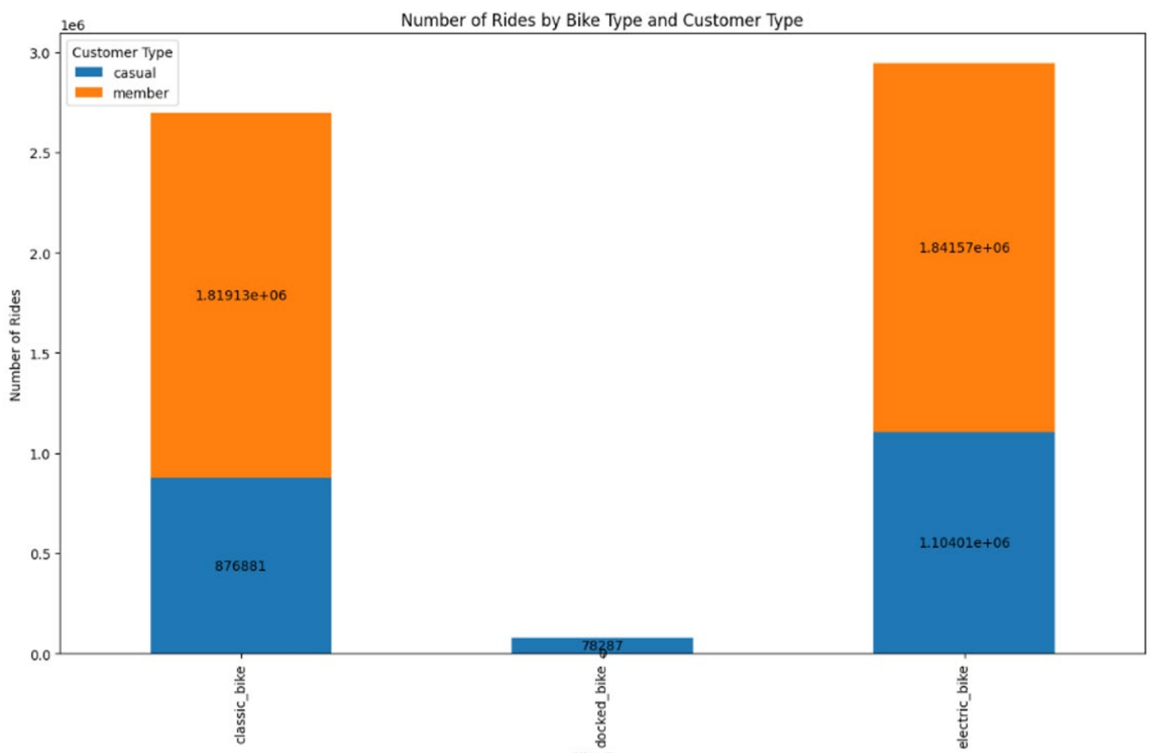

All Customers
Electric
51%
Classic
48%
1%
Classic Bike
Docked Bike
Electric Bike


Casual
Electric Bike
53.6%
Classic Bike
42.6%
Docked Bike
Classic Bike
Docked Bike
Electric Bike

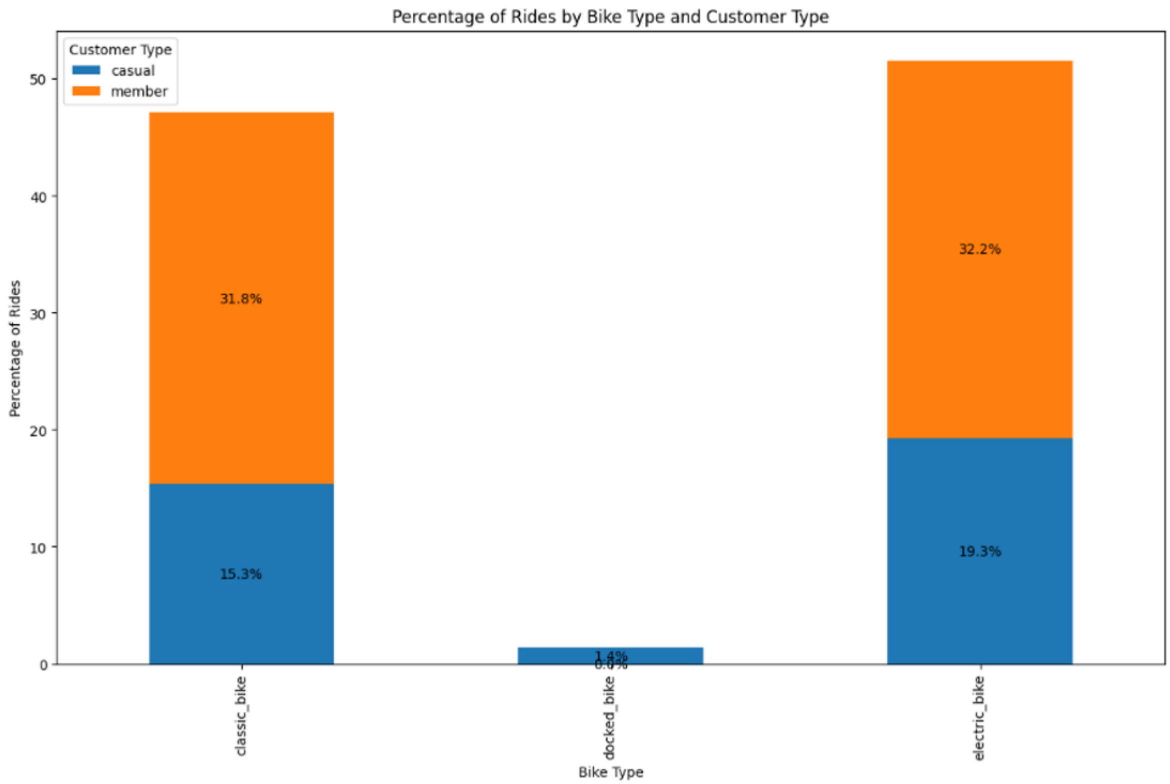
## MEMBER



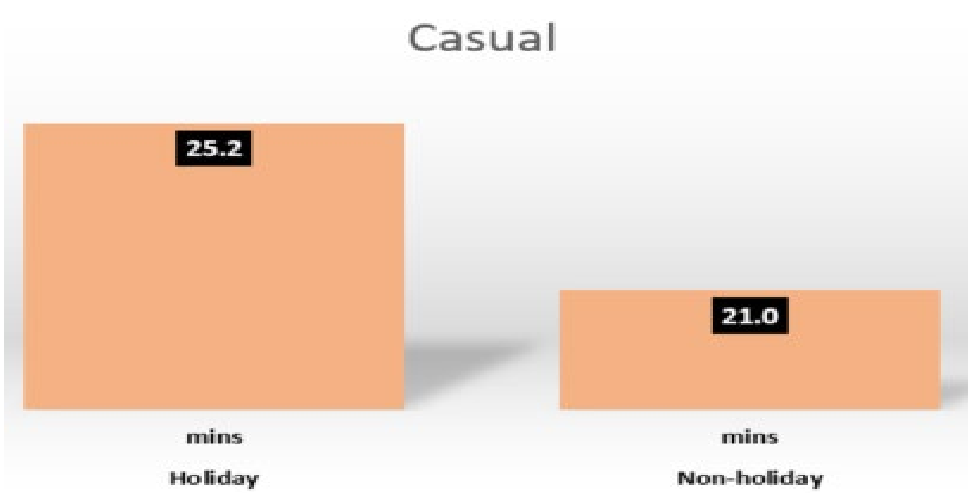
Electric Bike
50%
Classic Bike
50%
Docked Bike
0%

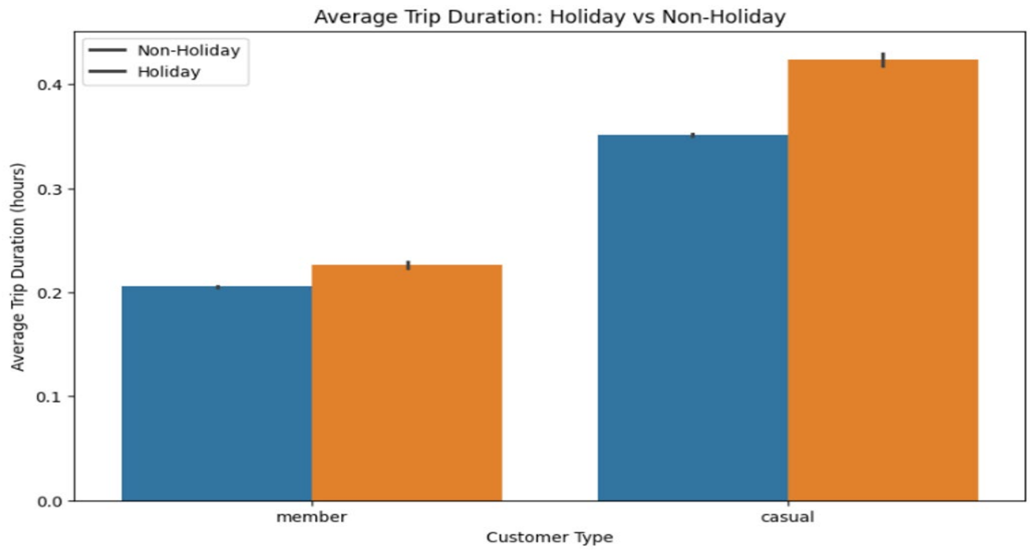



Number of Rides by Bike Type and Customer Type
1e6
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Number of Rides
Customer Type
casual
member
1.81913e+06
876881
78287
0
1.84157e+06
1.10401e+06
classic_bike
docked_bike
electric_bike
Bike Type

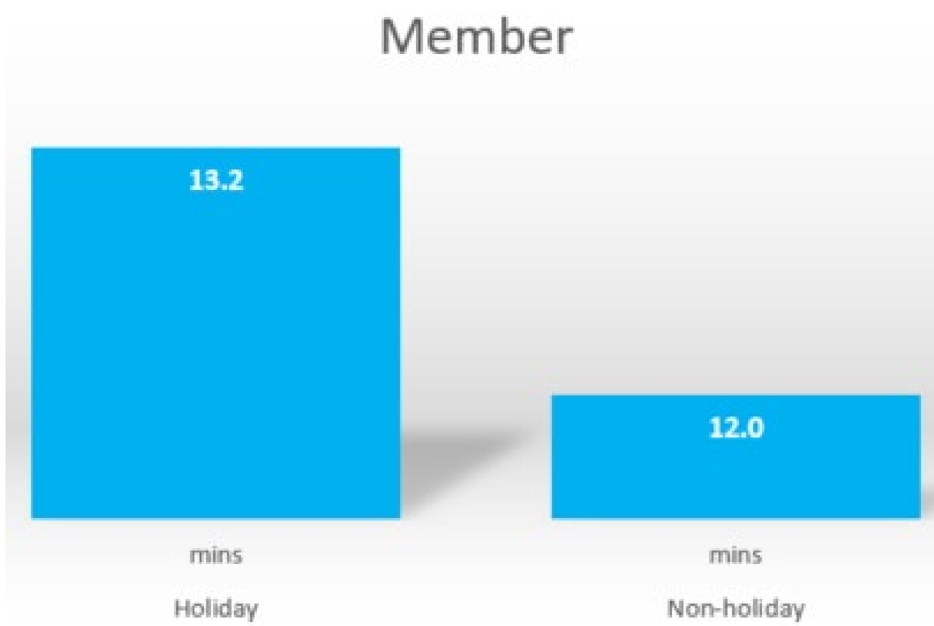



16. Trip Duration Holiday vs Non-Holiday
Member (Holiday 13.2 mins Non-Holiday 12 mins)
Casual (Holiday 25.2 mins Non-Holiday 21 mins)

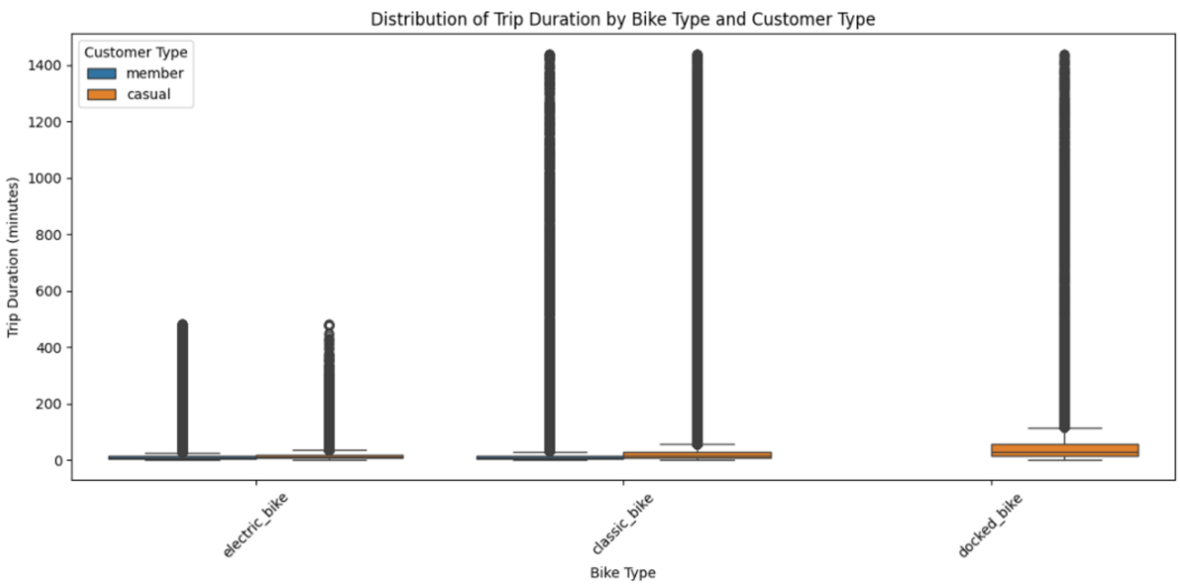




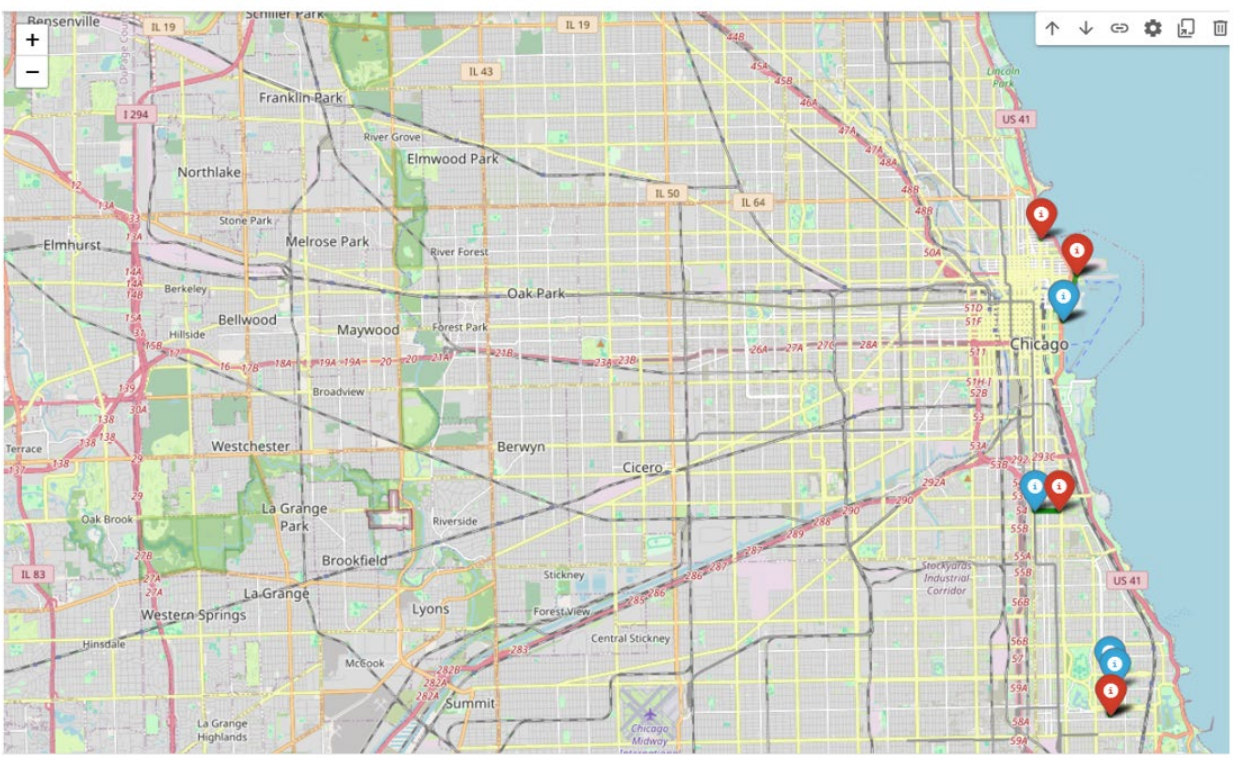
Average Trip Duration: Holiday vs Non-Holiday
Non-Holiday
Holiday
Average Trip Duration (hours)
0.0
0.1
0.2
0.3
0.4
member
casual
Customer Type

Member
13.2
12.0
mins
mins
Holiday
Non-holiday




## 17. Geographical Area of Popular Bike Routes



## Step 5 – Share

### Objective of Sharing the Findings

To communicate findings and insights to decision makers using a compelling narrative in a way that the audience will understand and be able to make informed decisions.

### Key Strategies

1. Organize findings into a coherent narrative that clearly illustrates the key insights.
2. Ensure the presentation answers the original question.
3. Identify and understand the audience including their level of expertise, interest, and expectations.
4. Choose the most appropriate means of conveying the information e.g. report, in-person presentation, PowerPoint presentation, or Dashboard.
5. Ensure the presentation is easily understandable by the target audience by considering use of jargon, visual aids, and data visualization best practices.
6. Craft the presentation in an engaging manner that creates and maintains audience interest.

For purposes of this case study, a PowerPoint presentation was prepared in addition to this report. This is the link to the presentation https://tinyurl.com/Cyclistic2023CaseStudyGChegge

### Key Findings and insights

1. **Customer Distribution**:
   - 64% of users are annual members, and 36% are casual riders.
2. **Trip Duration**:
   - Casual riders account for 56% of total trip duration, while members represent 44%.
   - Casual riders tend to have longer trip durations (on average, 21–25 minutes), while members have shorter rides (12–13 minutes).
3. **Usage by Day of the Week**:
   - Casual riders are more active on weekends, particularly Saturdays (15.4% of rides).
   - Members use bikes more evenly throughout the week, with Monday being the least active (12.8%).
4. **Usage by Time of Day**:
   - Peak usage for both groups is between 3 p.m. and 6 p.m.

- Casual riders have the lowest usage early morning (5 a.m. to 8 a.m.), while members have their lowest usage between midnight and 4 a.m.

5. **Seasonality**:
    - Casual riders' usage drops during the winter months (January–March) but peaks during summer (June–August).
    - Members have a more consistent usage pattern throughout the year but slightly lower in the summer.
6. **Bike Type Preferences**:
    - Members use both electric and classic bikes equally, while casual riders favor electric bikes for longer trips.
    - Docked bikes, though less used (1% of trips), had the longest trips.
7. **Trip Distribution by Time of Year**:
    - The busiest months are June to September for both groups.
    - The slowest months for casual riders are January to March, correlating with winter weather.
8. **Trip Duration on Holidays vs Non-Holidays**:
    - **Casual Riders**: The average trip duration for casual riders on holidays was **25.2 minutes**, which is longer than their average on non-holidays (**21 minutes**).
    - **Annual Members**: For members, the average trip duration on holidays was **13.2 minutes**, slightly longer than their average trip duration on non-holidays (**12 minutes)**
9. **Popular Routes**:
    - There were different preferences for popular routes between casual riders and members.
    - **Top 5 most popular routes** for **casual riders** were identified separately from the **top 5 routes for members**, indicating that each group has distinct riding patterns. Casual riders likely use routes associated with more scenic or tourist areas, while members tend to use routes linked to commuting or regular transportation.
10. **Trip Starting Locations**:
    - Casual riders and members also had different **starting locations**. Casual riders tended to start their trips in locations that are likely associated with recreational or tourist activities, while members started trips from more practical, commute-related areas.
    - **Members** often started trips near residential or work areas, suggesting they use bikes as part of their daily routines.
11. **Trip Ending Locations**:
    - Like starting locations, the **end trip locations** differed between the two user types. Casual riders often ended their trips at locations different from

where they started, indicating that their usage is more exploratory or for leisure.

- **Members**, on the other hand, were more likely to return bikes to locations close to where they started, aligning with more utilitarian, point-to-point usage.
- **Casual** riders are more likely to take scenic or exploratory routes, and **members** use the bikes as part of their daily routine

### Top 4 Findings and Insights

- **Trip Duration**: Casual riders take longer trips than members, likely due to different usage motivations, such as leisure versus commuting.
- **Time and Day Preferences**: Casual riders prefer weekends and afternoons, whereas members use bikes more uniformly across the week, suggesting that casual users may engage in recreational activities while members likely use bikes for commuting or daily transportation.
- **Seasonal Patterns**: Casual riders are more sensitive to weather conditions, with a noticeable drop in winter, while members show more consistent usage year-round.
- **Bike Types**: Casual riders prefer electric bikes for longer rides, while members have no strong preference, indicating that members may prioritize convenience over speed.

## Step 6 – ACT

### Objective of Taking Action

As the saying goes, "Plans do not become reality just because they are good and desirable, we must put them into action!" While the findings and insights may be valuable, it is the responsibility of the stakeholders to act on them. This might involve raising additional questions or requesting further analysis, but ultimately, the goal is to implement a well-informed action plan that addresses the core problem. Only through effective execution will the true purpose of data analysis be fulfilled, transforming insights into meaningful outcomes.

### Conclusion

- **Above findings and insights effectively demonstrate the differences in how the two groups of customers, Annual Members and Casual Riders, use Cyclistic's services.**
- **The insights are adequate to help guide marketing strategies to convert Casual Riders into Annual Members**

## Room for Improvement

As the saying goes, "The largest room in the world is room for improvement!" Following are areas which could offer opportunities for further analysis and collection of additional insights.

☐ **Depth of Interpretation:** While the case study identifies key insights, some areas, like **geographical analysis or weather impacts**, could be explored in more depth to enrich the findings.

☐ **Missing Contextual Factors:** Certain external factors like **socio-economic factors, transportation infrastructure, or bike availability** were not considered, which could provide further insights into rider behavior.

……………………………………………………………………**END**…………………………………………………………………..