Citi Bike Analytics Case Report

By: Alex Reyes

Background:

I was recently hired on by New York Citi Bike Program, the largest bike sharing program in the United States, as the new lead analyst to assist in generating regular reports for city officials. Since 2013, the Citi Bike Program has implemented a robust infrastructure for collecting data of the program's utilization infrastructure. Because city officials are looking to publicize and improve the city program, the following "CITY BIKE ANALYTICS CASE REPORT" was created to share Citi Bike's program utilization via monthly collected data from January 2020 to June 2020.

Note: This Report was written/edited with Tableau Public, Python, Pandas, Microsoft Word, and Microsoft Excel.

https://public.tableau.com/views/Tableau_challenge_citibike_TopStartingStationsMonth2020/Story1?:language=en&:display_count=y&:origin=viz_share_link

Task 1: Unexpected Phenomena

- Top 10 Starting Stations:

    1. Grove St PATH
    2. Newport Pkwy
    3. Liberty Light Rail
    4. Hamilton Park
    5. Sip Ave
    6. Harborside
    7. Columbus Dr at Exchange Pl
    8. Marin Light Rail
    9. Newport PATH
    10. City Hall

- Top 10 Ending Stations:
    1. Grove St PATH
    2. Liberty Light Rail
    3. Newport Pkwy
    4. Hamilton Park
    5. Harborside
    6. Columbus Dr at Exchange Pl
    7. Sip Ave

8. Marin Light Rail
9. Newport PATH
10. City Hall

- Higher Male User % than Female.

- Alarming # of Gender Misinformation from Non-Subscribers.

- More Citi Bike Utilization during Morning (7-9AM) & Afternoon (4-7PM) Rush Hour.

- More Millennial Citi Bike users than ANY other Generation.

- Bike Maintenance is HEAVIER for "Bike Trip Count" bike units than "Bike Trip Duration" bike units.

Task 2: Tableau Dashboards

- Introduction: Citi Bike Analytics

- Citi Bike Map Start Station: January - June 2020 / Monthly Visual.

- Citi Bike Map End Station: January - June 2020 / Monthly Visual.

- Citi Bike Map Start & End Station Table Summary: January - June 2020.

- Citi Bike Monthly Trip & Gender User Count Comparison.

- Citi Bike Gender Count & User Type Comparison: January - June 2020.

- Citi Bike User Type Comparison (Subscriber vs. Customer): January - June 2020.

- Citi Bike User Type Utilization by Month Date: January - June 2020.

- Citi Bike Utilization Unit Count by 24 Hour Time Period: January - June 2020.

- Citi Bike Age Demographics: January - June 2020.

- Citi Bike Maintenance Scatter Plot Recommendations: January - June 2020.

- Citi Bike Maintenance Bar Graph Recommendations: January - June 2020.

- End of Report

Task 3: Visualizations - Dynamic Maps (by month & year)

- Dynamic Maps (by month & year) are shown in Git Hub Repository.

Task 4: Phenomenon Analysis

- It would be highly recommended for the Citi Bike Program to install more resources and customer service representatives near these locations. I believe it will greatly help to retain our current subscribers and encourage new customers by having Citi Bike personnel in these "high traffic" stations. Also, our Citi bikes must be in top shape in these locations. The last thing we want is for our WORST maintenance bikes to be at our BEST starting and ending location stations.

- It appears that there is a significantly higher population of Male Citi Bike users than Female users. Nearly 3x the rate from January to March 2020 & 2x the user rate from May to June 2020. Also, please notice a % increase in the "unknown" gender users who do not truthfully input their gender.

- An increasing % of customers are not truthfully inputting their gender. This could be due to gender classification (no longer labeling sexuality as male/female), customer fear of sharing personal information with Citi Bike, unease of surveillance, dismay of location awareness (user tracking), technologically illiterate, etc. We need to find out why they are doing this!!!

- Although there has been an increase of female ridership from the Citi Bike gender outreach, there is still a significant gap between female/male subscribers and customer/subscriber user type in general. What additional outreach programs can we create to get more subscribers? What additional outreach programs can we instill to promote female ridership?

- On average there seems to be a higher rate of Citi Bike users during the middle of the month than towards the end of the month. A similar correlation is seen between Customers and user Subscribers. I would recommend having a promotional outreach program on the first day of each month. Maybe something like a "Freeride Monday"? We could use the monthly promotion to increase our overall usage rate / user.

- On average there seems to be a higher usage rate of Citi Bikes during the morning and afternoon rush hour. Specifically, there is an upward trend from 7am to 9am and another spike from 4pm to 7pm. This upward trend is most likely due to users getting to and from work. Although this shows the starting time usage from January to June 2020, it would be interesting to see the comparison to daylight savings time (more sunlight) and the summertime period (summer break). Also, how can we increase our Citi Bike usage rates during the nighttime?

- There seems to be a high percentage of Citi Bike users born from 1980 to 1995 for both Male and Female users. This age demographic is significant toward our future marketing campaigns. The Millennial generation is are currently our highest users and Generation X users come in second. This could be due to Millennials wanting to be more environmentally cautious than other age demographics. We should also look at market strategies to attract future Generation Z users.

- Maintenance Recommendation: The following Citi Bikes should be maintenance and/or replaced due to trip duration distance and utilization trip count.

1. Bike ID: 42197 - Highest Trip duration; 42535 - 2nd Highest; 38355 - 3rd Highest; 42532 - 4th Highest; 42228 - 5th Highest.

2. Bike ID: 42296 - Highest Trip Count; 42545 - 2nd Highest; 42134 - 3rd Highest; 42298 - 4th Highest; 42258 - 5th Highest.

Dataset

- [Citi Bike Data](#) webpage.
- 202001-citibike-tripdata.csv.zip
- 202002-citibike-tripdata.csv.zip
- 202003-citibike-tripdata.csv.zip
- 202004-citibike-tripdata.csv.zip
- 202005-citibike-tripdata.csv.zip
- 202006-citibike-tripdata.csv.zip

Dataset Index

- Trip Duration (seconds)
- Start Time and Date
- Stop Time and Date
- Start Station Name
- End Station Name
- Station ID
- Station Lat/Long
- Bike ID
- User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member)
- Gender (Zero=unknown; 1=male; 2=female)
- Year of Birth