# **Cyclist bike share project**

# Data available at: <https://divvy-tripdata.s3.amazonaws.com/index.html>

# Overview

Cyclistic is a bike-share company in Chicago, that is trying to improve their business, and it relies on data and their analysts to guide them with solution. Cyclistic's finance analyst came to realization that annual members generate more profit than casual riders. Head of Marketing believes that maximizing the number of annual members will be key to the future growth, and that there is a very good chance that current casual riders can be converted into members, using the campaign that targets current – casual riders rather then new customers.

Key business task behind this analysis is to identify how to annual members and casual riders use Cyclistic differently. To do this, I will analyze the Cyclistic historic bike trip data to identify trends. This analysis has a task of understanding casual riders, their approach to Cyclistic bikes, and how to target and convince them to switch to annual membership. In this analysis I will rely completely on R programming language, for both analysis and visualization.

A bike-share program that features more than 5,800 bicycles and 600 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can’t use a standard two-wheeled bike.

* In the plot below, we can see that the majority of users are annual members (63%), while casual users account only for 37%.

Chart, pie chart

Description automatically generated

* The plot below shows us that on average, casual riders drive longer than the annual members. This is probably due to the fact that annual members are more flexible in terms of leaving and renting the bike again. Also, the plot shows us the average ride time by the day of the week. Expectedly, people are driving longest on Sunday (7) and Saturday (1).

Chart, scatter chart

Description automatically generated

* This screenshot from RStudio shows the same results as the graph above. I included it so we can see in numbers the mean, max, and min ride time by the user type and day of the week. Once again, we can clearly see that casual drives much longer than annual members.

Text

Description automatically generated

* The graph below ignores outliers (rides above 1 hour) se we can see the distribution of ride time. We can conclude that most of the rides are relatively short, between 1 and 20 minutes for both user groups.

Chart, histogram

Description automatically generated

* Boxplot below shows the same results as the histogram above. It is used to show distribution of ride time between two groups. We can conclude casuals drive longer (also in more variation).

Chart, box and whisker chart

Description automatically generated

* This plot tells a different story. Even though we have previously concluded that Casual riders drive longer, here we can clearly see that annual members drive more frequently. In other words, they feel free to leave and take a new bike anytime they want, since there won’t be an additionally charged for that.

Chart, scatter chart

Description automatically generated

* Boxplot below is used to show distribution of number of rides between two groups. This graph tells the same story as the one above. Hence members drive more frequent (also in lass variation).

Chart, box and whisker chart

Description automatically generated

* Finally, we can se which type of bike does the members and casuals prefer.

Chart, scatter chart

Description automatically generated

From the analysis of this dataset we can conclude following:

* Most riders are members (~63%), and the rest are casual users.
* Even though annual members use bikes more frequently, casuals rides them for a longer period of time.
* We can see that casuals rides bikes mostly during the weekend, while members ride constantly, and even more during workdays.
* Members use mostly classic bike and then electric bike. Casuals mostly ride electric bike, than classic and the last is docked bike.
* There are some users that rides bikes over 10h a day, and dose are the outliers in the data. Most of does rides happens on the weekend and by casual users.
* Ignoring outliers we can conclude that most of the rides are up to 20min long.
* According to the boxplots, members are more consistent with the usage time and usage rate, while that highly varies with casual users.
* Top starting stations are: Streeter Dr & Grand Ave, Ellis Ave & 60th St, University Ave & 57th St.

\*some more insights and conclusions can be find above each graph.