Email Exchange with Client and Draft Proposal

To the Apocalytics Data Group,

It was great to meet with you and chat at Techcrunch Disrupt NYC. We'd love to take the next steps to see if working together is something that would make sense for both parties.

As we mentioned we have some reliable tips on a potential terrorist attack and we are interested in analyzing potential targets for a zombie virus outbreak. We have reason to believe that Dr. Evil will unleash a zombie virus into the NY MTA subway system. We want to be ahead of the curve on this and figure out the most optimal way to allocate our resources if/when the virus is released.

We'd like to solicit your engagement to use MTA subway data. We intend to use our staff and police officers as a first line of defense. We have a limited number of officers that we can devote to managing this threat. As a result we need to position them where and when they will be most effective. Since we do not know exactly when the attack will occur we will need to know how to position our resources throughout the day.

The ball is in your court now—do you think this is something that would be feasible for your group? From there we can explore what kind of an engagement would make sense for all of us.

Best,

Thomas R. Frieden

Director of the U.S. Centers for Disease Control and Prevention (CDC)

Dear Director Frieden -

Hope you are well! Again, we are sad to hear about a potential zombie apocalypse. We pulled the mta data from the public data portal and ran some initial feasibility analysis. Check out the attached document. With the MTA data alone, we can definitely help you to optimize the placement of officers and staff to maximize your team's response efficiency in case of an attack. We also included a few more ideas of different scales that would greatly improve the effectiveness our analyses could have toward achieving your goals.

Let us know what you think—we can set up a call at your earliest convenience since this is a matter of national security.

Looking forward to chatting again,

Apocalytics Group

"specialized data analytics for global scale crises"

Email Exchange with Client and Draft Proposal

Draft Proposal:

Center for Disease Control: *Potential Apocalypse*

Problem statement:

The CDC needs to optimize protocol for police and staff in case of a potential zombie outbreak planned by Dr. Evil in the coming weeks. The goal is to have staff in the areas most likely to be attacked at all hours of the day to minimize spread of the disease.

MTA data obtained from the NYC Data Portal provides rich information about the travel patterns of New Yorkers -- using this information, we create a street team deployment plan that will minimize the effects of any outbreaks and possibly prevent the extinction of the human race.

Preliminary results:

Here are some preliminary results we found, as a sample of what we can provide.

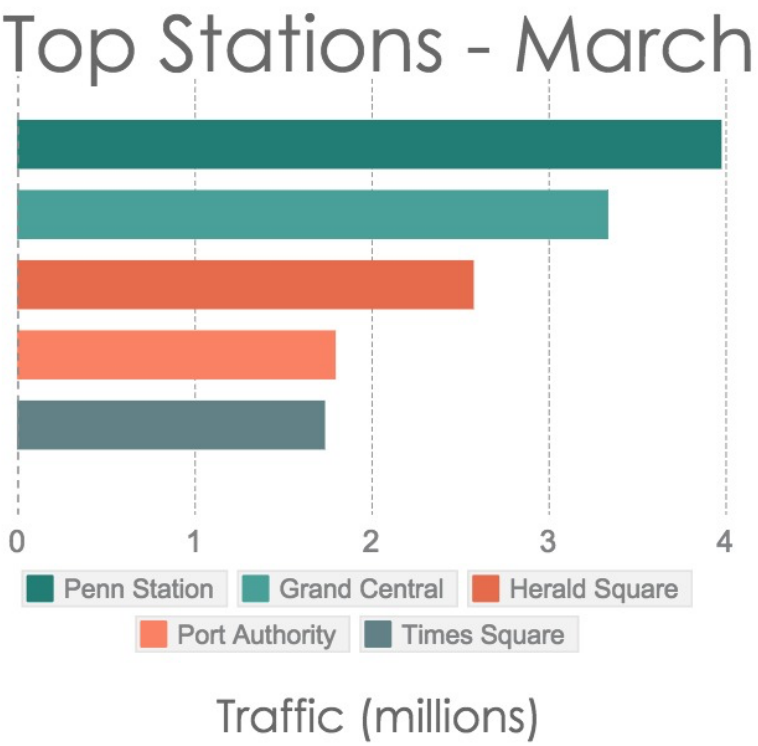
Highest Traffic Areas:

Let's start at the beginning and simply see where the most traffic is. This has a lot of assumptions baked into it, so we'll examine these and refine our recommendations as we continue.

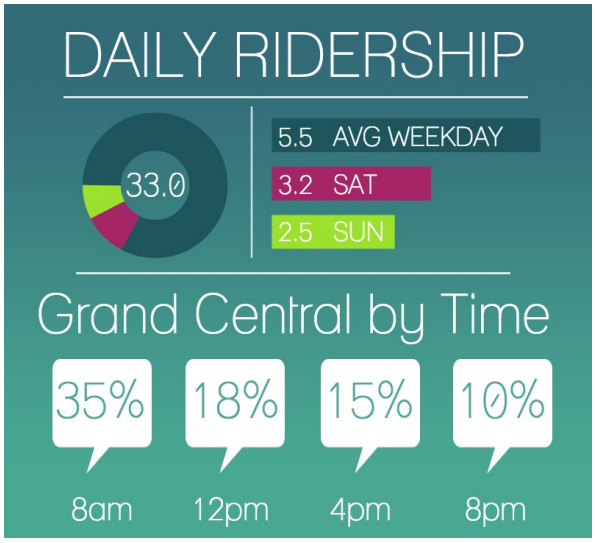
Top ten stations in terms of average daily turnstile entrances:

i. 34th St - Penn Station
ii. 42nd St - Grand Central Station
iii. 34th St - Herald Square
iv. 42 St - PA Bus Terminal
v. 42nd St - Times Sq
vi. 14 St - Union Square
vii. 86 St and Lex
viii. Main St in Queens
ix. Rockefeller Center
x. 59th St and Lex

Email Exchange with Client and Draft Proposal



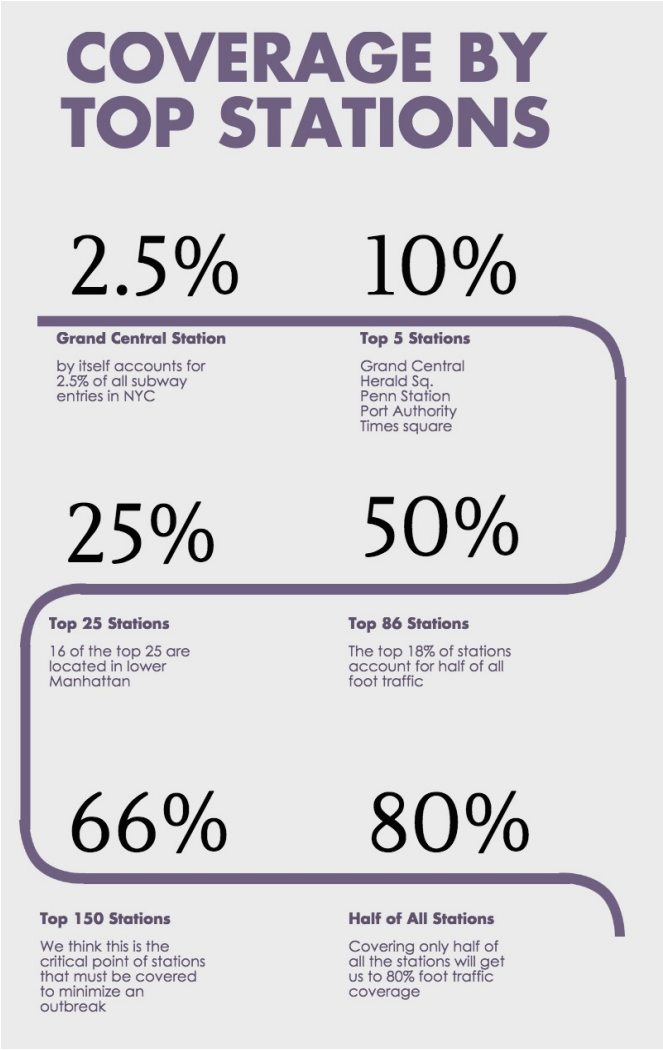
We identified in the data a signature of commuter travel: stations with high and consistent usage on weekdays, and low utilization on the weekend. So depending on the day, officers must be positioned very differently to account for this disparity.

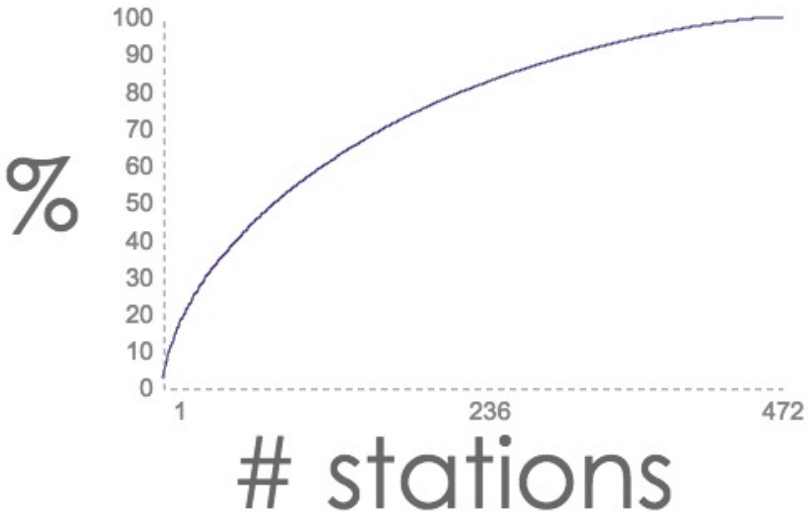


*Figures for ridership in millions unless otherwise labelled

Email Exchange with Client and Draft Proposal

We were also able to identify plot how many stations you would need to man to cover the highest percentage of riders. We believe this data will be very useful in coming up with efficient assignments based on areas.





Email Exchange with Client and Draft Proposal

What else we could do:

Obviously with more time and a commitment on your end to obtain the necessary city permissions to get real time data, we would create real time updates for each data representation. By creating a real time API, this can be utilized by the necessary parties to make the necessary decisions to ensure the safety of both NYC, the rest of the US, and the world.

This data can lead to the creation of a real time Risk Heat map of NYC along with an interface for quick predictions of spread.

Adding more data sources (etc)

-While the MTA turnstile data is limited, we can only use it to make approximations about station usage based on patterns. We would need access to real time dynamic data.
-We can also use real time street traffic data to plot events above ground
-Twitter, FB, Instagram can also be tapped to track real time events
-The latest population density maps from either census data or municipal data can help to hone our model
-Weather data may also play an important role depending on the type of virus and implications to ridership
-We would need all available data on the virus and it's properties, especially in regards to how it spreads

Conclusion:

Based on your interests and scope, we can go ahead with anything discussed above, or meet to discuss any of these ideas. This is a very important threat to national security and we are prepared to staff all the resources necessary to complete the project as soon as possible. We look forward to working with you.