Finding Optimal Stations for NYY Hat Sales

BY: WILLIAM PAULSON

Introduction

➢ Hat Hawkers House (HHH) wants to increase revenue

  ➢ Perform exploratory data analysis on New York MTA turnstile data

  ➢ Find best stations

  ➢ Deploy Hat Hawkers efficiently to maximize sales potential

Methodology

Focus on Exit Data for Weekday Night Games

➢ Find Exit Traffic Information for Summer 2019 (May-August)

➢ Isolate days during timeframe when fans would be most likely to travel

➢ Differentiate the data between Game days and regular days

➢ Analyze the results

➢ Data
  ➢ New York MTA turnstile Data
  ➢ NY Yankees 2019 Schedule

➢ Tools
  ➢ Python
    ➢ Pandas
    ➢ Datetime
    ➢ Matplotlib
  ➢ SQLite
    ➢ SQLalchemy

Biggest "Game Day" Effect

	STATION	Avg Game Exits	Avg Regular Exits	Game Effect
232	GRD CNTRL-42 ST	40303.562500	12688.757143	27614.805357
78	50 ST	22362.411765	9859.807692	12502.604072
21	161/YANKEE STAD	12791.791667	4808.615385	7983.176282
355	TWENTY THIRD ST	6522.354680	180.067265	6342.287415
117	9TH STREET	3464.938053	263.040000	3201.898053
47	231 ST	3304.666667	582.576923	2722.089744
17	14TH STREET	2533.234568	180.088803	2353.145765
6	116 ST	5392.875000	3307.441860	2085.433140
341	SPRING ST	4218.750000	2830.243902	1388.506098
46	23 ST	11127.833333	10253.711538	874.121795

|	STATION	Avg Game Exits	Avg Regular Exits	Game Effect
33	181 ST	2748.612903	3597.000000	-848.387097
278	METS-WILLETS PT	1020.875000	2142.692308	-1121.817308
124	ATL AV-BARCLAY	14750.230769	15895.192308	-1144.961538
352	TIMES SQ-42 ST	21818.370370	24057.615385	-2239.245014
269	LEXINGTON AV/53	3998.146341	6970.192308	-2972.045966
176	CATHEDRAL PKWY	2926.625000	5931.076923	-3004.451923
212	EXCHANGE PLACE	240.825088	4149.578447	-3908.753359
251	JAMAICA CENTER	7166.526316	11456.583333	-4290.057018
316	PAVONIA/NEWPORT	701.979920	9812.293233	-9110.313313
59	34 ST-HERALD SQ	21437.394737	34095.692308	-12658.297571

Game Day Effect as Percentage

	STATION	Avg Game Exits	Avg Regular Exits	Game Effect	Game Effect %
17	14TH STREET	2431.893701	180.011364	2251.882337	1250.966768
117	9TH STREET	3331.788136	261.686391	3070.101745	1173.198858
47	231 ST	3193.880000	582.114286	2611.765714	448.668892
350	THIRTY ST	366.905747	107.714519	259.191228	240.627941
232	GRD CNTRL-42 ST	39347.740000	13576.545455	25771.194545	189.821443
21	161/YANKEE STAD	12739.640000	4585.342857	8154.297143	177.833968
78	50 ST	22134.096154	9687.000000	12447.096154	128.492786
343	ST. GEORGE	1.120000	0.680851	0.439149	64.500000
167	BROOKLYN BRIDGE	1322.780000	882.510000	440.270000	49.888387
6	116 ST	5393.000000	3684.673077	1708.326923	46.363053

Best Station – 161st-Yankee Stadium

	STATION	DATE	REAL_EXITS
0	161/YANKEE STAD	05/06/2019	11006.0
1	161/YANKEE STAD	05/07/2019	11398.0
2	161/YANKEE STAD	05/08/2019	12044.0
3	161/YANKEE STAD	05/09/2019	11413.0
4	161/YANKEE STAD	05/10/2019	3569.0
5	161/YANKEE STAD	05/13/2019	6499.0
6	161/YANKEE STAD	05/14/2019	5190.0
7	161/YANKEE STAD	05/15/2019	10202.0
8	161/YANKEE STAD	05/16/2019	4635.0
9	161/YANKEE STAD	05/17/2019	11751.0
10	161/YANKEE STAD	05/27/2019	2260.0
11	161/YANKEE STAD	05/28/2019	10441.0
12	161/YANKEE STAD	05/29/2019	3847.0
13	161/YANKEE STAD	05/30/2019	5164.0
14	161/YANKEE STAD	05/31/2019	15138.0
15	161/YANKEE STAD	06/10/2019	9102.0
16	161/YANKEE STAD	06/11/2019	15969.0
17	161/YANKEE STAD	06/12/2019	4887.0
18	161/YANKEE STAD	06/13/2019	3535.0
19	161/YANKEE STAD	06/14/2019	3613.0

	game_date	game_time
0	05/06/2019	18:30:00
1	05/07/2019	18:30:00
2	05/08/2019	18:30:00
3	05/09/2019	18:30:00
4	05/17/2019	19:00:00
5	05/28/2019	18:30:00
6	05/31/2019	19:00:00
7	06/11/2019	19:00:00

Plot of Exits over Time

Daily Exits for 161st Street-Yankee Stadium
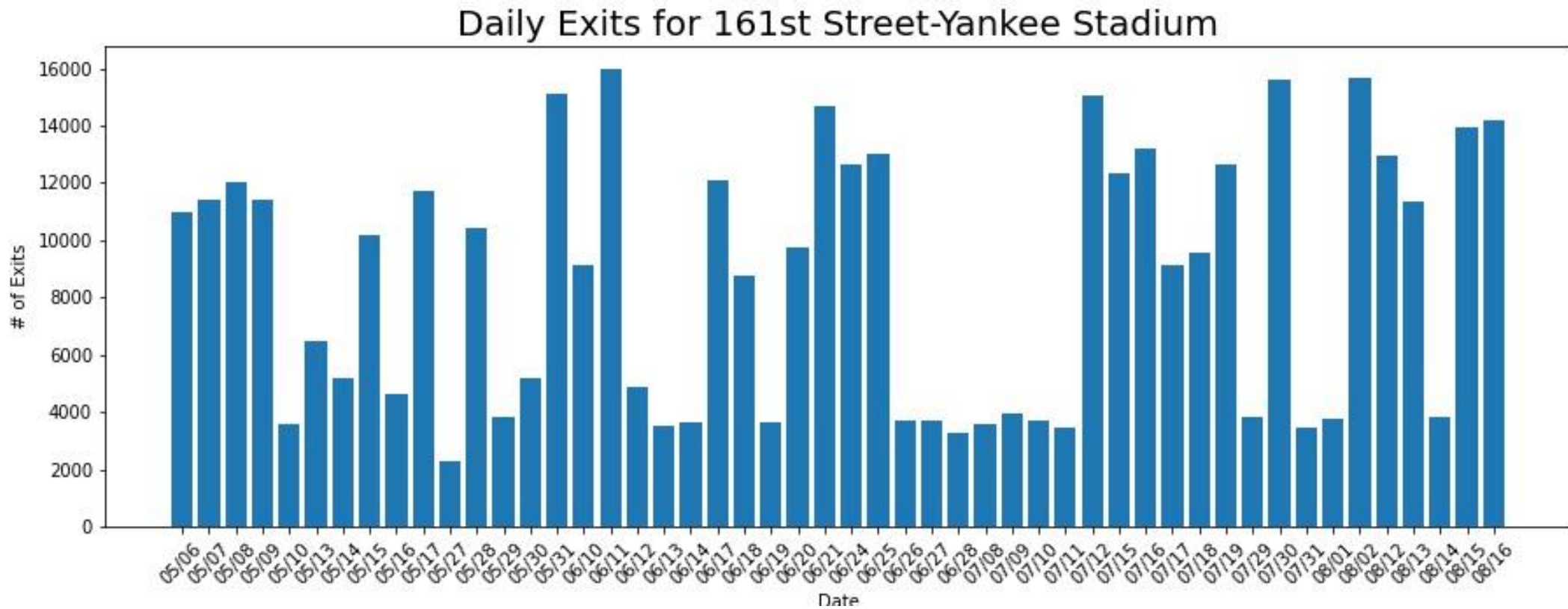
Traffic by Day of the Week

Average Exits for 161st Street-Yankee Stadium by Day of Week
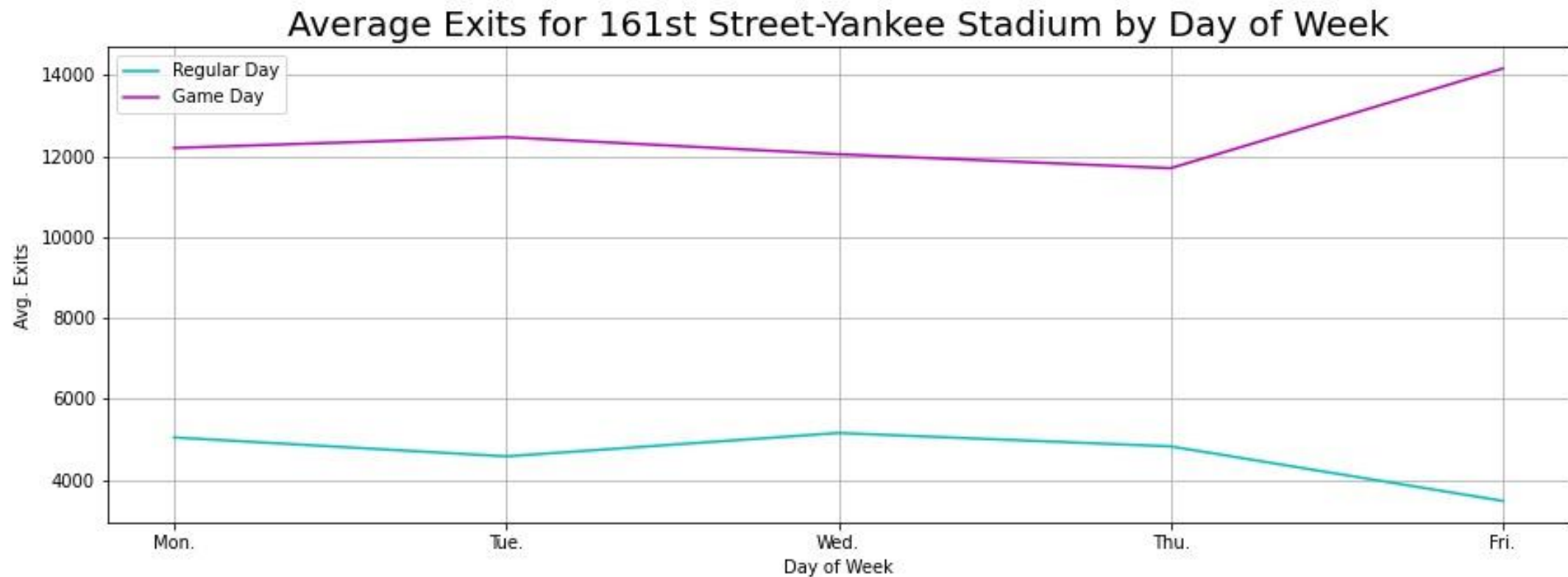
Conclusions

➢ Best Places to Sell
  ➢ Grand Central Station- 42nd St
  ➢ 161st St-Yankee Stadium

➢ Worst Places
  ➢ Jamaica Center
  ➢ Pavonia-Newport

➢ Best Time to Sell
  ➢ Fridays showed 14% higher traffic

Future Work

➢ Effect of double-headers

➢ Wednesday Day Games

➢ Weekend Games

➢ Effect of weather on subway traffic

Appendices

Highest Exit Traffic

	STATION	Avg Game Exits
232	GRD CNTRL-42 ST	40303.562500
78	50 ST	22362.411765
352	TIMES SQ-42 ST	21818.370370
14	14 ST-UNION SQ	21445.916667
59	34 ST-HERALD SQ	21437.394737
216	FLUSHING-MAIN	19158.375000
61	34 ST-PENN STA	17282.023810
68	42 ST-PORT AUTH	15966.125000
124	ATL AV-BARCLAY	14750.230769
256	JKSN HT-ROOSVLT	14063.416667

Outlier Skew

	STATION	DATE_TIME	REAL_EXITS
48	231 ST	2019-08-15 20:00:00	66095.0
27	231 ST	2019-06-26 20:00:00	914.0
0	231 ST	2019-05-06 20:00:00	673.0
13	231 ST	2019-05-30 20:00:00	659.0
32	231 ST	2019-07-10 20:00:00	647.0
16	231 ST	2019-06-11 20:00:00	643.0
7	231 ST	2019-05-15 20:00:00	642.0
35	231 ST	2019-07-15 20:00:00	640.0
8	231 ST	2019-05-16 20:00:00	639.0
17	231 ST	2019-06-12 20:00:00	635.0
2	231 ST	2019-05-08 20:00:00	632.0
30	231 ST	2019-07-08 20:00:00	632.0
1	231 ST	2019-05-07 20:00:00	630.0
43	231 ST	2019-08-01 20:00:00	624.0