

Mask Distribution

In New York stations to reduce
Covid spread



Objective

- To inform MTA about opportunities to optimize distribution of masks in New York MTA stations to mitigate COVID as an endemic and reduce its transmission.

Data

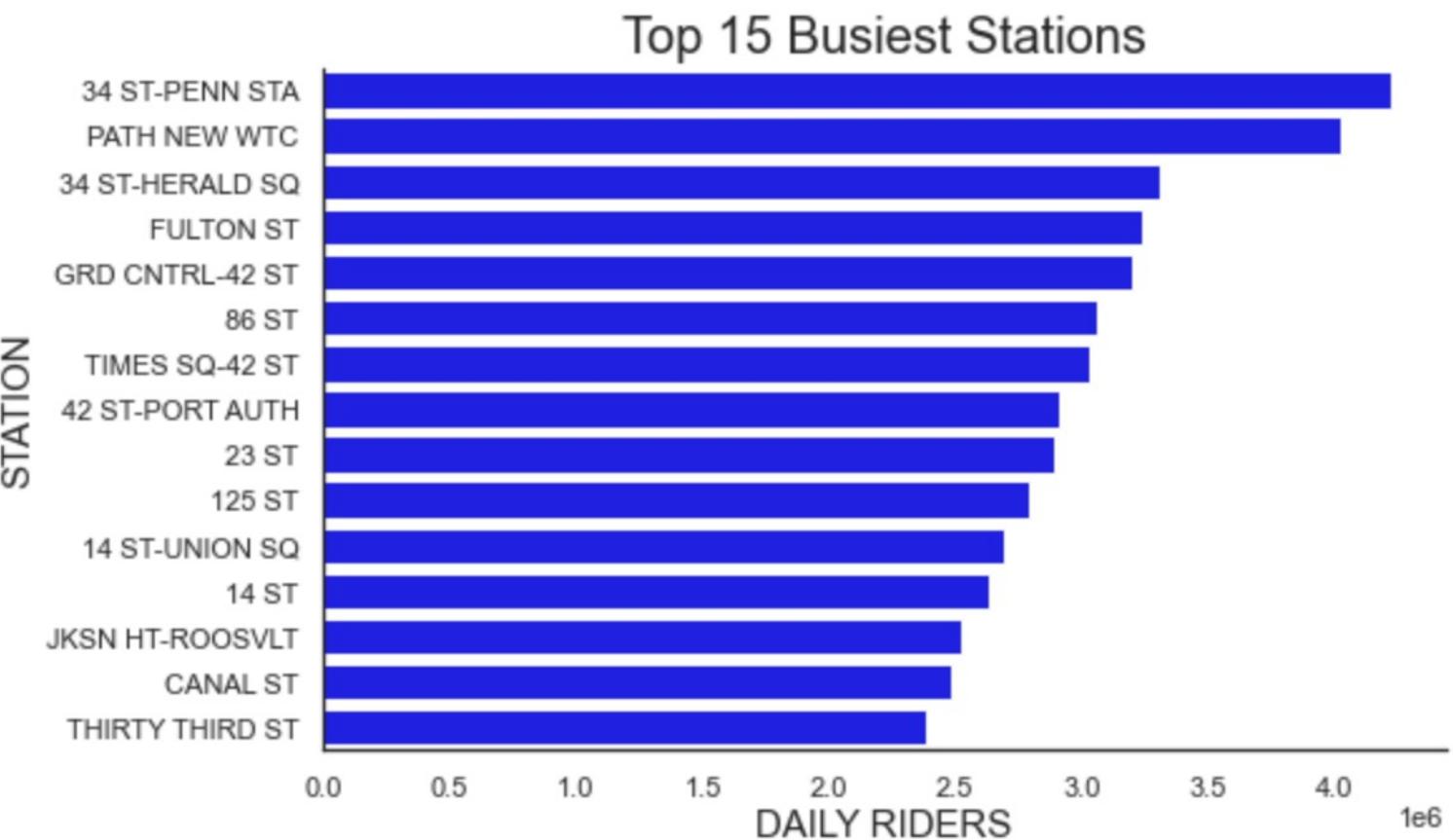
Data sources

- MTA Data
- Covid-19 cases from NYC Open Data
- 3 months – Nov. 2021 to Jan. 2022

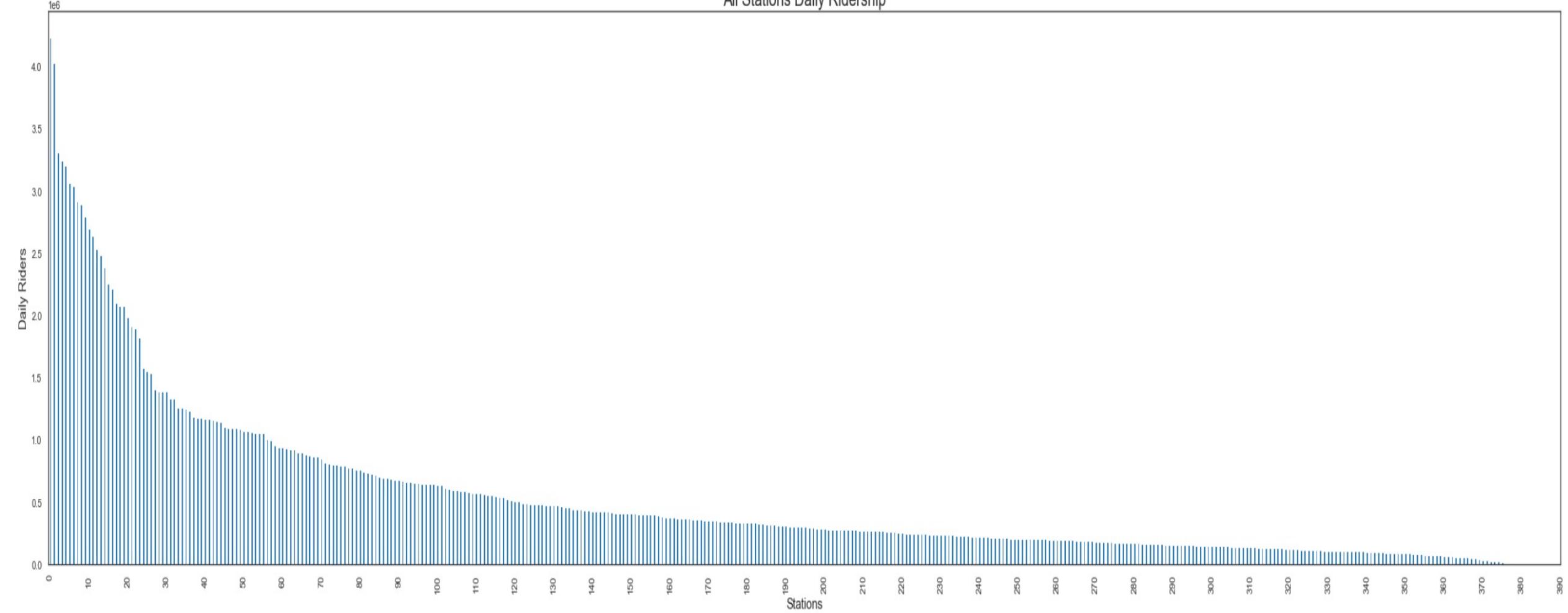
Datasets created

- MTA top busiest stations.
- Daily count of riders in all the stations combined and individual stations
- Daily Covid cases count
- Daily Riders and Covid cases merged

Top 15 stations
contribute to
21% of total
MTA riders
(378 stations)



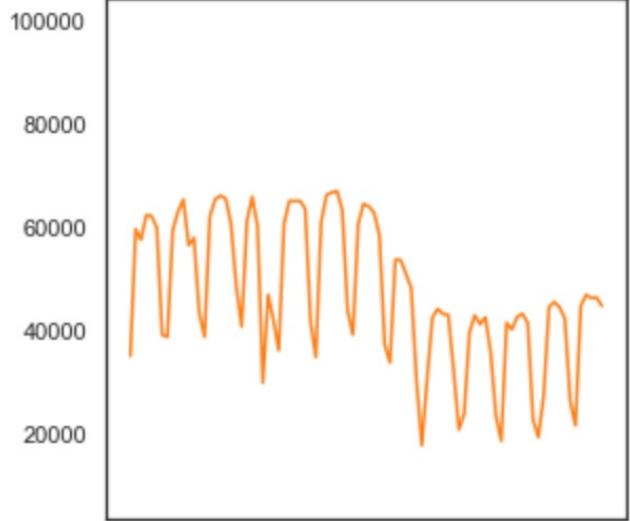
All Stations Daily Ridership



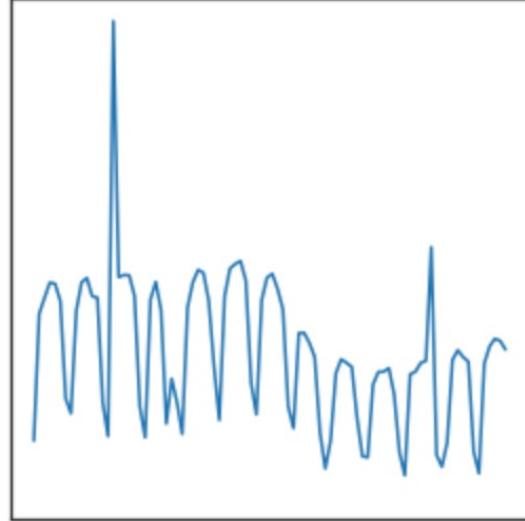
By choosing top 80 heavily populated stations, we can cover 60% of total MTA riders. This graph helps us to choose number of stations to include in order to cover a desired percentage of MTA population.

EDA: The ridership time trends indicate that the effects of COVID (declines) were largely uniform across stations

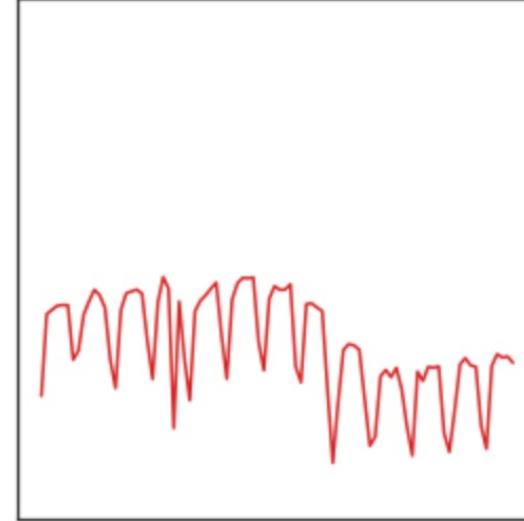
34 ST-PENN STA



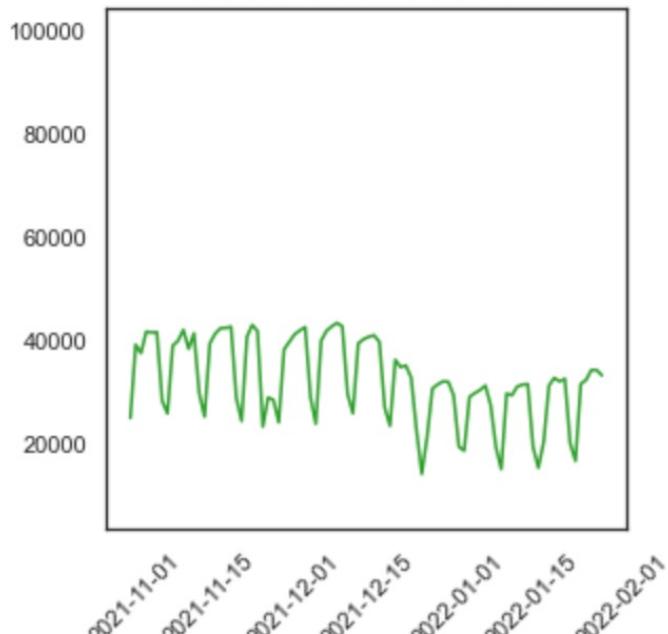
GRD CNTRL-42 ST



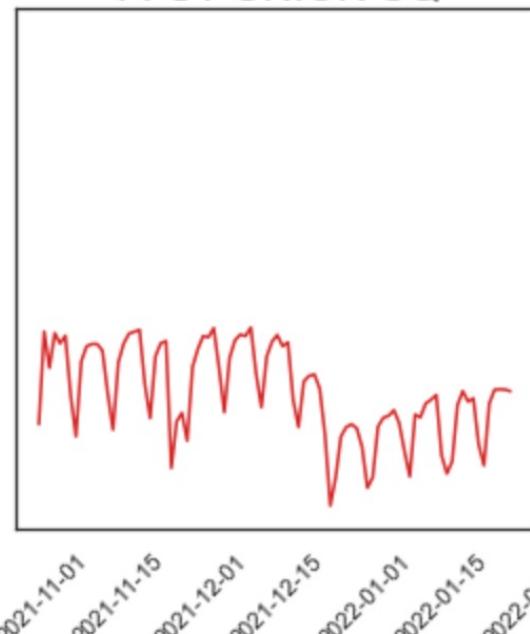
34 ST-HERALD SQ



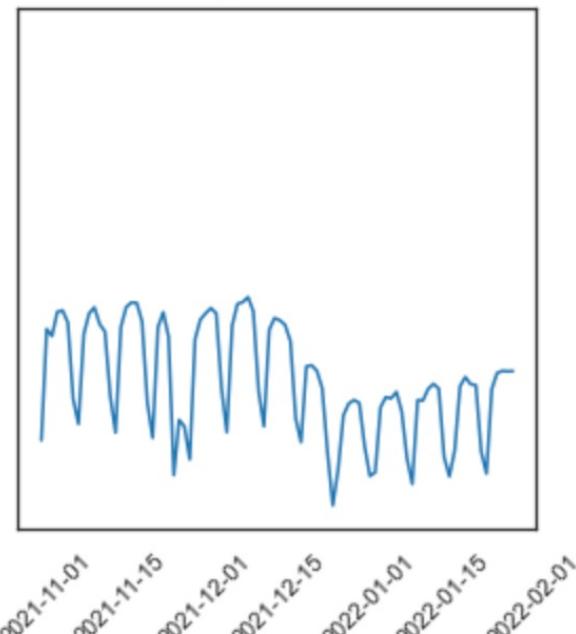
42 ST-PORT AUTH



14 ST-UNION SQ

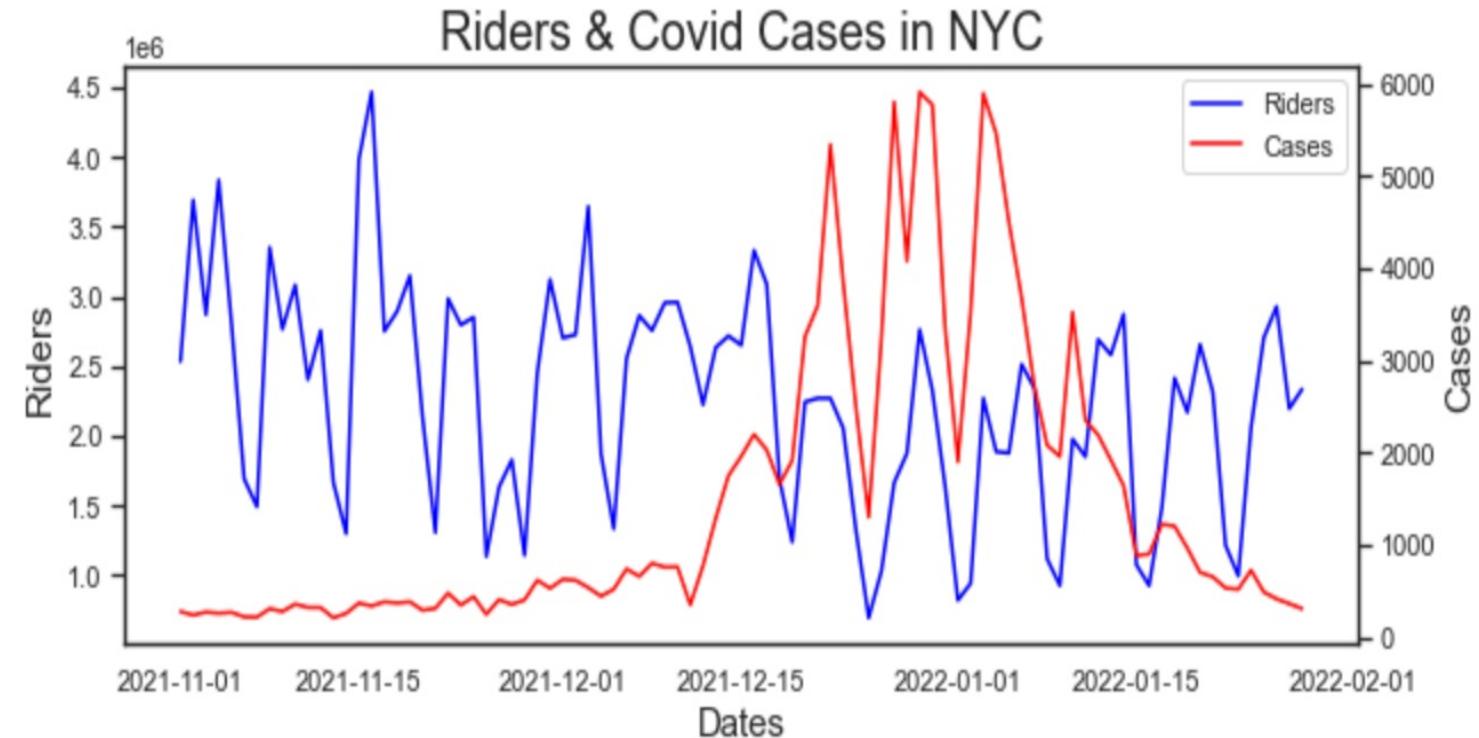


23 ST



Insights

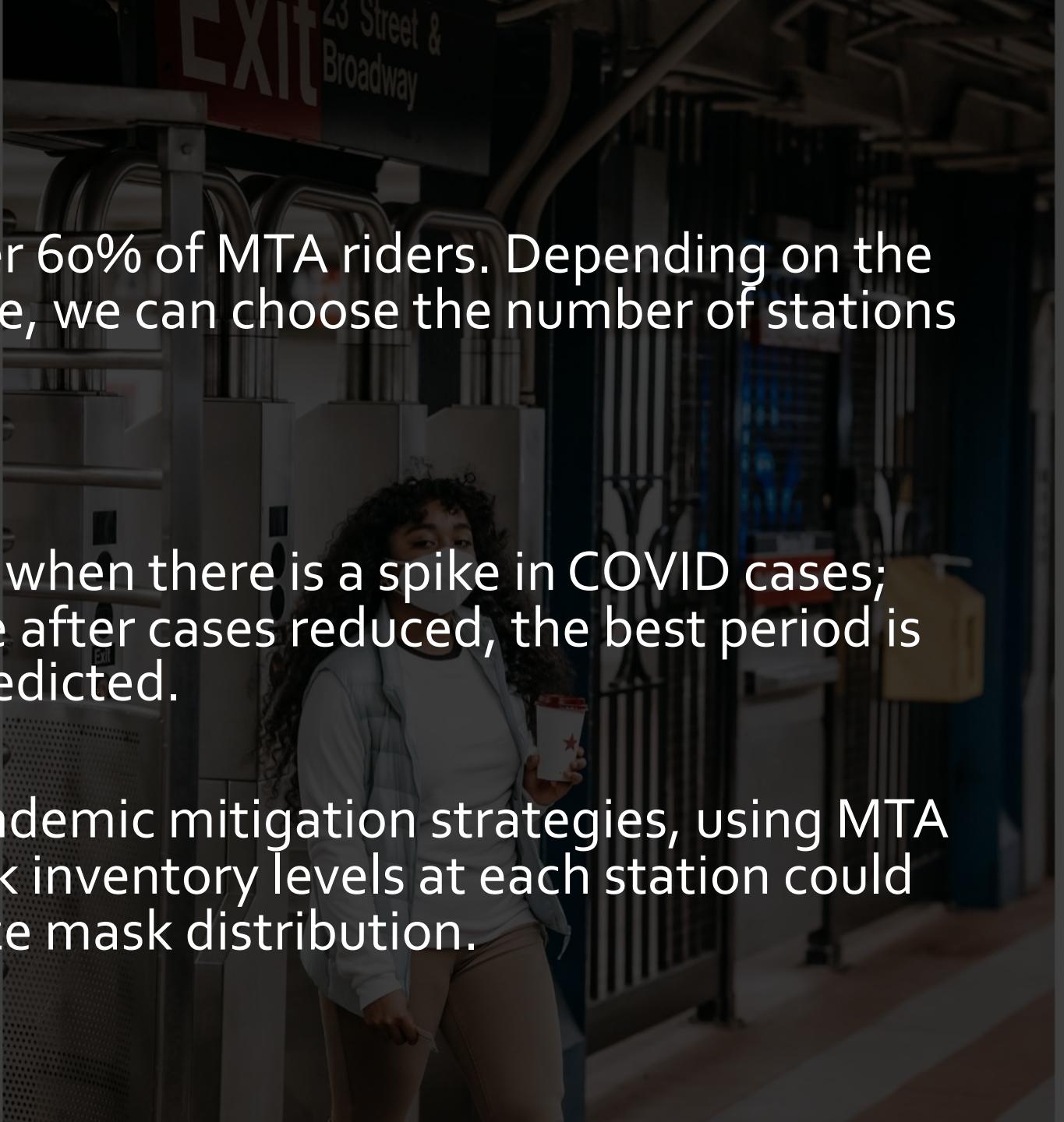
1. Average ridership dropped by 27% as COVID wave spiked between 12/20 – 1/15.
2. Once the COVID wave subsided, average ridership still showed a drop of 21% as compared to pre-spike dates.



Conclusion

- **Where to place masks?**
 - Top 80 busiest stations cover 60% of MTA riders. Depending on the desired percentage coverage, we can choose the number of stations accordingly.
- **When to place masks?**
 - Since the ridership dropped when there is a spike in COVID cases; increased at a very slow rate after cases reduced, the best period is anytime before a spike is predicted.

Next steps: For longer-term endemic mitigation strategies, using MTA ridership data to forecast mask inventory levels at each station could be a good next step to optimize mask distribution.



Appendix

- MTA Data source link
<http://web.mta.info/developers/turnstile.html>
- COVID 19 Data source link
<https://data.cityofnewyork.us/Health/COV-ID-19-Daily-Counts-of-Cases-Hospitalizations-an/rc75-m7u3>

Appendix

- Code:

```
#Calculating the percentage of riders belonging to top n  
stations compared to all the stations  
  
total_riders_top =  
Densely_Pop_Station.DAILY_ENTRIES[:80].sum()  
  
total_riders = sum(Densely_Pop_Station.DAILY_ENTRIES)  
  
Per_top_over_total = (total_riders_top/total_riders)*100  
print(Per_top_over_total)  
# it is coming to 59.4%
```

Appendix

Code:

```
#Calculating average ridership Drop at the begining of Covid spike and at the end  
date_before = np.datetime64('2021-12-20')  
date_during = np.datetime64('2022-01-15')  
mean_before_wave = station_by_date[station_by_date['DATE2'] < date_before].mean()  
#2597296.06
```

```
mean_during_wave = station_by_date[(station_by_date['DATE2'] > date_before) &  
(station_by_date['DATE2'] <= date_during)].mean()  
#1891555
```

```
mean_after_wave = station_by_date[station_by_date['DATE2'] > date_during].mean()  
#2031262.0
```

```
drop_percent = ((mean_before_wave-mean_during_wave)/mean_before_wave)*100  
#27.17%
```

```
current_percent = ((mean_before_wave - mean_after_wave)/mean_before_wave)*100  
#21.79%
```