Project Proposal

◊ Question/need:

The model we are planning to build using this data is designing a place in the most crowded station where entertaining shows can be performed. The entertaining shows can be used for advertising or for other purposes. The benefits of this model could include many aspects. Such as slowing traffic on metros. Other aspects can be improving the quality time of people waiting. In addition, the advertising shows can increase the profitability of metro stations.

◊ Data Description:

The data has been selected for us, which is the MTA turnstile data. It is obtained from the turnstile unit counting the entries and exits. We have decided to use all metro stations data including all units. In addition, the data we aim to use will be at least three months of data. The data contains eight columns including C/A, unit, and subunit channel position. These three combined represent a unique turnstile. The data include station, date, time, entries, and exits as well. All data will be useful in our case.

◊ Tools:

Technologies:

- Python: Using Jupyter notebook
- SQL
- Tableau

Libraries:

- Pandas
- Matplot
- Seaborn