# Cyclistic

## Bike-Share Analysis Case Study

Understanding the Behavioral Differences Between Annual Members & Casual Riders

- by Sachin Kanchan | LinkedIn
- Last Updated – 9th Oct, 2024

# Objective

Analyze the differences in how annual members and casual riders use Cyclistic bikes.

The goal is to uncover insights that will guide the development of a targeted marketing strategy to convert casual riders into annual members, thus driving higher profitability for Cyclistic.

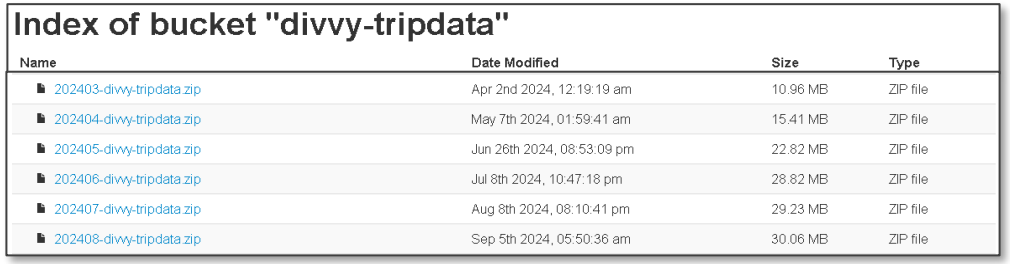
## Stakeholders

- Marketing Team
- Executive Team

# Data Collection & Prep

## Data Sources Used

The data used for the analysis was obtained from the publicly available Divvy bike-share system, which provides historical trip data for Cyclistic bikes. The data spans the last 6 months and is stored as monthly CSV files, accessible through Divvy Trip Data.



**Index of bucket "divvy-tripdata"**

| Name | Date Modified | Size | Type |
|---|---|---|---|
| 202403-divvy-tripdata.zip | Apr 2nd 2024, 12:19:19 am | 10.96 MB | ZIP file |
| 202404-divvy-tripdata.zip | May 7th 2024, 01:59:41 am | 15.41 MB | ZIP file |
| 202405-divvy-tripdata.zip | Jun 26th 2024, 08:53:09 pm | 22.82 MB | ZIP file |
| 202406-divvy-tripdata.zip | Jul 8th 2024, 10:47:18 pm | 28.82 MB | ZIP file |
| 202407-divvy-tripdata.zip | Aug 8th 2024, 08:10:41 pm | 29.23 MB | ZIP file |
| 202408-divvy-tripdata.zip | Sep 5th 2024, 05:50:36 am | 30.06 MB | ZIP file |

This data includes key variables such as:

- Ride ID: Unique identifier for each bike trip.
- Rideable Type: Type of bike used (e.g., classic, electric).
- Start/End Timestamps: Details of when and where each ride started and ended.
- Station Information: Names and IDs of the stations where trips began and ended.
- Geolocation Data: Latitude and longitude coordinates for start and end locations.
- User Type: Classification of riders as either ‘casual’ or ‘member’.

| ride_id | rideable_type | started_at | ended_at | start_station_name | start_station_id | end_station_name | end_station_id | start_lat | start_lng | end_lat | end_lng | member_casual |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6F2A48192B2F2921 | classic_bike | 00:43.0 | 11:51.0 | California Ave & Altgeld St | 15646 | Damen Ave & Pierce Ave | TA1305000041 | 41.92669 | -87.697668 | 41.90939601 | -87.67769193 | member |
| B929FF5BF24D4F83 | classic_bike | 00:54.0 | 03:33.0 | University Library (NU) | 605 | Sheridan Rd & Noyes St (NU) | 604 | 42.052939 | -87.673447 | 42.058239 | -87.677432 | member |
| 05ADF1DA01BBB6BF | electric_bike | 01:15.0 | 06:14.0 | Eckhart Park | 13289 | Ogden Ave & Chicago Ave | TA1305000020 | 41.89644992 | -87.66100669 | 41.89636246 | -87.65406127 | member |
| C555B5538B0C461B | classic_bike | 01:31.0 | 05:44.0 | Clark St & Newport St | 632 | Sheffield Ave & Wellington Ave | TA1307000052 | 41.94454 | -87.654678 | 41.93625348 | -87.6526621 | casual |
| 45F5D93322B356A7 | classic_bike | 01:46.0 | 27:00.0 | 900 W Harrison St | 13028 | Clinton St & Polk St | 15542 | 41.874754 | -87.649807 | 41.87146652 | -87.64094913 | member |

The data is sourced from Motivate International Inc., ensuring it adheres to licensing agreements and privacy protocols by excluding any personally identifiable information. This dataset provides sufficient detail for an in-depth analysis of rider behaviour trends, helping answer the key business question about usage patterns between casual riders and annual members.

# Data Processing

## Data Import and Initial Review

- Imported 6 months of data (Mar-2024 to Aug-2024) into SQL tool
- Performed initial data structure and content review

## Data Cleaning and Feature Engineering

- Dropped records - 211 duplicate rows, 159 rows where start time was after end time
- Created Table with additional date time columns to help with analysis

## Performance Optimization

- Created non-clustered indexes on key columns

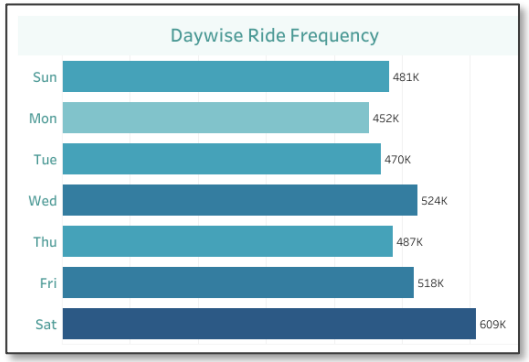
## Final Checks

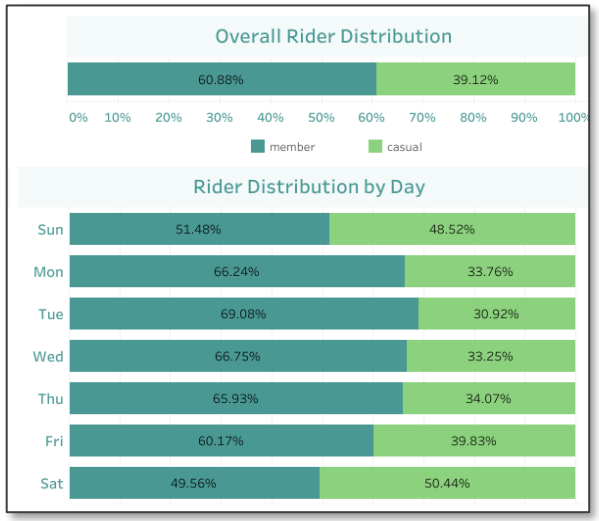
- Ensured no duplicates
- Verified data types and null values
- Reviewed extreme values and sorted data for visual inspection

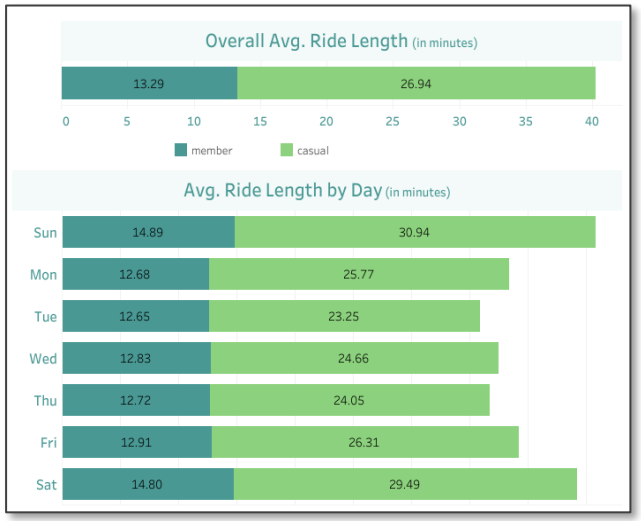
# Analysis

## Observations

- **Membership split**: Members (60.88%) have shorter trips, while casual riders (39.12%) take longer rides.
- **Weekly dominance:** Members dominate weekdays (66-69%), casual riders surge on weekends (~50%).
- **High-traffic locations:** Stations like "Streeter Dr & Grand Ave" are ideal for targeted marketing.
- **Ride variability:** Casual riders show more variability, including extreme outliers (up to 26-hour rides).



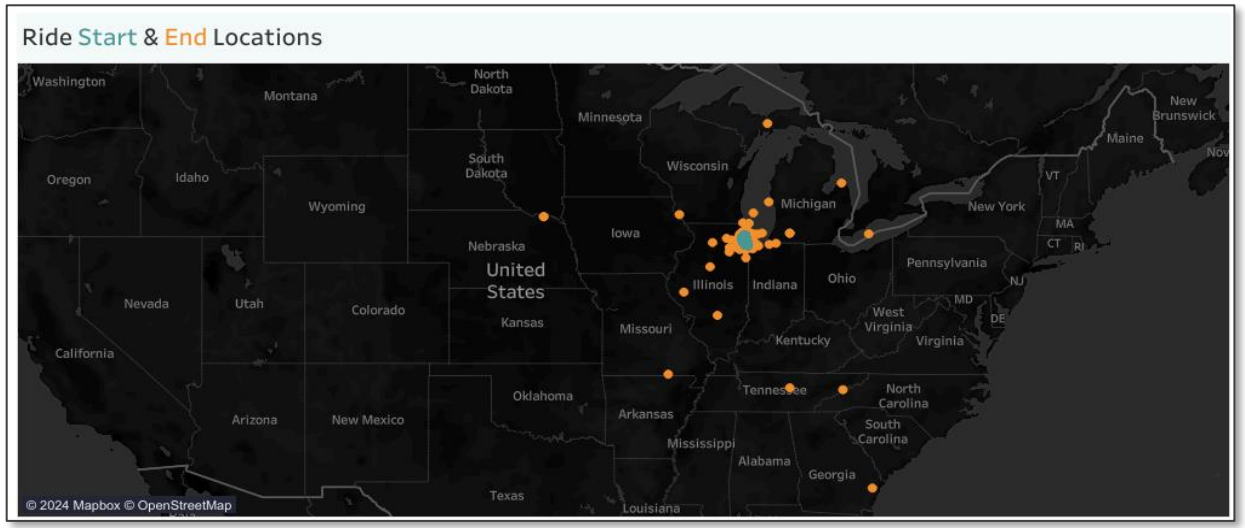




### Ride Length Summary

| | |
|---|---|
| Min Ride Length: | 0.02 mins |
| Median Ride Length: | 10.43 mins |
| Max Ride Length: | 1,559.93 mins |
| Avg Ride Length: | 18.63 mins |

### Top 10 Longest Rides

| Ride Id | Rideable Type | Member Casual | Start Station Name | End Station Name | Ride Length | Day Name |
|---|---|---|---|---|---|---|
| 2E7A26B0B6513B10 | classic_bike | casual | Green St & Madison St | Null | 25:59:56 | Saturday |
| 446B8A4213AAD837 | classic_bike | casual | Franklin St & Chicago Ave | Null | 25:59:55 | Saturday |




## Hypothesis

- Casual riders prefer weekend leisure rides, while annual members use bikes for weekday commuting.
- Casual riders engage in more recreational or tourist activities, while members use bikes for practical purposes.

# Recommendations

## Action & Next Steps

Based on our analysis, here are three key recommendations:

1. Targeted Weekend Conversion Campaign: Launch a marketing campaign aimed at converting casual riders to members, focusing on weekends when casual ridership peaks. Highlight the cost-savings of membership for frequent weekend riders.
2. Introduce Tiered Membership Options: Develop membership tiers that cater to different usage patterns. For example, a "Weekend" membership for those who primarily ride on weekends, or a “Vacation" membership for users who prefer extended rides.
3. Location-Based Marketing: Implement targeted marketing strategies at high-traffic stations, particularly "Streeter Dr & Grand Ave". This could include on-site promotional events, digital advertising geotargeted to these areas, or special offers for first-time members who sign up at these locations.

# APPENDIX

1. **Cyclistic (Divvy) Trip Data** - Original dataset available here .
    - Monthly trip data from March 2024 to August 2024 used in this analysis.
2. **Licensing Information –**
    - Project Code License - This project is licensed under the MIT License.
    - Cyclistic Data License Agreement - The dataset is provided by Divvy and is governed by the Divvy Data License Agreement.
3. **Tools Used –**
    - SQL Server
    - Tableau
    - PowerPoint

For questions about this project, please contact via the repository's issue tracker.

# Thank You