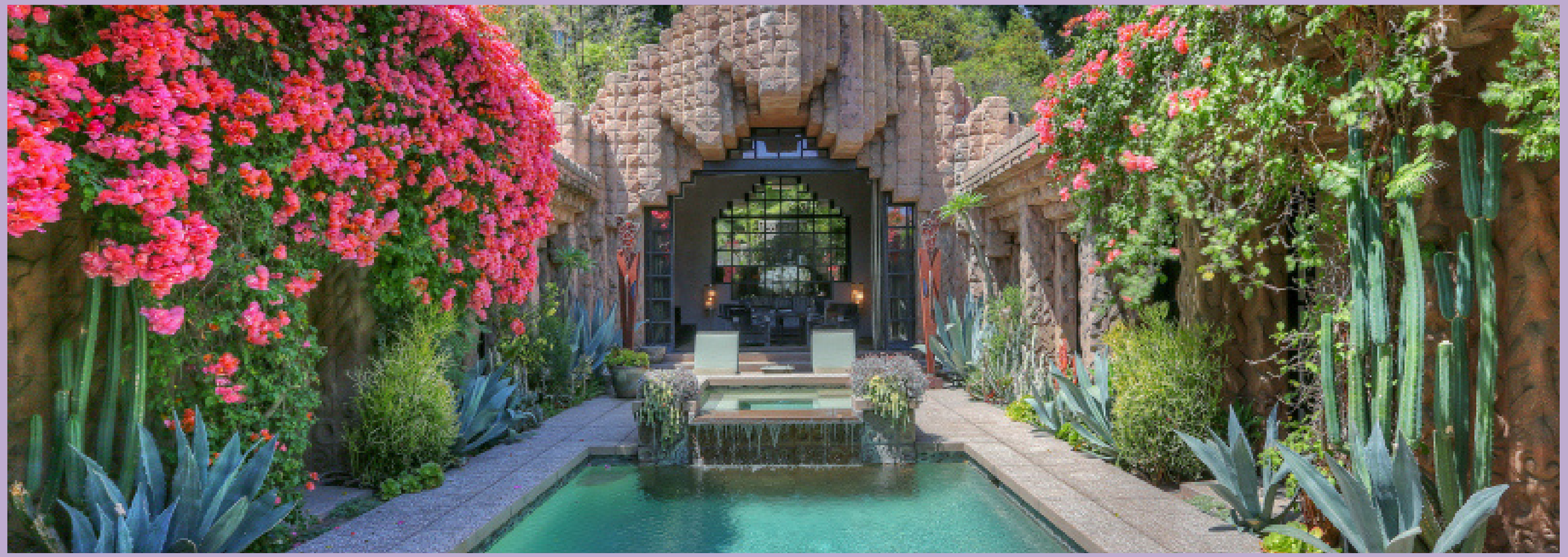
Exploratory data analysis MTA turnstile in NYC



Advertising campaign for new resort in nyc

Back Story - Client Email

Hello

We are Emily and Henry
Marketing Team of Palma Resort Company

A new resort that provides the family with entertainment, relaxation and great times.
We will open the resort on April 1 and will conduct an advertising campaign pre and during first week opening
We have a budget of 5 billboards that we want to distribute in the most crowd subway stations in different divisions where it will cover a wide geographical area
As we said, we will start opening the resort's doors on April 1
The AD shows that the resort is a beautiful destination to spend wonderful time with the family during the spring break

In your opinion, as a data analyst, can you help us with:
determine most crowded stations in different divisions where the billboards will be distributed

The Target :

determine most crowded stations in different divisions where the billboards will be distributed

Data

The dataset contain MTA turnstile data with 3 months worth of data for january ,february and march .

Algorithms

Perform a thorough Exploratory Data Analysis of the MTA turnstile data; clean, explore, aggregate, and visualize the data as appropriate to address the client's needs.

Tools

Numpy and Pandas for data manipulation Matplotlib and Seaborn for plotting , SQLalchemy

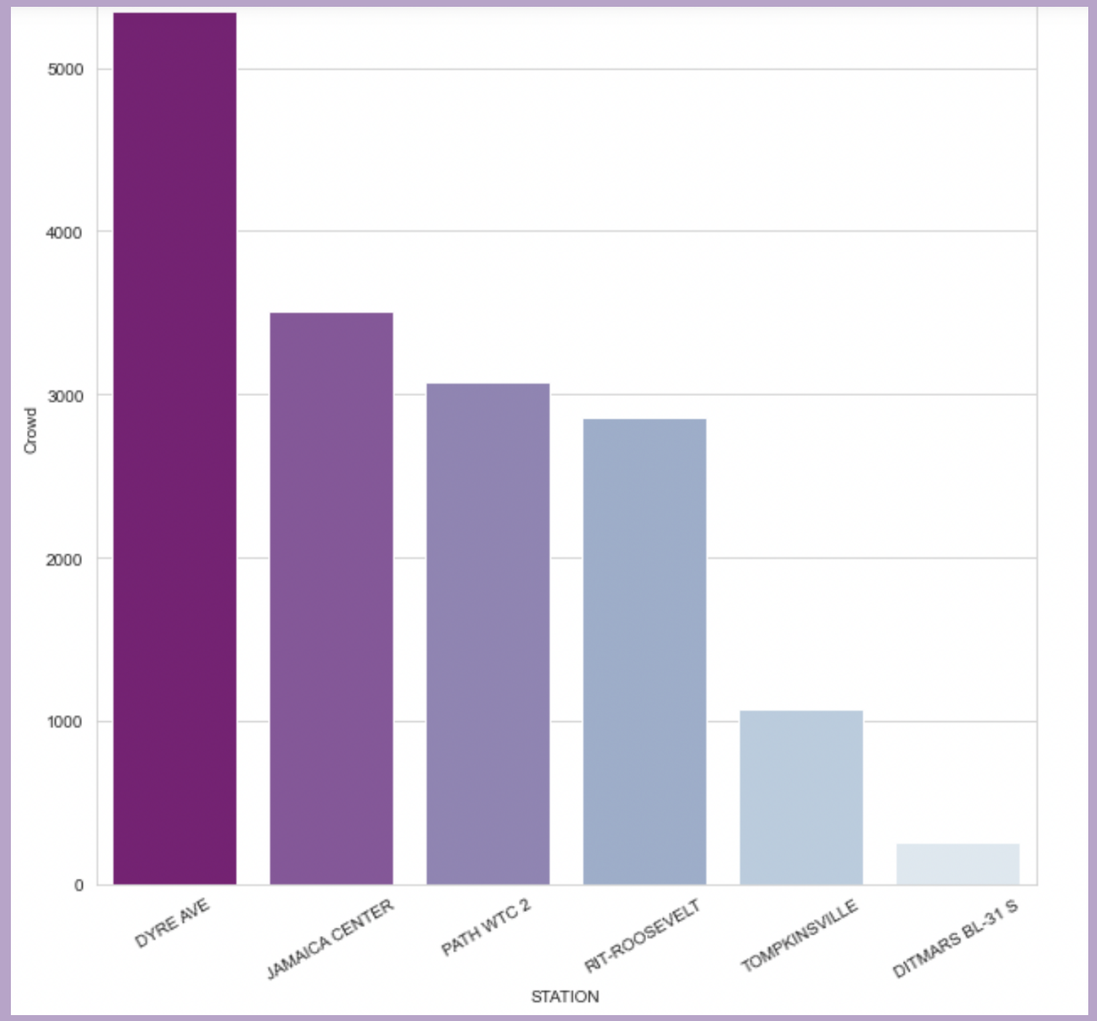
checking crowded stations
by adding daily entries to daily exits

This is most crowd stations but some of it share same division !!!
then i did not reache the goal yet

By using sqlalchem :

STATION	mean	std
1 AVE	1584.208308	24121.989981
103 ST	2365.642734	17614.003373
103 ST-CORONA	2450.142261	12603.745876
104 ST	2502.814128	14441.033563
110 ST	3387.198944	25870.718312

Most Crowded Stations in Unique Divisions



The main point in this exploratory data analysis I looking for most crowded stations in different divisions:
the resulte is :

- **DYRE AVE** from IRT
- **JAMAICA CENTER** from IND
- **PATH WTC 2** from PTH
- **RIT-ROOSEVELT** from RIT
- **TOMPKINSVILLE** from SRT

By Exploratory data analysis i reach my main target and and deliver it to Palma Resort Company

THANK YOU for listening