Abstract:

The goal of this project is to understand and analyze the dataset provided by the Metropolitan Transportation Authority (MTA), to devise a working plan for installing Wi-Fi network across the Subway stations in New York State. The objective of the MTA is to provide the most convenient experience to the daily commuters; keeping them well connected to the Internet during their usual ride, as this is at the top of the MTA priorities, so the role of the Data Science Team is to pinpoint what are the most congested Subway stations, in order to begin labor work at these stations.

The MTA dataset provides information about the daily Entries and Exits for each station, which helps in segmenting the stations according to high and low traffic.

Data:

The data used in the Exploratory Data Analysis is the MTA Turnstile dataset for the three months spanning from July to September of 2021.

Tools:

- Numpy and Pandas for data manipulation
- SQL for creating a local database with the MTA Dataset
- Matplotlib and Seaborn for plotting

Communication:

Project Deliverables will be communicated to the MTA via PowerPoint Presentation.