Cyclistic Bike-Share: Analysis Summary

# Objective

I am a junior data analyst working in the marketing analyst team at Cyclistic, a fictional bike-share company in Chicago. The objective of this Cyclistic bike-share analysis project is to gain insights into bike-sharing usage patterns and user behavior to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, our team will design a new marketing strategy to convert casual riders into annual members as the director of marketing believes the company’s future success depends on maximizing the number of annual memberships.

This project will break down user types between casual and members then each group will focus on analyzing usage patterns and user behavior trip durations, and station utilization patterns.

The project aims to provide actionable recommendations and strategies aimed at converting casual riders into annual members. To do that, however, the marketing analyst team needs to answer the key business questions shareholders have:

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

The analysis will provide insights and recommendations tailored for the Cyclitic executive's team who must approve your recommendations, Lily Moreno the director of marketing, and my manager, who will use the analysis to design marketing strategies.

# Key metrics

Key metrics for the Cyclistic bike share dataset include:

1. **Ride\_Id**: A unique identifier for each trip.
2. **Rideable\_type**: The type of bike used for the trip, which includes classic bikes, electric bikes, and docked bikes.
3. **Started\_at and ended\_at Time**: The date and timestamp indicate when each trip started and ended.
4. **Start\_station\_name and End\_station\_name**: The names of the stations where each trip originated and ended.
5. **Start\_station\_id and end\_station\_id**: The IDs of the stations where each trip originated and ended.
6. **Start\_lat and end\_lat**:the latitude coordinate of the starting and ending point of a bike trip.
7. **Start\_lng and end\_lng**: the longitude coordinate of the starting and ending point.
8. **member\_casual**: The type of user, such as "Member" or "Casual," indicating whether the user is a regular subscriber or a one-time customer.
9. **Ride\_length**: A difference in minutes between the start and end time of each trip that indicates the duration of each trip.
10. **Trip\_route**: A combination of start and end station names that represents the specific route taken during the trip.
11. **Peak\_hours**: The busiest hours when the bike usage is highest.

These key metrics mentioned above provide essential information about the Cylistics bike-sharing dataset. They provide patterns, user behavior, station usage, trip durations, and other important aspects of a bike share system. By analyzing these metrics, I will gain insights into user preferences, popular routes, peak usage times, and common and differences between the user types, and importantly answer the business question by shareholders of Cyclistic so they can design marketing strategies.

# Key variables

1. User type: The type of user, such as "Member" or "Casual," indicating whether the user is a regular subscriber or a one-time customer.
2. Total trips: The total number of bike trips recorded in the dataset
3. Rideable type: The type of bike used for the trip, which includes classic bikes, electric bikes, and docked bikes.
4. Ride length: A difference in minutes between the start and end time of each trip that indicates the duration of each trip.
5. Trip route: A combination of start and end station names that represents the specific route taken during the trip.
6. Peak hours: The busiest hours when the bike usage is highest.
7. Days of the week:
8. Popular stations: The stations with the highest number of trips originating or ending
9. Peak hours: The periods with the highest bike trip volumes
10. Popular ridable type: the most preferred or used kind of bike
11. User Type Distribution: The percentage or count of members and casual riders

# Statistical Analysis

## Users Type Breakdown

**Members** are regular users of the bike-sharing service of Cyclitic company who have a subscription to an annual membership.

**Casual Riders** are occasional users of the Cyclitic bike-sharing service who don't have subscriptions to an annual membership.

The total number of Cyclistic bike-sharing users in 2022 amounted to 5667717.

* Member users were (59%)
* Casual users were (41%)

The analysis reveals that the majority of bike-sharing users were annual members, accounting for 59% of total trips, while casual users accounted for 41%.

Understanding the distribution of user types is crucial for the Cyclistic bike-sharing company since the goal of the marketing director is to design marketing strategies aimed at converting casual riders into annual members.

The almost 50:50 proportions suggest there is a strong base of loyal customers who may have subscribed to long-term memberships and most importantly a base of non-members who have recurring usage patterns. The significant presence of casual riders indicates the marketing strategy won't be wasteful if done correctly. In addition, there is an opportunity for attracting more occasional riders to convert to annual members now or in the future.

Understanding the distribution of user types is crucial for Cyclictic to tailor and develop marketing strategies aligning with their usage preference and pattern, and most importantly to confirm the strong base of casual riders does exist and that tailoring market strategies would be profitable.

### 

The following analysis will explore the difference in usage patterns and preferences between the two user types, members and casual riders.

### Most preferred rideable type

Casual:

* Electric\_bike - 53.97%
* Classic\_bike - 38.39%
* Docked\_bike - 7.64%

member:

* Classic\_bike - 51.10%
* Electric\_bike - 48.90%
* Docked\_bike - 0%

The analysis revealed that among members, the classic bike shows as the most preferred rideable type, constituting 51.10% of their total rides of choice. On the other hand, casual riders showed a clear preference for electric bikes, with 53.97% of their rides utilizing this rideable type.

Cyclists may prefer one over the other based on their specific needs, preferences, and the type of cycling experience they desire.

Based on the analysis, Casual members make use of the types of ride bike that are commonly used for commuting especially long distances and easily tackles uphills as they offer a boost in speed and reduced physical effort. With this, we can hypothesize casual riders mainly used bikes to commute to work.

On the other hand, the majority of members make use of classic bikes, which are commonly used for leisure and physical fitness purposes. These users are often fitness enthusiasts or riders who prioritize exercise during their commute or leisure rides.

However, it should be noted that nearly 8% make use of docked bikes. these are users who value the convenience of a reliable and regulated bike-sharing service, with designated docking stations for pick-up and drop-off. During the year, casual riders used bikes, they are popular for short-term or one-way trips within a city or urban environment. When developing market strategies, it should be wise to take note of this usage as docked bike systems typically offer ease of access, predictable bike availability, and well-maintained bikes.

### Popular stations

The most popular start station within each user type

* Casual: Streeter Dr & Grand Ave
* Membe: Kingsbury St & Kinzie St

The most popular end station within each user type

* Casual: Streeter Dr & Grand Ave
* Membe: Kingsbury St & Kinzie St

The analysis reveals that the most popular station amongst casual users is Streeter Dr & Grand Ave, while for members is Kingsbury St & Kinzie St.

The director of marketing aimed at converting casual riders into annual members, this piece of insight is important for designing the market strategies tailored towards casual riders.

The users consider many factors into choosing the stations. The attributions might be; Convenience, riders often consider the convenience of accessing and returning the bikes. They opt for bike stations located near their starting point and destination, ensuring easy access and availability of bikes. Bike Availability, riders will prefer stations that have a sufficient number of bikes available, especially during peak hours or in high-demand areas.

### Most used days

The common days for average trips per year by each user type:

* Casual: Thursday
* Member: Saturday

The analysis reveals that Thursday has the highest average number of trips among casual users. And Saturday for members. Understanding the average day of the week is crucial to determine the most popular day for biking for each user type.

Thursday being popular for casual riders could indicate certain patterns or factors that make Wednesdays more favorable for bike rides, such as favorable weather conditions, specific events or activities happening on Thursday, or commuting patterns that are prevalent on that day. With no customer feedback data, it is not certain the reason but to hypothesize it commuting patterns during weekdays as the pattern shows a constant peak on weekdays.

With members, the busiest days are on the weekends, with Saturday having the highest frequency of trips among all the days of the week. This may be due to leisure that happens over weekends or for fitness usage.

### Peak Hours

The peak hours when bike usage is highest

* The overall peak hours: 5 PM
* Members: 5 PM
* Casual: 5 PM

Based on the analysis, it is evident that the peak hours for bike-sharing rides occur around 5 PM, and this trend is consistent for both member and casual riders.

With both members and casual users show a peak around 5 PM indicating a common usage pattern among different user segments during the late afternoon. As stated in a briefing, Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day. This suggests that various casual riders, including commuters, students, and leisure riders, are utilizing the bike-sharing service to meet their transportation needs as the above analysis shows the most used days are on weekdays whilst the members' busiest days are weekends thus are to engage in recreational activities later in the day.

### Trip Distance

Casual riders:

* The **median trip distance** is 2.1838, indicating that 50% of casual users have a trip distance less than or equal to this value.
* The **interquartile range (IQR)** is the range between the lower hinge (2.0240) and the upper hinge (2.2175), capturing the middle 50% of the trip distances

Members:

* The **median trip distance** is 2.1443, indicating that 50% of members have a trip distance less than or equal to this value.
* The **interquartile range (IQR)** is the range between the lower hinge (2.0240) and the upper hinge (2.2175), capturing the middle 50% of the trip distances

The median trip distance for casual users (2.1838) is slightly higher than that of members (2.1443), indicating that casual users tend to have slightly longer trip distances on average than members.

However, the difference in median trip distances between casual users and members is relatively small, indicating that their overall trip distance preferences are quite similar. But it should be noted, casual riders' trips exhibit greater variability in their trip distances. It implies that among casual users, there is a wider spread of trip lengths, with some users taking shorter trips and others taking longer trips in comparison to those members who tend to have more consistent trip distances.

It is important to note the outlier value of 1.7 among casual users suggests that there might be instances where a casual user opted for a very short trip, even though the majority of casual users have trip distances within the typical range observed for this user group. Further investigation and analysis would be necessary to understand the factors contributing to this outlier and its implications for the bike-sharing system and its users

Overall analysis reveals that members typically have slightly shorter trips with less diversity. On the other hand, casual riders have a long, wider range of trip distances, casual users may have more diverse trip patterns and behaviors compared to members. This understanding can be valuable for bike-sharing operators and policymakers in tailoring their services and offerings to accommodate the needs and preferences of casual users with different trip distance requirements.

### Total trips

Members took a total of trips 2611171

Casual took a total of trips was 1758189

The overall total number of trips taken during the 2022 year is 4369360.

### Trip Duration

The average ride length:

* Casual: 29 minutes and 14 seconds
* Member: 13 minutes 11 seconds

The common rides length:

* Casual: 6 minutes
* Member: 5 minutes

The average ride duration of casual users lasts longer than that of members approximately 2 times. On the surface, this indicates casual riders travel long distances than members. However, the trip distance analysis reveals the difference in median trip distances between the two groups of users is relatively small (approximately 0.04 units), indicating that the disparity in travel distance is not substantial. This suggests that the long trip duration may be influenced by various factors, including traffic conditions, bike-sharing station availability, user behavior, and the purpose of the trip that causes a trip to substantially last longer than those of members. With no customer feedback data to can to explore it's hard to confirm an exact reason. Therefore, it would be wide before tailoring the market strategy for casual riders to conduct the survey.

On the other hand, members indicate that their trip last shorter. They may be choosing to ride bikes for shorter distances as a means of incorporating physical activity into their daily routines, promoting an active lifestyle, leisurely around, and avoiding sedentary modes of transportation.

### Geographical Density Analysis

* Casual riders: higher concentration along the coastal line and moving inwards the urban
* Members: high concentration spreading along the Chicago coastline line.

The analysis reveals the concentration of both casual riders and members at the beginning of the trip is similar, as dense patterns move along the Chicago coastline line. While the end trip's density of members remains along the coastal line and one of the casual riders spreads inwards the city urban.

This suggests that members might have a preference for commuting or traveling along these specific routes. This implies that they are more likely to use the Cyclictic bike-sharing services for leisure activities, such as enjoying coastal views, exploring beach areas, or engaging in outdoor recreational activities around the coastline.

On the other hand, causal riders' density of end trips spreads inwards the city. This suggests the group had a wide range of users, the ones who commute or travel along the coasts, for work and leisure purposes. There are also great riders who might be using the bikes for short urban trips, such as commuting to work or running errands within the city. This suggests that casual riders are more likely to utilize bike sharing as a convenient mode of transportation for urban mobility.

2. Temporal Analysis:

- Mental Model: Explore how trip distance and duration vary over time, identifying any temporal trends or patterns.

- Analysis Techniques: Plot the average trip distance and duration over different periods (e.g., hourly, daily, weekly, monthly) to observe any patterns or variations. Use line charts or heatmaps to visualize changes in trip characteristics across time.

3. Spatial Analysis:

- Mental Model: Investigate the relationship between trip distance/duration and geographical factors.

- Analysis Techniques: Plot the trip start and end points on a map to visualize the spatial distribution of bike trips. Analyze whether trip distances or durations vary based on specific geographic features such as neighborhoods, landmarks, or bike station locations. Consider aggregating trip data at different geographical levels (e.g., city, district) to identify any spatial patterns.

4. Segment Analysis:

- Mental Model: Analyze trip characteristics based on different segments or subsets of the dataset.

- Analysis Techniques: Segment the dataset based on relevant variables such as user demographics (age, gender), membership type, day of the week, or time of day. Compare trip distance and duration across these segments to identify any differences or preferences. Use bar charts or stacked column charts to visualize and compare trip characteristics among different segments.

5. Outlier Analysis:

- Mental Model: Identify and analyze any extreme or unusual trip distances or durations.

- Analysis Techniques: Detect outliers by examining trip distances and durations that significantly deviate from the majority of the data. Consider using statistical methods such as box plots, z-scores, or interquartile ranges to identify outliers. Investigate these outliers further to understand the potential reasons behind them, such as system errors, special events, or unique user behaviors.

6. Correlation Analysis:

- Mental Model: Explore the relationship between trip distance and duration.

- Analysis Techniques: Calculate the correlation coefficient between trip distance and duration to understand the strength and direction of their relationship. Use scatter plots to visualize the relationship between these variables. Consider segmenting the data by other factors (e.g., user type, time of day) and examine the correlation within each segment.

Remember, these are just some analysis approaches, and you can adapt them based on the specific goals and context of your analysis. Exploring these different angles will help you uncover meaningful insights about the trip distances and durations in your Cyclitic bike dataset.

### Seasonal Analysis

#### Trip distance

The trip distance of casual riders remains relatively the same throughout the year with slight changes then drop significantly traveling short distances in November and December.

This suggests that casual riders maintain relatively consistent trip distances throughout the year with slight changes indicating a general preference for longer rides. This could be due to various factors such as commuting needs, or using bikes for around-the-city errands and they have established their preferred riding habits and continue to use the bikes to travel similar distances throughout the year. The notable drop in trip distances during November and December reveals a change in behavior during the holiday season. Casual riders may be less likely to use bikes for transportation activities during this period, potentially due to adverse weather conditions, opting for other modes of transportation, or the holiday season decreasing the need for transportation to commute.

On the contrary, the trip distances of members gradually become longer as the year progress, reaching its peak in July and then remaining consistent throughout the year till its drop in September. They travel long distances as the seasons get warmer. The peak in trip distances for members in July could be attributed to the summer season when outdoor activities are more popular. Longer daylight hours, warmer weather, and potential vacation periods during this time might contribute to members embarking on more extensive bike trips. The noticeable drop in trip distances for members in September could be influenced by factors such as the end of summer vacation periods, back-to-school routines, or a decrease in leisure time.

#### Peak hours

The peak hours for both groups of riders remain relatively consistent throughout the year regardless of the seasons.

#### Trip duration

Members' trip durations stayed consistent throughout the year with a slight drop at the beginning of September when early cold weather periods start and slight drops continues until December. This suggests that members' rides last relatively longer throughout the year than become short at the beginning of September. September marks the end of summer in many regions. As the weather starts to transition towards cooler temperatures and shorter days, people's outdoor activities, including bike riding, may naturally decrease. Members might be less motivated to take longer trips during this time, resulting in a drop in trip durations. September is a month when people often return to their regular routines after the summer vacation period. As work and school schedules resume, riders may have less time for leisurely bike trips, leading to shorter durations. This could indicate a shift in priorities and a decrease in overall free time available for bike riding. However, this did not have a large impact on the ride duration as we hypothesized throughout the analysis that members make use of Cyclistic bike-sharing services majority for leisure purposes.

In contrast, the trip duration of casual riders seems to be viable throughout the year depending on the month. This suggests the weather affects the ride length of casual riders. The rides last significantly high early in the year enjoying warm and hot seasons or might be additional factors such as marketing campaigns, events, or any significant changes in the bike-sharing service or its offerings during that period. Then the duration dropped significantly in September. Although the weather conditions might still be pleasant in September, there could be subtle changes such as cooler mornings or evenings, which might discourage casual riders from opting for longer trips. The slight decrease in temperature or the perception of approaching colder weather might lead to shorter ride durations as people become more cautious. In addition, the drop may be due to changes in commuting patterns, school schedules, or social activities, which might indirectly impact the duration of bike trips.

The overall season analysis shows that members' ride length is not significantly correlated with seasons or time of the year. This suggests that members' usage of bike-sharing services is consistent and not strongly influenced by seasonal variations. On the contrary casual riders' ride length is viable throughout the year; higher ride lengths are observed in spring and summer, and lower ride lengths in autumn and winter. This suggests that seasonal factors have a noticeable influence on the ride lengths of casual riders.

By understanding these patterns, the market director can tailor his strategies to convert casual riders to subscribe to membership by catering to the different preferences and behaviors of casual riders.

## Correlation Analysis

### Trip Distance vs Rideable Type

The analysis shows there is a positive correlation between the traveled distance and the preferred rideable type. This suggests that the riders who travel short distances opt for classic bikes and then docked bikes and for longer trip distances they opted electric bikes.

This might be attributed to many factors; riders who travel shorter distances may find classic bikes and docked bikes to be more efficient and comfortable for their needs as they are often preferred for shorter commutes, leisurely rides, or when riders prioritize physical activity and fitness. They allow riders to enjoy the benefits of exercise while covering moderate distances.

These bikes provide a convenient mode of transportation without the need for additional features or assistance. The long-distance riders, on the other hand, opted for electric bikes as they have an integrated motor that provides assistance while pedaling, enabling riders to cover longer distances with less effort and at higher speeds. This makes them suitable for riders looking for a faster and less physically demanding mode of transportation.

### Trip Duration vs Trip Distance

The analysis suggests there is no correlation between trip distance and duration traveled.

This lack of correlation between variables could have several implications:

##### Usage pattern

There might be different factors contributing to taking longer trips that don't necessarily involve covering more distance. The riders might take detours, make stops along the way, stuck in traffic during peak hours, or ride at a slower pace, resulting in longer trip durations without significantly increasing the distance traveled. In addition, the riders might have different motivations for using the bike-sharing service, such as commuting, exercise, or leisure leading to variations in trip durations that are not necessarily influenced by the distance covered.

##### Usage Preferences

Riders may use bike-sharing services for a leisurely and relaxed riding experience, taking their time to enjoy the surroundings or engage in other activities during the trip. These riders may have longer trip durations but cover relatively shorter distances. The riders may also prefer a route that offers scenic views over the shorter route thus increasing the duration of a trip

These findings highlight the different behaviors and preferences of members and casual riders based on seasons.

# Interpretation

**Members** are regular users of the bike-sharing service of Cyclitic company who have a subscription to an annual membership.

**Casual Riders** are occasional users of the Cyclitic bike-sharing service who don't have subscriptions to an annual membership.

### The summary of the above analysis:

##### Preferred Bike:

Members prefer classic bikes, while casual riders opt for electric bikes, indicating different preferences for traditional or motorized cycling experiences.

##### Most Used Days:

Members have higher usage on Saturdays, suggesting a preference for weekend bike rides.

Casual riders use bike-sharing services more on Thursdays, indicating a preference for weekday rides.

##### Peak Hours:

Both members and casual riders have peak usage at 5 PM, indicating a common trend of bike usage during the late afternoon or early evening hours.

##### Trip Distance:

Members tend to have shorter trip distances, suggesting a preference for shorter commutes or leisurely rides.

Casual riders have longer trip distances, indicating a preference for more extensive or adventurous rides.

##### Trip Duration (Average):

Members have slightly longer average trip durations, potentially indicating a more relaxed riding style or additional stops during their trips.

Casual riders have slightly shorter average trip durations, suggesting a more direct or faster riding approach.

##### Geographical Density:

Members are concentrated along the coastal line, indicating a preference for riding near the coast or waterfront areas.

Casual riders are more geographically dispersed, with a presence spanning throughout the city, indicating a broader range of riding destinations.

##### Seasonal Trends

Members' bike usage remains relatively constant throughout the year, suggesting a consistent preference for bike sharing regardless of the season.

Casual riders show a seasonal trend, with higher usage in warm and hot seasons and lower usage during cold weather, likely influenced by weather conditions and outdoor activities.

# Conclusion

In summary,

**Casual Riders:**

Casual riders exhibit distinct preferences and patterns in their bike-sharing behavior. They primarily opt for electric bikes, indicating a preference for motorized cycling experiences. Their usage peaks on Thursdays, suggesting a preference for weekday rides. Furthermore, casual riders tend to embark on longer trips, indicating a desire for more extensive or adventurous rides. They also have slightly shorter average trip durations, suggesting a more direct or faster riding approach. Geographically, casual riders are dispersed throughout the city, indicating a broader range of riding destinations. Additionally, their bike usage shows a seasonal trend, with higher activity in warm and hot seasons and lower usage during colder weather, likely influenced by weather conditions and outdoor activities.

**Members:**

Members, on the other hand, exhibit different preferences and behaviors compared to casual riders. They prefer classic bikes over electric ones, indicating a preference for a more traditional cycling experience. Their usage peaks on Saturdays, indicating a preference for weekend bike rides. Members tend to have shorter trip distances, suggesting a preference for shorter commutes or leisurely rides. They also have slightly longer average trip durations, potentially indicating a more relaxed riding style or additional stops during their trips. Geographically, members are concentrated along the coastal line, indicating a preference for riding near the coast or waterfront areas. Unlike casual riders, their bike usage remains relatively consistent throughout the year, reflecting a consistent preference for bike sharing regardless of the season.

The summary in a table format:

| **Variables** | **Members** | **Casual Riders** |
| --- | --- | --- |
| 1. Preferred bike | Classic | Electric |
| 2. Most used days | Saturday(weekends) | Thursday (weekdays) |
| 3. Peak Hours | 5 PM | 5 PM |
| 4. Trip Distance | Short | Long |
| 5. Trip Duration(average) | 29 minutes and 14 seconds | 23 minutes and 11 seconds |
| 6. Geographical Density | Coastal line | Coastal line and span out in a city |
| 7. Seasonal trends | Remain relatively constant throughout the year | High in warm and hot seasons and low in cold weather |

These findings provide valuable insights into the preferences, behaviors, and usage patterns of both members and casual riders of Cyclitic. Cyclitic stakeholders can utilize this information to optimize their services, target marketing efforts, allocate resources effectively, and enhance the overall user experience for each user group.