



MTA Turnstile Data Analysis

T5 Data Science Bootcamp
Proposal by: Sarah Alabdulwahab

Tuesday, September 21, 2021

Table of Contents

1.	Introduction.....	3
2.	Data Description	3
3.	Tools	3
4.	MVP Goal	3
5.	References.....	4

1. Introduction

The Metropolitan Transportation Authority (MTA) is North America's largest transportation network, covering a 5,000-square-mile travel area surrounding New York City through Long Island, southeastern New York State, and Connecticut [1]. The MTA is planning on adding more turnstiles to reduce the crowds of people in the busiest stations, therefore, my goal would be to find the best place to add them.

2. Data Description

The MTA Turnstile dataset will be obtained from the MTA website that stores weekly data that contains the following fields:

- **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.

Additionally, I want to create new fields, as follows:

- **ENTRIES_PER_STATION:** Represents the total of entries per station at a certain moment.
- **EXITS_PER_STATION:** Represents the total of exits per station at a certain moment.
- **WEEKDAY:** Represents the name of the day (Sunday, Monday...).

3. Tools

To meet the requirement of the project, the Python libraries I will be using are SQLAlchemy, Pandas, Numpy, Matplotlib, and Seaborn.

4. MVP Goal

The expected outcome is a database containing all the necessary fields, a report of the analysis and its conclusion, and finally, a presentation highlighting all the main points from beginning to end.

5. References

- [1] "The MTA," Metropolitan Transportation Authority, [Online]. Available: <https://new.mta.info/about-us>. [Accessed 20 September 2021].