Best Place to Survive a Zombie Apocalypse in NYC

By Eduardo Beltrán Herrera

Oct. 22, 2020

1. Business Problem

If 2020 has taught us anything is that **things can always get worse**. The global pandemic has demonstrated that even as advanced as our civilizations are, we are still not ready to face the worst possible scenarios that may be put in our way. That's why preparation is key, and it's up to us to make the right decisions, with the right data, when the time comes to confront the problems that plague humankind.

As unlikely as it might be, a **zombie apocalypse** is one of these scenarios we should prepare for. This project will try to solve one of the first problems one should face in such a situation: **where should I bunker myself when the zombie breakout has begun?**

We will analyze the neighborhoods of New York City (where, as we all know, all world-changing catastrophes begin) to find the one that better accommodates for the following basic needs:

1. Food and water (a.k.a. Grocery Stores)
2. Medicine & First Aid resources (Pharmacies)
3. Construction supplies for bunker-building (Home Depot, Lowe's...)

Other venue categories might be considered as they appear on the data. Neighborhoods will be clustered using the **k-means** algorithm and later described so that we can determine the best cluster.

2. Data

The **Foursquare API** will be used to obtain the data relating to the venues in each neighborhood. The Foursquare database allows us to gather information of venues close to a particular location, providing additional data like the category of the venue (this will tell us if the venue is a grocery store, coffee shop, etc.).

Foursquare also allows us to see the ratings given to each venue, but since we are planning for a survival situation, the quality of the place does not concern us. The **location and category of the venue** should be enough for us to determine how good the neighborhood is to meet our needs during the apocalypse.

3. Methodology

A listing of New York neighborhoods is imported the same way it was done in Week 3. Foursquare queries are executed for each neighborhood to find all venues in a determined radius.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Table 1. *Sample neighborhood geodata*

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
2	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
3	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop

Table 2. *Sample venue information*

From a quick analysis of all venue categories, the ones that will be considered relevant for zombie apocalypse-survival will be:

- Pharmacy
- Convenience Store
- Grocery Store
- Supermarket
- Furniture / Home Store

These venue types were filtered and counted on each neighborhood. A total of **508** venues were distributed around all neighborhoods.

Venue Category	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
0	Allerton	0	0	1	1	2
1	Arden Heights	0	0	0	1	0
2	Arlington	0	0	1	0	0
3	Arrochar	0	0	0	1	1
4	Astoria	0	0	1	0	0
...	...	...	...	...	...	...
210	Wingate	0	0	1	1	0
211	Woodhaven	0	0	0	2	1
212	Woodlawn	0	0	1	0	0
213	Woodrow	0	0	1	2	0
214	Woodside	0	0	6	1	0

Table 2. *Sample venue count per neighborhood*

All neighborhoods were clustered with the **k-means** algorithm into **5** different groups to determine their defining characteristics.

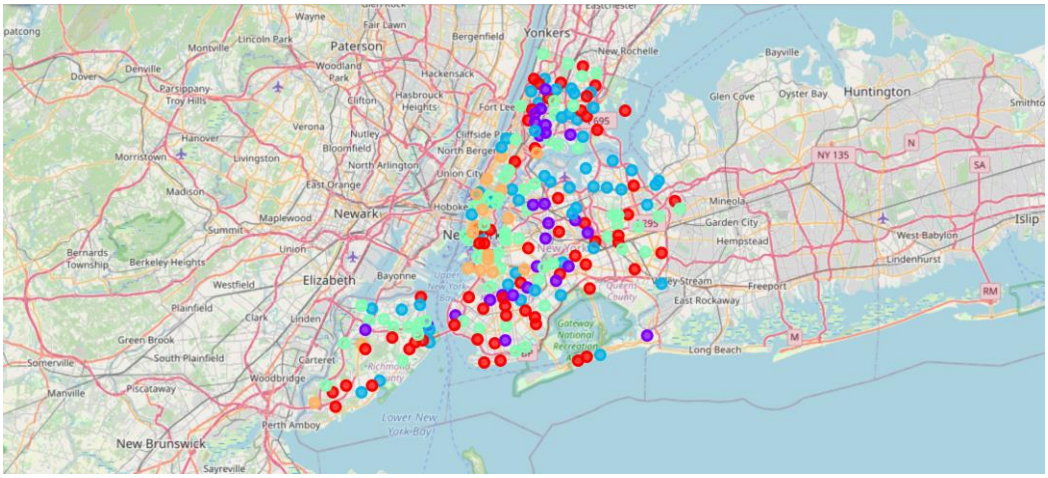

Figure 1. New York City neighborhoods colored by cluster

4. Analysis

All five clusters are briefly described:

Cluster 0

Cluster 0 contains neighborhoods with a couple pharmacies, with some supermarkets appearing now and then.

Venue Category	Cluster Labels	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
1	0	Arden Heights	0	0	0	1	0
7	0	Bath Beach	0	0	0	2	0
12	0	Bayside	0	0	0	1	0
16	0	Belle Harbor	0	0	0	1	0
20	0	Borough Park	0	0	1	2	0
...	...	...	...	...	...	...	...
200	0	Wakefield	0	0	0	2	0
206	0	Westchester Square	0	0	0	2	1
210	0	Wingate	0	0	1	1	0
211	0	Woodhaven	0	0	0	2	1
213	0	Woodrow	0	0	1	2	0

Cluster 1

Cluster 1 neighborhoods contain one or two grocery stores, and not much else.

Venue Category	Cluster Labels	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
9	1	Bay Ridge	0	0	2	2	0
27	1	Bulls Head	1	0	2	2	1
39	1	City Line	0	0	2	1	0
41	1	Claremont Village	0	0	3	0	0
50	1	Concourse	0	0	4	1	0
71	1	Erasmus	1	1	2	1	1
72	1	Far Rockaway	0	0	2	1	2
81	1	Forest Hills Gardens	0	0	2	1	0
102	1	Homecrest	0	0	3	1	0
110	1	Kensington	0	1	3	1	1
113	1	Kingsbridge Heights	0	0	3	1	0
127	1	Maspeth	0	0	2	1	0
128	1	Melrose	0	0	2	3	1
137	1	Morrisania	0	0	2	1	0

Cluster 2

Cluster 2 seems to contain neighborhoods with between 1 and 3 supermarkets, with a fair number of pharmacies and grocery stores as well.

Venue Category	Cluster Labels	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
0	2	Allerton	0	0	1	1	2
3	2	Arrochar	0	0	0	1	1
5	2	Astoria Heights	0	0	0	0	1
6	2	Auburndale	0	0	0	1	1
10	2	Bay Terrace	0	1	0	1	3
11	2	Baychester	1	0	0	0	1
13	2	Bedford Park	1	0	1	2	2
14	2	Beechhurst	1	0	0	0	1
32	2	Castle Hill	0	0	0	1	1
37	2	Chinatown	0	1	0	0	1
44	2	Clinton	0	0	0	0	1
48	2	College Point	1	0	0	1	1

Cluster 3

In cluster see we have neighborhoods with very spread-out venues, with one or two at most.

Venue Category	Cluster Labels	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
2	3	Arlington	0	0	1	0	0
4	3	Astoria	0	0	1	0	0
8	3	Battery Park City	1	0	1	0	0
15	3	Bellaire	1	0	1	0	0
17	3	Belmont	0	0	1	0	0
...	...	...	...	...	...	...	...
201	3	Washington Heights	0	0	2	0	0
207	3	Westerleigh	1	0	0	0	0
208	3	Williamsburg	0	0	1	0	0
209	3	Windsor Terrace	0	0	2	0	0
212	3	Woodlawn	0	0	1	0	0

Cluster 4

Cluster 4 contains neighborhoods with the most grocery stores, with some pharmacies and supermarkets.

Venue Category	Cluster Labels	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
19	4	Boerum Hill	0	3	2	0	0
35	4	Charleston	0	1	1	0	0
40	4	Civic Center	0	1	0	0	0
59	4	Downtown	0	1	2	0	0
75	4	Flatiron	0	2	0	0	0
90	4	Gowanus	0	3	1	0	0
96	4	Greenpoint	0	1	0	0	0
104	4	Hudson Yards	0	1	1	0	1
115	4	Lincoln Square	0	1	1	0	0
122	4	Manhattan Valley	0	1	1	0	0
143	4	New Springville	0	1	1	1	0
153	4	Park Slope	0	1	1	0	0

Because of the high volume of supermarkets, grocery stores and pharmacies, we conclude that Cluster 2 has the best neighborhoods for our survival needs. Particularly, **Kingsbridge** contains a good distribution of food resources, medicine, and a very sought-after Home Store for all our bunkering needs.

Venue Category	Cluster Labels	Neighborhood	Convenience Store	Furniture / Home Store	Grocery Store	Pharmacy	Supermarket
112	2	Kingsbridge	0	1	1	2	2

5. Results and Conclusions

We have concluded that **Kingsbridge** is probably the safest neighborhood to survive the zombie apocalypse. However, a more complex analysis is recommended to include other factors that will impact our survival odds, such as population density, and transportation methods.

Overall, our analysis will be a handy guide for keen survivalists to further expand their plans, in case the worst is yet to come...