**The Effects of COVID-19 on the Bike-Sharing System of New York City**

Clifford Cele

# Project Description

Citi Bike is New York City’s (NYC) public bicycle sharing system. Created in 2013, the private company has over 20,000 bikes and 1,300 stations available in the Bronx, Brooklyn, Manhattan, and Queens. According to their website: “[the system] consists of a fleet of…bikes that are locked into a network of docking stations throughout the city. The bikes can be unlocked from one station and returned to any other station in the system, making them ideal for one-way trips. People use [the] bike share to commute to work or school, run errands, get to appointments or social engagements, and more.” As the coronavirus disease 2019 (COVID-19) hit the United States (US) in 2020, city dwellers were forced into lockdown between March 2020-June 2020 where all schools were closed, and non-essential workers were required to stay home. During this period and with the rise of social distancing as a measure to mitigate the spread of the disease, people seemed to turn to cycling as a form of mobility and exercise. On the 18th of May 2020, Christina Goldbaum published the New York Times article “Thinking of Buying a Bike? Get Ready for a Very Long Wait” where they discuss the shortage of bicycles available in the US and the desire to be outside amidst the pandemic. Researchers Padmanabhan et al. (2021), studied the effects of COVID-19 on bike sharing systems in three US cities (NYC, Boston, and Chicago) during the first wave (March-June 2020). They concluded that the number of bike trips were negatively impacted by the number of COVID‐19 cases, but trip durations increased. As such, the goal of this project is to explore and visualize how cycling behaviors have changed in NYC due to the COVID-19 pandemic between 2019 and 2021.

# Data Description

Due to data processing constraints, only data between 1 March and 30 June from 2019, 2020, and 2021 are used for this project.

## Citi Bike Trip Histories

The Citi Bike dataset contains 21,765,814 observations across 15 variables. Each observation represents a bike ride, and each variable (in **BOLD**) contains information about the bike ride.

* **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**

According to the Citi Bike website, their data was processed to remove trips taken by staff as they service and inspect the system, trips taken to/from any “test” stations, and trip that were below 60 seconds in length (potentially false starts or users trying to re-dock a bike to ensure its secure) were removed. Due to large size of this dataset, stratified sampling was conducted where 15% all datapoints per year were analyzed.

## Weather

Daily summaries of weather patterns in NYC were obtain from the US National Oceanic and Atmospheric Administration’s (NOAA) National Centers for Environmental Information (NCEI). For this analysis, the variables **1) max air temperature (Fahrenheit), 2) precipitation (inches), and 3) snow depth (inches)** are used. These measurements were taken from the Central Park weather station in Manhattan.

## COVID-19

The NYC Department of Health and Mental Hygiene (NYC Health) maintains a GitHub repository of COVID-19 data for the city. In this analysis, the following daily counts are used: **1) cases, 2) hospitalizations, 3) deaths, and 4) vaccinations** (regardless of whether it is a single dose, first dose, or second dose vaccination). Although Staten Island does not have the Citi Bike system, and therefore using citywide counts would overestimate each variable, data on vaccinations by borough is not available and NYC Health states that the “sum of counts [by borough] may not match values in citywide tables because of records with missing geographic information.” As the primary interest is in the changing trends of these variables, citywide counts were deemed to be sufficient.

# Research Questions

1. **How has the annual usage of the Citi Bike system changed before and during COVID?**

Baby sleep cycle: <https://www.reddit.com/r/dataisbeautiful/comments/5l39mu/my_daughters_sleeping_patterns_for_the_first_4/>

1. **Where are people going and what times of day are popular for using the Citi Bike system throughout the city on a typical workday before and during COVID?**

This question is based on a blog post by Todd W Schneider, where they created a visualization of Citi Bike trips for one full day (Wednesday, 16 September 2015). It was assumed that every trip followed the recommended cycling directions from Google Maps and excluded trips that started and ended at the same station. They illustrated the

1. **What does Citi Bike usage look like in the different boroughs?**
2. **How do Citi Bike trips between Manhattan and the outer boroughs change between 2019, 2020, and 2021?**

This question is based on a blog post by Todd W Schneider, where they created a plot of the average number of weekday Citi Bike trips by hour of day between Manhattan and the outer boroughs. It illustrates that in the morning there are more riders entering Manhattan than leaving, and the reverse occurs in the afternoon, illustrating commuter patterns similar to rush hour (when people are heading to work/school).

1. **Is there a relationship between waves of COVID, the vaccine rollout, and bike usage?**

This question is a continuation of the study by Padmanabhan et al. (2021) who studied the effects of COVID-19 on bike sharing systems in three US cities. The idea is to compare pre-COVID biking trends to the first wave of the pandemic (March-June 2020), and the period where vaccines became widely available to adults in NYC (March-June 2021).

# Visualization Development & Results

## Design Space

### **Exploratory Data Analysis**

### **Design Ideas**

Exploration of design space:

* Provide online references of visuals made by others for the same or similar datasets
* Results of own design exercises

## Final Design

Discuss what each visualization is showing and how to read it

## Implementation

What did you use to build designs d3???

## Insights

## Video & Code

* Links to a 5-min video and code