Optimizing WTWY Street Team Work

Introduction

01 Motivation: To spread awareness of women in tech and to increase attendance at the WTWY gala.

02 Objective: To optimize canvassing work conducted by WTWY street teams

03 Goals: To determine the busiest subway stations and their corresponding busiest days of the week.

Methodology: The Data

	ORIGIN	DESCRIPTION	FORM
mta_data.db	Metropolitan Transportation Authority (MTA) Developers Resources	Turnstile data was acquired for three months from 00:00:00 on 05/29/2021 to 23:59:56 on 08/27/2021.	SQL Database via SQLite3

Metrics

- Entry data by individual turnstile
- Exit data by individual turnstile
- Entry data by station
- Exit data by station

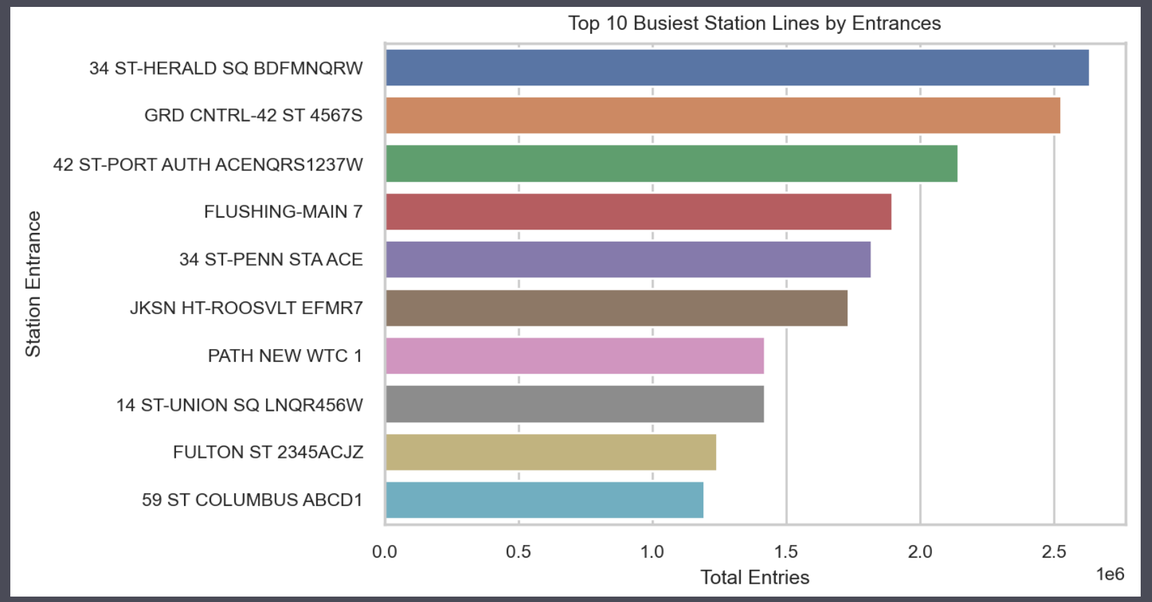
METHODS

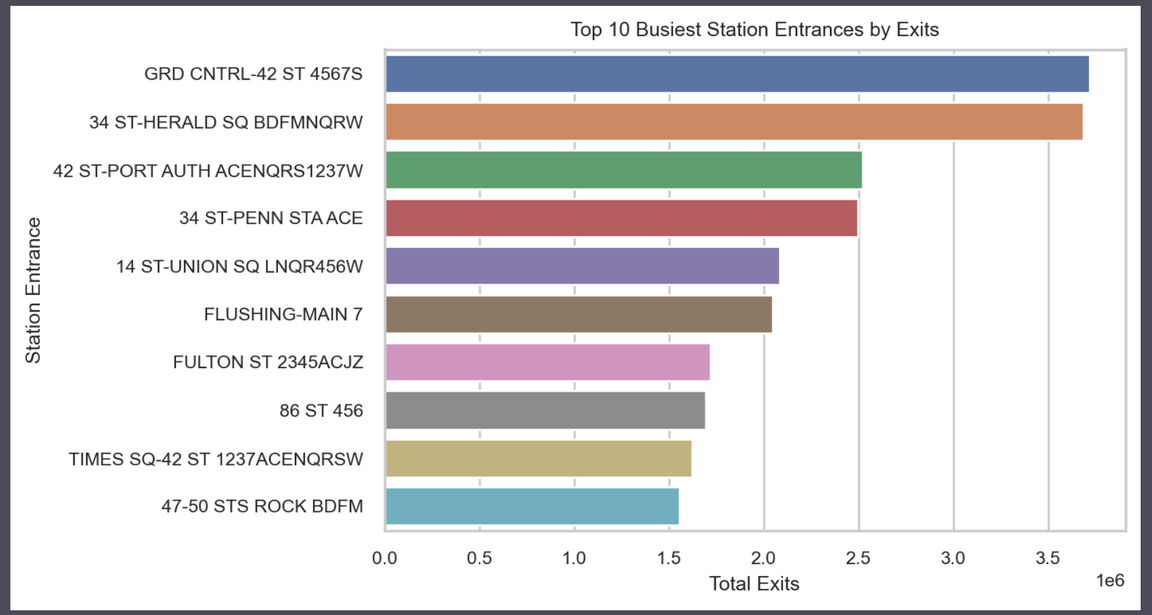
Tools

- Sqlite3 for data acquisition
- SQLAlchemy for data retrieval and exploration
- Pandas for data manipulation
- Matplotlib and Seaborn for plotting and visualizations

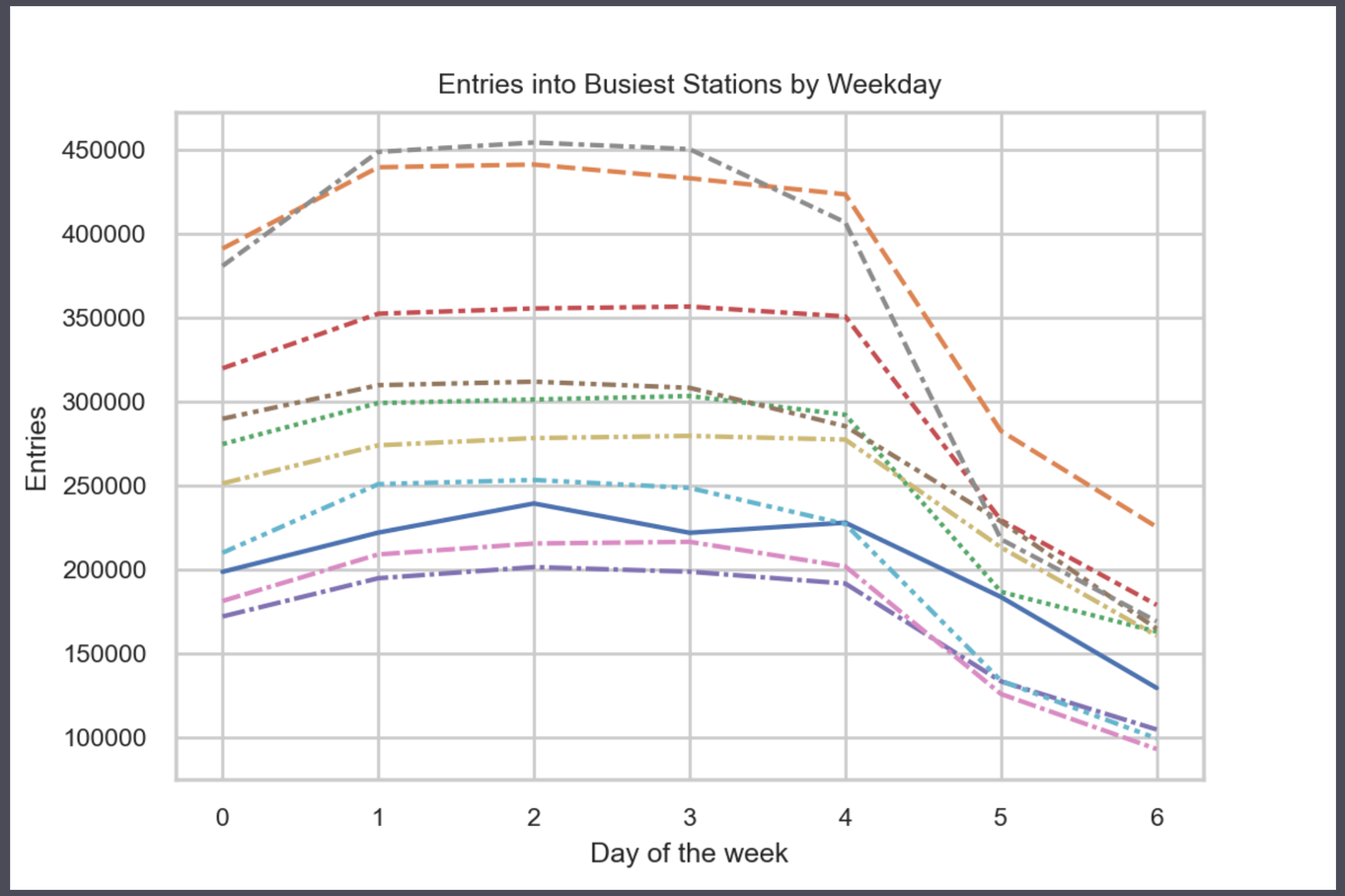
Results

Top ten busiest stations by entry and exit data.





Results: Entries for the top ten busiest stations by day of the week



0-6 represent the days of the week, with 0 representing Monday and 6 representing Sunday on the x-axis.

Conclusion

To increase the efficiency of the street teams, I recommend placing street teams at the top busiest stations by entrance in the morning and the top busiest stations by exit in the afternoon/evening to reflect the commute to and from work.

INTERESTING INSIGHTS

The busiest days of the week across all stations were Tuesday, Wednesday, and Thursday. Therefore, these would be optimal days for street teams to canvas on.

FUTURE WORK

To increase the accuracy of the stations designated as the busiest, I would ensure the station and line names accurately represented the subway network. For instance, I would ensure all rows for Time Square-42 Street station were under the same line name.

STATION	LINENAME
TIMES SQ-42 ST	1237ACENQRS
TIMES SQ-42 ST	1237ACENQRSW
TIMES SQ-42 ST	ACENQRS1237W

In addition to further cleaning the data, I would incorporate additional data to exclude popular stations for tourists and commuters to target New York City residents.

Appendix

	station_id	ENTRIES
89	34 ST-HERALD SQ BDFMNQRW	2634766
312	GRD CNTRL-42 ST 4567S	2526630
102	42 ST-PORT AUTH ACENQRS1237W	2141962
289	FLUSHING-MAIN 7	1897147
93	34 ST-PENN STA ACE	1819080
339	JKSN HT-ROOSVLT EFMR7	1733084
404	PATH NEW WTC 1	1421847
28	14 ST-UNION SQ LNQR456W	1421764
302	FULTON ST 2345ACJZ	1241989
123	59 ST COLUMBUS ABCD1	1195711

	station_id	total_exits
312	GRD CNTRL-42 ST 4567S	3721429
89	34 ST-HERALD SQ BDFMNQRW	3686060
102	42 ST-PORT AUTH ACENQRS1237W	2524626
93	34 ST-PENN STA ACE	2496843
28	14 ST-UNION SQ LNQR456W	2084005
289	FLUSHING-MAIN 7	2047983
302	FULTON ST 2345ACJZ	1721308
156	86 ST 456	1694518
448	TIMES SQ-42 ST 1237ACENQRSW	1624381
106	47-50 STS ROCK BDFM	1558607

	STATION	LINENAME
0	34 ST-PENN STA	123
1	34 ST-PENN STA	123ACE
2	34 ST-PENN STA	ACE