

SCENARIO

NYC MTA stations want to place a first aid station in all of their stations, but due to their tight budget, they have been allowed to place a first aid station in only five stations.

This problem case was presented to us in order to suggest to them which stations there will be a first aid center.

NEED

This analysis will assist in resolving the problem of the New York City subway stations by identifying the top five crowded stations to provide a first aid post in.

QUESTIONS

- Are the top five crowded stations identified?
- Are the stations crowded during the whole week days?
- Is the presence of a first aid center in the selected stations effective?

DATA DESCRIPTION

We need dataset from MTA data for the subways station NYC dataset, we use the features of ENTRIES, EXITS to find out the number of passengers on each trip and In addition to the features of the DATE to ensure that the station operates during the days of the week.

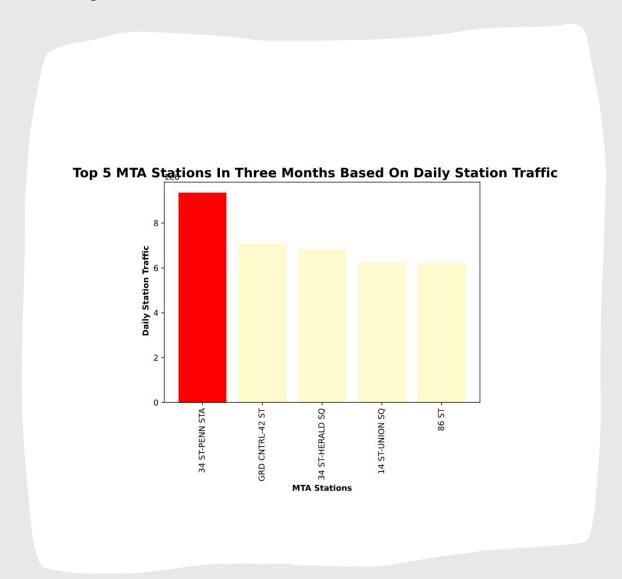
The analysis will be carried out over a period of three months start from 1 May 2021 to 25 September 2021.

TOOLS

Jupyter Notebook, Python, Pandas, Matplotlib, SQLAlchemy and Seaborn

MVP GOAL

The project's goal is to show the top five crowded stations of New York City in order to put a first aid center in it.



To start exploring this goal, I used a Bar Chart model to describe the number of passengers generated over a three-months period at each station.

The figure drawn shows the top 5 stations with number of passengers at each station.

This result indicates that the presence of a first aid center in the busiest stations may have a positive impact on passengers.