T5 - SDAIA

EDA MTA DATA ANALYSIS

SAFERIDE SOLUTION

Work by: Muntaha Jaber

Instructor: Dmitry Denisov

T5 - SDAIA

CONTENTS

Project Overview
Problem statement
Solution
Data workflow
Results
Conclusion
Recommendation

PROJECT OVERVIEW

The easyTech company is working on strategies to prevent congestion at New York railway stations during the Covid 19. The company is building an application and platform (SafeRide) that will connect visitors, workers, and others to the station in order to minimize congestion. Visitors should be aware that one of the station's entrances will be quite congested at a certain period of time and date and suggest to them other stations and entrances based on their tickets and destination. Also, to manage social separation and Adhere to precautionary measures, staff will be able to distribute themselves based on the number of projected traffic jams at each station entrance.

PROBLEM STATEMENT

- Which stations are the busiest?
- What days of the week and times of day have the most traffic?
- How can visitors to the station avoid congestions?
- How can station employees effectively manage and monitor social distancing?

T5 - SDAIA

SOLUTION

- Analyzing MTA data and highlighting the busiest stations, days, and times. As a result, station employees may have a better understanding of the most common congestion.

- Using the daily congestion summary, station visitors will be notified of severe congestion based on their tickets and destination.

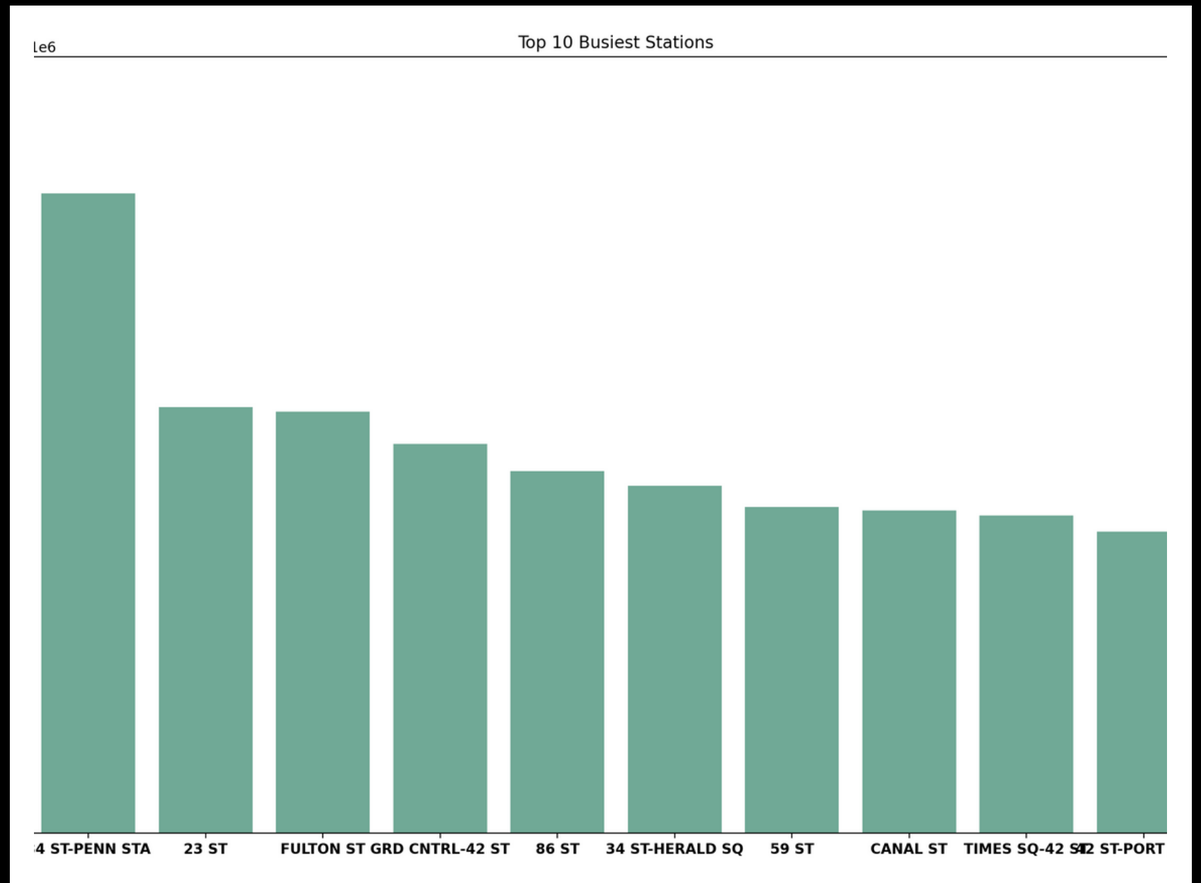
T5 - SDAIA

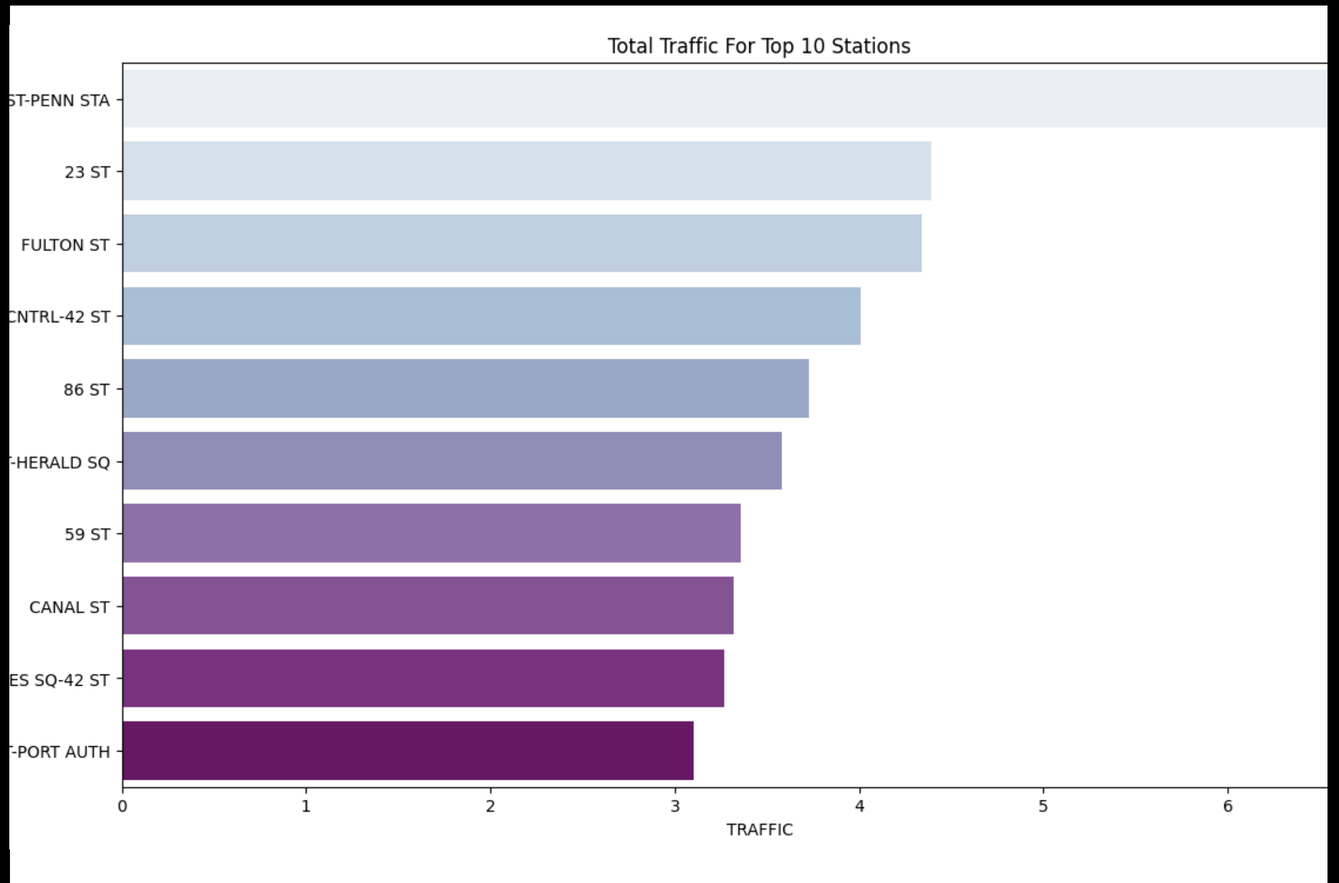
DATA WORKFLOW

- AWS Database creation - Mysql
- Database connection (user name - port - password - endpoint)
- Google colab - Python
- Data from Jun 2021 - Sep 2021
- Data cleaning
- Data upload to database - to_sql()

RESULTS #1

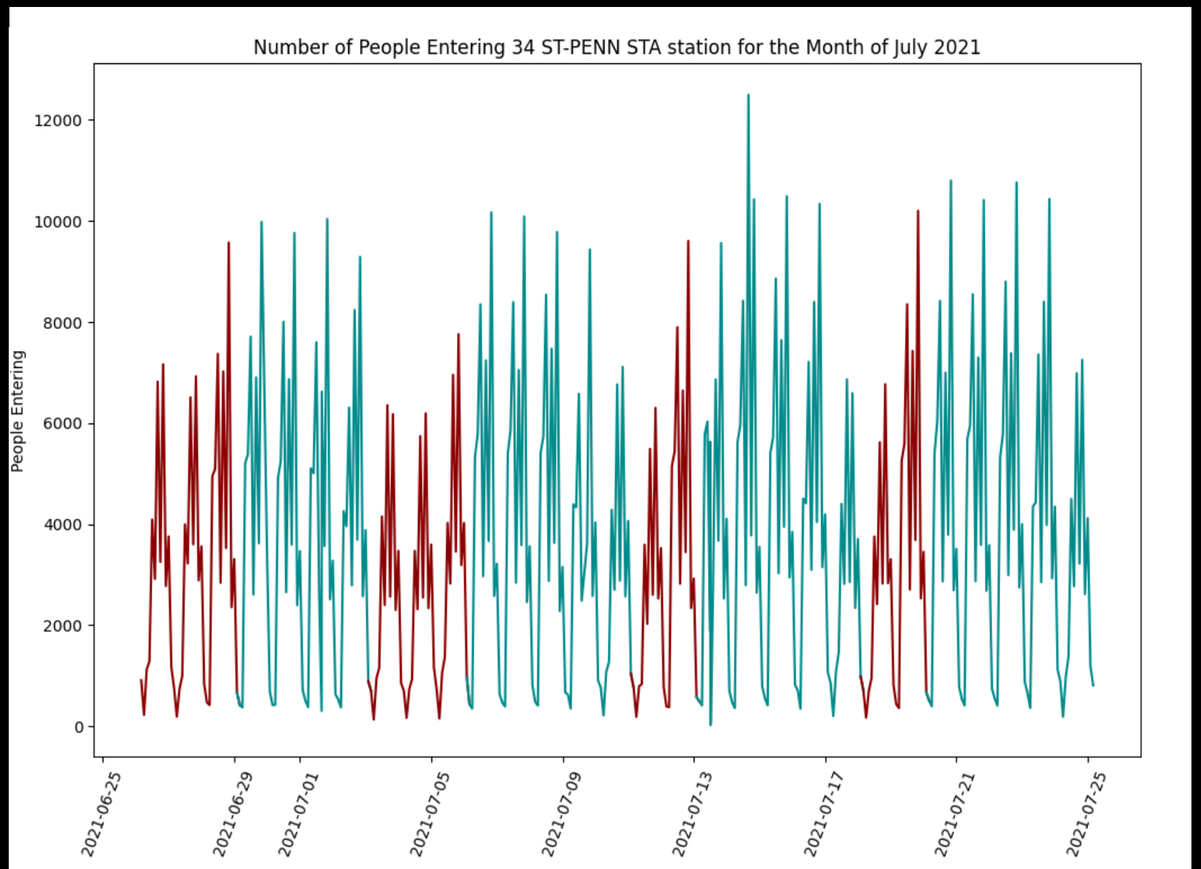
- Graph 1: Top busiest stations based on total entries per station
- Graph 2: Total Traffic For Top 10 Stations (entries + exits)

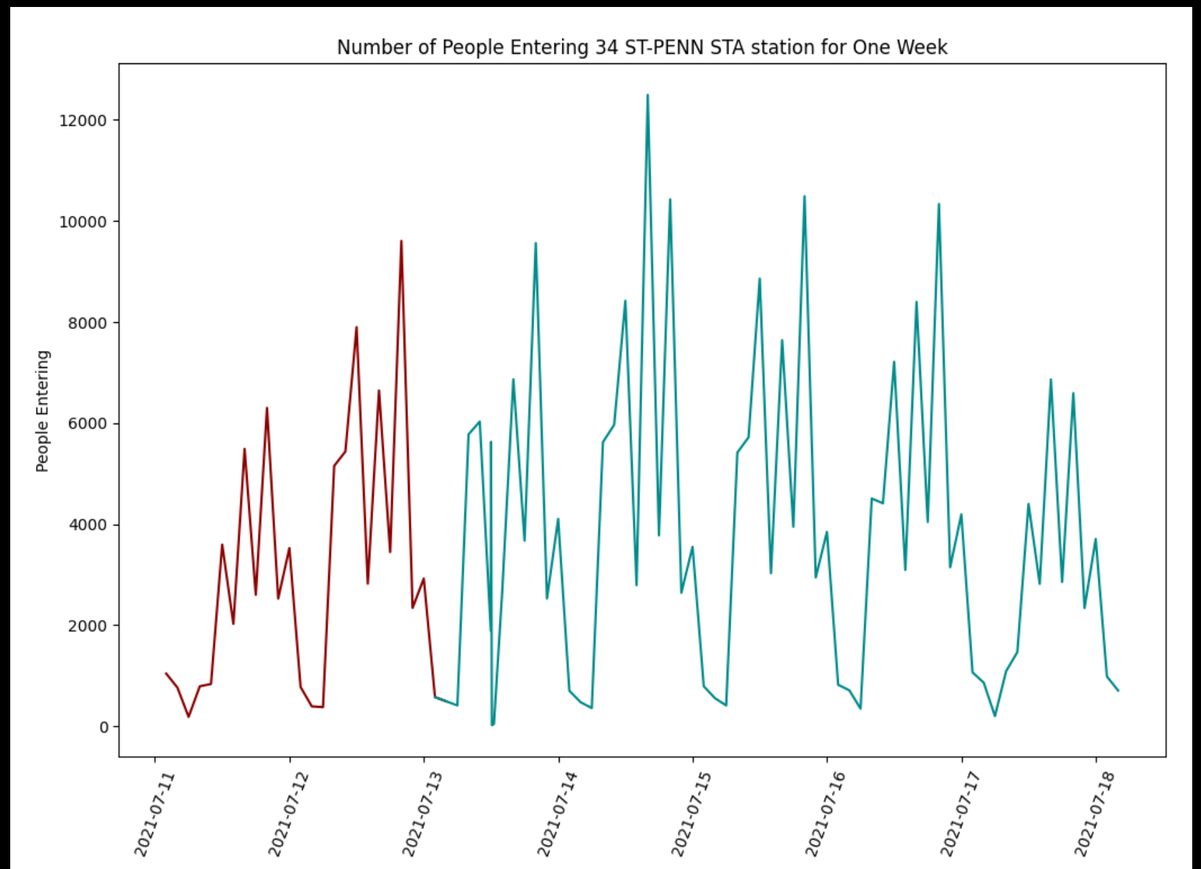




RESULTS #2

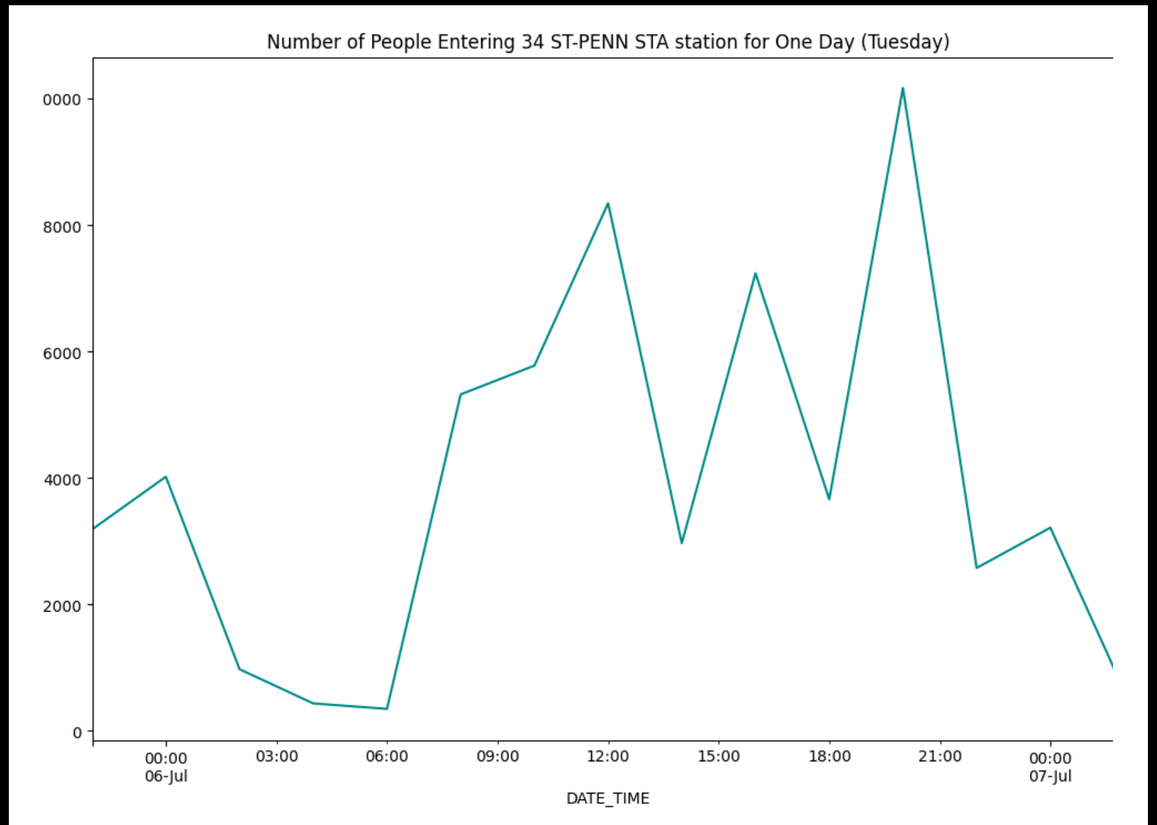
- Graph 1: Number of People Entering 34 ST-PENN STA station for the Month of July 2021
- Graph 2: Number of People Entering 34 ST-PENN STA station for One Week

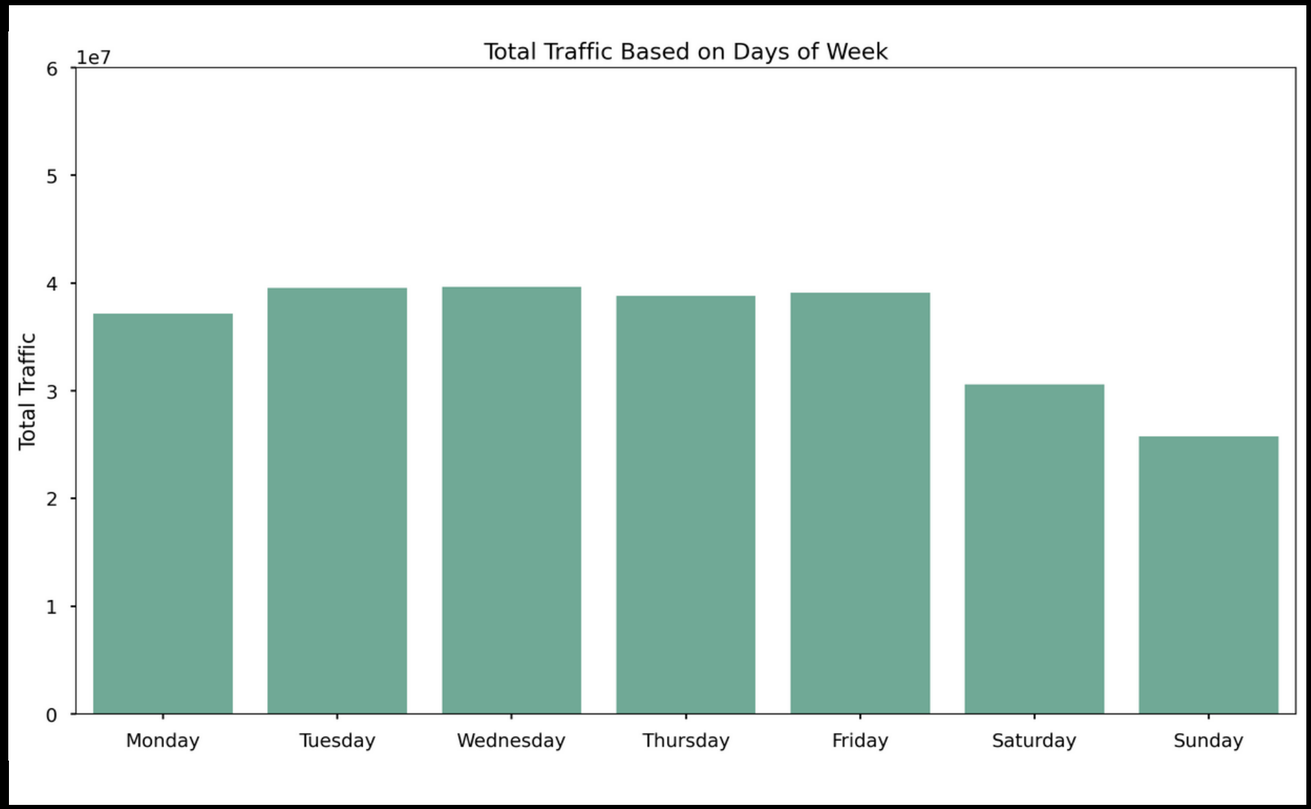




RESULTS #3

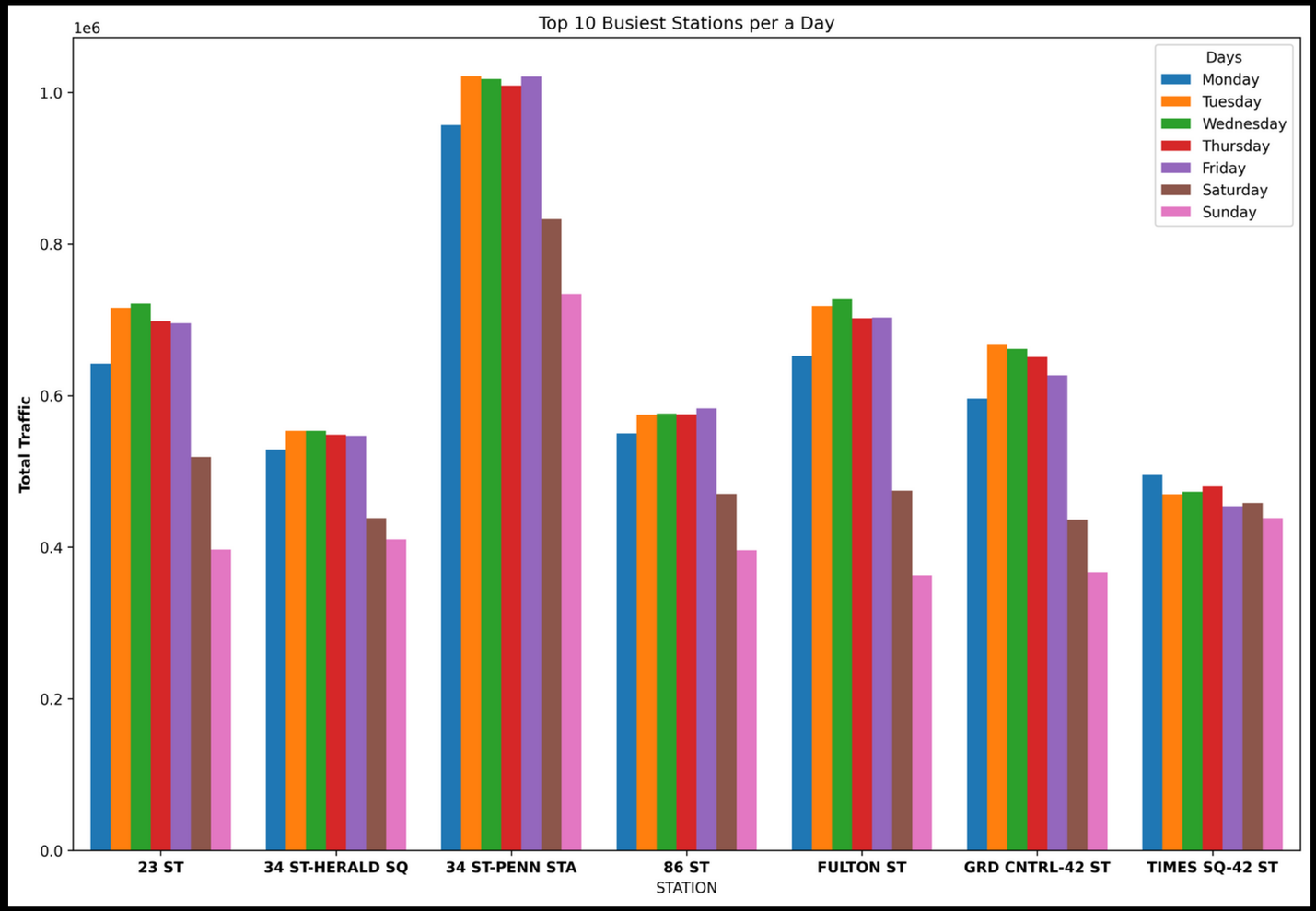
- Graph 1: Number of People Entering 34 ST-PENN STA station daily
- Graph 2: Total Traffic Based on Days





RESULTS #4

- Graph 1: stations traffic based on time

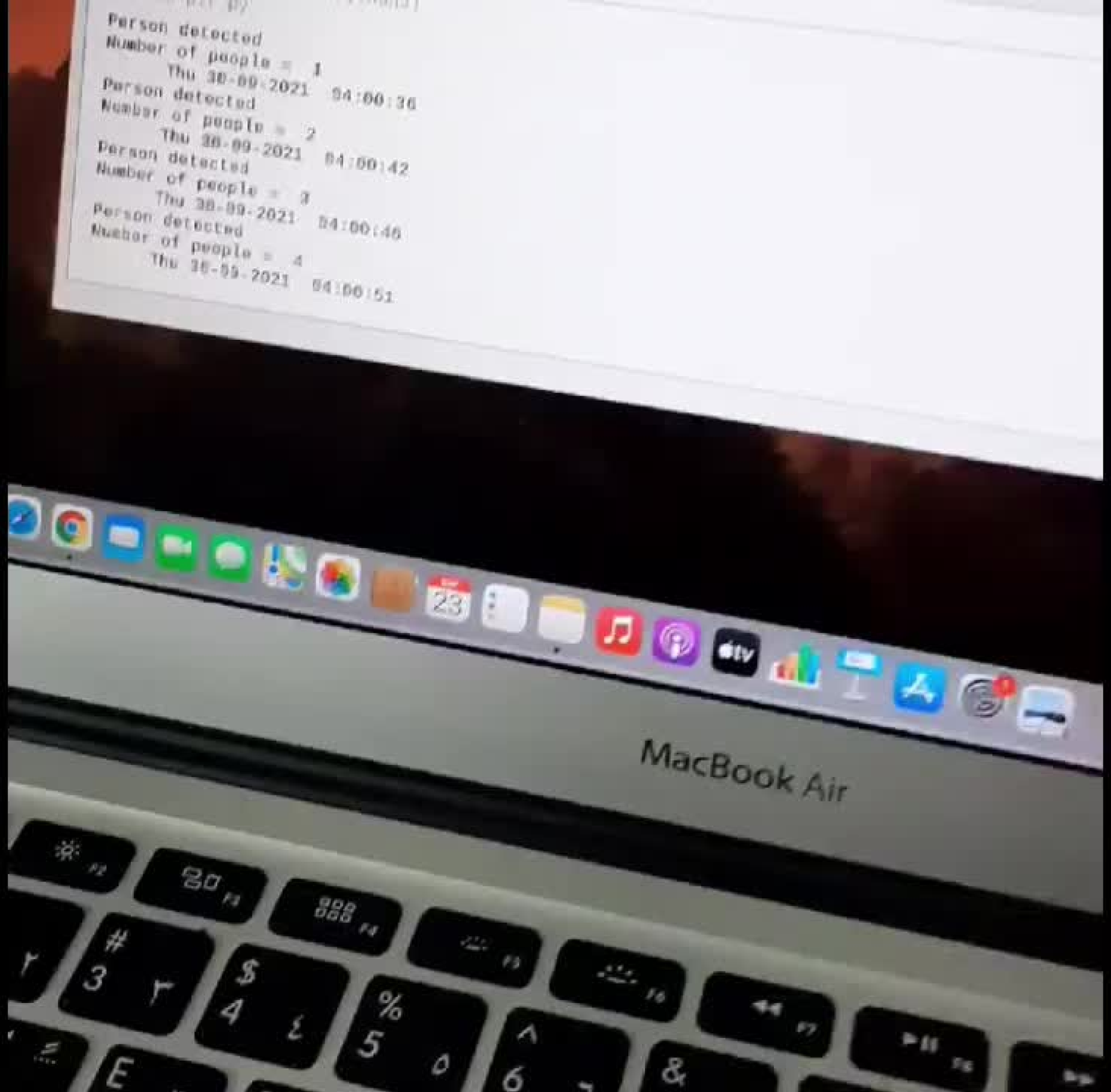


CONCLUSION

- Workers at the station will be able to better distribute themselves and control traffic.
- Travel on weekends before 6 AM to avoid excessive station congestion.
- People will be notified on Saturday which considers the most congested day of the week.
- Workers will be ready every day from 6 a.m. to 5 p.m., especially on Saturday, which have the most traffic.

RECOMMENDATION

- Analyzing stations congestion based on holidays, and other events.
- Using a different dataset for Covid 19 to analyzing the percentage of daily and weekly infliction, and how station congestion will affect this percentage.
- Using real-time datasets provided by IoT sensors to collect, analyze, and share data on a daily basis.


Person detected
Number of people = 1
Thu 30-09-2021 04:00:36
Person detected
Number of people = 2
Thu 30-09-2021 04:00:42
Person detected
Number of people = 3
Thu 30-09-2021 04:00:46
Person detected
Number of people = 4
Thu 30-09-2021 04:00:51
MacBook Air

T5 - SDAIA

THANK YOU

MUNTAHA JABER