

Analysis of Student Frequencies at Metro Stations

Recommendation on display advertising in New York Subway

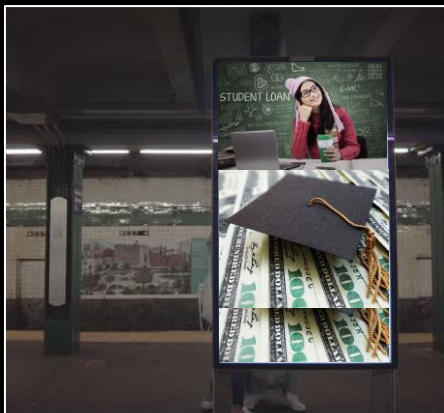
METIS - EDA

Fabian Paul, 23/03/2022

Introduction

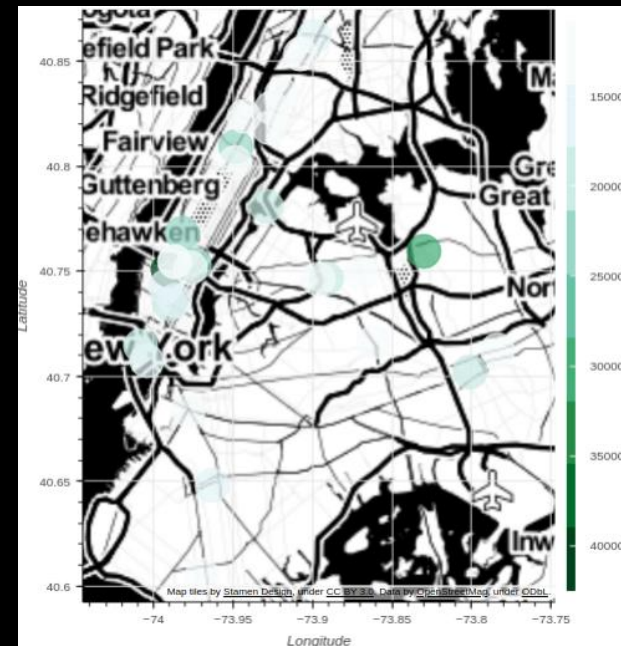
Business Problem

- Fintech firm: Buy display advertising on New York subways
- Offering: Loans to actual and prospective college students
- More than 600m cumulative subway station entries in 2021



Objective

- Which stations to launch display ads?
- Which time to advertise?
- Which metro lines to focus on?



Methodology

Turnstile



Daily data on entries per unit, control-area and station

Fare



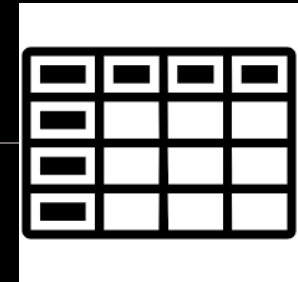
Weekly data on metrocard swipes per station

Coordinates & Lines



Longitude, latitude and metro lines per station

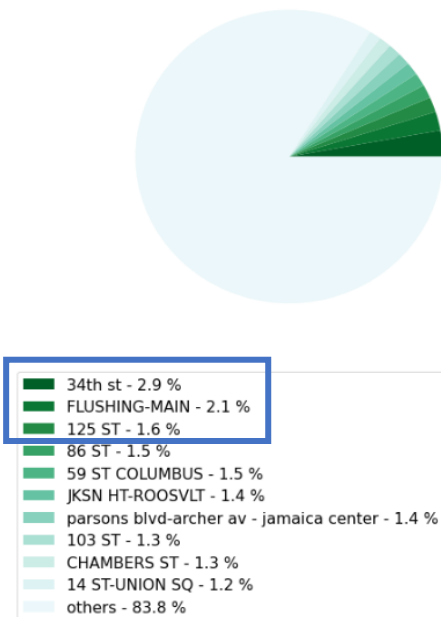
- 1) Calculate %-share in total of CUNY and HS in fare data ✓
- 2) Calculate daily entries in turnstile data ✓
- 3) Merge by weeknum with turnstile data ✓
- 4) Clean data and calculate daily entries ✓
- 5) Calculate number of different customer groups by multiplying daily entries with %-shares ✓
- 6) Merge with coordinates and lines ✓
- 7) Produce output tables and visualizations ✓



Results

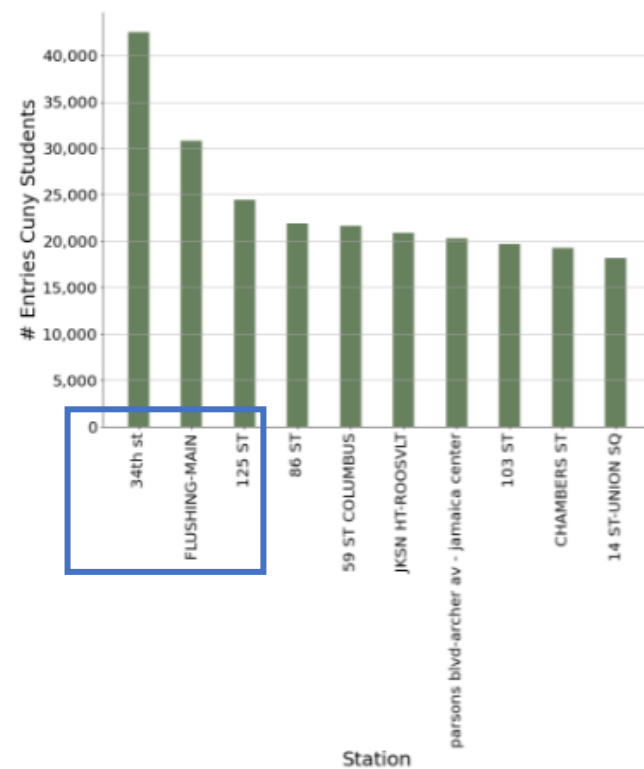
Stations

Top 10 Cuny Stations by all Cuny entries 2021



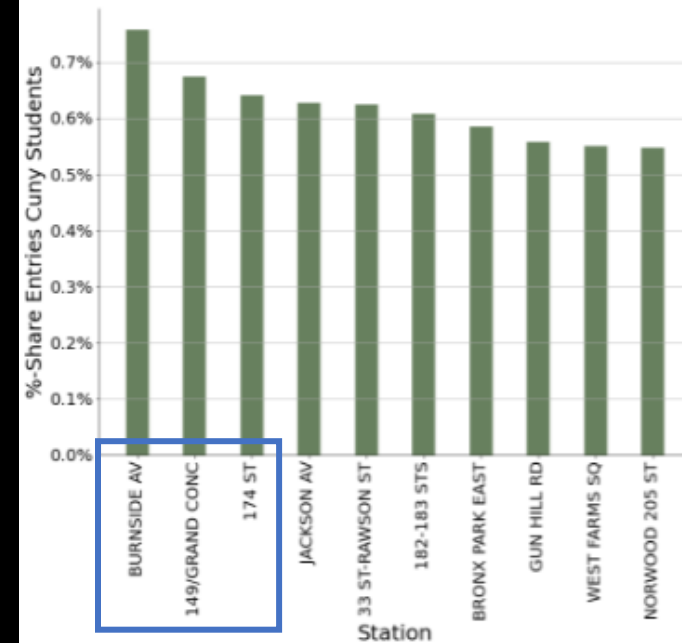
Recommendation with unlimited Budget

Top 10 Cuny Stations 2021 by total entries



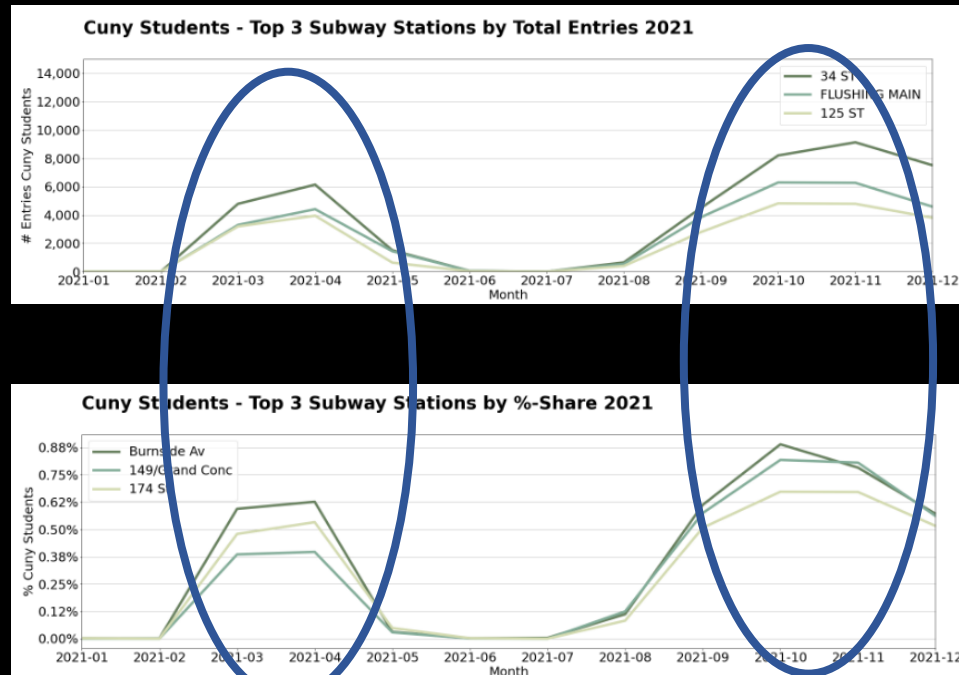
Recommendation with limited Budget

Top 10 Stations 2021 by %-Share Cuny Students



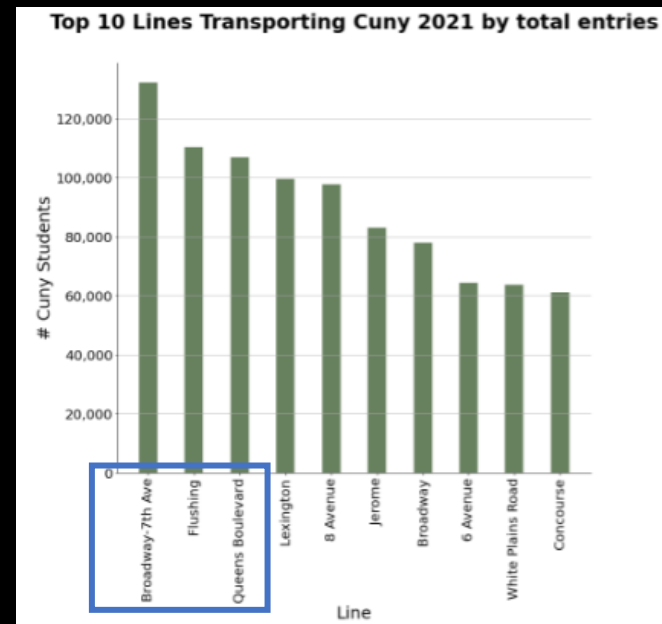
Results

Time

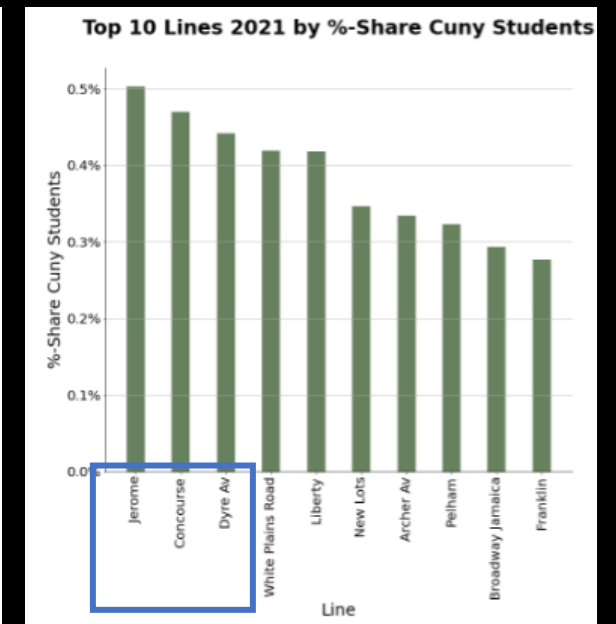


Metro Line

Recommendation with unlimited Budget



Recommendation with limited Budget



Conclusions

Results

- Top 3 stations to reach highest total no. of college students if ...
 - budget is unconstrained: 34th Main, Flushing-Main, 125 St
 - budget is unconstrained: Burnside Av, 149 / Grand Conc, 174 St
- Top 3 stations to reach highest total no. of college students if ...
 - budget is unconstrained: IRT Broadway 7th Av, IRT Flushing, IND Queens Boulevard
 - budget is unconstrained: IRT Jerome Av, IND Concourse, IRT Dyre Av
- Time to increase ad intensity: Spring & Fall/Winter

Assumptions & Limitations

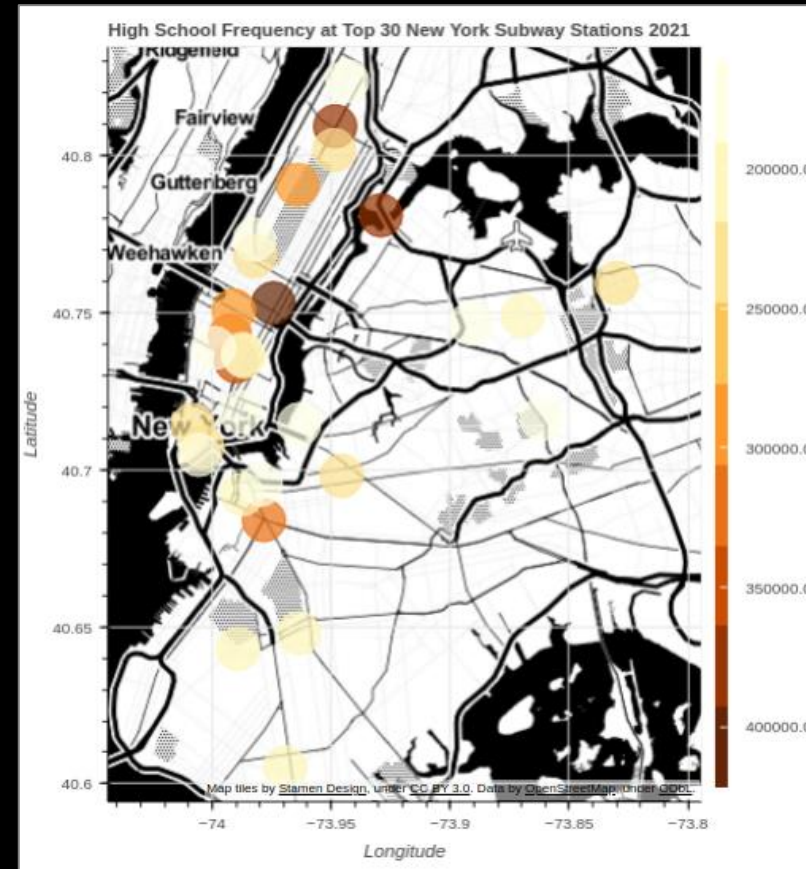
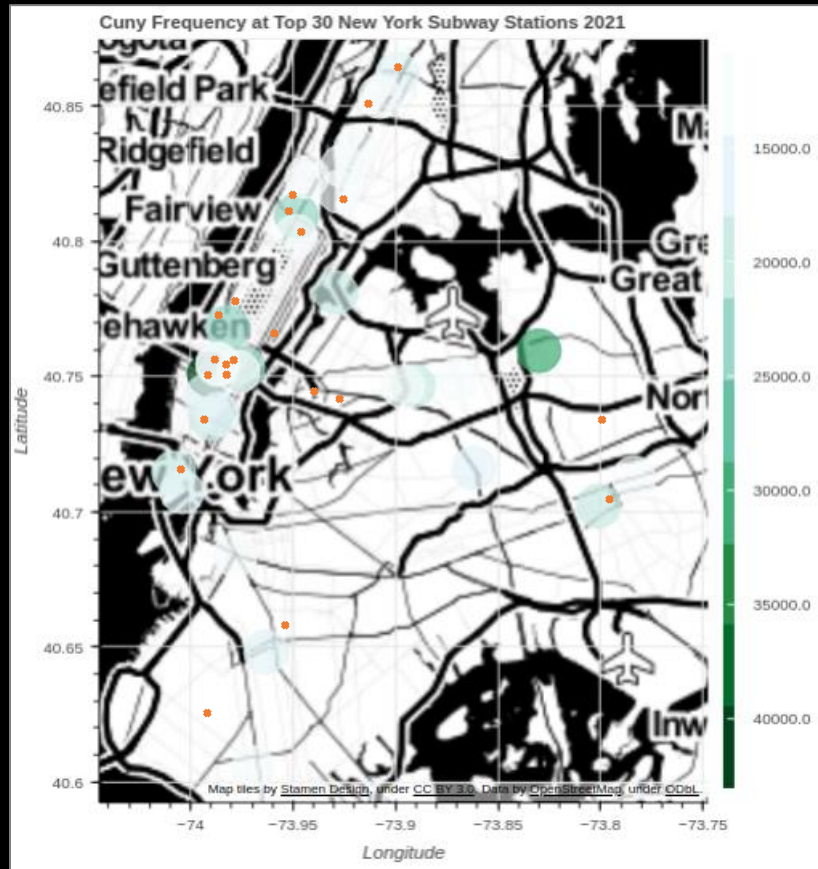
- Metrocard swipes as a proxy for Cuny student population in New York
- Representativeness of Cuny students to overall college student population in New York
- Calculation of Metro Line traffic: Each line receives equal share of entries at each metro station

Future Work

- ① → Connect station coordinates to zip-codes and analyze largest student frequencies by zip-codes
- ② → Find optimal point between “unlimited-budget” and “limited-budget” recommendations by introducing a combined metric
- ③ → Refine line information on student transportation by combining highest lines with highest stations (ads in which line between which stations?)
- ④ → Compare Entries in MTA turnstile data and Metrocard Swipes in MTA fares data and how analyze how they (don't) match
- ⑤ → Development over time for years 2018-2020 to check pre-Covid levels

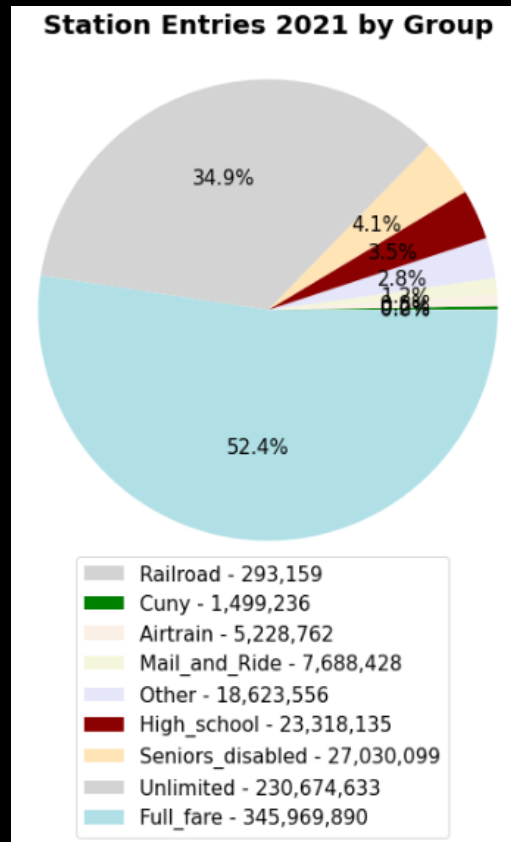
Appendix

Appendix | Map - Top Metro Stations



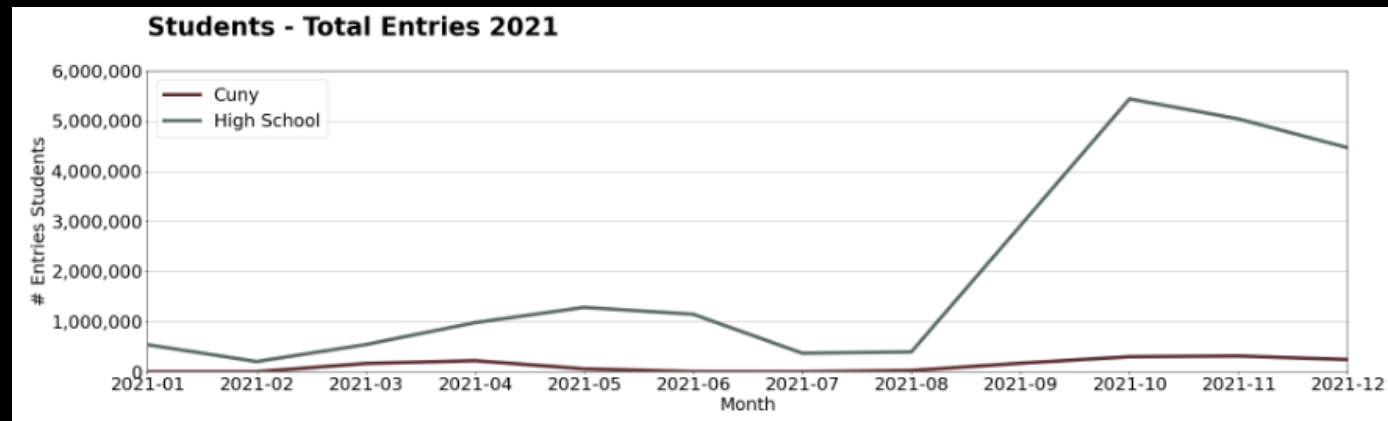
✦ College of City University of New York

Appendix | Volumes of Cuny and HS



Small percentages but large total entry volumes in 2021 may strengthen assumption of representativeness of sample

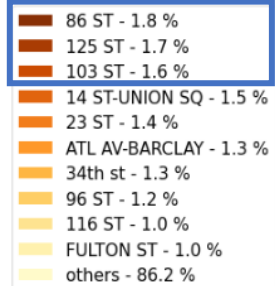
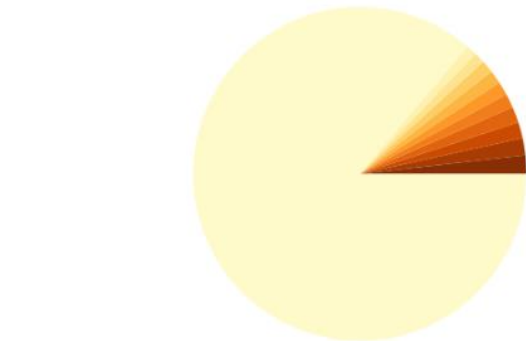
Fall/Winter seems to be even more preferable for advertising than spring/summer



Appendix | High School

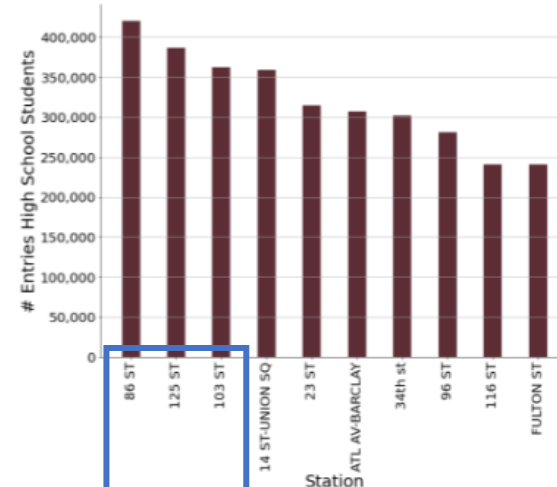
Stations

Top 10 High_School Stations 2021 by all HS entries 2021



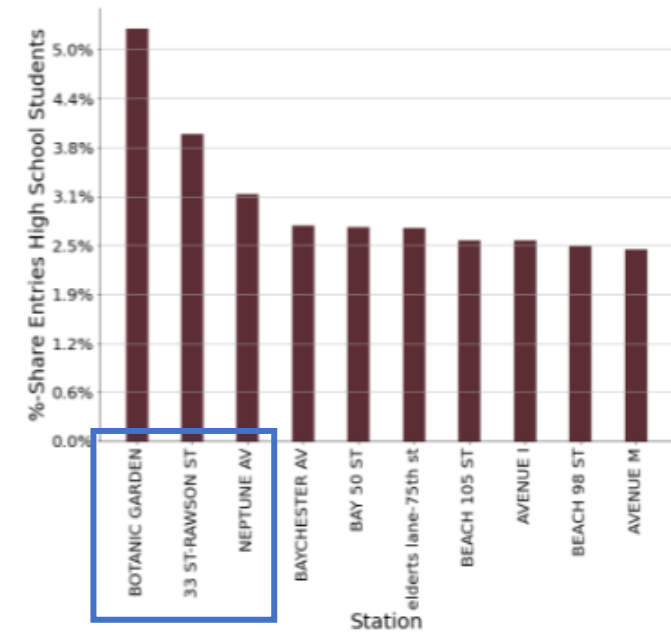
Recommendation with unlimited Budget

Top 10 High School Stations 2021 by total entries



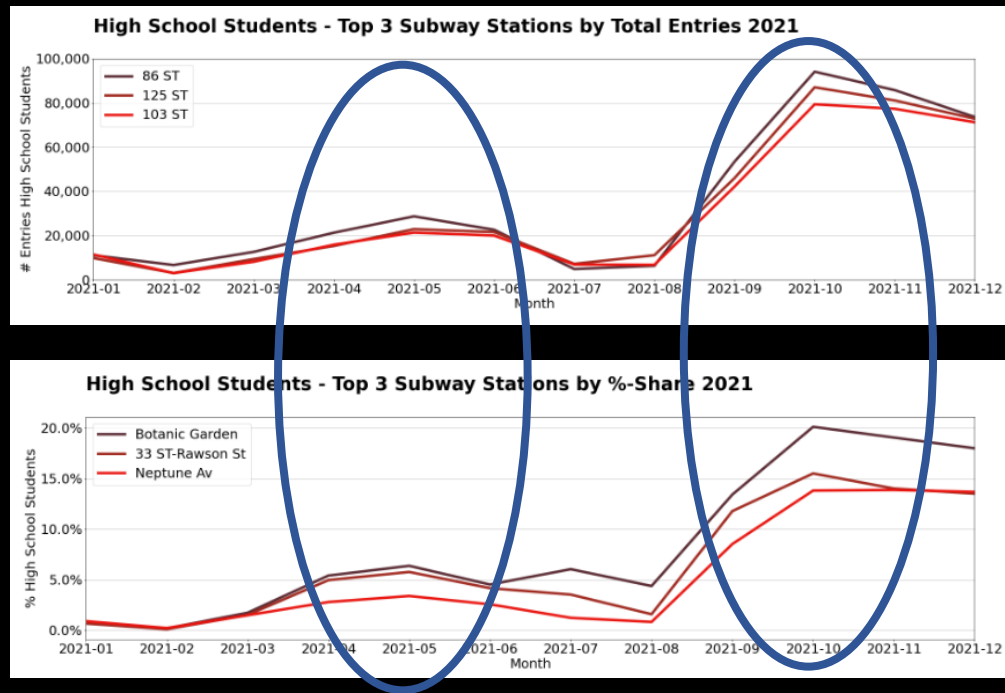
Recommendation with limited Budget

Top 10 Stations 2021 by %-Share High School Students



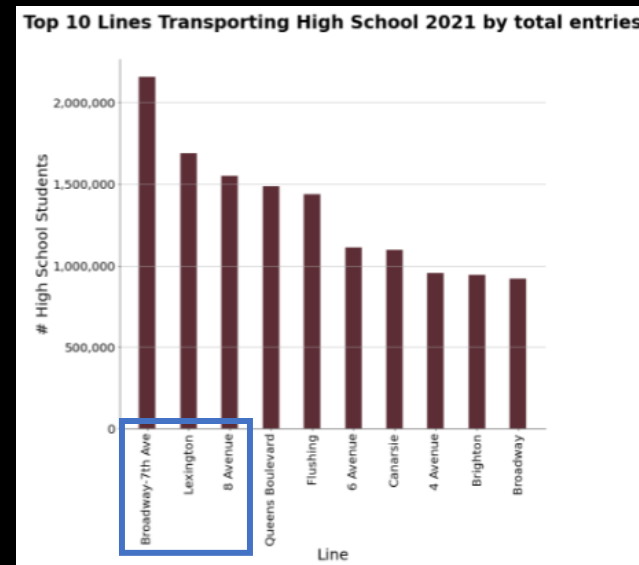
Appendix | High School

Time

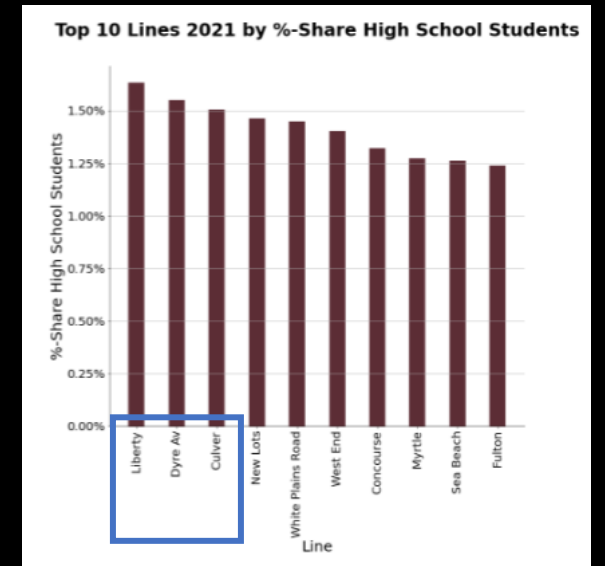


Metro Line

Recommendation with unlimited Budget



Recommendation with limited Budget



Appendix | Descriptive Statistics

Cuny

	station	daily_entries_cuny_students
0	34th st	43065.506188
1	FLUSHING-MAIN	30771.969448
2	125 ST	24636.949909
3	86 ST	22199.494041
4	59 ST COLUMBUS	21649.750541
5	JKSN HT-ROOSVLT	20899.989383
6	parsons blvd-archer av - jamaica center	20292.222351
7	103 ST	19781.781274
8	CHAMBERS ST	19540.334737
9	14 ST-UNION SQ	18165.886189

```
entries_stations_21['daily_entries_cuny_students'].describe()
count      377.000000
mean       3976.753899
std        4811.056507
min         0.000000
25%        1306.501260
50%        2345.693518
75%        4692.625836
max        43065.506188
Name: daily_entries_cuny_students, dtype: float64
```

High School

	station	%_high_school
0	BOTANIC GARDEN	0.210797
1	33 ST-RAWSON ST	0.157053
2	NEPTUNE AV	0.126034
3	BAYCHESTER AV	0.110322
4	BAY 50 ST	0.109223
5	elderts lane-75th st	0.108928
6	BEACH 105 ST	0.102677
7	AVENUE I	0.102446
8	BEACH 98 ST	0.099603
9	AVENUE M	0.097865

```
entries_stations_21['daily_entries_high_school_students'].describe()
count      377.000000
mean       61851.816098
std        64737.950301
min         0.000000
25%        23407.880399
50%        41844.516064
75%        75304.366233
max        420424.547426
Name: daily_entries_high_school_students, dtype: float64
```

	station	daily_entries_high_school_students
0	86 ST	420424.547426
1	125 ST	386570.856872
2	103 ST	362823.856778
3	14 ST-UNION SQ	359055.873826
4	23 ST	314849.911767
5	ATL AV-BARCLAY	307023.550579
6	34th st	302423.819916
7	96 ST	280873.027634
8	116 ST	240906.600936
9	FULTON ST	240902.676475

Thank you for your attention!