# Covid 19 Vaccination Plan 2021

Hernan Trujillo

## Objective

## Analysis purpose:

How to increase the Covid-19 vaccinated population in low-income neighborhoods in NY with a vaccination plan throughout the MTA system.

## Who benefits from exploring this question or building this model/system?

Client: we have been asked to deliver a preliminary analysis for a nonprofit organization called Foundation for Morristown Medical Center, who will improve tremendously its resources effectiveness on organizing the logistic of their "Vaccination points" in NY by identifying the busiest stations by day an hour to get as many vaccinated populations as possible in low-income communities.

The Foundation for Morristown Medical Center (FFMMC) is a standalone foundation, which is a not-for-profit fundraising organization that solicits funds in its general appeal for the benefit and support of Atlantic Health System, Inc. and AHS Hospital Corporation (the Corporation), and all subsidiaries thereof. Morristown Medical Center is the principal recipient of the fundraising carried out by FFMMC.

Low-income population: we will decrease the mortality rate on low-income communities due to Covid-19 pandemic.

## Data

To do this, I took advantage of the free, accessible data about the patterns of transit traffic in New York City: MTA turnstile data. http://web.mta.info/developers/turnstile.html. I considered 11 weeks of the data of March, April and May 2021.

Also, I imported some extra data in order to complete my analysis. Data from Kaggle with socioeconomic data in NY such as Income, borough among others.
www.kaggle.com/muonneutrino/mapping-new-york-city-census-data

Besides to that, during the exploratory process I founded important data related to Covid 19 statistics which was very useful to complete my analysis.
https://www1.nyc.gov/site/doh/covid/covid-19-data.page
https://data.cityofnewyork.us/browse?category=Health&q=covid
.

# Algorithms

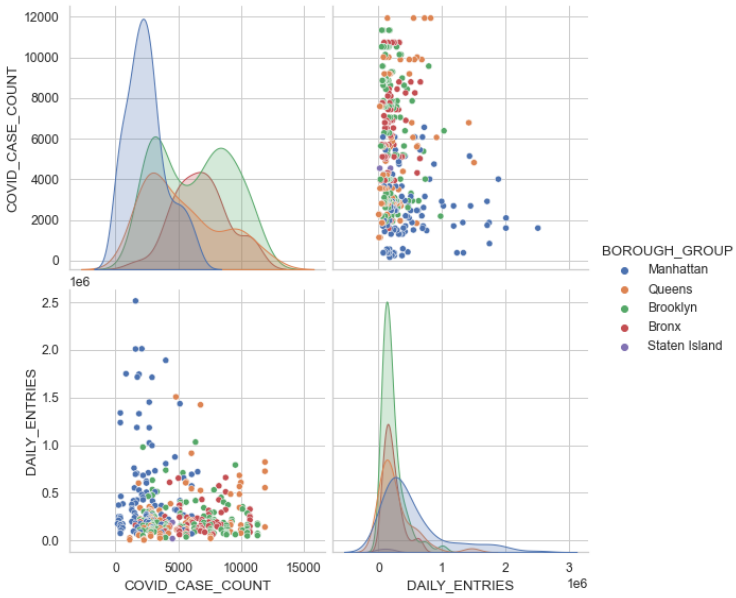
*EDA*

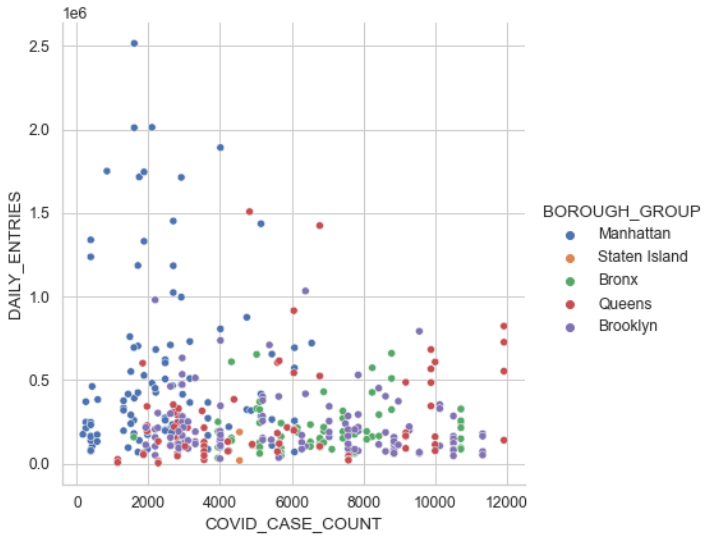
1. Collect and cleaning the data
2. Exploratory data analysis (EDA) using Pandas, SQL
3. Loading data into Pandas
4. Summarizing data by column or by aggregate group
5. Handling missing values
6. Selecting portfolios of data by index or using a mask
7. Manipulating data in Pandas by applying functions, creating columns, etc
8. Sorting, grouping and aggregating
9. Merging multiples dataframes together
10. Joining a dataframe with itself
11. Communicating results with graphics that explain my findings. For plotting analysis, I used Matploblib and Seaborn.

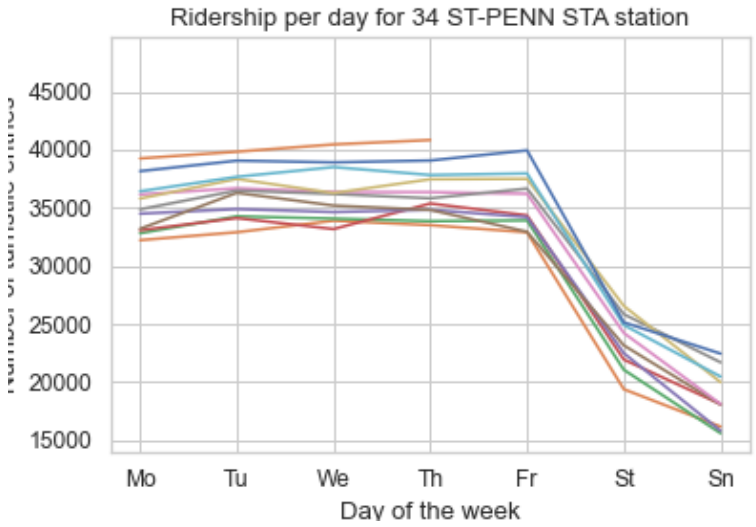
# Tools

- Numpy and Pandas for data manipulation
- Matplotlib and Seaborn for plotting

# Communication


COVID_CASE_COUNT
DAILY_ENTRIES
1e6
BOROUGH_GROUP
Manhattan
Queens
Brooklyn
Bronx
Staten Island


1e6
DAILY_ENTRIES
COVID_CASE_COUNT
BOROUGH_GROUP
Manhattan
Staten Island
Bronx
Queens
Brooklyn


Ridership per day for 34 ST-PENN STA station
45000
40000
35000
30000
25000
20000
15000
Mo
Tu
We
Th
Fr
St
Sn
Day of the week