



أكاديمية سدايا
SDAIA Academy

EDA OF MTA TURNSTILE DATA

Prepared by: Asma Althakafi

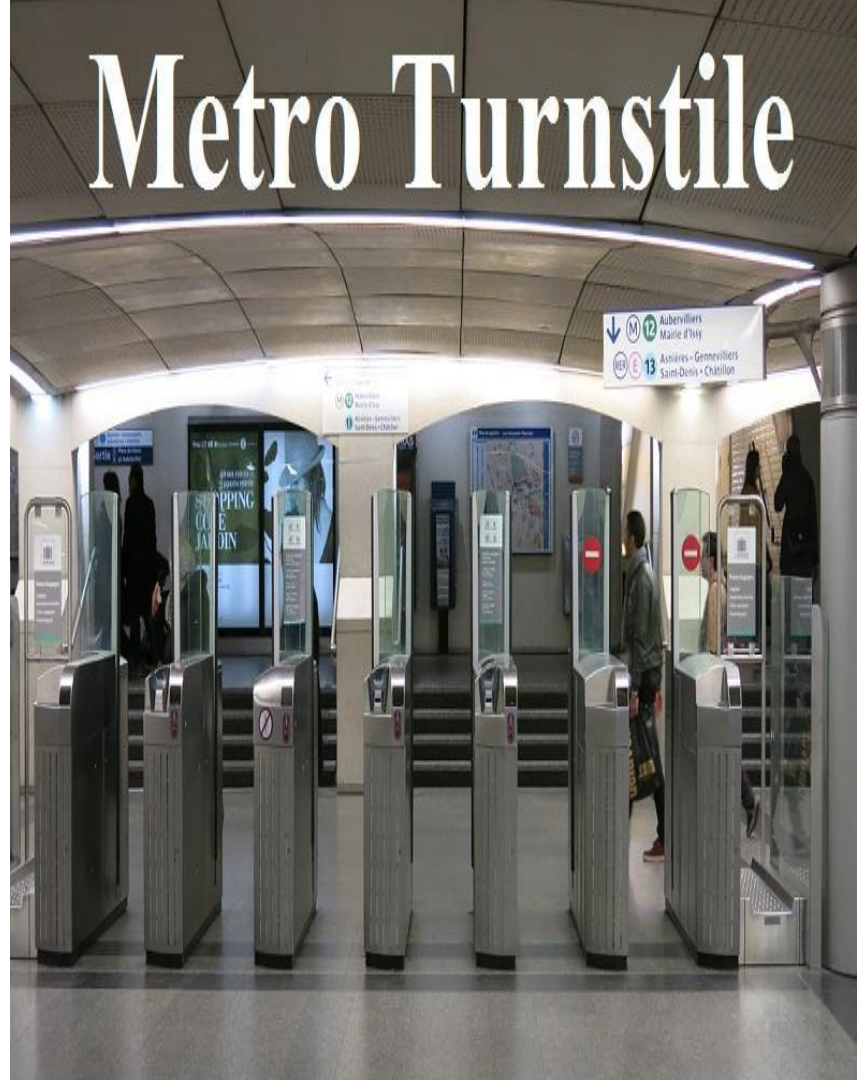
PROBLEM

Café chain wants to open a new branches in New York



SOLUTION

Analyze the MTA turnstile data to see which stations have a lot of people



MTA TURNSTILE DATA

- 3 months from 2021: June, July, August
- Contains 2,753,240 entries with 11 features.
- The main features: Station, Date, Time, Entries, Exits

PROCESSING STEPS

1

**STORING DATA
IN A DATABASE**

2

**CLEANING AND
PREPROCESSING**

3

**ANALYZING AND
VISUALIZATION**

I- STORING DATA IN A DATABASE

SQLite database

- Created database using SQLAlchemy in Python
- Read files and stored them to database

2- CLEANING AND PREPROCESSING

FEATURE ENGINEERING

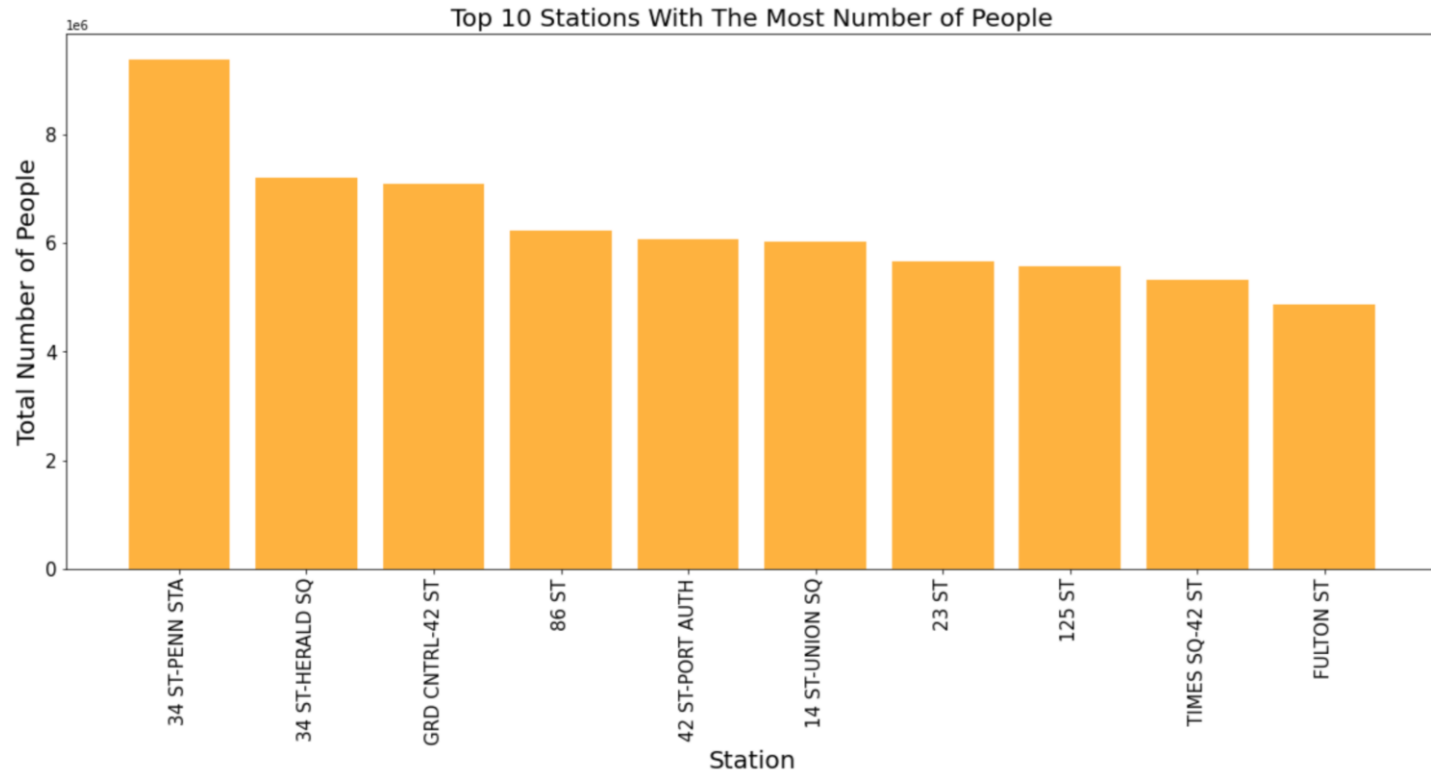
date-time, daily entries, daily exits, sum of daily entries and exits

REMOVE NULL VALUES

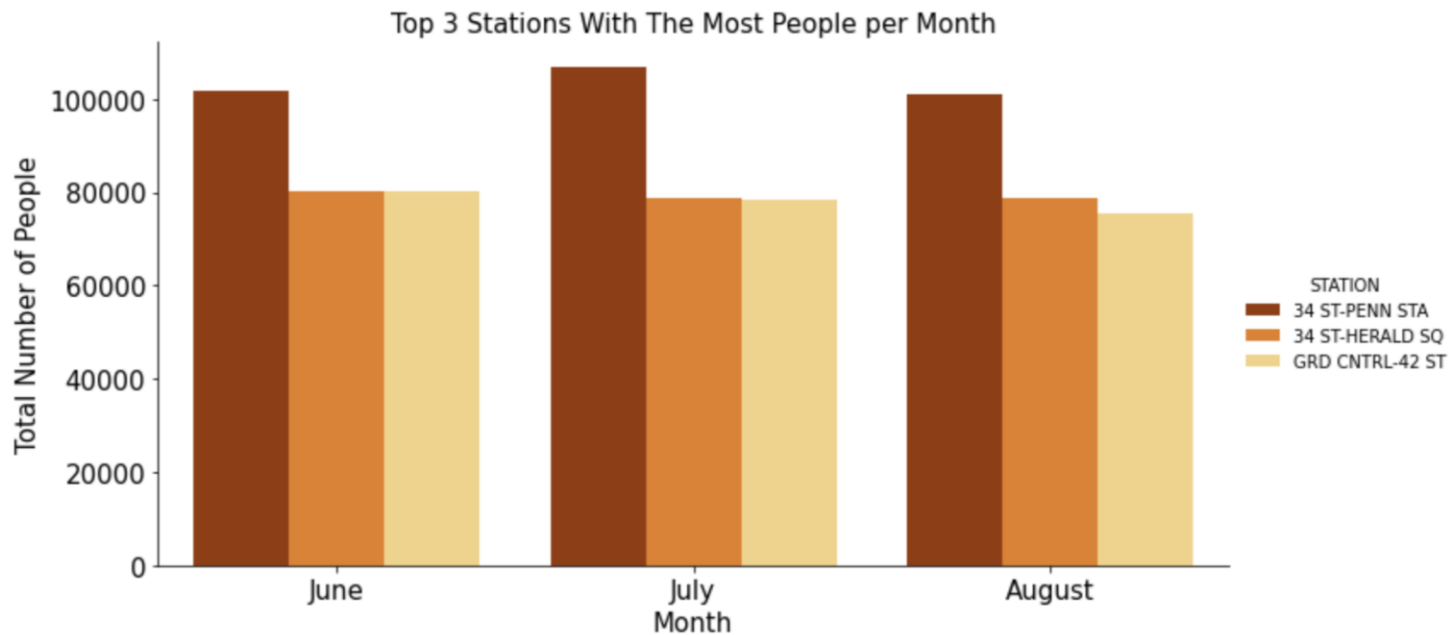
REMOVE DUPLICATE ROWS

3- ANALYZING AND VISUALIZATION

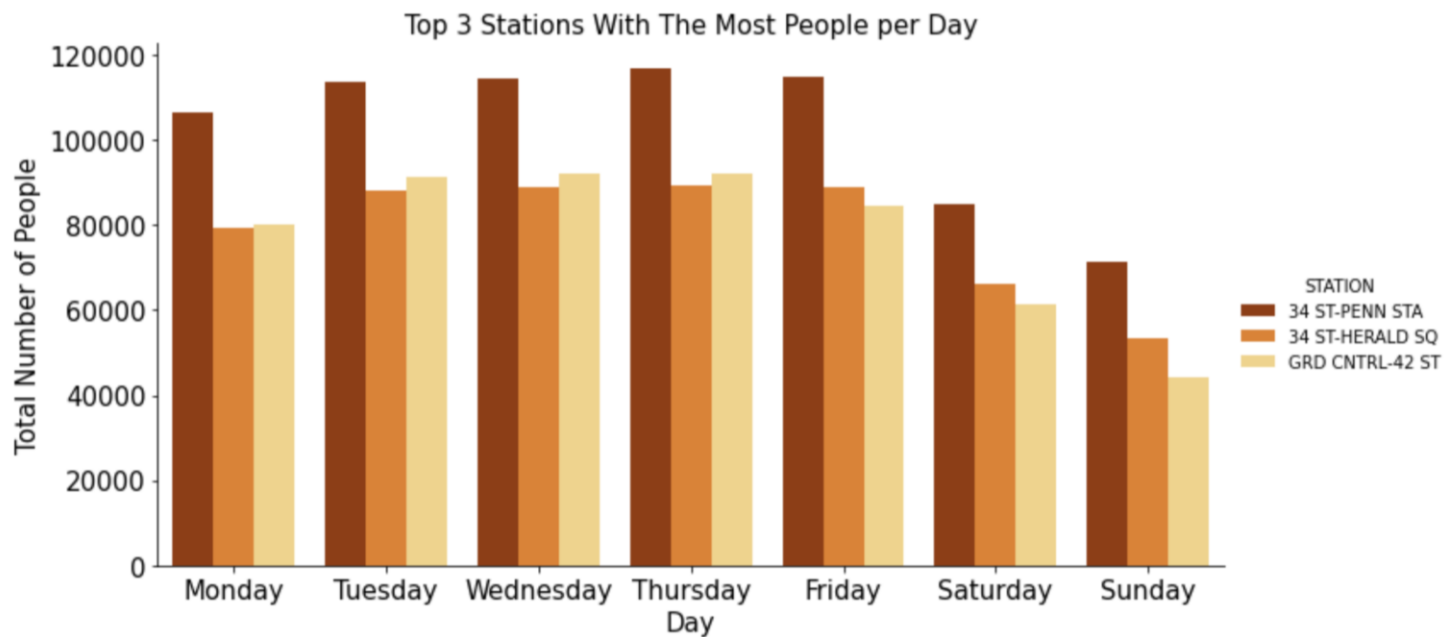
TOP 10 STATIONS WITH THE MOST NUMBER OF PEOPLE



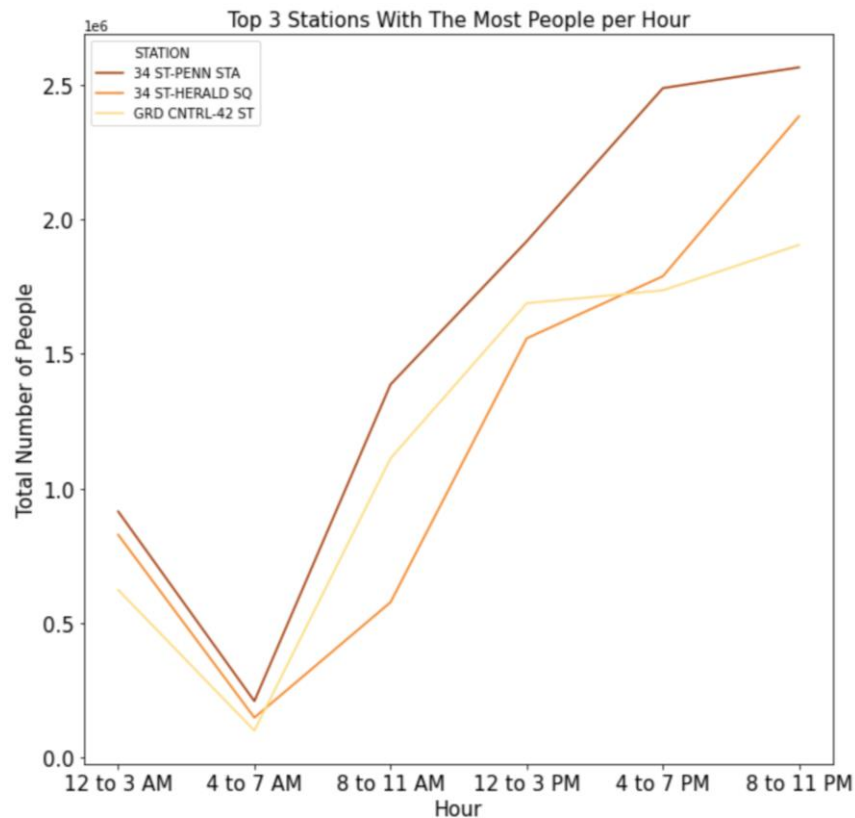
TOP 3 STATIONS WITH THE MOST PEOPLE PER MONTH



TOP 3 STATIONS WITH THE MOST PEOPLE PER DAY



TOP 3 STATIONS WITH THE MOST PEOPLE PER HOUR



DASHBOARD

Exploratory Data Analysis of MTA Turnstile Data Dashboards

Total Number of People

All Stations

Top 3 Stations

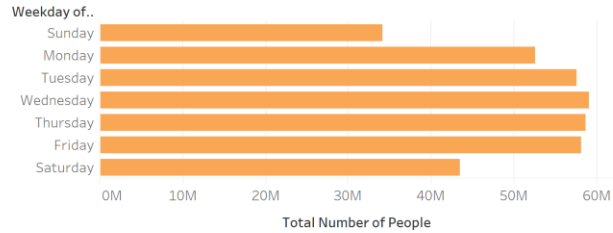
Total Number of People

Total Number of People per Month

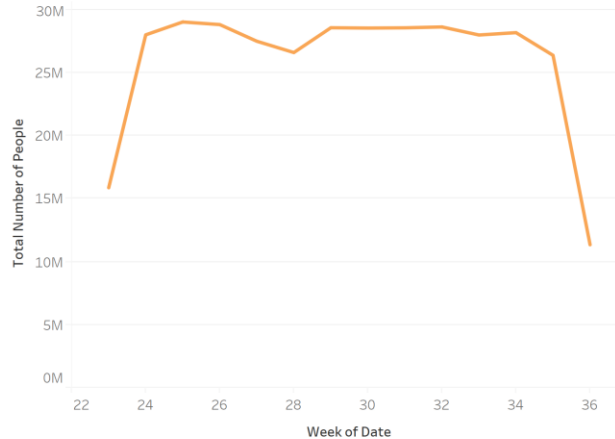


Month
June
July
August

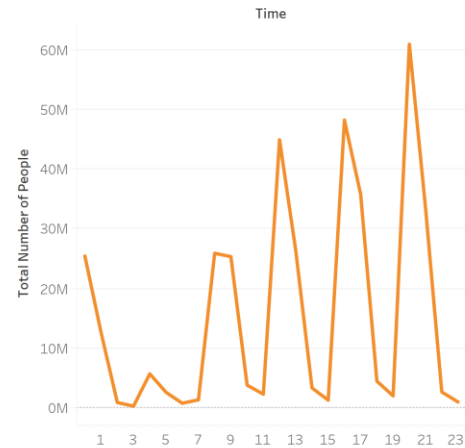
Stations With Total Number of People per Day



Total Number of People per Week



Total Number of People per Hour



Exploratory Data Analysis of MTA Turnstile Data Dashboards

Total Number of
People

All Stations

Top 3 Stations

All Stations

Station

All

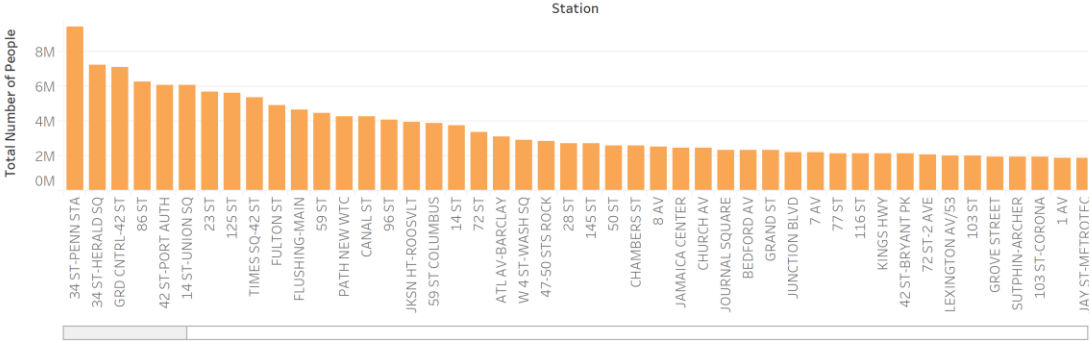
Month

All

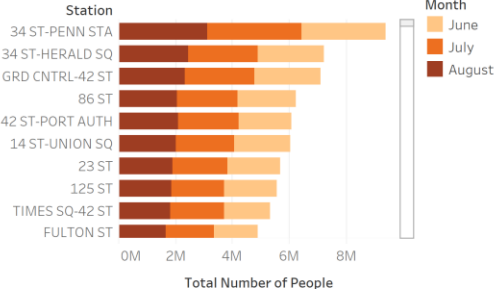
Day

All

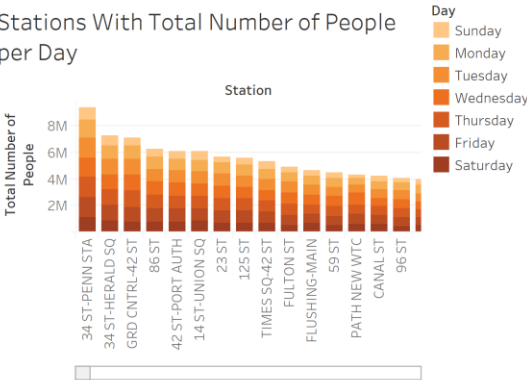
Stations With Total Number of People



Stations With Total Number of People
per Month



Stations With Total Number of People
per Day



Exploratory Data Analysis of MTA Turnstile Data Dashboards

Total Number of
People

All Stations

Top 3 Stations

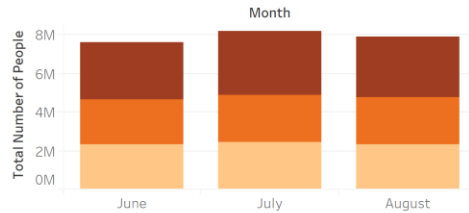
Top 3 Stations

Station ■ 34 ST-PENN STA ■ 34 ST-HERALD SQ ■ GRD CNTRL-42 ST

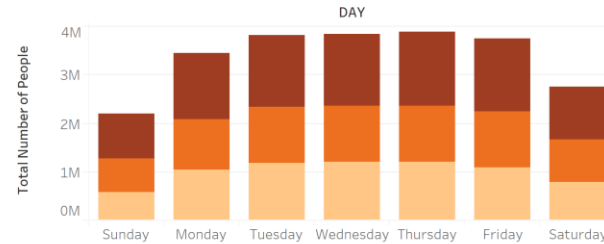
Top 3 Stations With The Most People



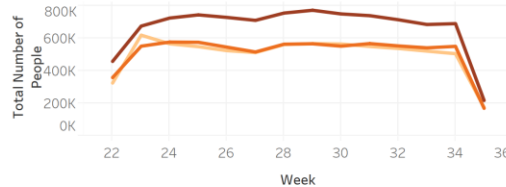
Top 3 Stations With The Most People per Month



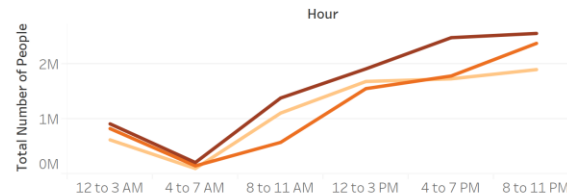
Top 3 Stations With The Most People per Day



Top 3 Stations With The Most People per Week



Top 3 Stations The Most People per Hour



CONCLUSION

The top 3 stations

- 34 ST-PENN STA, 34 ST-HERALD SQ, GRD CNTRL-42 ST.

Recommendation

- Open a new branch near to 34 ST-PENN STA.
- Day of opening on Thursday.
- Employed more employees between time 4 to 11 PM.

**THANK YOU FOR
LISTENING**