

# Planes, Trains, and Afflictions

Agent-based modeling for the spread of disease through transportation networks

NS Group 3

Faculty of Computer Science  
Higher School of Economics  
Moscow

27 JUN 2020

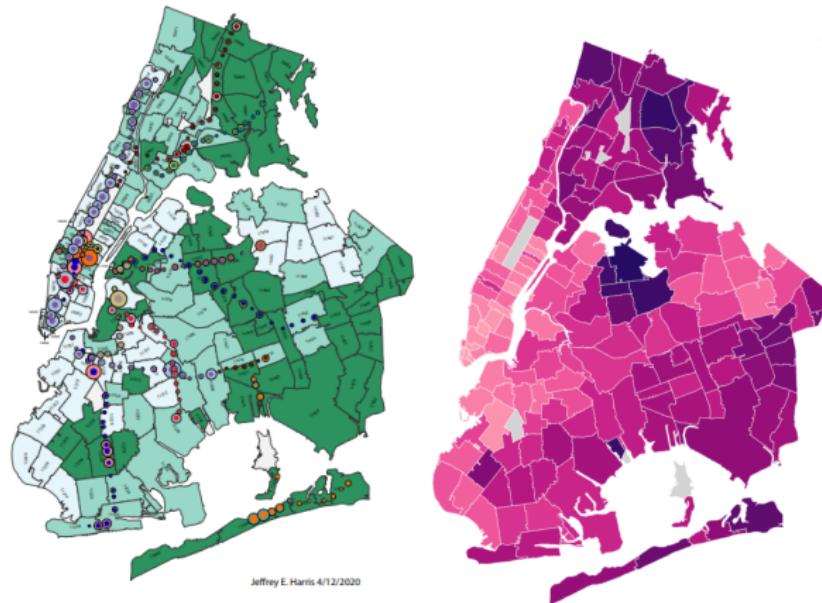
# Table of Contents

- Motivation
- Prior Research
- ABM Framework
- Examples
- Deep Dive: New York City (NYC) Subway
- Discussion
- Conclusion
- Credits

# Motivation

Why were certain areas of NYC more affected by COVID-19?

How were the subways involved?

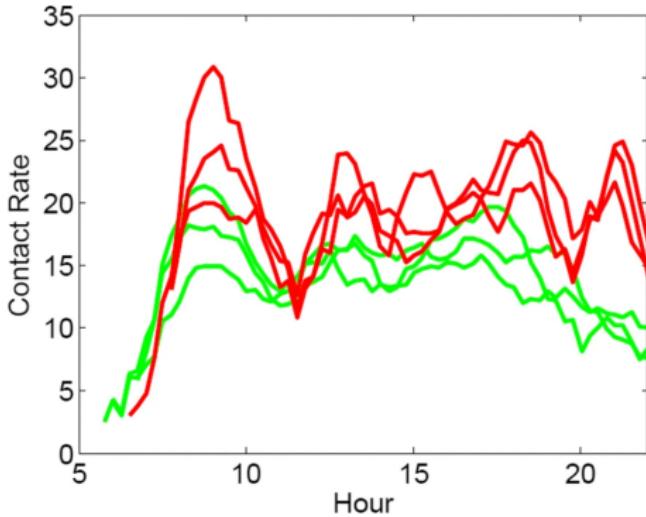
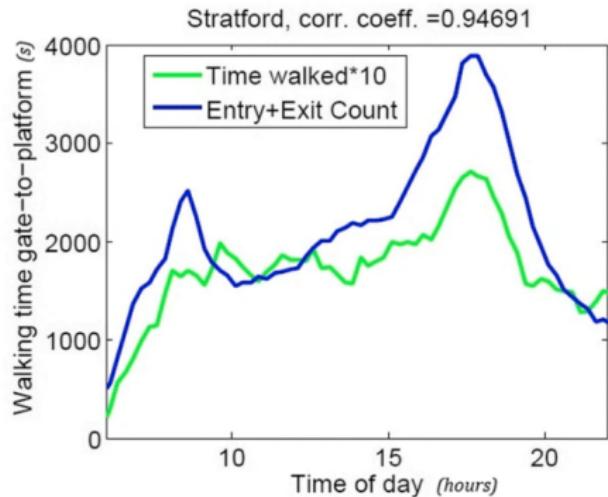


**Figure:** Left: Case Rate And Subway On April 12, 2020 [7]

Right: NYC COVID Case Rate As Of June 23, 2020 [1]

## Prior Research (London Underground)

Data: Oyster (Card), CASA (Timetable), PHE Data (Demographics)



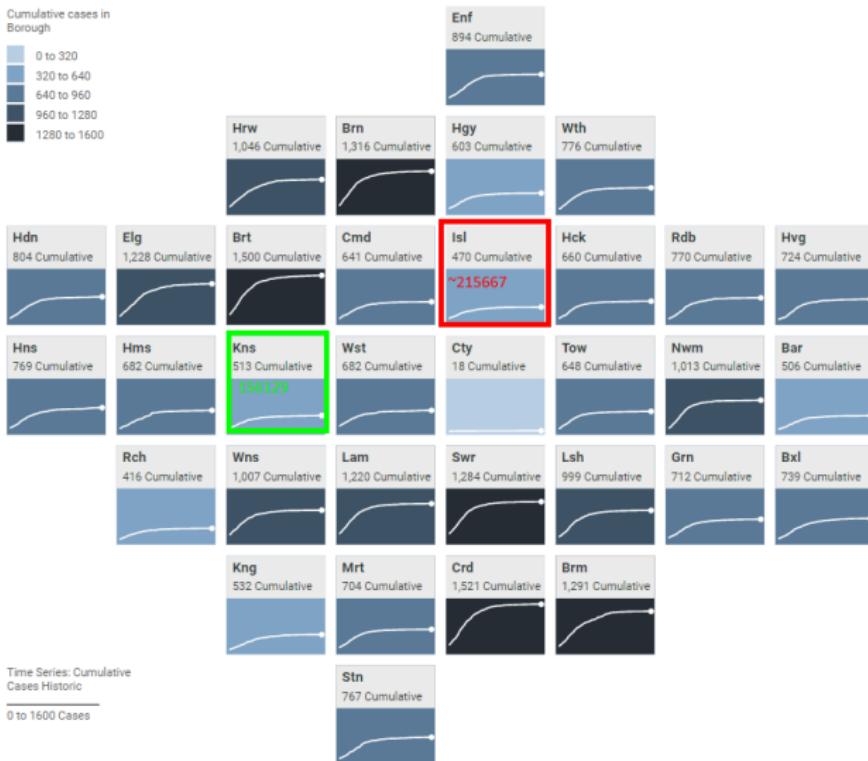
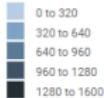
*Results show a correlation between the use of the underground and ILI cases in London, specifically they show that higher numbers of ILI cases arise in those boroughs where the population spend more time in the Underground and/or incur in a higher number of contacts when travelling. [6]*

# COVID in London

Covid-19 Cases by London Borough (2020-03-23 to 2020-06-21)

Displaying cumulative count (all historic) at: 2020-06-21

Cumulative cases in  
Borough

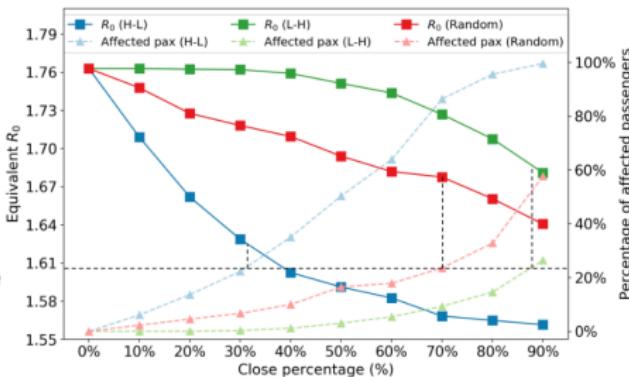
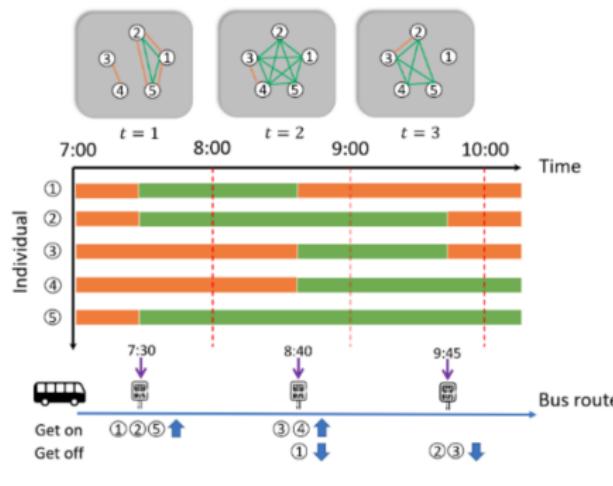


Source: <https://coronavirus.data.gov.uk/> - Note: Data for most recent 5 days may be incomplete.

Graphic by GLA City Intelligence | London Squared Format by After The Flood

Speaker's Remark: 🔍🔍🔍

# Prior Research (Singapore Buses)

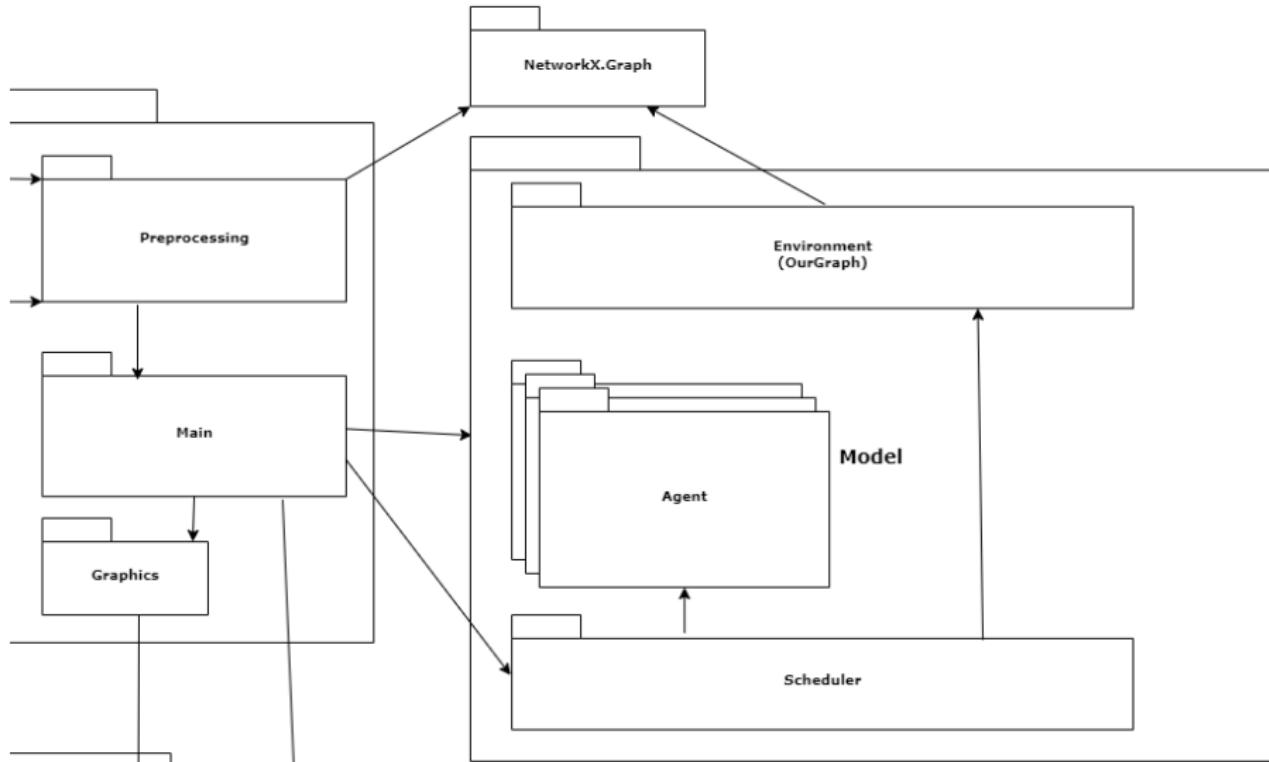


*The direct contact in trains is, however, difficult to obtain from smart card data because the transactions are recorded at the station level [5]*

*Speaker's note:*

- We show these things not to embarrass ourselves, but the depth of research available even just looking at one system.
- List some other research

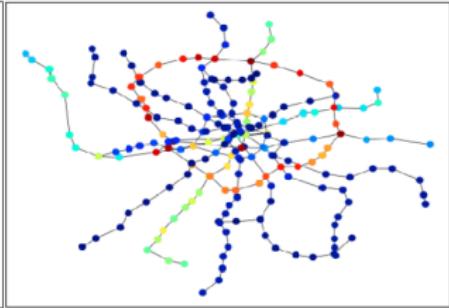
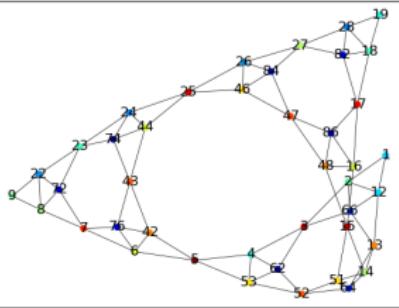
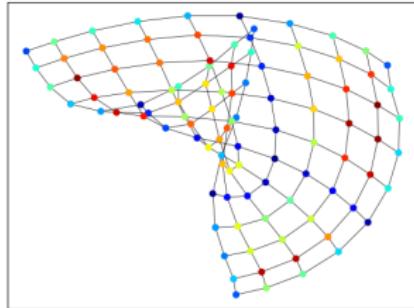
# ABM Framework



Components: Python, MESA, Networkx

Base Classes: Transportation Model, SEIR Agent

# Simple Geometries



- Grid
- Sierpinski's Triangle
- Moscow

# World Airline Network (Passenger Flow) - Challenges

Obtaining accurate data about world passenger flow is a bit of challenge, one has 3 options:

- ① Pay for it
- ② Using superior math to extrapolate or estimate
- ③ Extrapolate based on what I know; We opted for option 3

## Algorithm 1 to estimate WAN (Passenger Flow)

- ① Find total number of passengers going through a specific airport
- ② Total number of destinations from said airport
- ③ Ratio of traffic from each airport
- ④ Split population of passengers into various destinations based on ratio

Major challenge with this is the fact that we need to find data to answer each of these questions in order to make an estimate (ratios data)

## Algorithm 2 to estimate WAN (Passenger Flow)

Using Data from Wikipedia and Openflights

- ① Get the total population of data passengers that go through a specific airport.
- ② Using HITS algorithm on airport pairs.
- ③ Split population of passengers into various destinations based on hubs obtained from HITS

Major challenge with this is the fact that the graph is undirected hence the HITS results were unreliable. Also if the data was split into major known hubs the data returned the same value from HITS.

# World Airline Network (Passenger Flow) Algorithm

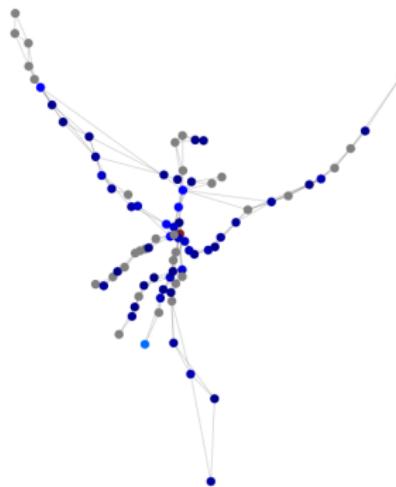
# Madrid Commuter Trains (Central Hubs)

How important are central transportation hubs to disease spread?

Data Sources:

- Madrid Renfe GTFS: <https://crtm.maps.arcgis.com/home/item.html?id=1a25440bf66f499bae2657ec7fb40144>
- Madrid Renfe Turnstile data: <https://data renfe com/dataset/volumen-de-viajeros-por-franja-horaria-madrid>

# Madrid Commuter Trains



**Figure:** Madrid Cercanías map 2020 (urban trains)

# Madrid Commuter Trains

Method: Agent Based Model (ABM)

Agents: passengers

Flow based on weights  $w$  which depend on the turnstiles data: *in-weight* and *out-weight* (incoming and outcoming passengers respectively).

# Madrid Commuter Trains - Passengers flow

At each time step  $t$ ...

$N$  passengers join the network randomly:

- ① Join at a random time.
- ② Origin station depending on the *in-weight*.
- ③ Platform choice based on the weight of the next platforms.
- ④ The passengers wait in the platform until a train arrives.

Passengers who are in a train:

- Leave the network (destination arrival) depending on the *out-weight* of the station.
- Leave the train and go to another platform depending on the weights of each platform.

# Madrid Commuter Trains - Parameters

In order to simulate the spread of an infection:

- Passengers (in a day): 500,000.
- Infected passengers: 50,000 (10% of them).
- Infection probability: 0.005.

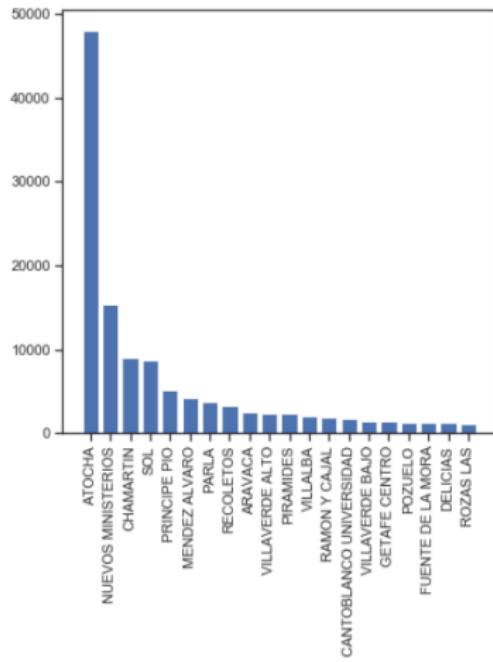
On average, an infected passenger will infect to other 5.04 people.

# Madrid Commuter Trains - Results



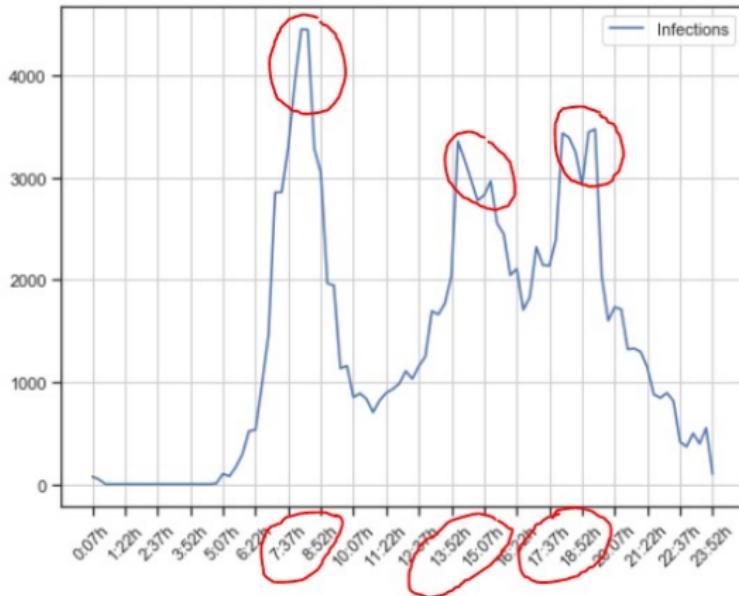
**Figure:** Top 10 stations with more infections

# Madrid Commuter Trains - Results



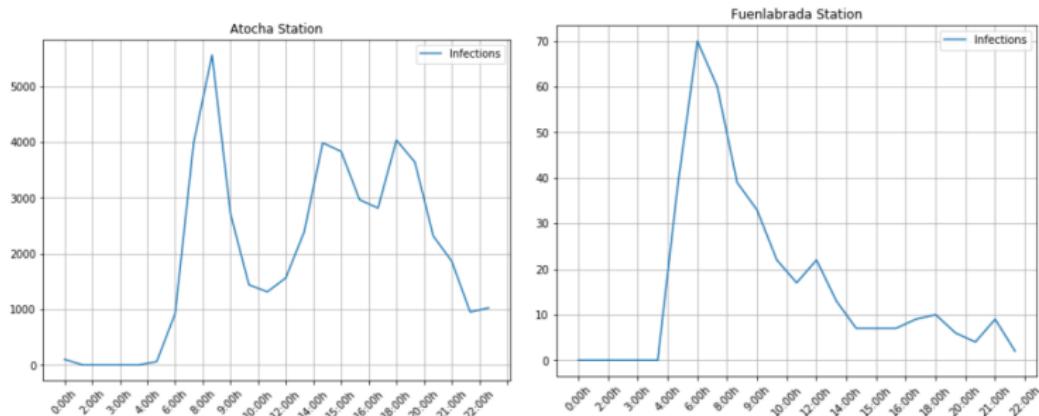
**Figure:** Top 20 stations with more infections

# Madrid Commuter Trains - Results



**Figure:** Infections in the whole network during a day.

# Madrid Commuter Trains - Results



**Figure:** Infections in a central hub (left) and a station in a suburb (right).



# NYC Demographics And COVID Timeline

*NYC residents working in Manhattan primarily travel by subway. This is also true for residents of the Bronx, Brooklyn, and Queens [2]*

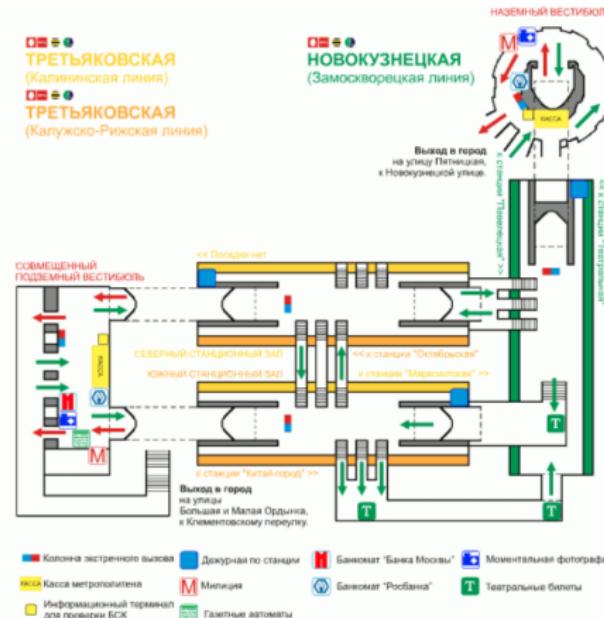
- March 9 - Mayor holds press conference and notes that there have been 16 confirmed cases. (106 cases)
- March 12 - Mayor declares a local state of emergency. (687 cases)
- March 15 - Schools officially close. (2,986 cases)

Borough - A geographical region. NYC has 5 boroughs.

MODZCTA - Modified Zip Code Tabulation Areas.  $\propto$  postal codes.

# Subway Systems

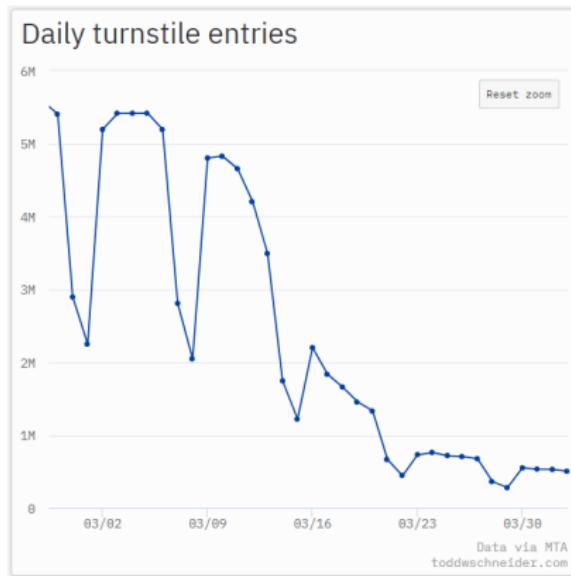
## Complex, Station, Line, Route



**Figure: Floor Plan of Novokuznetskaya Metro Station [4]**

# NYC Subway Data

Stations, Map, Turnstiles (GTFS not considered)



**Figure:** NYC Subway daily turnstile entries for March 2020 [3]

167,167,A32,IND,8th Av - Fulton St,W 4 St,M,A C E...

167,167,D20,IND,6th Av - Culver,W 4 St,M,B D F M...

# NYC Subway Modeling

---

## Algorithm 1 Simulation of Disease Spread on Subways

---

```
1: for  $i = 1; i < TIMESPAN; i++$  do
2:   Check conditions ( $i$ , number of infected) to see if we should deploy
      COUNTERMEASURES
3:   for Station in SubwayModel.Environment.Nodes do
4:     Calculate 'Local Exposure' from infected and commute time.
5:     Calculate 'Route Exposure' from infected on the same route.
6:     Calculate 'General Exposure' due to city-wide infected.
7:     Update 'Exposure' at station based on above conditions
8:   end for
9:   for Agent in SubwayNetwork.Agents do
10:    Get 'Exposure' At Location
11:    Get City-wide COUNTERMEASURES
12:    Get Percentage of commuters
13:    Calculate SEIR beta and gamma based on conditions
14:    Update SEIR numbers
15:  end for
16: end for
```

---

Subway Agent - SEIR Agent with additional exposure based on location.

Subway Graph - Nodes  $\propto$  Stations, Edges  $\propto$  **Lines, Complex**, Route Lookup



# NYC Model Parameters, Hyper-parameters

Default SEIR Rates:  $\beta, \alpha, \gamma = 1.75, 0.20, 0.50$

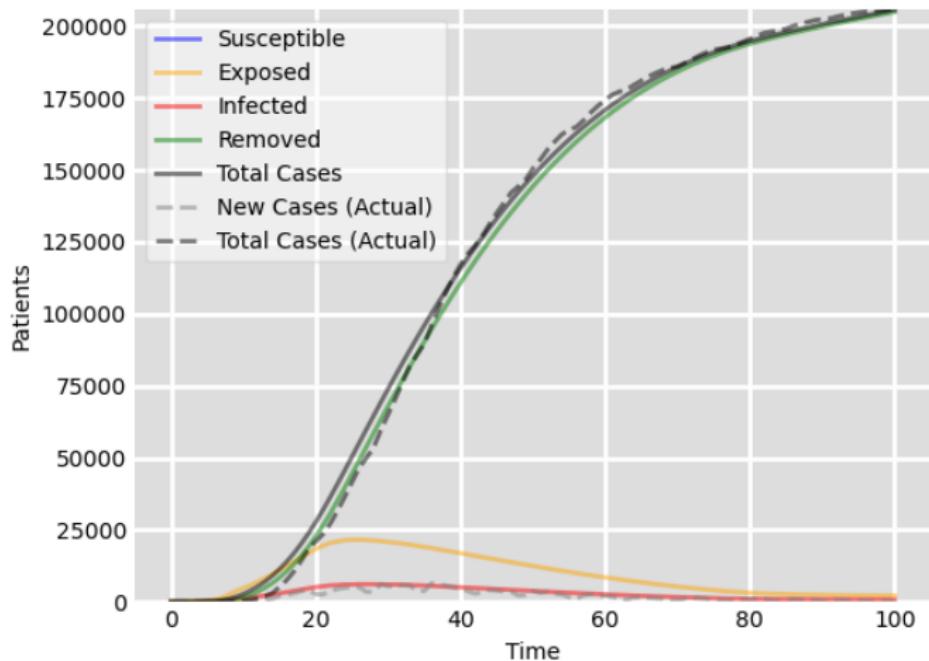
Countermeasures: Isolation, Recommendation, Awareness: = 5000, 500, 500

*Defiance*: 10% Local, 25% Global

*Global Exposure*: 0.7

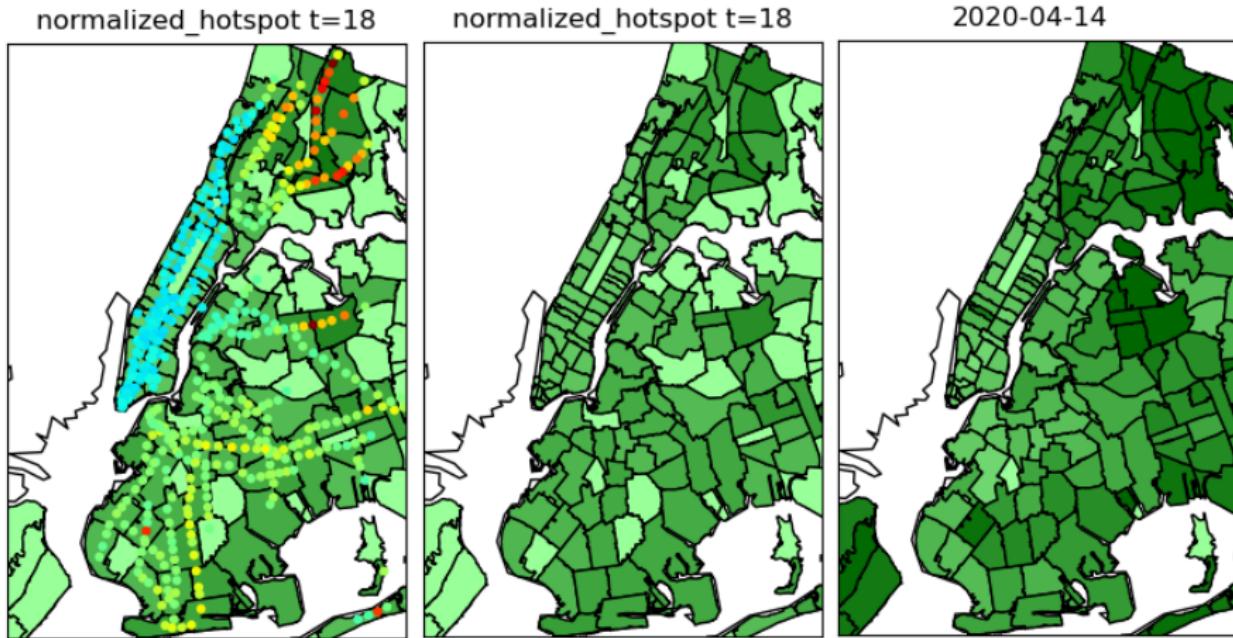
*Commute Time*: Squared shortest path from Grand Central or Times Square

# Compartmental Modeling Results (NYC)



**Figure:** SEIR Fitting to NYC Case Total.  $MAPE(t \geq 30)$ : 0.0145

# Results by MODZCTA (NYC)



New York City Timelapse

# Discussion

All models are wrong!

- Passenger Flow
- Outer suburbs phenomenon exists
- Simple GDP/capita data
- Erase some of these stupid parameters

# Conclusions

- Determining passenger flow is a difficult problem for all transportation networks at all granularities.
- 
-

## Credits, Links, References

Frank Acquaye - WAN, Passenger Flow

Ho Lum Cheung - NYC, Organization, Research, Testing

Dimas Muñoz-Montesinos - Madrid, Hotspots

Elie Wanko - Theory, Consulting

github link references used in slides

 Covid-19: Data.

<https://www1.nyc.gov/site/doh/covid/covid-19-data.page>.

 The ins and outs of nyc commuting.

<https://www1.nyc.gov/assets/planning/download/pdf/planning-level/housing-economy/nyc-ins-and-out-of-commuting.pdf>.

 New york city subway usage.

<https://toddwschneider.com/dashboards/nyc-subway-turnstiles/>.

 Plan of metro station novokuznetskaya.

<http://www.karta-metro.ru/stations/96/428/>