Citi Bike Analysis

My analysis focuses on two phenomena for the last three years (Jan 2019 – Dec 2021):

1. Ridership trends and an increasing number of Customers (casual riders) vs. Subscribers (members). This visual analysis is featured in the Rider Trend Dashboard (RiderTrendDASH).
2. Exploring differences in user types; in other words, finding differences in the riding habits of Customers versus Subscribers. This visual analysis is featured in the User Type Dashboard (UserTpeDASH).

**Ridership Trends**

The Rider Trend dash indicates that the number of rides by Customers has increased in the past few years.

* The Ridership Trends by User Type area chart reveals the number of Subscriber rides has stayed relatively consistent: Subscriber rides had an extended decline from September 2019 through April 2020 (exacerbated by the onset of COVID 19 Pandemic in March 2020). Rides increased during the summer of 2020, declined again during the fall and winter, before rebounding close to 2019 levels in 2021.

Customer rides followed a similar pattern in terms of seasonality; however, the number of Customer rides increased in 2020, and increased significantly in 2021. Per the Trend visual, customer rides accounted for roughly 50% of all rides during the summer of 2021.

* Per the Rides by Year bar chart, the number of trips has increased to 640,825 total rides in 2021 – an increase of 58% since 2019. Customer rides hit 289,166 rides in 2021, nearly a 7x increase since 2019, while subscriber rides have decreased 2.8% over the same period.
* Finally, Subscriber Ratio changes are consistent among the top 10 stations (top 10 based on number of rides). The Subscriber Ratio table shows that 9 of the stations have had a decrease in subscriber ratio each year, from 81 to 95% in 2019, to 53 to 72% in 2021.

**Subscriber vs. Customer Riding Habits**

The increase in Customer use leads us to look into how this will affect ridership using the User Type dashboard.

* In the Trip Duration scatter chart, we see that Customer Trip Durations are consistently higher than subscriber trip durations during each month of the year, particularly in early spring. Citi Bike may need to make more bicycles available at each station, given the longer usage by customers as well as the increasing number of rides by Customers.
* The Distance – Duration scatter plot highlights a key difference between Subscribers and Customers. Trip Distances are roughly the same between Customers and Subscribers. However, a Subscriber typically takes 5 to 15 minutes to complete most trips, with longer trips distances taking more minutes to complete. Meanwhile, Customers require more time to complete the trips, and have much more variability, with most trips taking anywhere from 15 minutes to 40 minutes. The R2 values of the two trendlines (Subscriber = 0.404 vs Customer = 0.008 ) reflect that it is much more difficult to predict how long it will take a Customer to complete a ride.
* Finally, Delayed Return chart indicates that customers are roughly 10 times more likely to end a trip more than 24 hours after starting (which indicates a delay in returning the bike to a station rather than a long distance trip). However, this only occurs roughly once every 1,000 trips. This is not likely to disrupt bicycle supply; however, continued analysis of this ratio is recommended.

**Map**

* For further understanding and spatial analysis of ridership changes at stations, please see the map, which shows Trips started and ended by Station by month.

**Conclusion**

Customer ridership is trending up at a fast rate, and with this increase in Customers comes longer and less predictable trip durations, which may lead to bicycle shortages across the station network. We recommend further statistical analysis to examine these trends to determine an appropriate action plan.