



EDA Project

RAGHAD ALARIFI

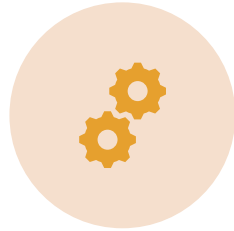
Agenda



BUSINESS
CASE



DATA SETS



TOOLS &
ALGORITHMS



DATA
ANALYSIS



RESULTS



MTA Maintenance Department

MTA Maintenance Department



Problem : Passenger delay caused by maintenance work at stations.



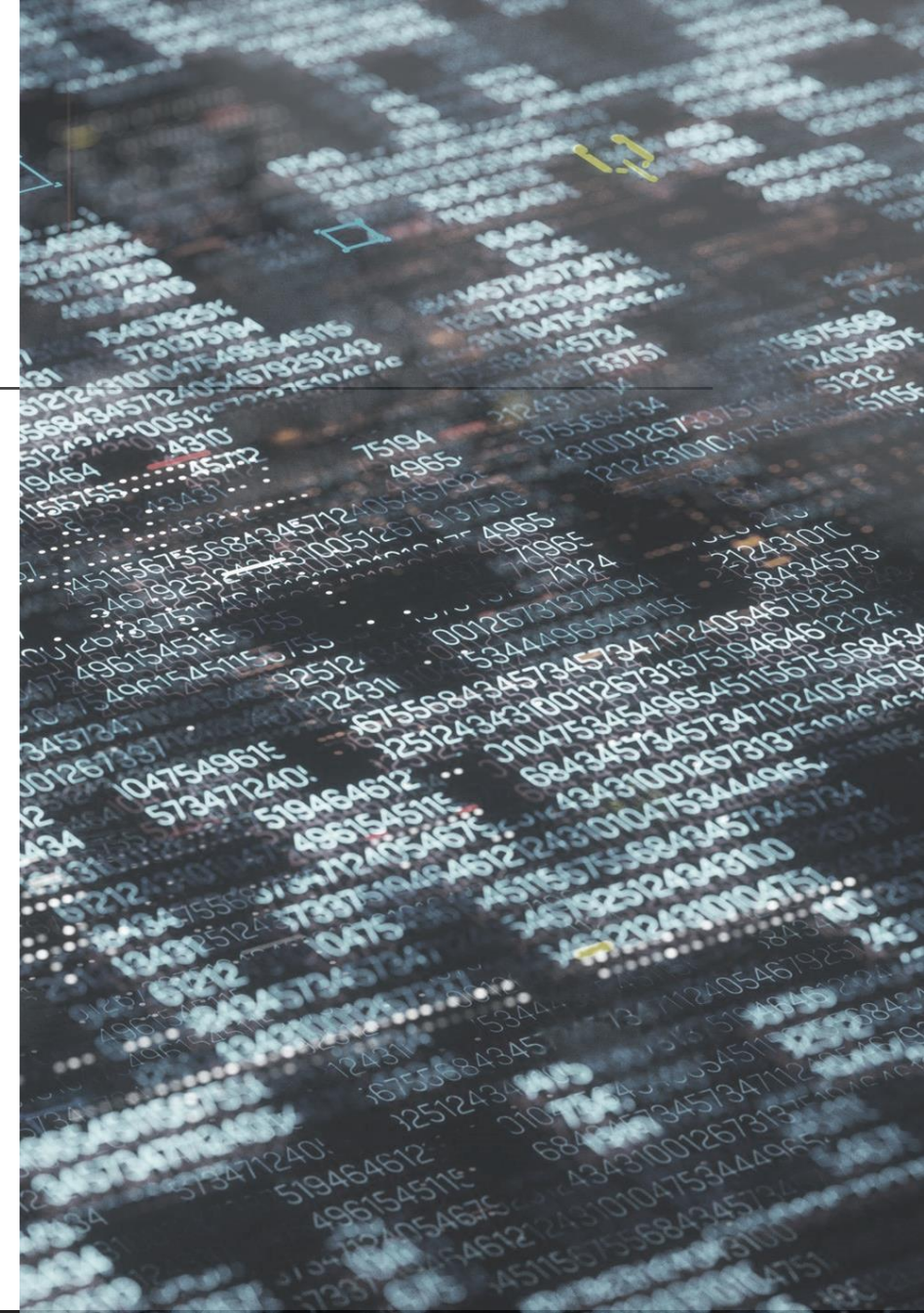
Project Goal: rescheduling maintenance work to minimize passengers delays.

2. DATA SETS

1- MTA Turnstile Data Set

2- MTA Equipment List Data Set

3- MTA Current outage data set



2.1 MTA Turnstile Data

Field Name	Description
C/A	Control Area (A002)
UNIT	Remote Unit for a station (R051)
SCP	Subunit Channel Position represents an 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 comulative entry register value for a device
EXITS	The cumulative exit register value for a device

C/A	UNIT	SCP	STATION	LINENAME	DIVISION	DATE	TIME	DESC	ENTRIES	EXITES
A002	R051	02-00-00	59 ST	BMT	BMT	06/25/2021	17:00:00	REGULAR	7508509	2557569

2.2 MTA Equipment List Data

Field Name	Description
station_name	Station Name the device is located at
equipment_id	Unique ID for the device
description	Device description and drop off location
borough	Name of borough the station located at
subway_lines	Train numbers the station serves
equipment_type	“EL” for Elevator, “ES” for escalator
ada_compliant	Is accessible to those with disabilities.
is_active	Is the equipment working
direction	Wither the device “one direction” or “two direction”
subset_lines	Train numbers the station serves
connection_to_ST	represents the device is connected to street

station_name	equipment_id	description	subway_lines	equipment_type	ADA	is_active	subset_lines	connection_to_ST
135 St	EL125	From A to B	A/B/C/D	EL	Y	N	A/B/C/D	Y

2.3 MTA Current Outage

Field Name	Description
station_name	Station Name the device is located at
equipment_id	Unique ID for the device
description	Device description and drop off location
borough	Name of borough the station located at
subway_lines	Train numbers the station serves
equipment_type	“EL” for Elevator, “ES” for escalator
ada_compliant	Is accessible to those with disabilities.
outagedate	The date and time of outage
estimatedreturntoservice	The estimated return to service
reason	Reason of outage

station_name	equipment_id	description	subway_lines	equipment_type	ADA	outagedate	ERS	reason	borough
135 St	EL125	02-00-00	A/B/C/D	EL	Y	10/7/2021 10:35:00 PM	10/8/2021 5:00:00 AM	Repair	MN

3. TOOLS AND ALGORITHMS

TOOLS:

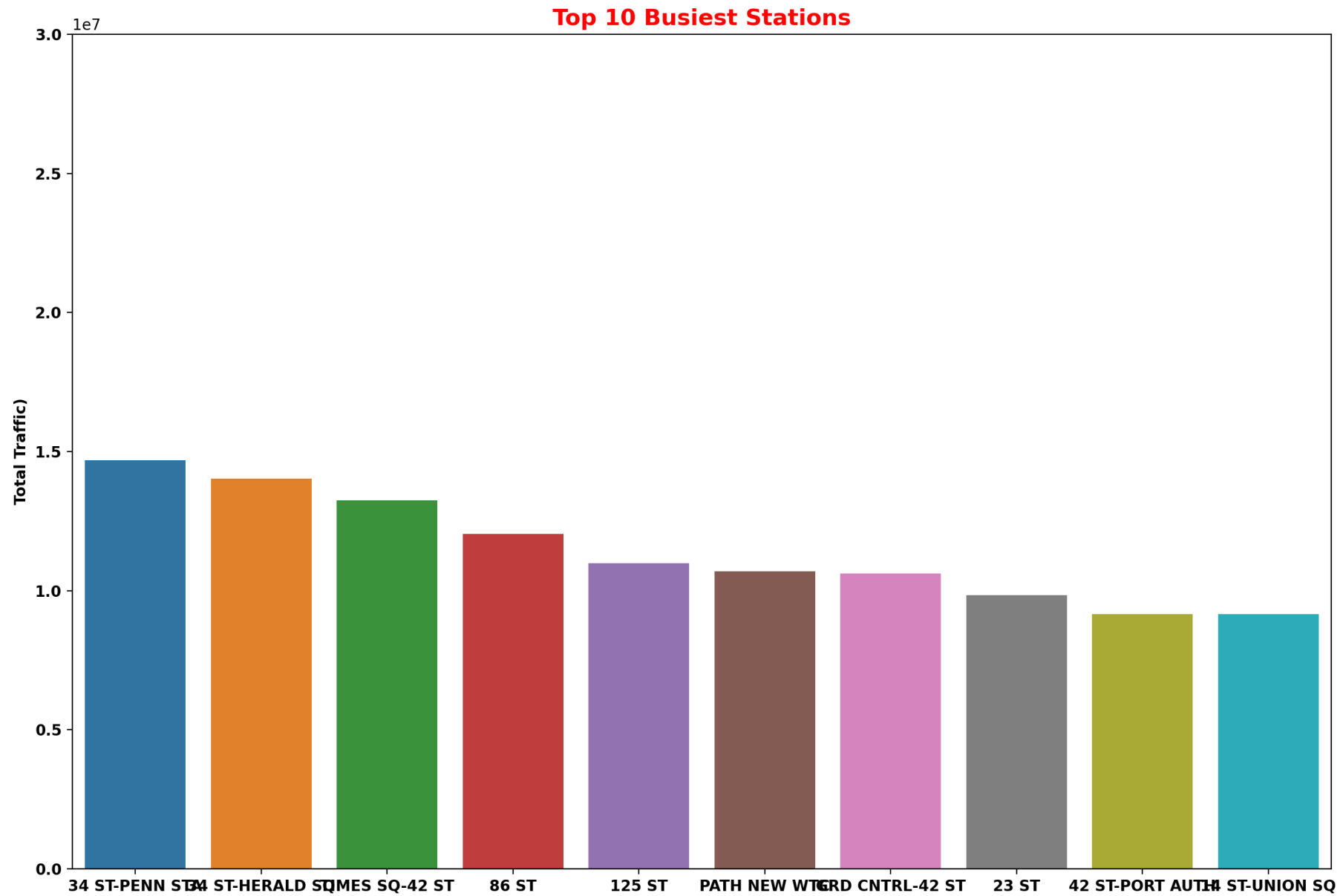
- SQLITE3
- Numpy, Pandas
- seaborn matplotlib
- Jupyter Notebook

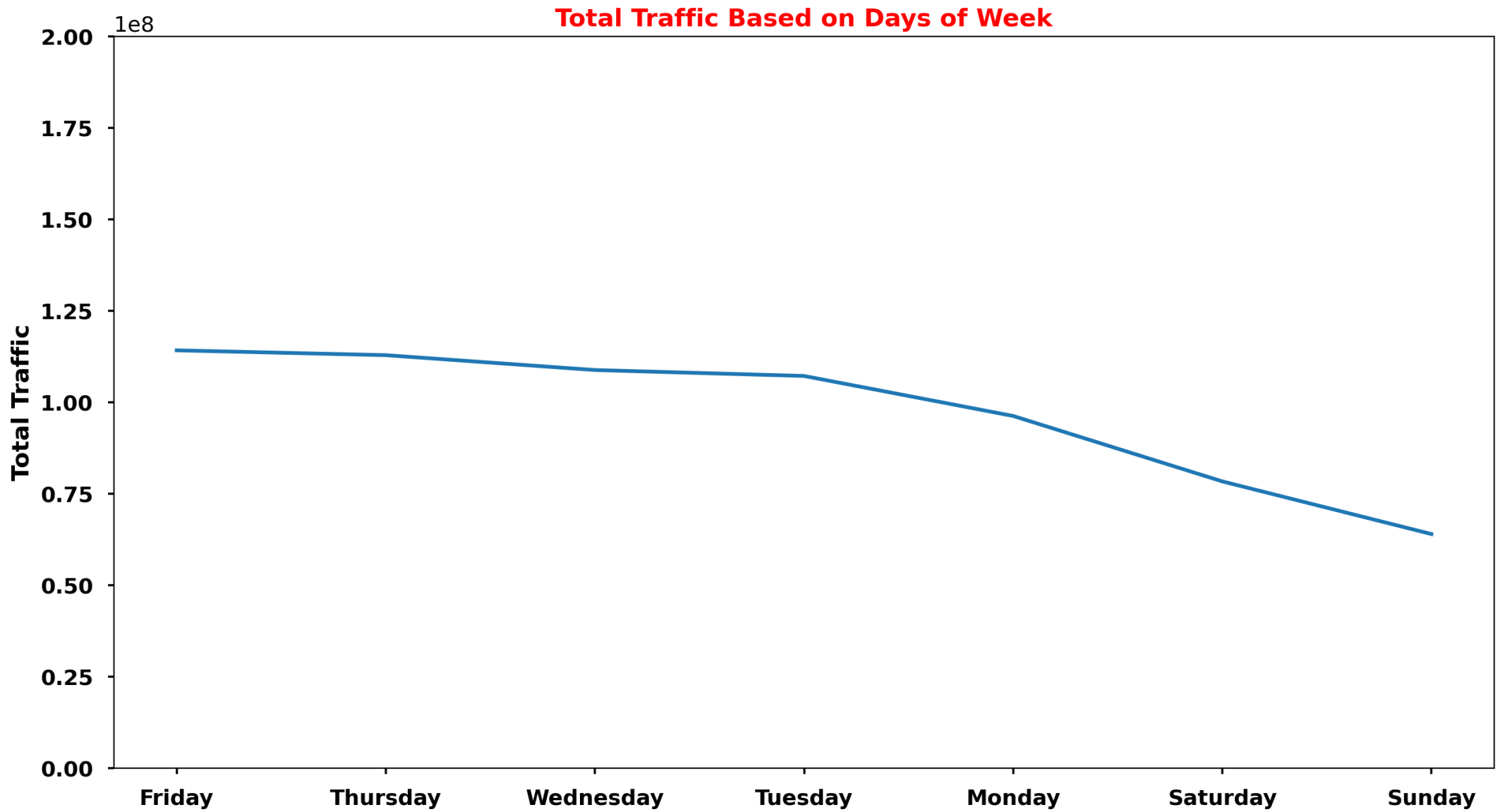
ALGORITHMS:

- Data Cleaning
 - Dealing with null values
 - Drop unnecessary column
- Feature Engineering
- Exploratory Data Analysis
- Data Visualization

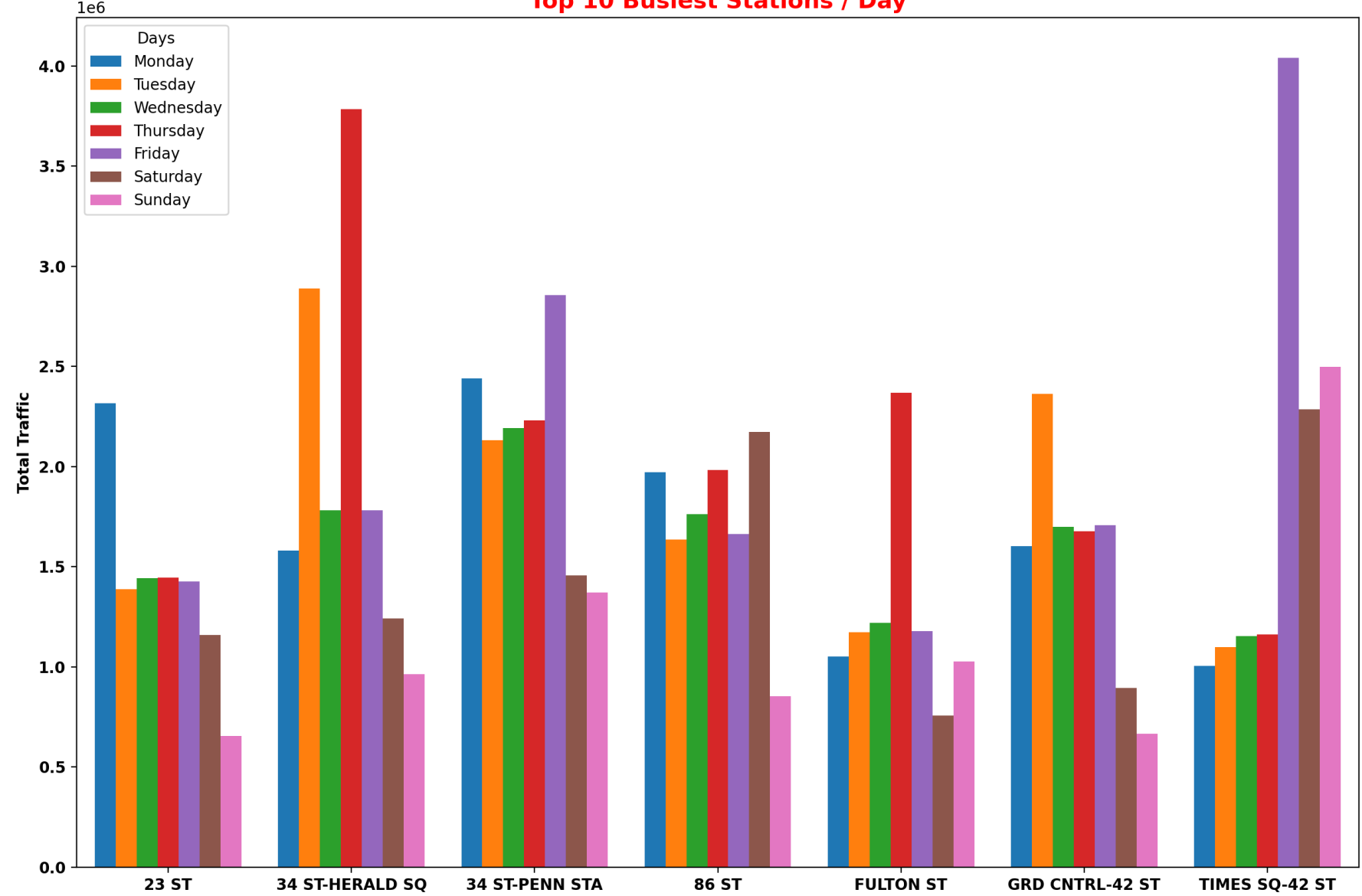


4. DATA Analysis

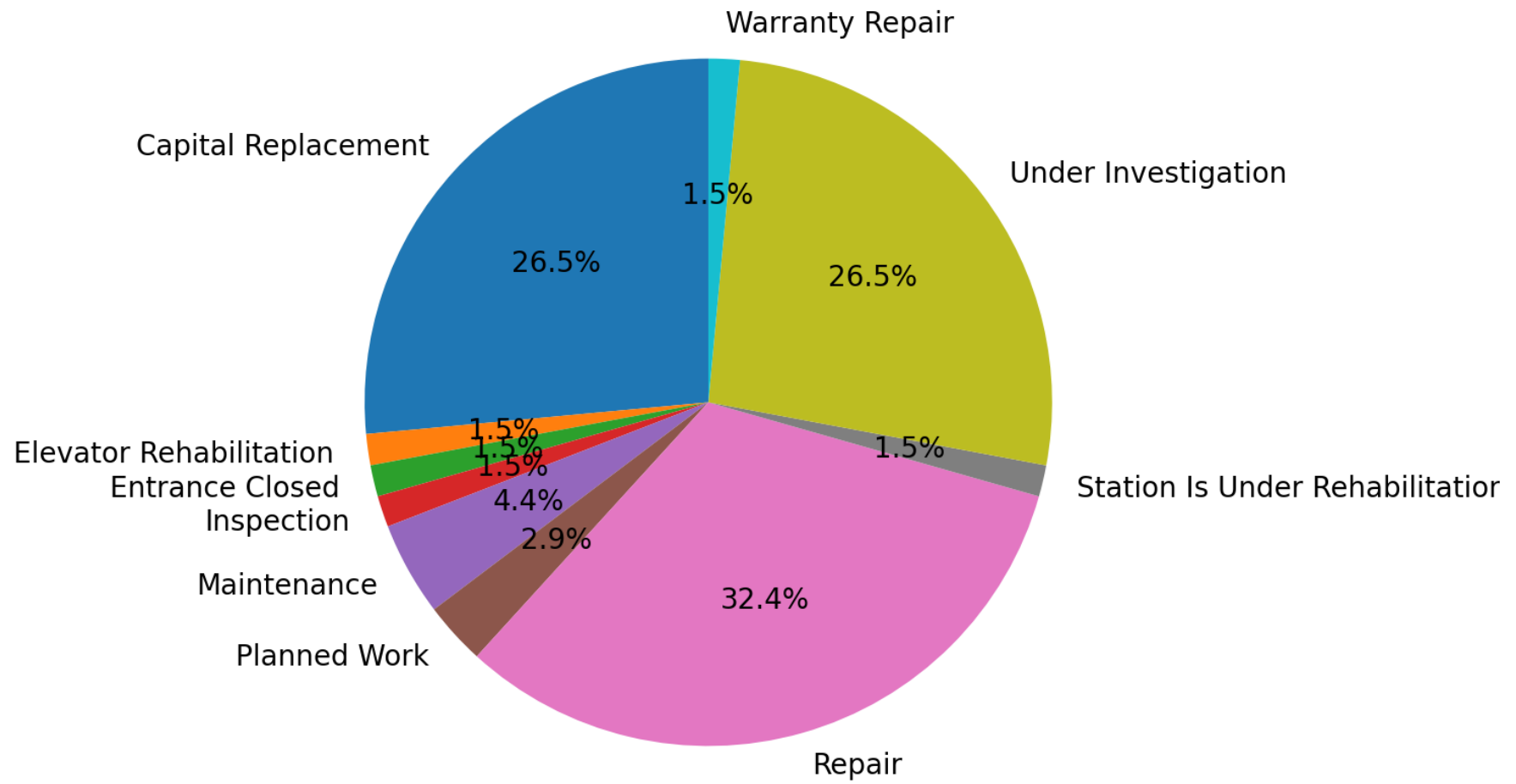


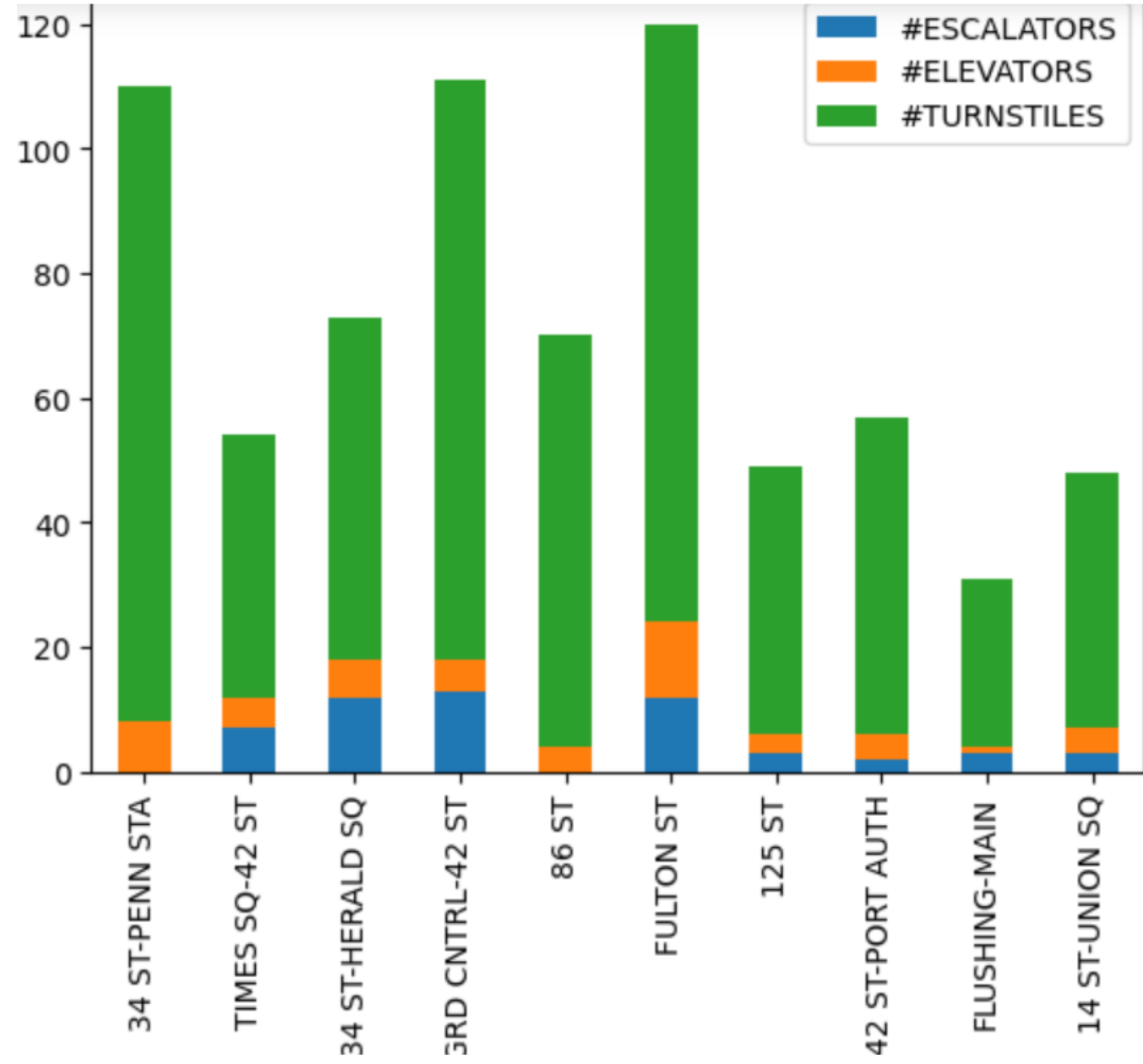


Top 10 Busiest Stations / Day



Reasons for current outages





Conclusion



Schedule the maintenance work at : Friday, Sunday



Prioritize maintenance work for devices located at busy stations.



Replace the frequently outage elevators and escalators

Any Questions?

Thank you for listening

A solid dark gray horizontal bar spanning the entire width of the slide at the bottom.