**Citi Bike Analytics Case Report**

By: Rebecca Melo

**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 Case Report was created to share Citi Bike's program utilization via monthly collected data from peak months in 2020 and 2021 from April to September months.

**Note: I was not able to upload my Tableau Workbook due to technical issues with publishing to Tableau Public. I kept getting stuck / frozen every time at the “ send data to the server” portion was executing. I have saved all of my screenshots from my Story / Dashboards as PNG files to follow along.**

**Task 1: Unexpected Phenomena**

* Top 10 Starting and Ending Stations in 2020:
  1. E 17 ST & BROADWAY
  2. W 21 ST & 6 AVE
  3. WEST ST & CHAMBERS ST
  4. BROADWAY & W 60 ST
  5. 7 AVE & CENTRAL PARK SOUTH
  6. 1 AVE & E 68 ST
  7. CLEVLAND PI & SPRING ST
  8. WEST ST & LIBERTY ST
  9. 12 AVE & W 40 ST
  10. CENTRAL PARK 5 & 6 AVE
* Top 10 Starting and Ending Stations in 2021:
  1. E 17 ST & BROADWAY
  2. W 21 ST & 6 AVE
  3. WEST ST & CHAMBERS ST
  4. BROADWAY & W 60 ST
  5. 7 AVE & CENTRAL PARK SOUTH
  6. 1 AVE & E 68 ST
  7. CLEVLAND PI & SPRING ST
  8. WEST ST & LIBERTY ST
  9. 12 AVE & W 40 ST
  10. CENTRAL PARK 5 & 6 AVE
* From 2020 to 2021, these TOP 10 Stations did not change that much. The same stations on both lists include:
  1. E 17 ST & BROADWAY
  2. W 21 ST & 6 AVE
  3. WEST ST & CHAMBERS ST
  4. BROADWAY & W 60 ST
  5. 7 AVE & CENTRAL PARK SOUTH
  6. 1 AVE & E 68 ST
  7. WEST ST & LIBERTY ST
  8. 12 AVE & W 40 ST
* 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.
* Growth for both Customer/ Casual and Subscriber (Member) have Increased in these Months and by Year. However, Subscribers percent increase is not as high as Customer.
  1. Percent Increase in Customer(Casual) : 63%
  2. Percent Increase in Subscribers(Members): 37%
* Monthly Trip Counts have increased by Month and by Year.
* More Millennial Citi Bike users than any other Generation.

**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.
* 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/ members and encourage new customers by having Citi Bike personnel in these "high traffic" stations.
* It appears that there is a significantly higher population of Male Citi Bike users than Female users. Also, an increasing % of customers are not 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, user error, etc. This should be looked into as to why this is happening.
  1. **Note: Citi Bike's 2021 data set did not include gender or age demographics to compare to 2020.**
* Although there has been an increase of ridership from the Citi Bike, there is still a significant gap between female/male riders. It would be beneficial to have additional outreach programs to create to get more subscribers, especially programs to promote female riders? Unfortunately, Citi Bike's 2021 data set does not include Gender and Age Demographics anymore. I am only inferring this through the 2020 data.
* On average there seems to be a higher rate of Citi Bike users during the end of the month than towards the beginning of the month. A similar correlation is seen between Customers(Casual) and Subscribers(Members). I would recommend having a promotional outreach program on the first day of each month to try to encourage riders to use Citi Bike and increase our overall usage rate / user.
* In 2020, both customer and subscriber growth went up exponentially. This could be due the fact that Covid was ramped during this year. When the weather was nice, I am sure a lot of riders, whether casual customers or subscribers, rode bikes more just to get out of the house and do something.
* In 2021, there is growth from April to September, and definitely higher than number count than 2020. However, there is a dip in member count from June to July. I believe the numbers are not increasing as fast as in 2020 due to the Covid spread slowing down around these months, along with people likely going away on vacation in the Summer months that members are not in the city using the bikes. In addition, you see the spike in September for both causal and member likely due to the School Season starting again for Fall semester.
* In 2020, monthly trip growth went up exponentially. This could be due the fact that Covid was ramped during this year. When the weather was nice, I am sure a lot of riders, whether casual customers or subscribers, rode bikes more just to get out of the house and do something.
* In 2021, there is growth from April to September, and definitely higher than number count than 2020. However, there is a dip in member count from June to July, and July to August. I believe the numbers are not increasing as fast as in 2020 due to the Covid spread slowing down around these months, along with people likely going away on vacation in the Summer months that members are not in the city using the bikes. In addition, you see the spike in September for both causal and member likely due to the School Season starting again for Fall semester.
* 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 April to September 2020-2021, it would be interesting to see the comparison to other months.
* 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 would be Citi Bike's significant target toward future marketing campaigns. The Millennial generation is currently Citi Bike's 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. Marketing should be looking into strategies to attract future Generation Z users.

**Dataset**

* [Citi Bike Data](https://www.citibikenyc.com/system-data) webpage.
* 2020-citibike-data-AprToSept.csv
  + Includes the following files converted together in Juypter Notebook:
    - 202004-citibike-tripdata.csv.zip
    - 202005-citibike-tripdata.csv.zip
    - 202006-citibike-tripdata.csv.zip
    - 202007-citibike-tripdata.csv.zip
    - 202008-citibike-tripdata.csv.zip
    - 202009-citibike-tripdata.csv.zip
* 2021-citibike-data-AprToSept.csv
  + Includes the following files converted together in Jupyter Notebook:
    - 202104-citibike-tripdata.csv.zip
    - 202105-citibike-tripdata.csv.zip
    - 202106-citibike-tripdata.csv.zip
    - 202107-citibike-tripdata.csv.zip
    - 202108-citibike-tripdata.csv.zip
    - 202109-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
* Ride ID
* Rideable Type
* User Type (Customer / Subscriber)
* Member Type (Casual / Member)
* Gender (Zero=unknown; 1=male; 2=female)
* Year of Birth