Pandemic Impact on NYC CitiBike Program

By: Hiam Debsi

January 15th, 2022

Introduction

CitiBike is the countrie's largest bike program, with 25,000 bikes, and over a 1000 stations across New York City. As of October 5th, 2020, NYC has reported over 250,000 COVID-19 cases and 23,861 deaths[(1)]. With this virus newly emerging in the city during the months of March and April, the city saw a surge that framed the entire nation's response to the pandemic. In the following sections, we will be looking at the effects the pandemic had on the flux and overall performance of the CitiBike Program. Particularly, we will look deeper into the months of March and April of both 2019 (pre-pandemic) and 2020.

Beginning Serge of Covid-19 Virus in New York City and It's Effect on the CitiBike Program

To investigate this phenomenon, we defined the time frame of the 'beginning serge' of Covid-19 as being the month of March 2020. Specifically, the first Covid case in NYC appeared on March 1st, 2020.

Lockdown Response to Covid-19 Virus in New York City and It's Effect on the CitiBike Program

To investigate this second phenomenon, we defined the time frame of 'Lockdown response' as being April 2020. In fact, in NYC, the strictest restrictions and emergency lockdown occurred on April 6th, 2020.

In the following section, we will collate the data required and analyse as needed in order to show the impact of both phenomena described above.

In Depth Analysis

Refer to Tableau Workbook or Tableau Public Link to see visualizations being analyzed

How Popular is CitiBike? (Sheet 1.1)
As a whole program, we can firmly say that people love it. It has a huge usage and immense trip duration for some bikes. No one has used a bike for less than an hour! In 2019, the top bike ID had a total of 3,018,100 minutes of trip duration. And it only increased (for that same bike ID) for 2020, with 4,100,282 minutes. But we want to look deeper into numbers of the who, when, where, and how of the CitiBike Program, to describe the effects of the pandemic in the months of March and April of 2020.

An Interesting Observation for the Year of 2020 (Sheet 1.2)
To further expand on the previous visual, we can see here that despite the start of the pandemic in 2020, the number of bike rides that took place are high. This visual is a combination of both months, but it still demonstrates an interesting observation.

Peak Hours of Bike Usage in March (Sheet 2.1)
In 2019, the peak hours on the weekend seem to be around 2-3PM. As for the weekdays, we can see a peak at around 8AM and 5PM, which is descriptive of the average population's work hours. And the same trend seems applicable in March of 2020. This is as expected given that when the virus first reached

cities, people were not as alarmed about it as they are today. We lacked information on the virus and didn't understand it's severity.

And we can see that subscribers are the most dominant types of users in March of both years, and the number of customers does slightly increase on weekends.

Peak Hours of Bike Usage in April (Sheet 2.2)
For April 2019, we can see that the trends weren't any different from March of both 2019 and 2020. Now, that quickly changes in April 2020. In terms of weekends, the overall peak hours have not changed, and there's in fact an increase in customer usage. As for weekdays, we clearly see the lack of peak morning hours and for some days, the afternoon. This is as a result of many workplaces, schools, etc., transitioning to online environments as a response to the city's announcement of stay-at-home policy, which took affect strictly on April 6th, 2020. This was as well expected, due to the limitations of travel that have been taking place in NYC at the time.

Gender Distribution Overview (Sheets 3.1 – 4.2)
How Distributed is the Program Across Genders?
In March particularly, the distribution was super similar from 2019 to 2020. There is a slight increase in Female users (~3%), a slight decrease in Male users (~4%), and a slight increase in unknown users (~3%). But generally, we can say the numbers did not change.

As for April, the numbers are a little greater in the change from 2019 to 2020. There is an increase of Female users of ~5%, a decrease of Male users of ~9%, and an increase in unknown users by ~3%.

Overall, we can say that the Program has a large Male target audience, and given their large regular presence, the effects of the restrictions in April were affecting them the most.

What is the Distribution of User Type amongst Genders for the CitiBike Program?
In March, we can see that, although not identical, the trend is quite similar from 2019 to 2020. However, we do clearly see a great increase in customers in 2020, and a slight decrease of subscribers in the unknown and male fields.

This quickly changes in April, there is an evident drop in both customers and subscribers for the male group, and an overall drop for all genders.

These results further emphasize that in March, although the virus was slowly spreading, people were not concerned yet. But when the government acted and started locking places down and restricting travel, we see the effects of it directly.

Age Distribution Overview (Sheets 6.1-7.2)
What Age Group Uses the CitiBike Program the Most?
Whether in March or April of either year, we clearly see that Adults are the main target group of CitiBike. They are the largest group to use bikes and particularly, the 50-52 years old range.

In March 2020, no change occurred from the previous year. However, as seen in other trends, April 2020 is where true impact occurred. We can clearly see a drop in the number of Bike IDs for the Adult group (yet there is still a peak in the 50 years old group), as well as all the Youths and the Seniors.

Which Age is the Most Active?

Although the Adult Group is the largest user of CitiBike, that does not necessarily mean they travelled the largest distances. In March 2020, we can see a spike in the average Trip Duration in minutes for the 17-18 years old. This could be because of Spring Break, when most students are off and have that time to go out. But generally, the trend for the other age groups is relatively the same.

If we investigate April, we see something a little different than other trends. There is an increase in average trip duration across all age groups, which is very interesting. Although the overall number of Bikes used at the same time was reduced, the time it was used for increased. This can be justified by the fact that many daily activities were paused due to lockdowns and restrictions in April 2020. People didn't have a lot to do, and outdoor physical activity was how many people spent their time. As we started to understand the virus more, and how it impacts the older population, some seniors were afraid of going out, which is observed in the 70-75 age group, as their trip duration decreased.

Let's Look into Average Distance Travelled (km) (Sheet 8.1)

We've been analyzing trip durations for users of the CitiBike Program.. how about the trip distance? Is there a similar trend? Are distances directly proportional to durations? And does the Pandemic have a similar impact? Well let's look into it...

Here we see the range of average distances travelled in km by bike users in both March and April of both years. The highest average distance was 8.423 km by the top Bike ID.
But let's take a deeper look to see the distribution of these averages amongst March and April, by age, gender, and user type.

Average Age vs Average Distance Travelled for March (Sheet 9.1)

In March, we can see that there's a concentration in specific Starting Stations, in both 2019 and 2020, for a dominant group of users between the ages of 30-40. Overall, customers do tend to be younger than subscribers. Given the R-value produced, there is no direct correlation between the distance travelled and the age of the biker, it is quite a sporadic area. The only certainty we have is that there are particular stations that are much more popular than other. In March 2020, the sporadicity the data is still there, and the concentration in particular stations is also still there. Although there isn't a direct correlation (according to the R-value obtained), there starts to be a slight decrease in age as you travel further along the graph.

Average Age vs Average Distance Travelled for April (Sheet 9.2)

In April, just like in March, we can see that there's a concentration in specific Starting Stations, in both 2019 and 2020, with a dominant group of users between the ages of 30-40. Overall, customers do tend to be younger than subscriber. Given the R-value produced, there is no direct correlation between the distance travelled and the age of the biker, it is quite a sporadic area. The only certainty we have is that there are particular stations that are much more popular than other. In April 2020, the sporadicity of the data is still there, and the concentration in particular station is also still there. However, we start to see a decrease in the number of subscribers that are travelling larger distances, and we start seeing a presence of members of the youth group. Customers are slightly increasing, and we see more of a dominance in females then males, which further justifies our previous findings about gender distribution in April of 2020.

The Top Start and End Stations for CitiBike (Sheets 5.1-5.4)

What are the Top Start/End Stations in March of 2019 and 2020?

Whether starting or ending the bike ride, **Pershing Square North Station** is the most accessed station, with a close second of **W21ST & 6 Ave Station.** And comparing 2019 to 2020, we can see that overall, the largest users are subscribers, and it doesn't necessarily alter from one year to another.

What are the Top Start/End Stations in April of 2019 and 2020?

Similarly to March, the top start and end station for both years is **Pershing Square North Station**. And the following 4 stations, whether the biker is starting or ending there, are the runner ups. The main change between March and April, is that we see a great decrease in the overall number of users in these top stations in 2020. In addition, there is a great decrease in the number of subscribers. This observation further emphasizes the impact that the restrictions that had begun in NYC had on CitiBike's performance as a whole.

Maps of Start and End Stations (Sheets 10.1-10.2)

Both maps were developed to have quick access to the following information: Station Location, Station Name, Count of Bike ID for each User Type, and the Average Age that accesses each station. As already mentioned, the station that has the largest frequency from Subscribers and Customers is Pershing Square North Station. Subscribers populate the centralized station a lot more frequently than those far away from the centre of the city. And generally, as previously analysed, the largest age group that uses CitiBike are Adults, particularly between 30-40 years old.

Concluding Remarks

1. The Beginning Serge of Covid-19 Virus in New York City did not greatly influence the flux of usage of bikes. There were slight changes from 2019, but not large enough to have indicated that the start of the spread of the virus caused people to stop using the CitiBike Program.
2. When New York City began to respond to the slowly increasing spread with lockdowns and restrictions, that played an important role in the usual flux in the CitiBike Program. It reduced the count of bike usage, number of male subscribers, average trip duration (min), average distance travelled (km), and the flow in the most popular stations.

Resources

(1) https://www.investopedia.com/historical-timeline-of-covid-19-in-new-york-city-5071986