## Design

Dunkin donuts started a breakfast menu camping to increase sales and expand their specialty beside coffee and doughnuts, The American company requested the data analysis team to give them best locations to open new branches in New York . the MTA (Metropolitan Transportation Authority) data set have been used to achieve this goal and get what are the best locations to open the branches based on number of entries in morning period for each station , and then study the variation for weekend , weekdays and how to maintain stable sales throw out the week .

## Data

Data set used for the analysis is MTA Turnstile from august - September 2021

containing 1.67 million observations with 11 features shown in the table below

| Name | Type |
|---|---|
| **C/A** | Text |
| **UNIT** | Text |
| **SCP** | Adress |
| **Station** | Text |
| **Line name** | Text |
| **Divisions** | Text |
| **TIME** | Time |
| **DATE** | Date |
| **DESC** | Text |
| **EXSIT** | Number |
| **ENTRIES** | Number |

## Algorithm

1- Importing Data and clean it by striping the column names and removing duplicate, adding datetime column
2- Group by and aggregate to get number of Entries for each turnstile with the corresponding time
3- Replacing unrealistic values (null, minus, outliers) with reasonable values
4- Data visualization and graphs that represent the data

## Tools

- Pandas and NumPy for data manipulation
- Matplotlib and seaborn for Visualization and data presentation
- SQLITE to move the data as database