MTA Turnstile Analysis

Juliana McCausland

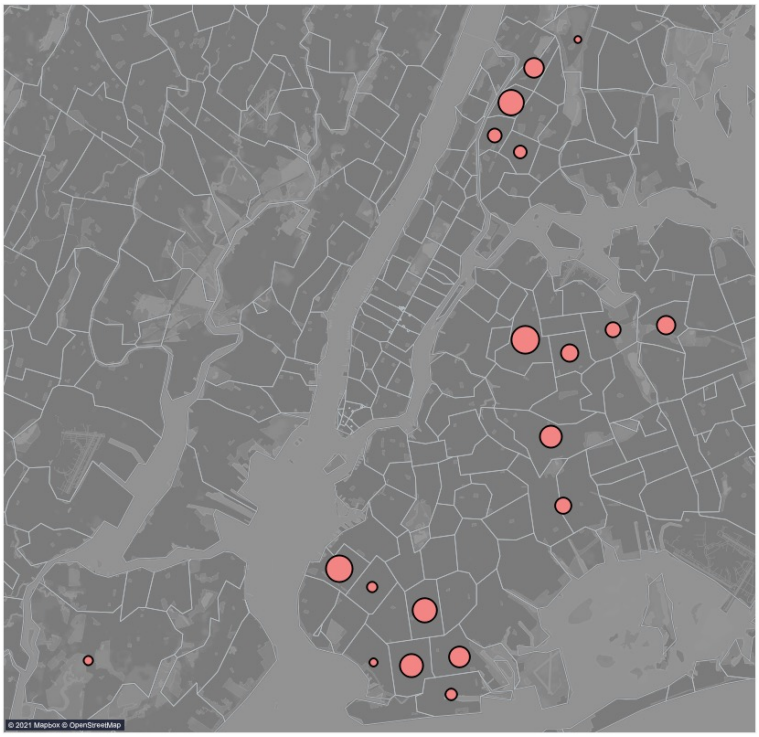
Background

- Health insurance company: can we utilize information about customer commutes to identify high-risk individuals who may need better coverage, particularly in crises similar to the current pandemic?

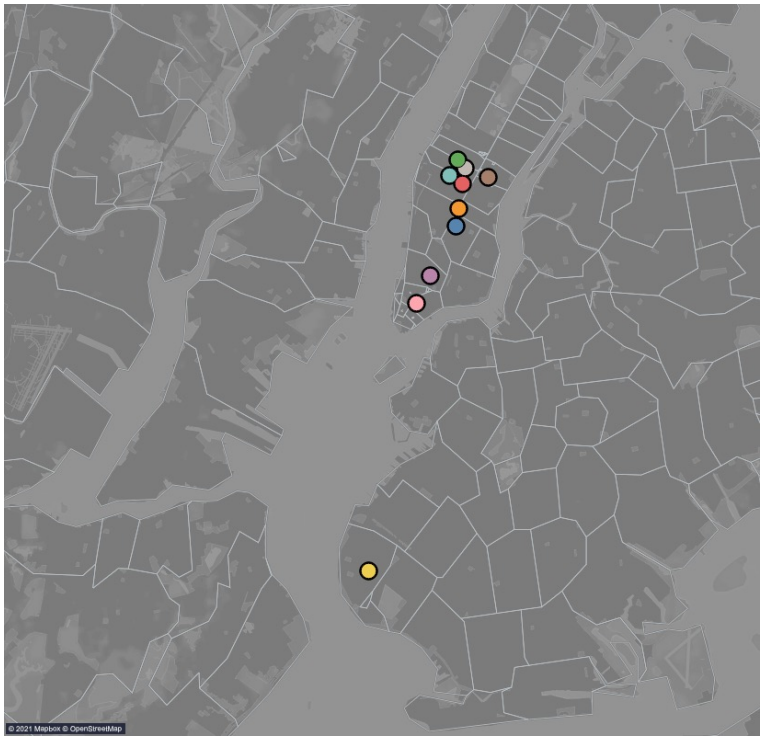
Methodology

- Analyze stations with high traffic and COVID hot spots
- Data:
  - MTA Turnstile Data: http://web.mta.info/developers/turnstile.html
  - MTA Stations with geographic coordinates: https://hub.arcgis.com/datasets/8c547512607847789e8b7648323dc462
  - COVID cases in NYC by zip code: https://github.com/nychealth/coronavirus-data/blob/master/totals/data-by-modzcta.csv

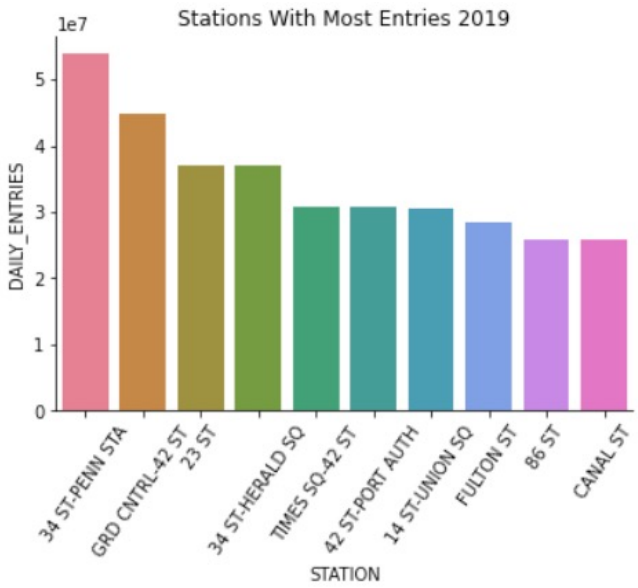
Locations with Highest COVID-19 Rates

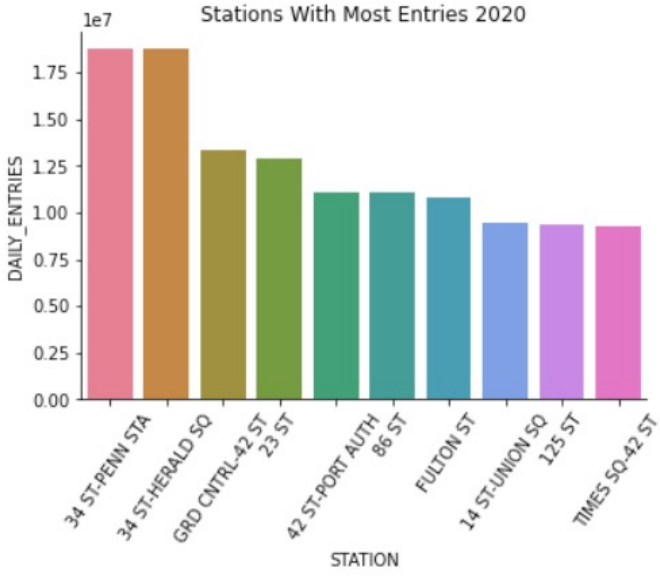


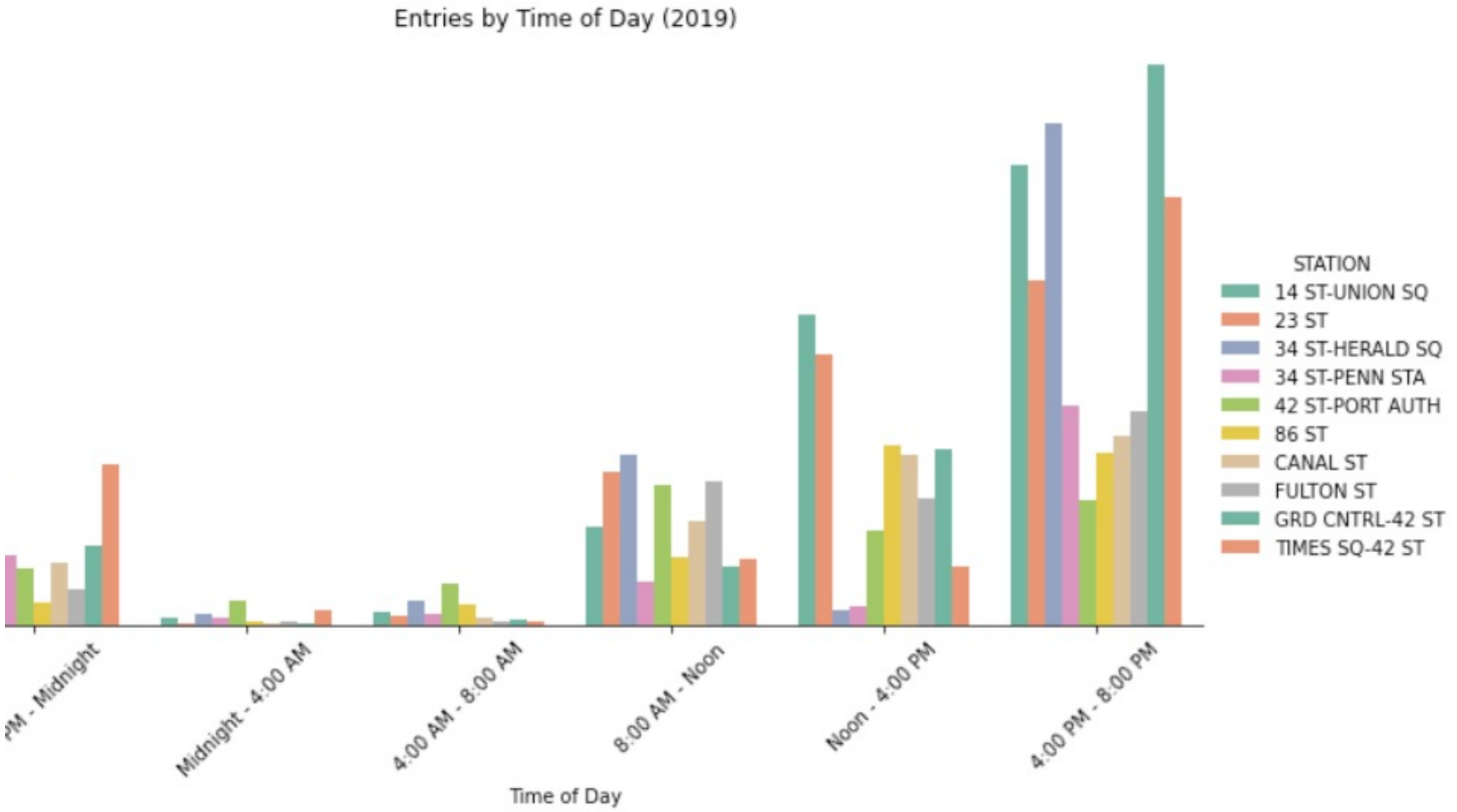
Most Populated MTA Stations

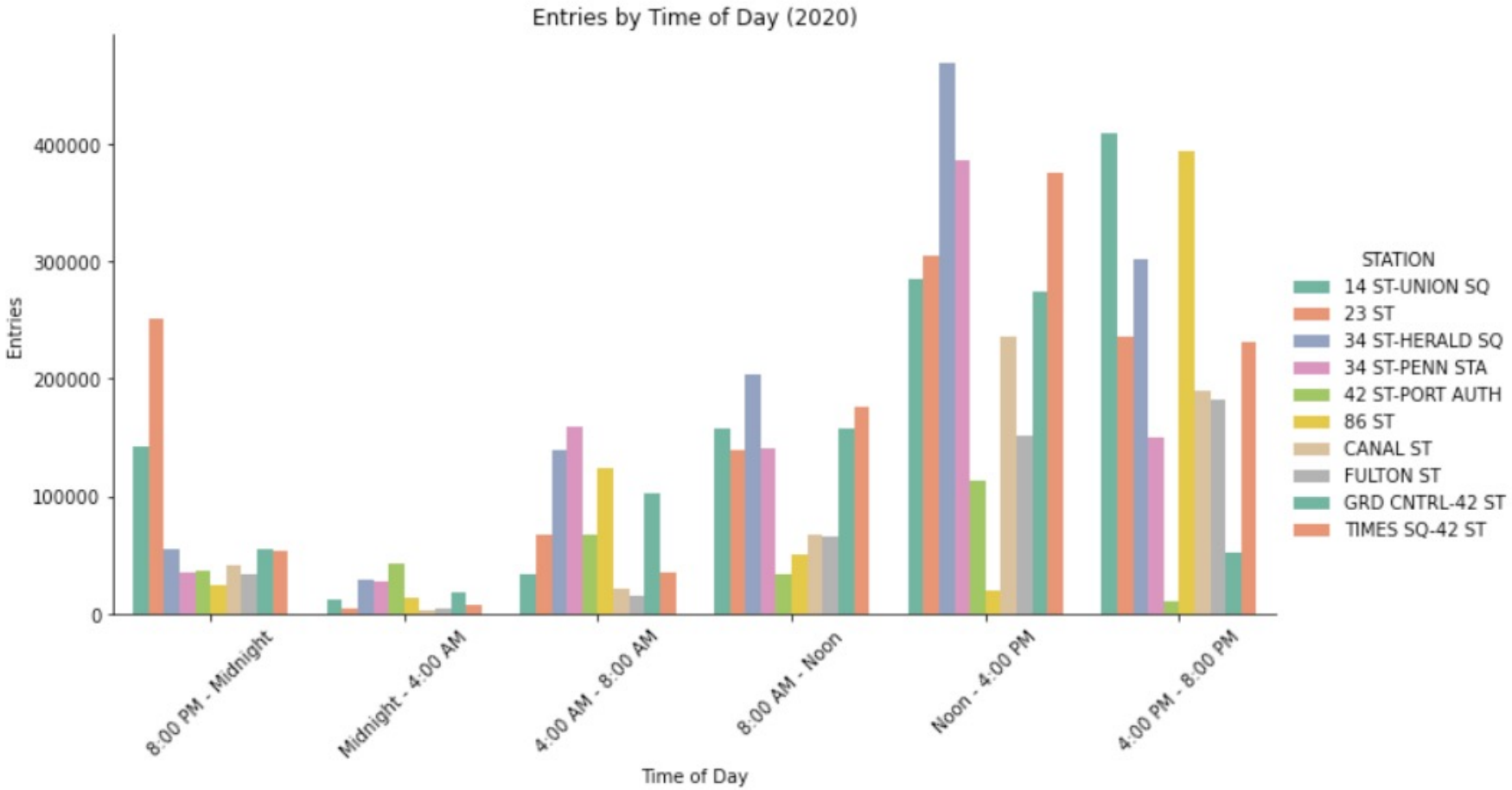


Most Populated Stations






Entries by Time of Day (2019)
STATION
14 ST-UNION SQ
23 ST
34 ST-HERALD SQ
34 ST-PENN STA
42 ST-PORT AUTH
86 ST
CANAL ST
FULTON ST
GRD CNTRL-42 ST
TIMES SQ-42 ST
PM - Midnight
Midnight - 4:00 AM
4:00 AM - 8:00 AM
8:00 AM - Noon
Noon - 4:00 PM
4:00 PM - 8:00 PM
Time of Day


Entries by Time of Day (2020)
Entries
0
100000
200000
300000
400000
8:00 PM - Midnight
Midnight - 4:00 AM
4:00 AM - 8:00 AM
8:00 AM - Noon
Noon - 4:00 PM
4:00 PM - 8:00 PM
Time of Day
STATION
14 ST-UNION SQ
23 ST
34 ST-HERALD SQ
34 ST-PENN STA
42 ST-PORT AUTH
86 ST
CANAL ST
FULTON ST
GRD CNTRL-42 ST
TIMES SQ-42 ST

Conclusions and Further Analysis

- Less MTA commuters overall – decreased risk for exposure in general
- Most populated stations are outside of the top 15 COVID hot spots
- A subset of people are still going into work – essential workers?

- A deeper analysis should include:
  - MTA exit data
  - Investigating average income of customers' neighborhoods