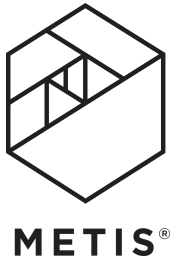


MTA Project

Recommendation on commute time and station

Introduction

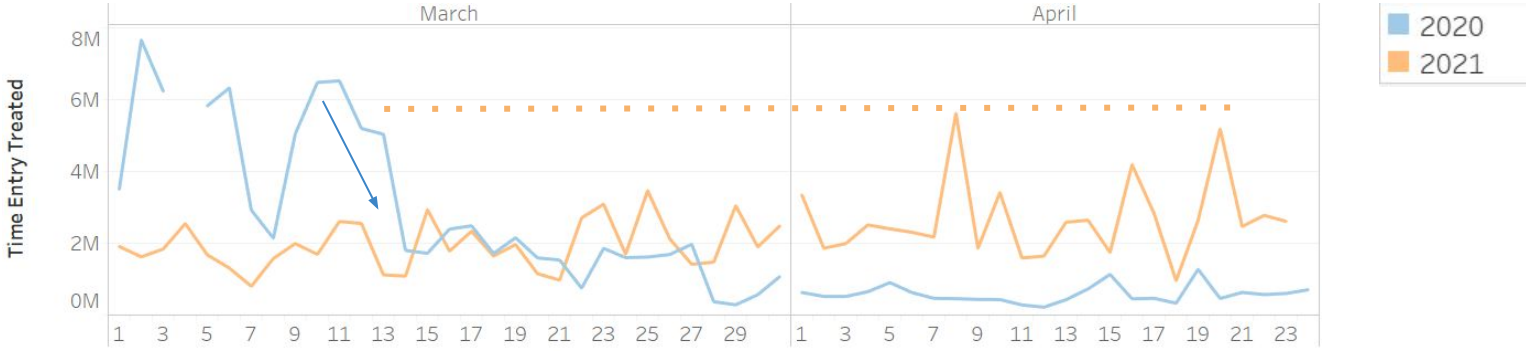
- **Motivation**: positive Covid cases is rising. The client - a small company - wants to advise their employees on how to minimize exposure to the virus while taking public transportation.

- **Goal**: per CDC guidance, one should avoid close contact (within about 6 feet) with other people, which translates to avoiding crowds in a closed small space. What we need to find out is the lower trafficked times and location.

- **Key questions to address:**
  - What time of the day do people use public transportation?
  - How many people visit a certain station on a daily basis?
  - Any stations to pay specific attention to? (e.g. popular stations)

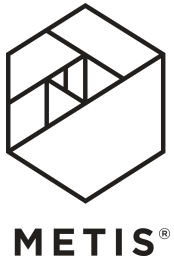
Motivation: Covid on the raise

New Positive Covid Cases



Daily Traffic

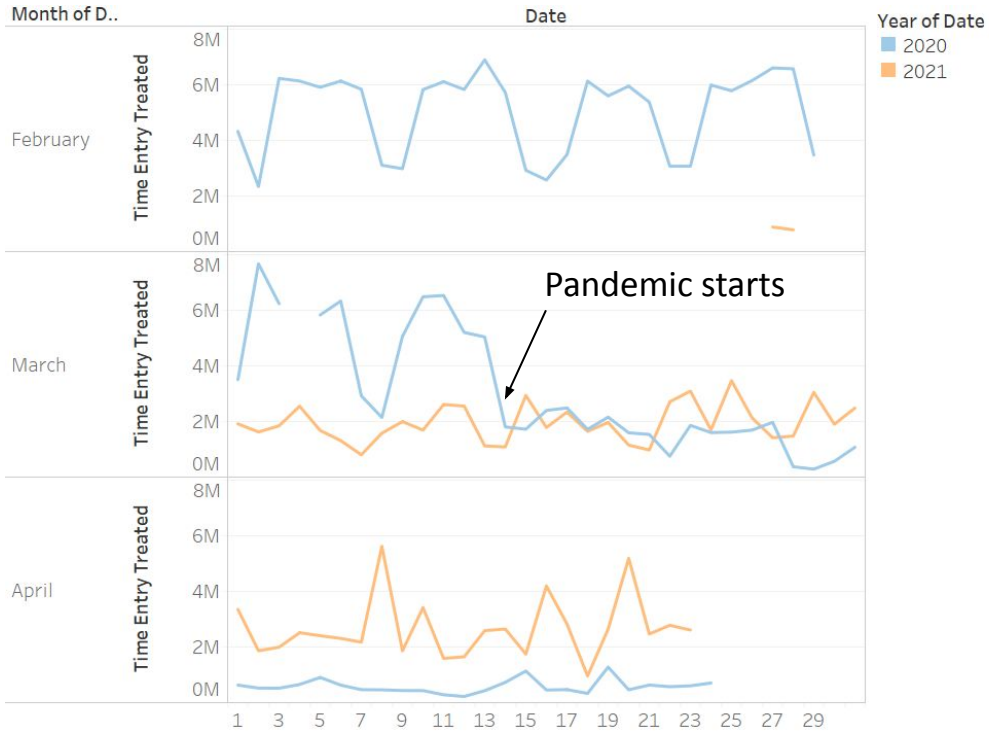
Methodology



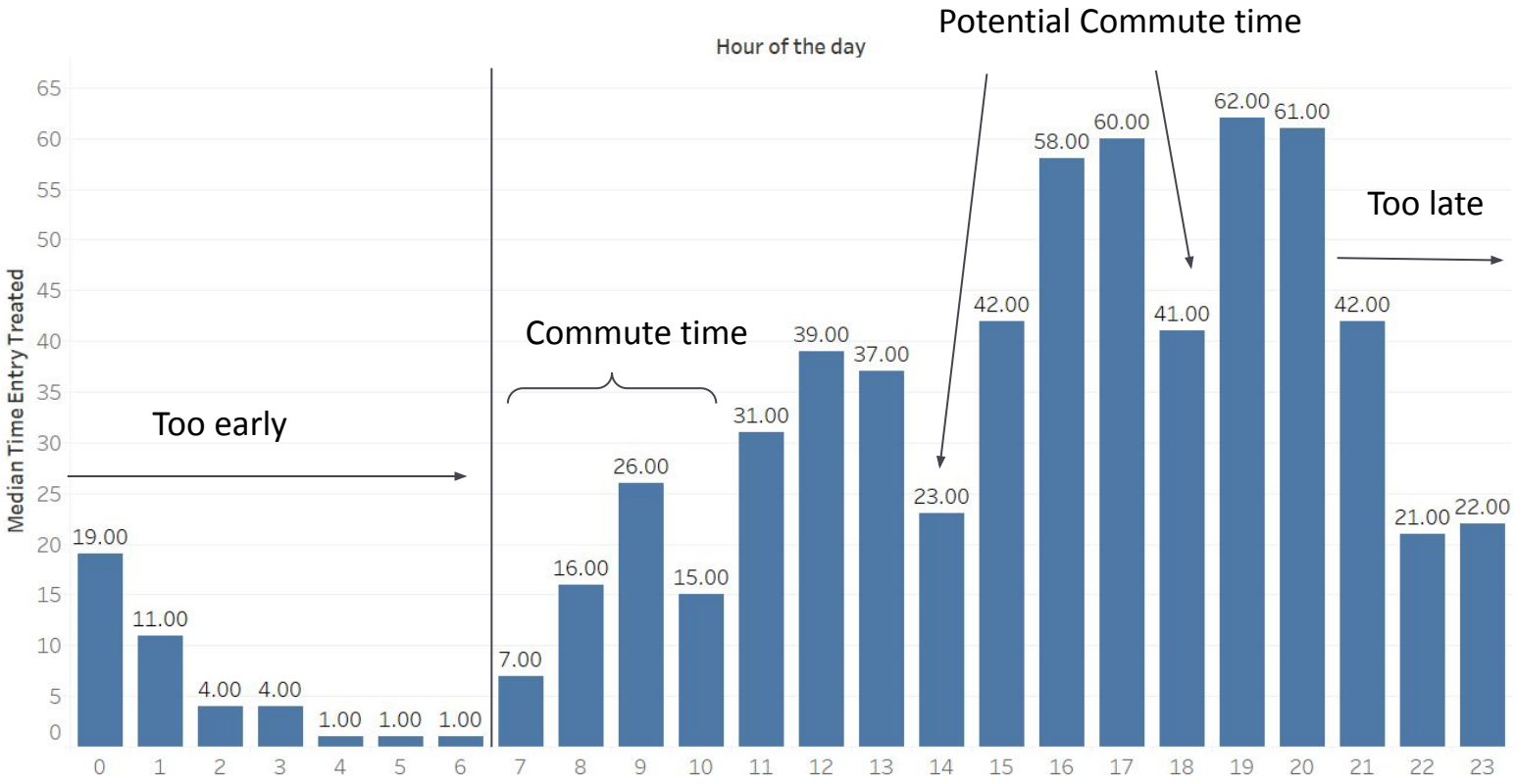
- **Data**: 2021 March to April MTA data, Covid positive test case data
- **Metric**:
  - At station level: total entry divide by count of turnstiles - this is an estimate of how crowded a station can get. (Assumption: more turnstiles means larger station)
  - At station level: total entry number
  - At hourly level: median entry number
  - At daily level: new positive covid cases and daily entries
- **Tools**: Tableau, Python, Seaborn, SQL
- **Data Treatment**: filter out 2020 and 2021 March and April data, solve reverse count turnstiles, take out extreme values(outliers), union 2020, 2021 data. Add covid positive test cases.

What time frame to use?

- Weekly pattern before pandemic
- Volume and pattern changed significantly since March 2020
- Conclusion: use 2021's data

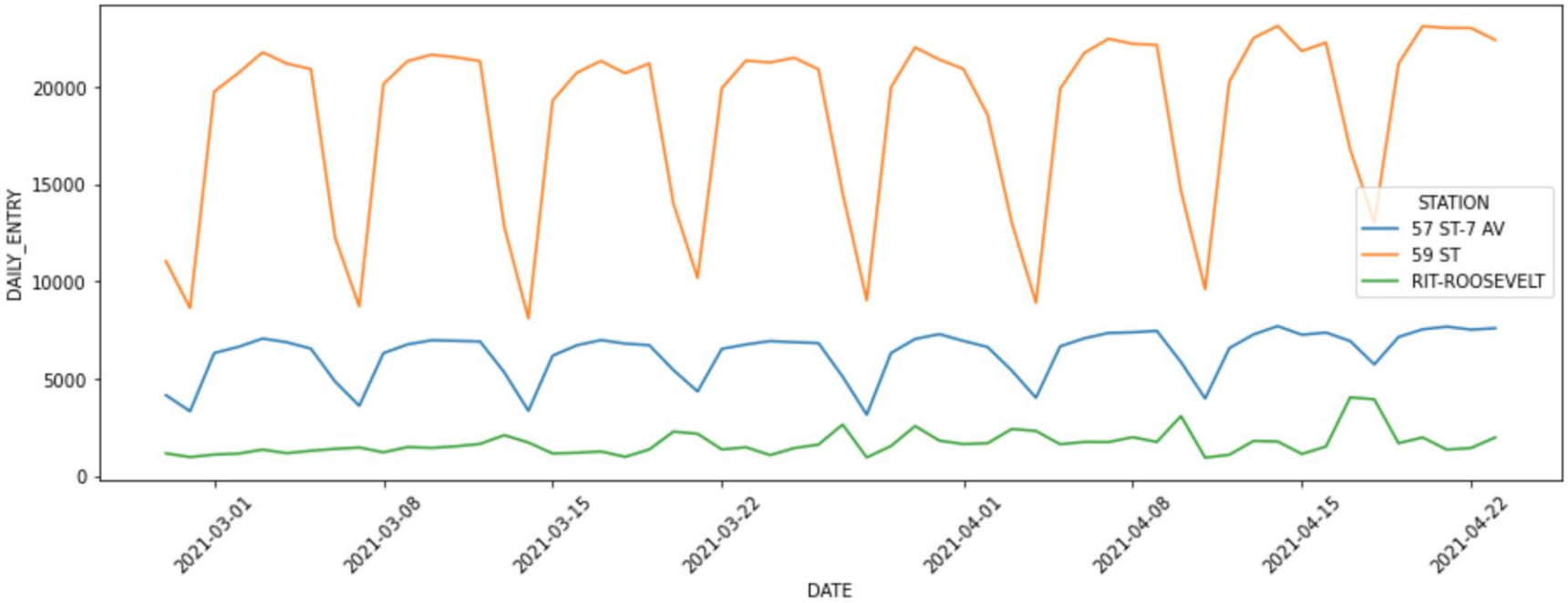


What time of the day should one travel?



Traffic varies between stations

Just knowing the time isn't enough to help the employee navigate, let's look at the traffic between different stations



Any popular station to avoid?

- Assumption: the more turnstiles, the larger the station
- Filtered for potential commute hour: 2pm and 6pm
- For more popular stations e.g. Kew garden, penn station, port authority, it's best to commute at 2pm.

Station	Date / Hour of the day: March 14	March 18	April 14	April 18
KEW GARDENS	1,392	1,547	1,826	2,063
34 ST-PENN STA	1,219	1,710	1,692	2,354
42 ST-PORT AUTH	1,045	1,809	1,435	2,552
5 AV/59 ST	619	2,389	905	3,160
5 AV/53 ST	431	1,627	615	2,221
JOURNAL SQUARE	527	484	464	378
WALL ST	317	833	444	1,108
CHRISTOPHER ST	324	760	287	524
GROVE STREET	353	441	287	353
NEWARK BM BW	372	463	280	368
THIRTY THIRD ST	263	807	222	626
14TH STREET	221	586	197	480
EXCHANGE PLACE	222	255	182	181

Time Entry/Co..
4 3,160

Conclusion

Recommendation:

1. The client should encourage employees to commute early between 7-10 am.

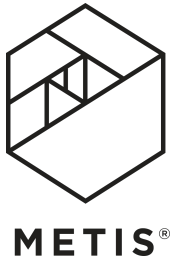
2. Since the traffic doesn't cool down until 10 pm, the best time to commute back is 2pm, followed by 6pm.

3. If one employee needs to take trains in a popular station (e.g. Kew Garden, Penn Station), it's highly recommended to leave before 2pm.

4. Assuming 8 working hours with 1 hour break, two possible work schedule:
    1. 7am - 2pm + 1 hr work/commute time
    2. 2pm to 10pm

Interesting finding: The overall 2021 traffic pattern is different from 2020, but when zoom into some specific stations, they still follow a weekly pattern.

Future Work



Potential deep dive area: geom data, correlation between number of positive covid cases and people using one popular station. (e.g. Queens)