

# MTA Foot-Traffic

Kevin Joo - DA Flex Student

# Introduction

**Question: Are there avenue for improvements in entering and exiting stations?**

**Objective: Analyze the foot-traffic of people using the MTA to travel within NYC**

**Goal: To inform MTA of possible improvements for future renovation to a given station on turnstile placements and/or directing riders to other turnstiles**

# Methodology

## **Data:**

- 3 month of MTA data (April'21 - July'21)**
- COVID vaccines were rolled out**

## **Metric:**

- Look at the entry and exit data to understand the foot-traffic for each station and turnstile**

## **Tools:**

- NumPy and Pandas for data manipulation**
- Matplotlib and Seaborn for data visualizations**

# Methodology

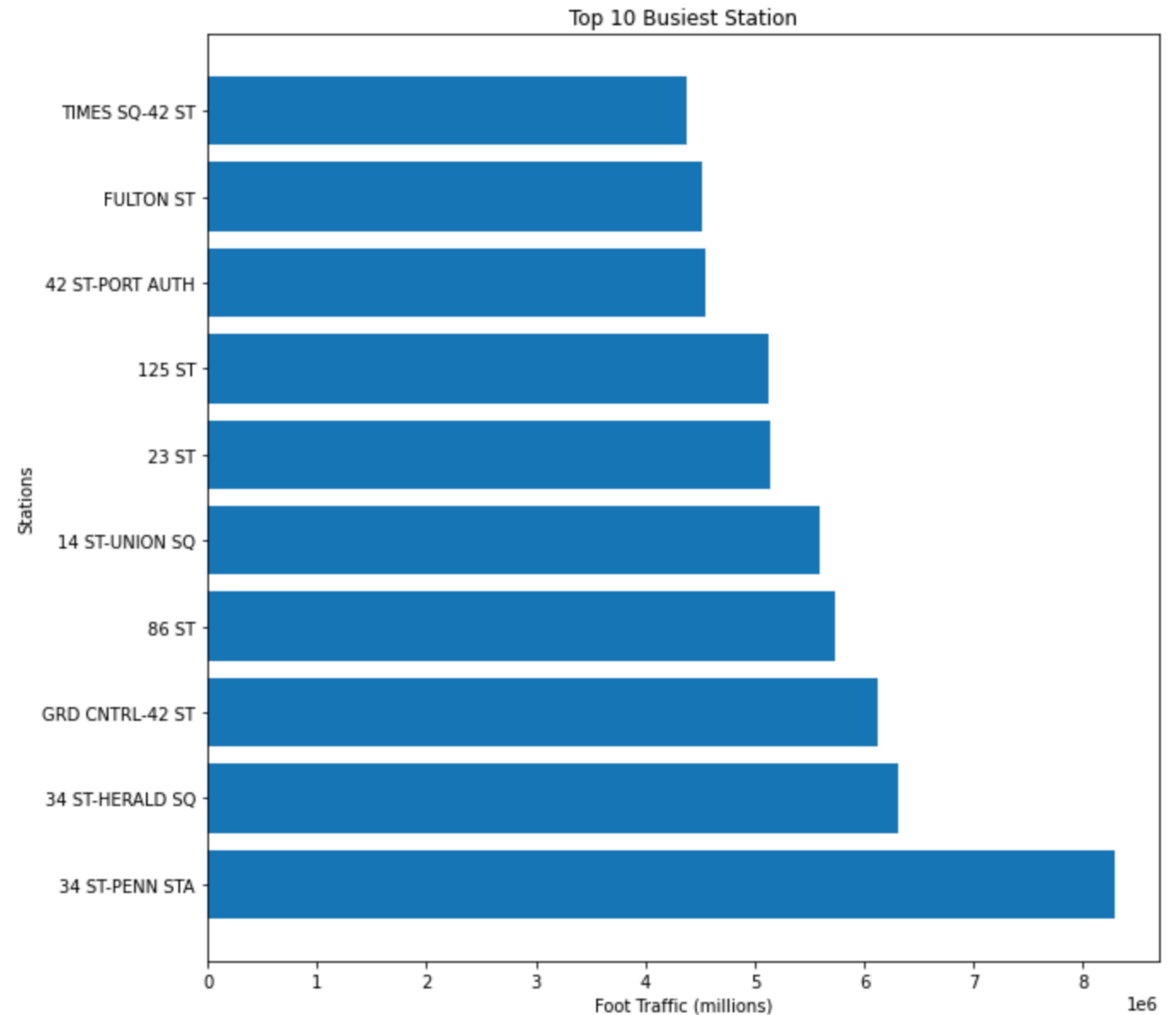
## Initial Steps:

- 1. Use different built-in functions to clean up the raw data**
  - a. .describe**
  - b. .groupby chained w/ .diff()**
  - c. mask**
- 2. Interpret and visualize data**

# Results

## Insight:

- Identified the top 10 stations with the most foot-traffic
- Penn Station: had the most foot-traffic and was selected to take a deeper dive into the foot-traffic pattern



# Results

- Insight:**
- There are 111 unique turnstiles placed through Penn Station
  - The average foot traffic during the 3month period for Penn Station was: 74.6K riders
  - Roughly 50% of the turnstiles are used above the mean
  - Suggests there are room to help direct riders to other turnstiles to help decrease the chance of bottle-necking at certain “popular” turnstiles

```
penn_station_df["avg_foot_traffic"] = penn_station_df["total_traffic"] > penn_station_df["total_traffic"].mean()
```

```
penn_station_df.head()
```

		turnstiles	total_traffic	avg_foot_traffic
0	R138R29300-03-04_34	ST-PENN STA	235635.0	True
1	N067R01200-05-00_34	ST-PENN STA	194128.0	True
2	N067R01200-05-01_34	ST-PENN STA	184551.0	True
3	N073R01302-00-01_34	ST-PENN STA	182315.0	True
4	R138R29300-00-02_34	ST-PENN STA	170209.0	True

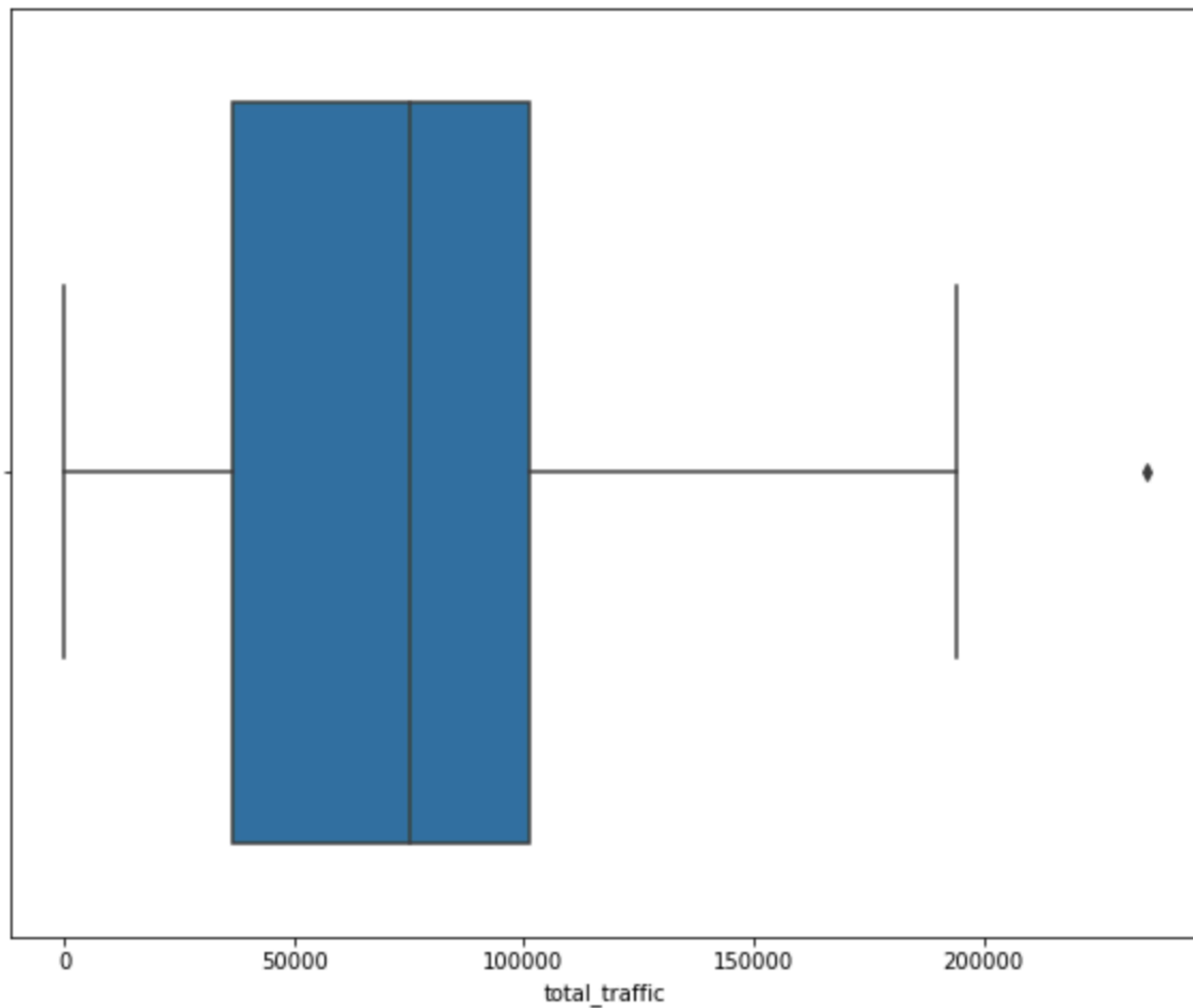
```
penn_station_df.tail()
```

		turnstiles	total_traffic	avg_foot_traffic
106	N071R01300-03-05_34	ST-PENN STA	0.0	False
107	N071R01300-03-04_34	ST-PENN STA	0.0	False
108	N071R01300-03-00_34	ST-PENN STA	0.0	False
109	N071R01300-03-02_34	ST-PENN STA	0.0	False
110	R138R29300-03-00_34	ST-PENN STA	-16.0	False

```
penn_station_df["avg_foot_traffic"].value_counts()
```

```
True      56
False     55
Name: avg_foot_traffic, dtype: int64
```

Penn Station turnstiles



# Conclusion

## **Recommendation:**

- Roughly ~50% of turnstiles are under-utilized, which means there are room for traffic-flow improvement**
- To reduce bottle-necking, there are 2 recommendations:**

**Economical method: inform riders of other turnstiles available to them**

**Costly method: build out new turnstiles to help direct foot-traffic away from “popular” turnstiles**

**Review the data again in the future**



# Future Work

**If I had more time:**

- Understand how emergency exits are tracked and see how much it is costing MTA**
- Understand the placements of turnstiles to see which areas are being over/under utilized**
- Incorporate Power BI to help create more visualization for deeper insight**

# Appendix

N071	R013	00-00-01	34 ST-PENN STA	ACE	IND	05/24/2021	08:00:00	REGULAR	2597061	2634738	2021-05-24 08:00:00	N071R01300-00-01_34 ST-PENN STA	0.0	0.0	0.0
N071	R013	00-00-01	34 ST-PENN STA	ACE	IND	05/24/2021	11:47:10	REGULAR	3500758	7791164	2021-05-24 11:47:10	N071R01300-00-01_34 ST-PENN STA	903697.0	5156426.0	6060123.0
N071	R013	00-00-01	34 ST-PENN STA	ACE	IND	05/24/2021	12:00:00	REGULAR	3500761	7791172	2021-05-24 12:00:00	N071R01300-00-01_34 ST-PENN STA	3.0	8.0	11.0
N071	R013	00-00-01	34 ST-PENN STA	ACE	IND	05/24/2021	12:00:00	REGULAR	3500761	7791172	2021-05-24 12:00:00	N071R01300-00-01_34 ST-PENN STA	3.0	8.0	11.0

R218	00-00-00	ELMHURST AV	MR	IND	07/26/2021	12:00:00	REGULAR	1037161	416028	2021-07-26 12:00:00	N325AR21800-00-00_ELMHURST AV	2.0	13.0	15.0
R218	00-00-00	ELMHURST AV	MR	IND	07/26/2021	16:00:00	REGULAR	16777220	13	2021-07-26 16:00:00	N325AR21800-00-00_ELMHURST AV	15740059.0	-416015.0	15324044.0
R218	00-00-00	ELMHURST AV	MR	IND	07/26/2021	20:00:00	REGULAR	16777220	13	2021-07-26 20:00:00	N325AR21800-00-00_ELMHURST AV	0.0	0.0	0.0

# End