

Finding NYC's neighborhoods which will benefit the most out of building a medical service nearby them

Ariel Nisenbaum

October, 2019

1. Introduction

A medical emergency is an acute injury or illness that poses an immediate risk to a person's life or long-term health. For emergencies starting outside medical care, a key component that can make a difference between life and death is the response time – i.e. the time frame in which the medical emergency start till the emergency medical services arrival. As a result, geographers have been increasingly involved in emergency medical-services planning. Of particular importance has been the application of location-optimization models that minimize ambulance response time. Therefore, it is necessary to geographers to know which neighborhood will be benefit the most out of building a medical service nearby them.

2. Problem

Focusing in New York City, our aim in this project is to find out which neighborhoods have the biggest response time, i.e. – finding the neighborhoods that their distance to a nearby medical service is the largest.

3. Required data

Since the main feature to investigate is distance, our data sets will only need to contain medical centers coordinates and neighborhoods coordinates (latitude and longitude) in New York City. For the neighborhoods data set we can use the following link:

- https://cocl.us/new_york_dataset

For the hospitals data set we can use the following data from Wikipedia:

- https://en.wikipedia.org/wiki/List_of_hospitals_in_Manhattan
- https://en.wikipedia.org/wiki/List_of_hospitals_in_the_Bronx
- https://en.wikipedia.org/wiki/List_of_hospitals_in_Brooklyn
- https://en.wikipedia.org/wiki/List_of_hospitals_in_Queens
- https://en.wikipedia.org/wiki/List_of_hospitals_in_Staten_Island

4. Methodology

4.1 Neighborhoods data set

The original data set was a json file from which the required data was borough, neighborhood, latitude and longitude. These were achieved by using the json package, which is a library to handle json files.

4.2 Hospitals\medical centers data set

The hospital\medical centers data set includes the hospital name, latitude and longitude. The last two were achieved by using geopy package.

4.3 Chosen model

In order to decide which neighborhoods will benefit the most out of building a nearby medical center, our model had to take into account two main characteristics: distance (as mentioned at the beginning) and **density**.

Moreover it needed to be a clustering model, since our data sets **were not labeled**. Hence, the most suitable model to the requirements above was the **DBSCAN model**.

By choosing only the latitude and longitude columns from the data frames in 4.1 and 4.2, combining them together and cluster them, I was able to distinguish between outliers, clusters which contain hospitals\medical center and clusters which were not contain hospitals\medical center.

5. Results

As mentioned in 4.3, by using the DBSCAN model, I was able to distinguish outliers, clusters which contain hospitals\medical center and clusters which were not contain hospitals\medical center. Hence, the outliers and clusters which didn't contain hospitals\medical center represent the neighborhoods that will benefit the most out of building a nearby medical center.

For choosing $\epsilon = 0.025$ and $n_{\text{samples}} = 5$, we got that the following neighborhoods are the ones that will benefit the most out of building a nearby hospital\medical center:

Project Results

The following NYC neighborhoods need a nearby hospital according to the DBSCAN model

```
[\"Douglaston\", \"Rosedale\", \"South Ozone Park\", \"Travis\", \"Rochdale\", \"Sea Gate\", \"Bath Beach\", \"Rockaway Park\", \"Breezy Point\", \"Bay Ridge\", \"Dyker Heights\", \"Chelsea\", \"Belle Harbor\", \"Laurelton\", \"St. Albans\", \"Neponsit\", \"Bayside\", \"City Island\", \"Springfield Gardens\", \"Bloomfield\", \"Fort Hamilton\", \"Rosbury\", \"Cambria Heights\", \"Bay Terrace\", \"Howard Beach\", \"Little Neck\", \"Brookville\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\", \"Far Rockaway\", \"Broad Channel\", \"Edgemere\", \"Arverne\", \"Rockaway Beach\", \"Somerville\", \"Hammels\", \"Bayswater\", \"Eltingville\", \"Annadale\", \"Woodrow\", \"Tottenville\", \"Hughenot\", \"Pleasant Plains\", \"Butler Manor\", \"Charleston\", \"Rossville\", \"Arden Heights\", \"Greenridge\", \"Sandy Ground\", \"Prince's Bay\", \"Richmond Valley\"]]
```

6. Observations

Out of the outliers of the model output, some are hospitals. i.e. – there are hospitals that their location is not ideal; they are located too far from the neighborhoods surrounding them. Moreover, from the results section above, it is seen that a substantial number of neighborhoods require a nearby medical center.

7. Conclusion

In this project, I wanted to determine which neighborhoods in New York City will benefit the most out of building a medical center nearby them. I identified New York City's neighborhoods, hospitals and medical centers, found their coordinates and model them into a DBSCAN model. This model can be very useful to geographers to decide where to build a new hospital\medical center and decide if hospitals are ideally located in relation to their surrounding neighborhoods.