أكاديمية سدايا
SDAIA Academy

MTA Exploratory Data Analysis: Subway Gathering Limitation Recommendations

Mohammed Osailan

Outlines

DATABASE

DATA CLEANING

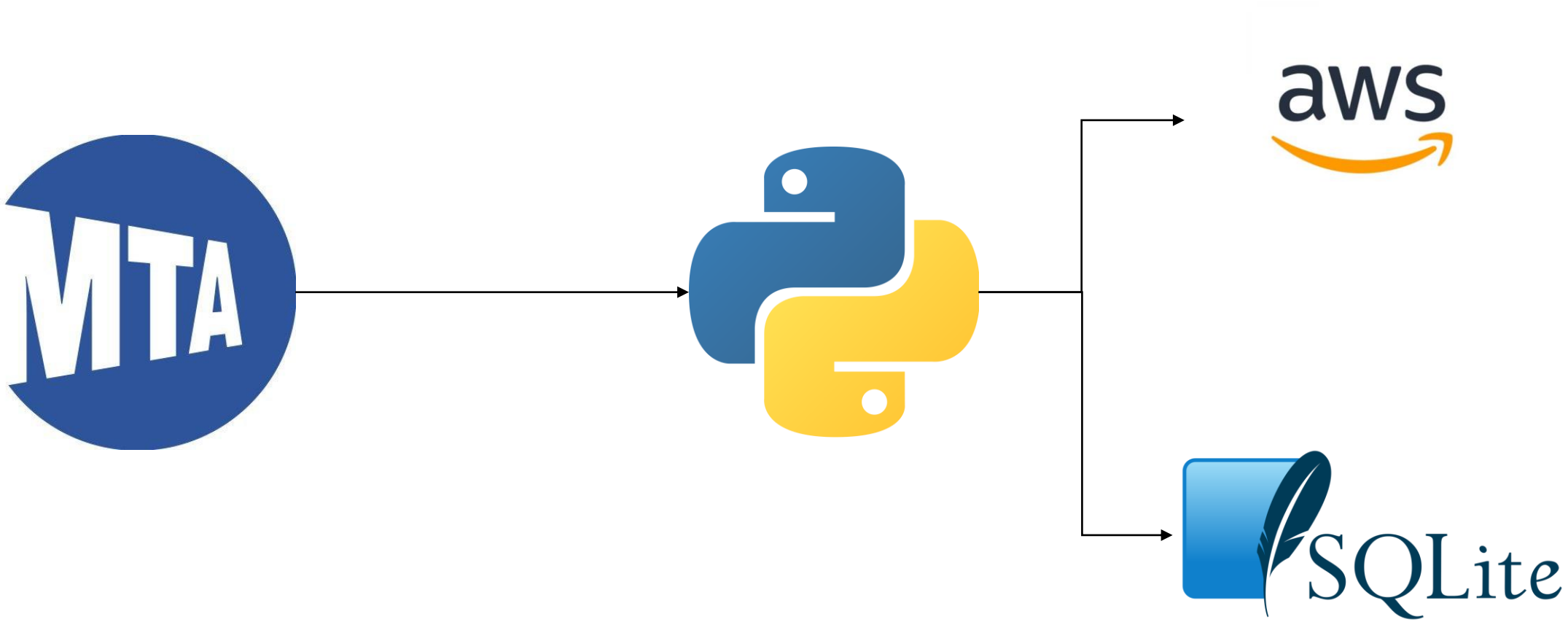
EXPLORATORY DATA ANALYSIS

CONCLUSIONS

Database

- First download the data:
- Two PostgreSQL databases were created. In a cloud using AWS. And locally using SQLite.

Database



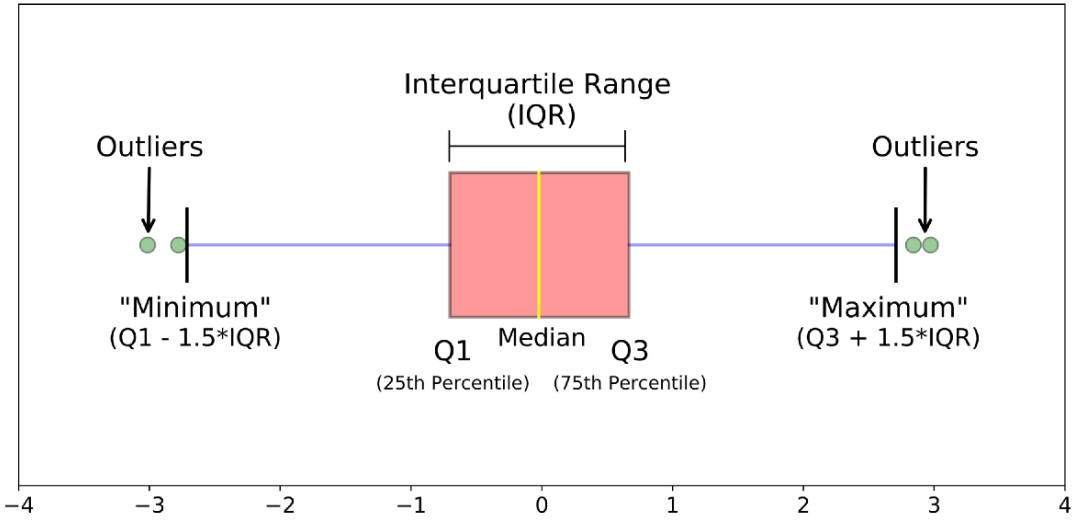
Database

- Data can be accessed in python by creating an **engine** (**create_engine**).

- Then SQL tasks can be performed using **pandas** (**read_sql**)

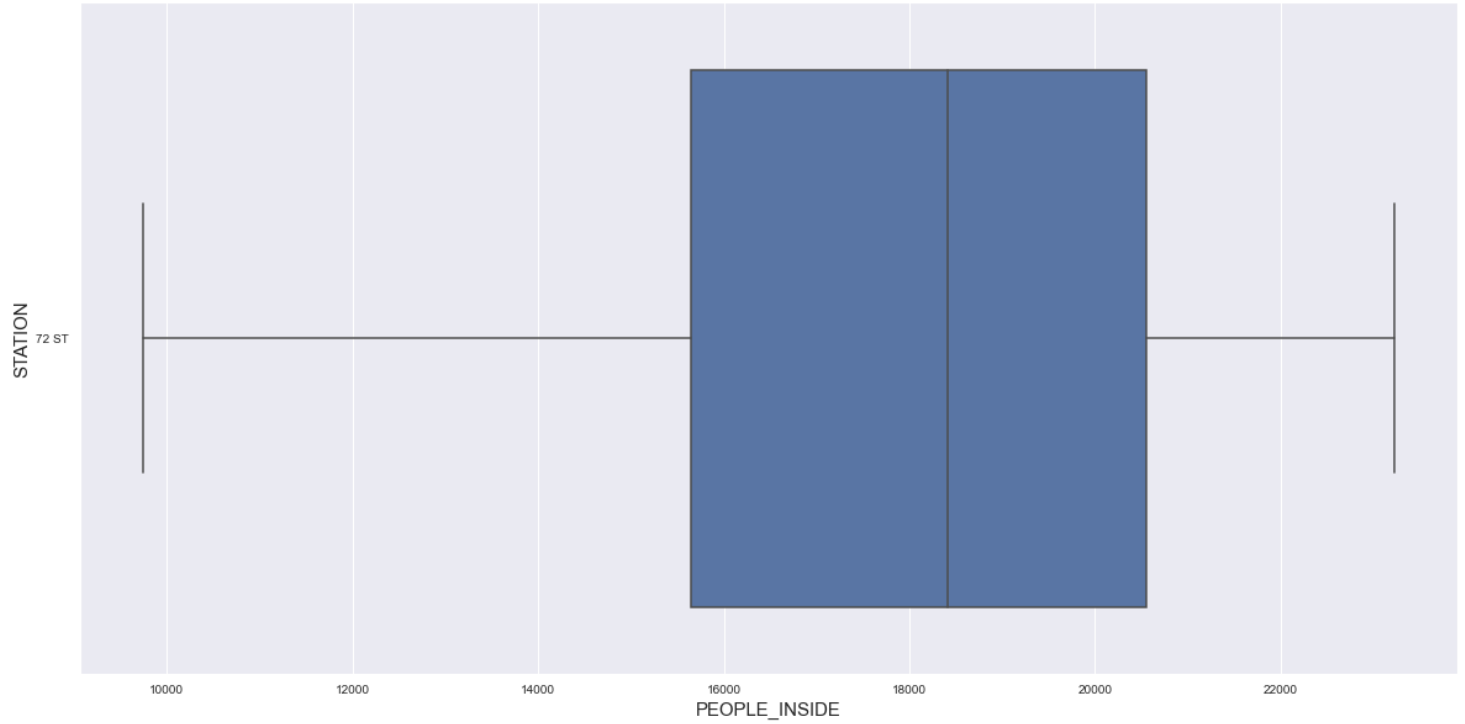
Data Cleaning

- Columns names.
- Dates type conversion.
- Dealing with negative values.
- Dropping outliers.



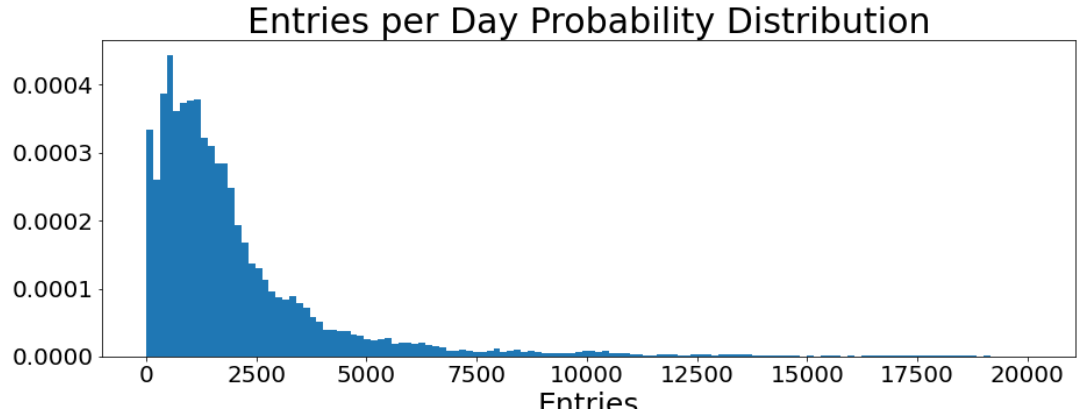
Data Cleaning

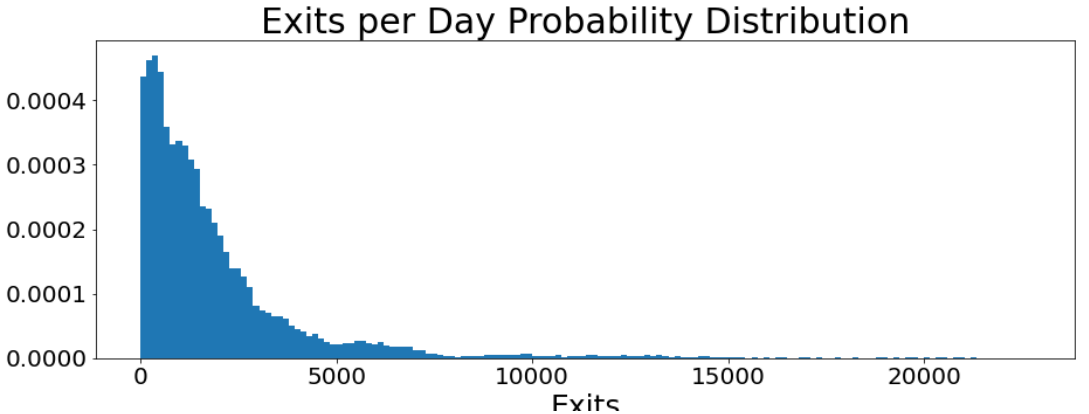
- Station ‘72 ST'

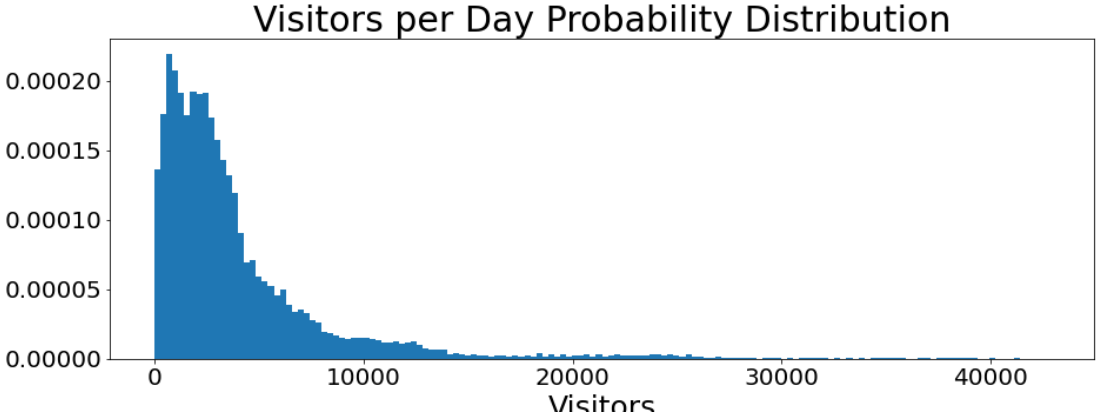


Exploratory Data Analysis

- Probability distribution of daily Exits and Entries data.

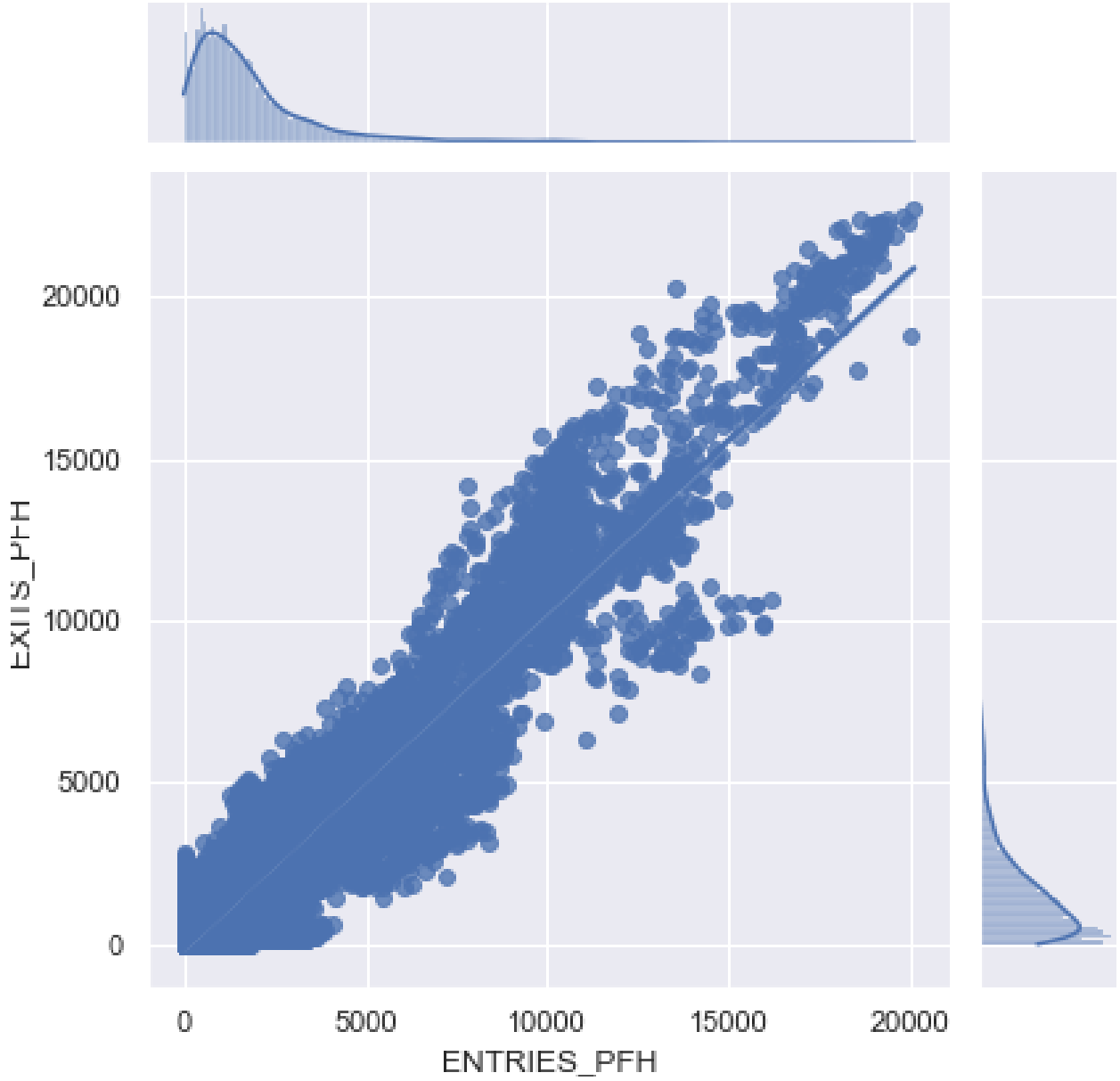






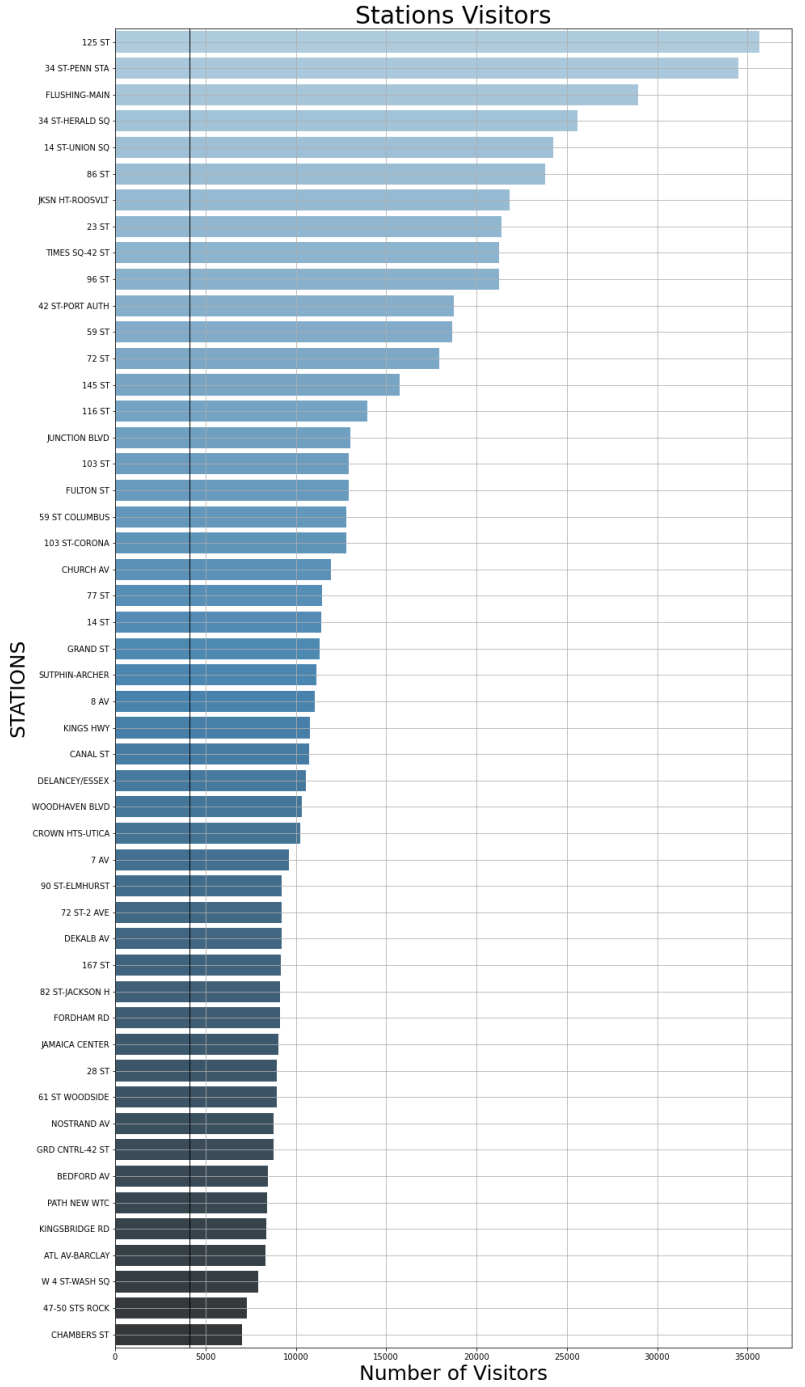
Exploratory Data Analysis

- Joint plot of daily Exits and Entries data.



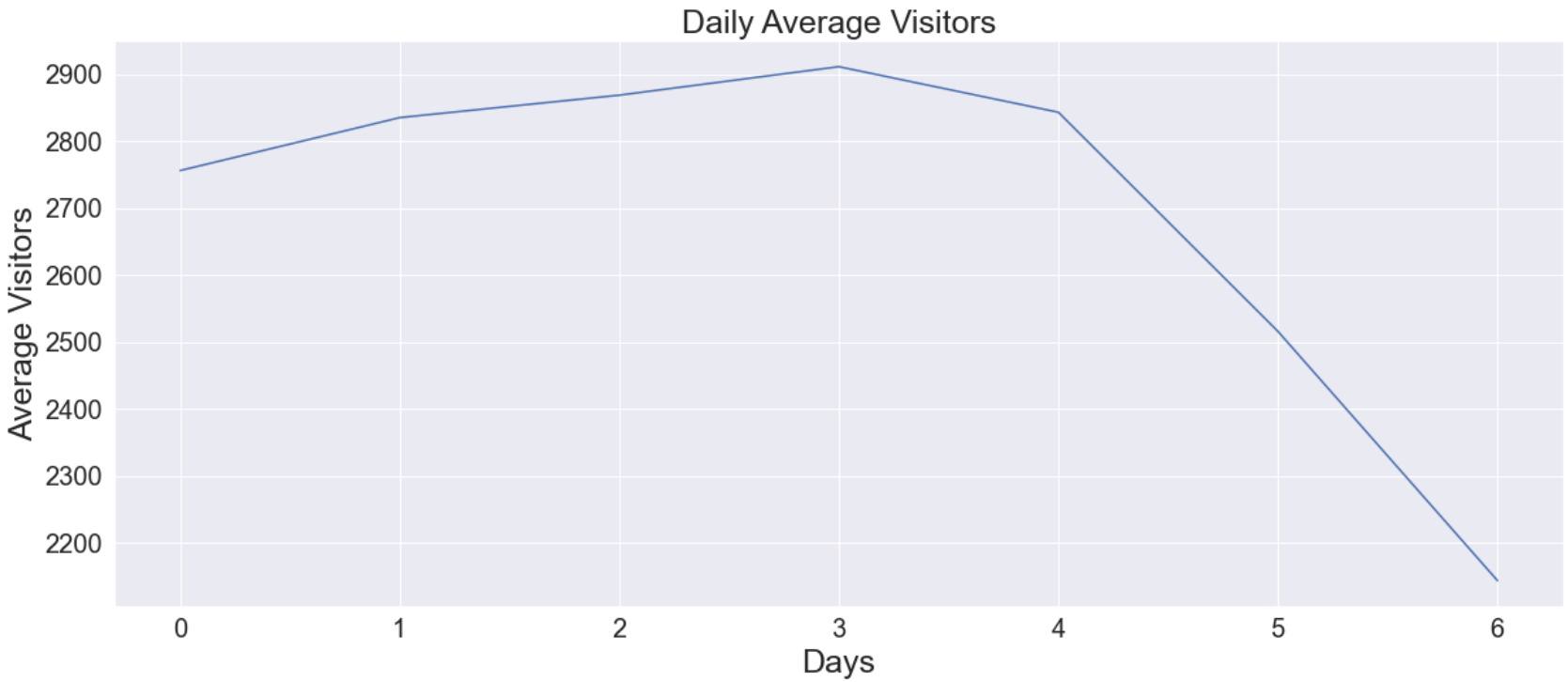
Exploratory Data Analysis

- Most 50 crowded stations.



Exploratory Data Analysis

- Average visitors per day.



Conclusion

- Focus on the 7 most crowded stations.
- Apply a strict gathering limitations regulations.
- Ease the entrance, exits processes.

Thank you