

The Best NYC Subway Stations to Raise #StopAsianHate Awareness and Provide Outreach

Sandy Weng

Abstract

The goal of this project was to perform EDA on the MTA Turnstiles and population datasets to find the best subway stations for the Stop AAPI Hate organization to set up 5 street teams. After cleaning and exploring the data, I aggregated the data to find the top ten busiest subway stations and the busiest time of the day for the top 5 busiest stations. Using this data, I found the racial demographics by zip code for the top ten busiest stations and made recommendations based on the demographics instead of just foot traffic.

Design

This project originates from Stop Asian Hate organization wanting to raise awareness for violence against the Asian American Pacific Islanders (AAPI) community and provide outreach for those in need. They want to set up 5 street teams in NYC that would have the highest outreach. The data is provided by NY Metropolitan Transit Authority and United States Zip Codes. Analyzing the datasets would enable the Stop Asian Hate Organization decide on which stations they want to allocate their resources and ensure that they can reach as many people as possible.

Data

I am using the most recent six months of the MTA Turnstile dataset as the primary data source, which contains all the MTA subway stations and 5,036,872 turnstile interval counts. The secondary data source is the Population dataset extrapolated from United States Zip Codes website. This dataset contains the subway stations and the racial demographics of the top 10 busiest stations.

Algorithms

I performed an Exploratory Data Analysis of the MTA Turnstile and Population data. I cleaned, explored, aggregated, and created visualizations of the data to present the findings to the organization.

Tools

- SQLite for data ingestion and querying from that database into Python via SQLAlchemy

- Pandas for exploratory data analysis
- Matplotlib and Seaborn for visualizations

Communication

- Presentation of slides and visuals
- Project write up summarizing my work