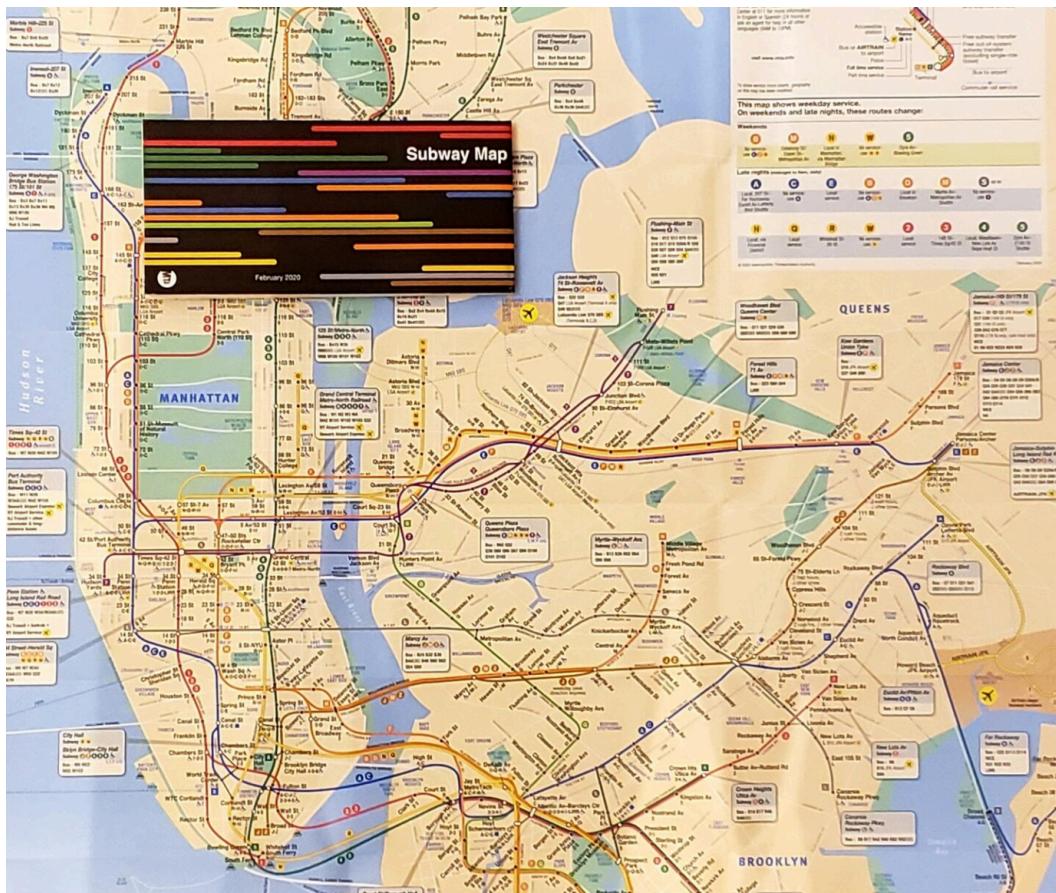


New York City Covid Vaccination Campaign



Project 1 Proposal

Prepared for: Metis Staff Instructors

Prepared by: Rahul D. Raju

June 16, 2021

EXECUTIVE SUMMARY

Objective

The immediate primary concern of most state and local municipalities is to insure a safe, orderly and expeditious return to socio-economic normalcy post-COVID. Much of this return is contingent on local populations achieving what the medical field refers to as “herd immunity.” This status is in turn a function of COVID vaccination rates. According to most medical experts the minimum rate in order to achieve this immunity status is 70%.

New York City currently has partial and full vaccination rates of 53% and 47% respectively, well below the 70% minimum benchmark. Vaccination campaigns have been challenging due to the size of NYC's population and its geographical dispersion over 5 fairly sizable boroughs. Another challenge facing NYC is the nature of its work force. Besides being highly transitory, approximately a quarter of NYC's work force commute from outside the 5 boroughs. This includes states such as New Jersey, Connecticut, Pennsylvania and other parts of New York State.

Solution

A potential option for NYC would be to use its vast subway system to act as an artery for the distribution of the vaccine throughout the body of the city. Subway commuter data could be used to determine which stations could serve as vaccination sites dependent on the amount foot traffic. Specific days and times could also be pinpointed in order to facilitate efficient deployment of resources and maximization of outreach.

Project Data

The primary data set that will be used in this analysis is turnstile data from the New York City Mass Transit Authority(MTA). The data is made freely available via the MTA's developers site, http://web.mta.info/developers/data/nyct/turnstile/turnstile_210612.txt. Each unit/row of data provides daily and hourly(grouped into 4 hour increments) of entry and exit foot traffic for specific stations. In addition to these features, the traffic of specific turnstiles will be analyzed in order to insure efficient placement of resources within the station itself. Additionally, COVID data will be used from OurWorldData.org as well the New York City Health Department's COVID data site, <https://www1.nyc.gov/site/doh/covid/covid-19-data-vaccines.page>. The data obtained from these resources will mostly focus on covid vaccination rates.

Project Tools

- 1). SQL: Primary uses will be the retrieval of raw data, aggregation and storage.
- 2). Pandas: The body of the cleaning and analysis will occur within this framework.
- 3). MatPlotLib & Seaborn: These data visualization modules will be used to highlight trends and insights.

Minimum Viable Product Vision

An MVP for this project would consist of the following:

1. A map of the New York City's subway system with stations of interest highlighted
2. A heat map of Covid vaccination rates across the city.
3. A list of the top subway stations by foot traffic as well other stations of note.
4. A table of Covid vaccination rates across the NY Tri-state area.