MTA Stations Remodeling Project

Abstract

The goal of this project is to redesign gates for the stations. To make the subway tunnels flood-proof by adding sliding doors. To protect the city subway system from flooding to avoid the high coasts of repairs after the damage is done. I worked with data provided by MTA

Design

After the least visit ten stations obtained we started design and constructions to install the sliding doors to it to prevent the station from being damaged by any future increasing of the water level or so

Data

Dataset contains 2094637 turnstile with 11 features for each, feature highlights include measurements of entries and exits, grouped the data by station to find the entries of every single station

Algorithms

Loaded the dataset by pandas and did some start the processing by dropping the duplicates for every these combination ["C/A", "UNIT", "SCP", "STATION", "DATE", “TIME"] , get absolute value of negative entries value to make sure there is no negative values in the entires, then replace any Entry number that higher than 1000000 with 1000000 to simplify the calculation, group by station name and date so we will have for every single station and date list of data, calculate daily entries by getting it’s difference from the previous day assuming that its a cumulative entries and drop first row after difference function because it has Nan, make sure there is not negatives after difference function, find the average daily entries for each station, use filter() to view STATION and ENTRIES only, and finally the plotting

Tools

- Pandas for data manipulation
- Matplotlib for plotting
- Python and Jupyter notebook