WHOLE IN 1

Overview:

We at WHOLE IN 1 consulting agency understand the challenges that new business owners face when trying to find the right locations for their business that can guarantee their success and we provide our services by conducting EDA's that can help them solidify their position in the market.

Question/need:

Finding out what are the best stations in the MTA that our client should invest on to start their business based on our predictions from EDA we shall find out the right location. Our client is a new Japanese on-the-*go food kiosk chain Called TakoRyu and they require the counsel of experts to select the right locations for their new business.*

Data Description:

We will use the MTA Dataset and pull the information that we want to analyze, the time period that we will operate on will be from July 2021 to September of 2021 and we will base our analysis on how high is the number of entries in each station, and we will evidently predict the highly utilized stations to recommend for our client.

This a table of the data provided to us by the MTA:

Field Name	Description
C/A	Control Area (A002)
UNIT	Remote Unit for a station (R051)
SCP	Subunit Channel Position represents a specific address for a device (02-00-00)
STATION	Represents the station name the device is located at
LINENAME	Represents all train lines that can be boarded at this station
DIVISION	Represents the Line originally the station belonged to BMT, IRT, or IND
DATE	Represents the date (MM-DD-YY)
TIME	Represents the time (hh:mm:ss) for a scheduled audit event
DESC	Represent the "REGULAR" scheduled audit event (Normally occurs every 4 hours)
ENTRIES	The cumulative entry register value for a device
EXITS	The cumulative exit register value for a device

The number of observations is: ~2514955

The data that we will use will be:

- STATION
- ENTRIES

Tools:

- Python
- SQL

Libraries:

- Pandas
- Matplotlib
- Seaborn
- SQLite