



MTA station: Source: Wikipedia

# Fundraising at the NYC subway stations

**Diego Duque Marquez - Data Analyst**

## Overview:

- American Cancer Society
- Recommendation for the best period of times and locations for fundraising at MTA stations.

## Requirements:

- Nov - Jan
- 5 weeks of ad budget
- NYC subway stations (in different neighborhoods)



## Execution:

1. Collect information from MTA's website.
2. Identify the 10 stations with the highest entering traffic.
3. Select the 5 stations with the highest rank but from different neighborhoods.
4. Analyze the 5 stations on monthly and weekly periods.
5. Recommendation of most suitable weeks and stations for fundraising.

# Methodology:

## Project workflow - technical:

### 1) Data Collection:

- SQL and Python

### 2) Data Preparation:

- SQL

### 3) Exploratory Data Analysis:

- SQL y Python Pandas

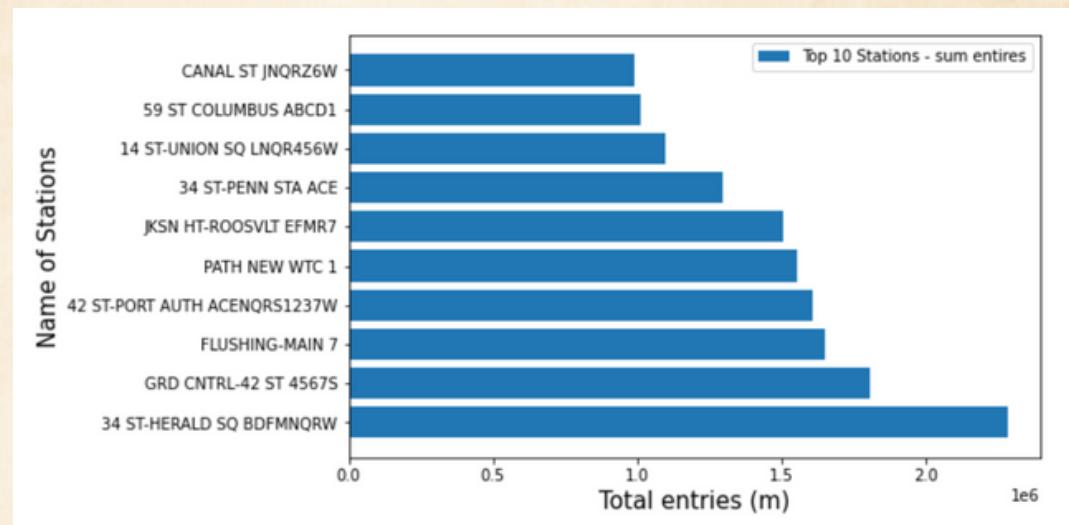
### 4) Data Visualization:

- Python Matplotlib

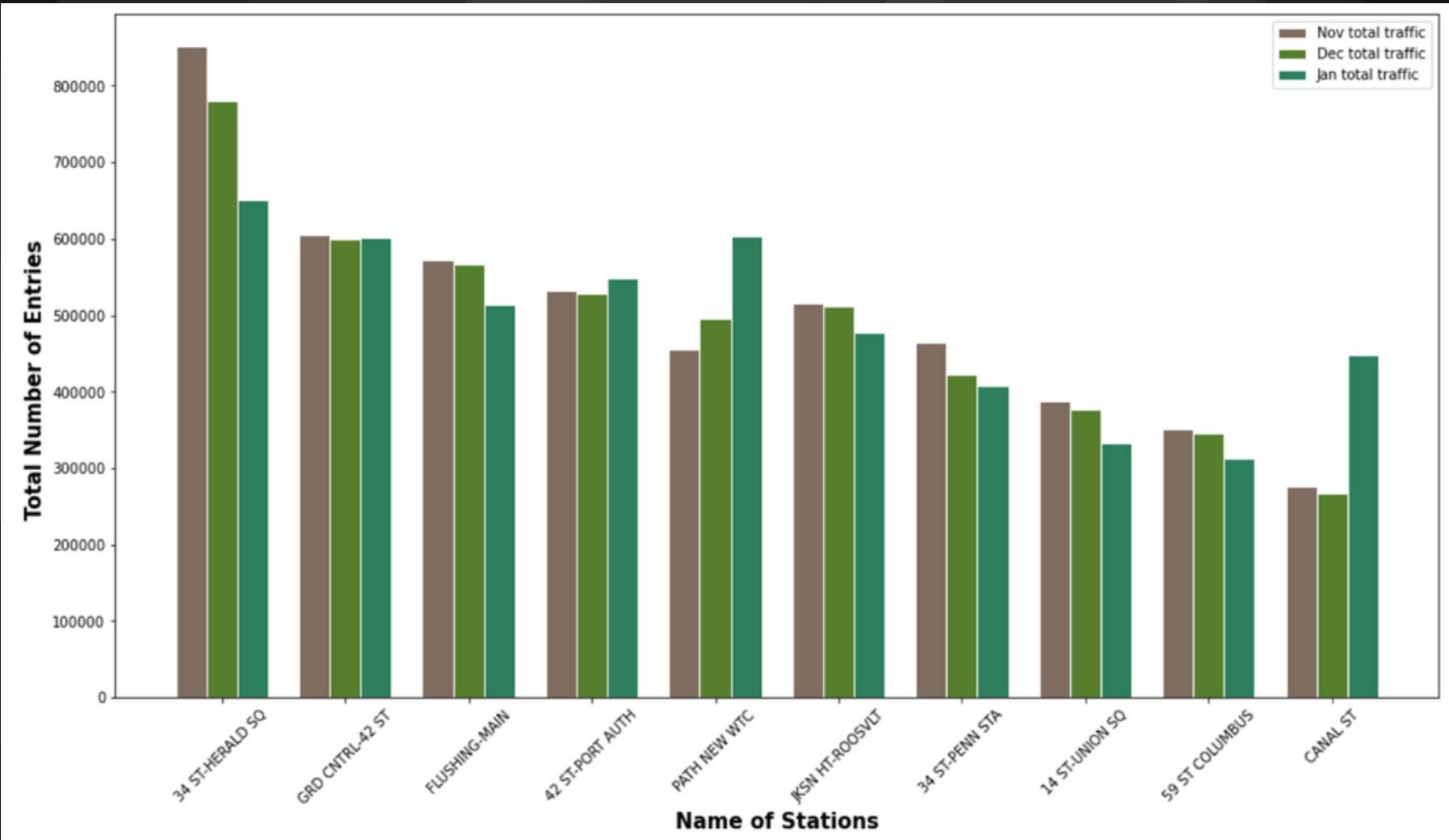
# Results:

Ten subway stations with the most traffic

1. 34 St - Herald Square
2. Grand Central - 42 St
3. Flushing - Main
4. 42 St - Port Auth
5. Path New WTC
6. JKSN HT - Roosevelt
7. 23 St Penn Sta
8. 14 St - Union Square
9. 59 St Columbus
10. Canal St



# Monthly total entries of the top 10 stations



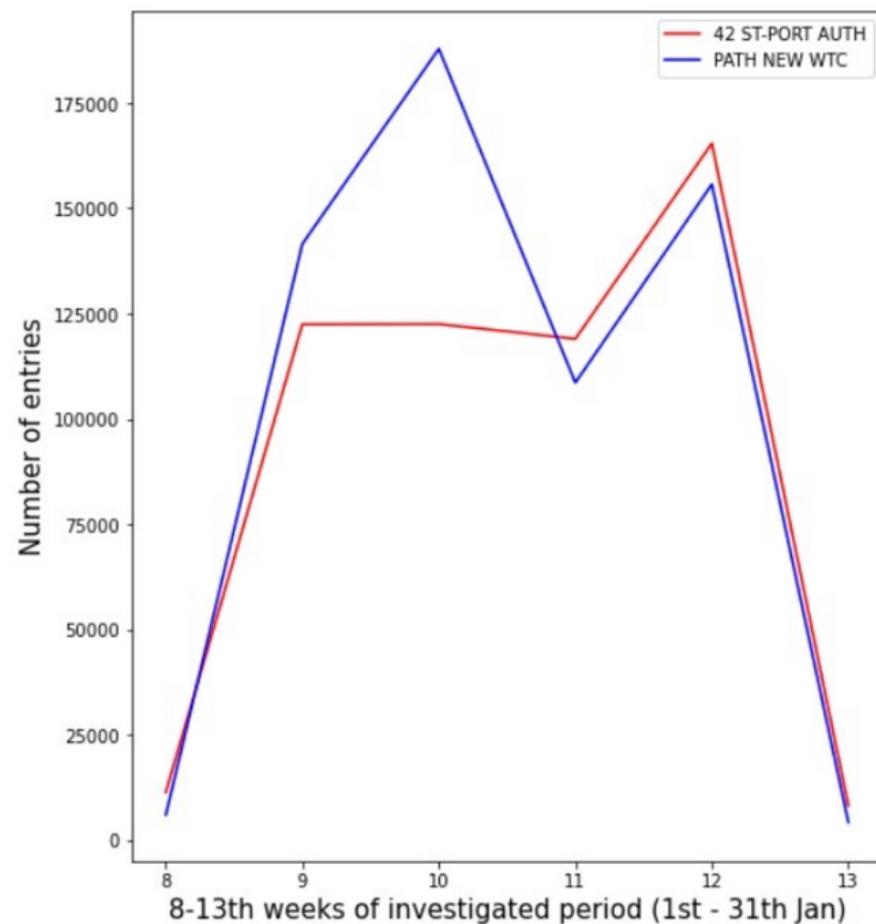
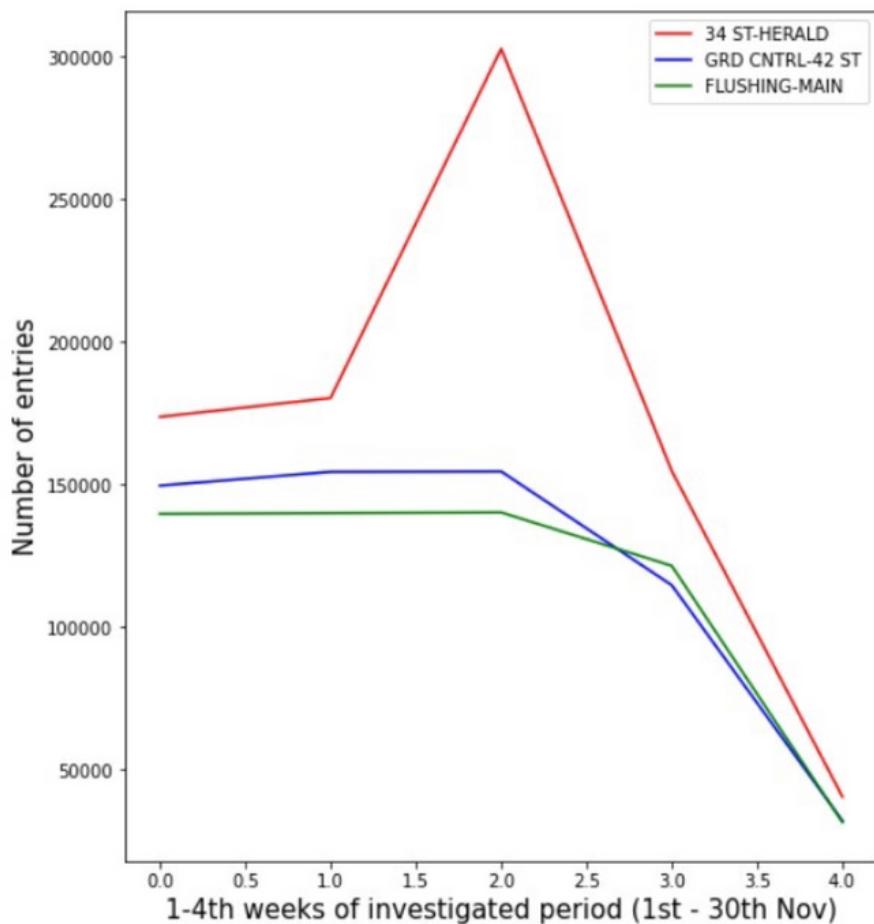


# Top 5 stations with neighborhoods:

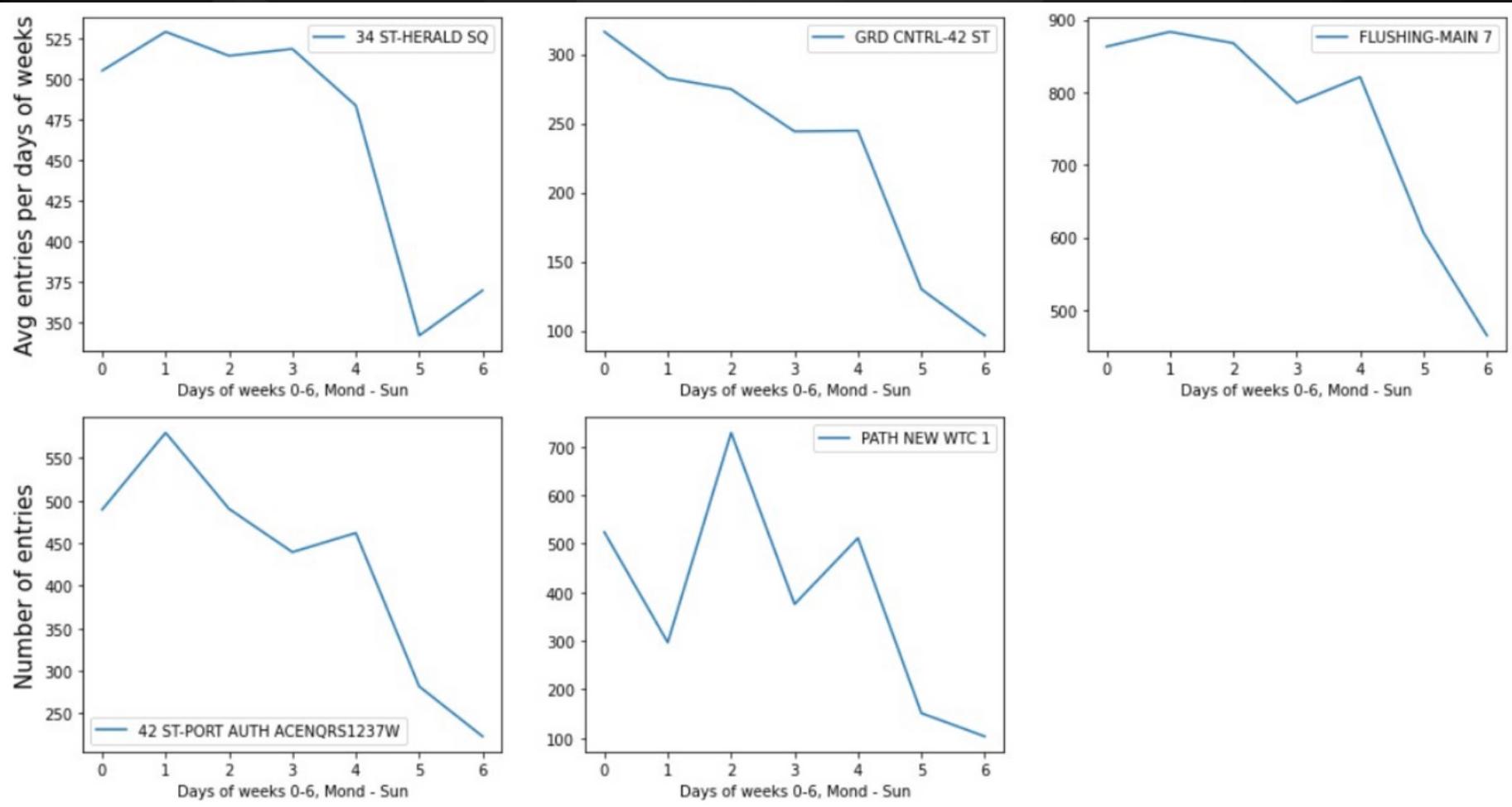
1. 34 St - Herald Square - Garment District
  2. Grand Central - 42 St - Murray Hill
  3. Flushing - Main - Queens/ Linden Hill
  4. 42 St - Port Auth - Hell's Kitchen
  5. Path New WTC - Lower Manhattan

MTA map. Source: MTA





The 5 selected station's weekly traffic for the month  
with the most entries



Average entries of the station per days of the week  
**(Between Nov 2020 - Jan 2021)**

# Conclusion:

The recommendation to the American Cancer Society for placement of their fundraising units are the following stations and days:

1. **34 St - Herald Square - 3rd week of November - Tuesday**
2. **Grand Central- 42 St - 3rd week of November - Monday**
3. **Flushing - Main - 3rd week of November- Tuesday**
4. **42 ST-PORT AUTH - 4th week of January - Tuesday**
5. **Path New WTC - 2nd week of January - Wednesday**

MTA station. Source: Wikipedia



## Appendix:

- RAPID Chart
- Gantt Chart
- SQL example
- Dashboard findings

# RAPID Chart

## American Cancer Society MTA fundraising project

	Which stations and days?	What is the timeline for the fundraising?
INPUT	<ul style="list-style-type: none"><li>• Program team</li><li>• Marketing director</li><li>• MTA Marketing department (partners)</li></ul>	<ul style="list-style-type: none"><li>• Program team</li><li>• Marketing director</li></ul>
AGREE	<ul style="list-style-type: none"><li>• n/a</li></ul>	<ul style="list-style-type: none"><li>• Senior Financial Analyst (responsible for all the fundraising financial analysis)</li></ul>
RECOMMEND	<ul style="list-style-type: none"><li>• Data Analyst (Diego Duque Marquez) recommended stations and days</li></ul>	<ul style="list-style-type: none"><li>• Marketing director</li></ul>
DECIDE	<ul style="list-style-type: none"><li>• Executive director (leadership team strives for consensus, but the executive director has the final decision)</li></ul>	<ul style="list-style-type: none"><li>• Director of programs</li></ul>
PERFORM	<ul style="list-style-type: none"><li>• Program team</li></ul>	<ul style="list-style-type: none"><li>• Program team</li></ul>

# Gantt Chart

# American Cancer Society MTA fundraising project

# SQL example

## SQLite running in Python via SQLAlchemy

Using SQLite during the Exploratory Data Analysis (EDA) SUM entries of some stations are too high, this was corrected.

```
pd.read_sql('SELECT STATION, SUM(ENTRIES) FROM mta_data GROUP BY STATION', engine)
```

	STATION	SUM(ENTRIES)
0	1 AV	275916927334
1	103 ST	79383543002
2	103 ST-CORONA	37472645330
3	104 ST	994725648708
4	110 ST	24245148515
...	...	...
373	WOODLAWN	19412440969
374	WORLD TRADE CTR	941968634938
375	WTC-CORTLANDT	433042636583
376	YORK ST	17424098719
377	ZEREGA AV	1582732655



Turnstile. Source: Wikipedia

## Main Findings:

- First two weeks of the month have the least traffic
- Most of the traffic are on Monday, Tuesday and Wednesday
- Some turnstiles count backwards
- In the top 5 busiest stations on Sundays the traffic decreases between 70-90%



MTA train. Source Wikipedia

**Diego Duque Marquez**  
Data Analyst  
[diegoduquem@icloud.com](mailto:diegoduquem@icloud.com)