**Clustering New York City hospitals**

**Seo Young Choe**

**April 9, 2020**

1. Introduction

1.1 Background

The transfer of a patient to another facility is a lesser known but important topic. The decision to transfer the patient is based on the benefits of care available at another facility against the potential risks involved. The need to transfer a patient should take into account the benefit of providing extra care on the management or outcome. The various contributors of need to transfer the patient include the presence of few centers which provide super-specialty care, non-availability of beds and funding of medical treatment.

One of the risk factors of intra-hospital transfer is distance. Patients with an acute ischemic stroke (AIS) due to large vessel occlusion often require transfer to a different hospital for treatment. However, an increasing distance between hospitals was shown to be associated with an increasing risk of mortality. Therefore, it is advantageous for hospitals to be aware of other nearby medical institutions, share capacity information and cooperate with each other in case of transfer.

1.2 Interest

Obviously, hospitals would be interested in cluster information about nearby hospitals, for communication and cooperation purposes. Other interested parties may include health officials and high-risk patients.

2. Data & Method

2.1 Data acquisition

This report uses the Foursquare.com resources for its geographical data source

2.2 Data selection & cleaning

The target data is hospitals in New York City area, specifically those within 32km radius of its center. There are 50 medical centers within the target area. Two institutions were excluded because its location data was missing. For the remaining 48 medical centers, name and location data was extracted and formed into a dataframe for further evaluation

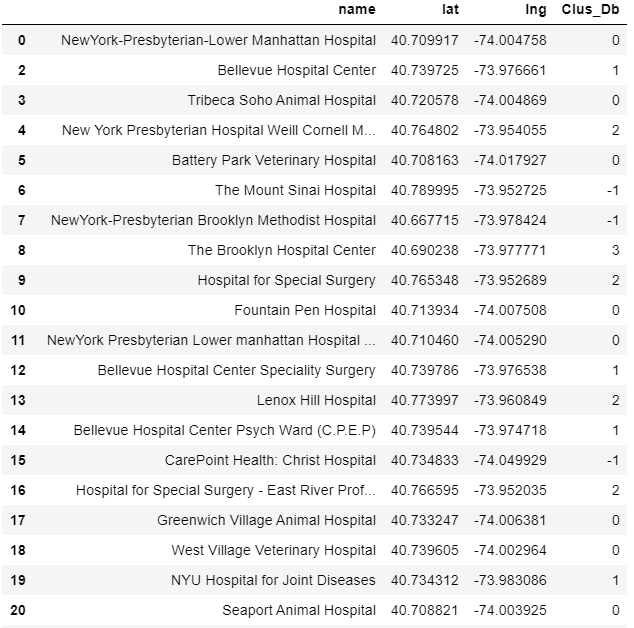
2.3 Data analysis

Density-based Clustering (DBSCAN) is a clustering method that locates regions of high density that are separated from one another by regions of low density. Density, in this context, is defined as the number of points within a specified radius. Density-based Clustering (DBSCAN) was used to cluster hospitals into groups according to their distance calculated by latitude and longitude. The clusters constitute of 3 or more hospitals within 30km (the distance an average car travels in 20minutes) radius of each other.

3. Results

3.1 Clusters

The 48 hospitals were grouped into 4 different clusters according to distance. Cluster 0 includes 14 hospitals. Cluster 1 includes 6 hospitals. Cluster 2 includes 6 hospitals. Cluster 3 includes 3 hospitals. There are 19 outliers

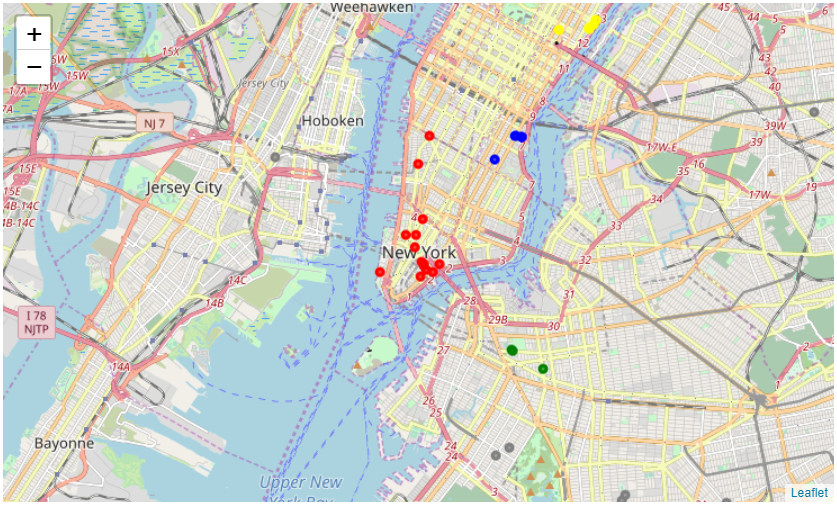






3.2 Visualization

The results were visualized as follows. Red circles indicate Cluster 0, blue circles indicate Cluster 1, yellow circles indicate Cluster 2, and green circles indicate cluster 3.



4. Discussion

This report chose New York City hospitals for evaluation, however, the results can be applied to other large cities with multiple hospitals.

5. Conclusion

A poorly organized and hastily done patient transfer can significantly contribute to morbidity and mortality. It is important for hospitals to be aware of other nearby medical institutions, share capacity information and cooperate with each other in case of transfer