



TOGO

FOOD TRUCK LOCATOR

SDAIA ACADEMY

T5 Boot Camp

EDA PROJECT REPORT

On: MTA Dataset

Done by: Ranya Alkhtani

Gmail: ranyaalkhtani@gmail.com

GitHub: ranya22

Supervised by: Mr. Mohamed El-Deeb

EDA module

9 September 2021



ABSTRACT

The goal of this project was to make EDA on MTA Dataset in order to solve existing problem or find an opportunity. I worked with Turnstiles Dataset provided by MTA. after exploring the data I come up with an opportunity and an idea of business. I wanted to make prove of concept using MTA dataset. I made assumptions and approve it within the data and some of searching. based on fact that Each Division represents a subway system. in each system there are services provided by that kind of system. I emphasis on that data would be helpful for TOGO business.

DESIGN

TOGO is made up company, the main feature that TOGO offer is telling the food truck owners a production about when, where and which station to go. to increase the sells.

DATA

MTA Turnstile Data is a subway data from New York City (NYC) Metropolitan Transit Authority (MTA). Turnstile is a device that record the number of entries and number exits. The dataset has 11 features and 2722623 record.

The scope of the project is time bounded between June and August at 2021. in total they are 3 month as suggested. all features are included in the EDA.

ALGORITHMS

- Data Collection: using SQLite to import the CSVs into database
- Pre-Processing: Handling duplicate, Handling outliers,
- Feature Engineering: calculate Traffic column.



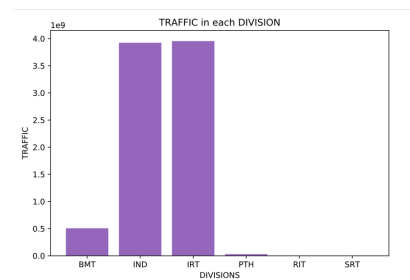
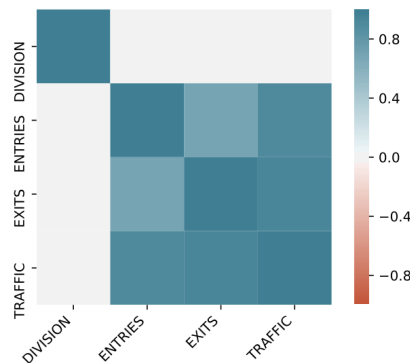
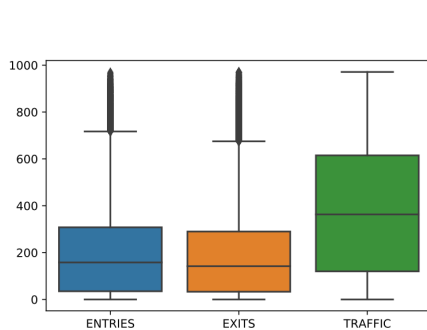
TOOLS

in this part of the report names of potential technologies and libraries are listed.

- Technologies: Python, Jupyter notebook, SQLite, Termina,
- Libraries: pandas, numpy, Matplotlib.

COMMUNICATION

in this part of the report I will show and document all of communication work has been shown both of visualisations and slide show:





COMMUNICATION

TOGO
FOOD TRUCK LOCATOR

9 September 2021

DIVISION DESCRIPTION
Each Division represents a subway system. In each system there are services provided by that kind of system. Below links that will describe each system are listed:

- IRT : A Division

Assumption
• more Traffic

EDA Module 3

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Data Collection

IMPORTING CSVS INTO THE DATABASE
sqlfah -mode csv
sqlfah -import _Turnstile_210605.csv Turnstiles

LINE	STATION	DATE	STATION	DATE	DATE	DATE	DATE	DATE	DATE
1	IRT	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20
2	IRT	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20
3	IRT	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20
4	IRT	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20
5	IRT	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20	10-10-20

EDA Module 4

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Feature Engineering

- New Columns Generated:
• TRAFFIC

INSIGHT
once I took June till August and they are not holidays season, so IRT Division has less traffic around all three months, no surprise that the A Division and B Divisions are much more crowded.

EDA Module 7

TOGO
FOOD TRUCK LOCATOR

9 September 2021

DIVISION DESCRIPTION
Each Division represents a subway system. In each system there are services provided by that kind of system. Below links that will describe each system are listed:

- IND and BMT, B Division

Assumption
• more Traffic

EDA Module 3

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Pre-Processing

- Remove non-regular Descriptions
- Handle Duplicate
- Handle Outlier
- Parse needed Text columns into numeric

EDA Module 5

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Conclusion

- we can give the food truck owner the best location based on the system (Division) used in the nearest station.
- TOGO food truck can increase its sells strongly in the locations with (IND & IRT) divisions.

EDA Module 8

TOGO
FOOD TRUCK LOCATOR

9 September 2021

DIVISION DESCRIPTION
Each Division represents a subway system. In each system there are services provided by that kind of system. Below links that will describe each system are listed:

- IRT :

Assumption
• less Traffic

EDA Module 3

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Feature Engineering

- New Columns Generated:
• DATE_TIME
• TRAFFIC
• Turnstiles_Name

EDA Module 6

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Challenges

- Unclearness in the Dataset.
- Time limitation.
- Resource limitation.

EDA Module 9

TOGO
FOOD TRUCK LOCATOR

9 September 2021

Reference

- A and B Division https://metro.fandom.com/wiki/List_of_New_York_City_Subway_inter-division_connections
- SRT : https://en.wikipedia.org/wiki/Staten_Island_Railway
- PTH : [https://en.wikipedia.org/wiki/23rd_Street_station_\(PATH\)](https://en.wikipedia.org/wiki/23rd_Street_station_(PATH))
- IND and BMT: [https://en.wikipedia.org/wiki/B_Division_\(New_York_City_Subway\)](https://en.wikipedia.org/wiki/B_Division_(New_York_City_Subway))
- IRT : [https://en.wikipedia.org/wiki/A_Division_\(New_York_City_Subway\)](https://en.wikipedia.org/wiki/A_Division_(New_York_City_Subway))
- RIT : <https://freetoursbyfoot.com/roosevelt-island-tramway/>

EDA Module 10

TOGO
FOOD TRUCK LOCATOR

9 September 2021

THANK YOU FOR LISTENING

EDA Module