180 Maiden Lane, 21st Floor
New York, NY 10038
(212) 676-0926

PCR Testing Center

6 October 2021

Overview

NYC health is one of the well-known centers in New York City to offer PCR tests for COVID-19, for the past months the center used to make the tests in their head office, in order to reach more clients and to control the spread of COVID-19 the center decided to establish ten mini examination centers in MTA stations, Thus NYC health is now in a mission to find the most crowded stations in New Your city based on analyzing the MTA stations dataset which will give a better understanding of the movements inside the stations and the persona.

Goals

1- Establish PCR testing centers in the top 10 crowded stations.
2- Help the healthcare providers to make informed decisions.
3- to increase revenue.

Specifications

By analyzing the dataset of MTA stations that contain the movements of people inside the stations from June to August 2021, we can get complete knowledge regarding the best stations to establish the new examination centers. The MTA dataset will be forwarded to the Data Analysis department at NYC Health; it will be analyzed using PYTHON and SQL to produce data required to establish the new ten examination centers.

Data Description:

1. **MTA Dataset:**

   The New York subway MTA turnstile data is a series of data files containing cumulative number of entries and exits by station, turnstile, date and time. Data files are produced weekly. Data records are collected typically every 4 hours with some exceptions.

   Sample size is 3 months from June to August 2021.

   The dataset has 2480677 rows and 11 columns. Column names are:

   - **C/A:** Control Area (A002).
   - **UNIT:** Remote Unit for a station (R051).
   - **SCP:** Subunit Channel Position represents a specific address for a device (02-00-00).
   - **STATION:** Represents the station name the device is located at. **LINENAME:** Represents all train lines that can be boarded at this station.
   - **DIVISION:** Represents the Line originally the station belonged to BMT, IRT, or IND.
   - **DATE:** Represents the date (MM-DD-YY).
   - **TIME:** Represents the time (hh:mm:ss) for a scheduled audit event.
   - **DESC**: Represent the "REGULAR" scheduled audit event (Normally occurs every 4 hours).
   - **ENTRIES:** The cumulative entry register value for a device.
   - **EXITS:** The cumulative exit register value for a device.

Tools:

- **Technologies**: SQLite, python, Jupyter notebook.
- **Libraries**: Numpy, Pandas, Matplotlib, Datetime, Sqlalchemy.