Optimizing City Planning
Using k-means to determine new hospital locations

Michael Gerakis | Patrick Gomes | Raphael Baysa

Introduction

Major cities in the United States are suffering from a space deficit. As the population in the world increases, our city planners need to become conscious of this and maximize efficiency in these cities. Boston is on the the cities suffering from this problem. Land spaces in boston are increasing as the city runs of the space it has to offer. While we can't remove the essential buildings, such as police stations, fire departments, or schools, we can determine how to better place them to minimize the number needed. The goal of our research was to determine how well a generic optimization problem could solve this issue for hospitals specifically.

Goals

The three main goals for our placement of any new hospital location are:

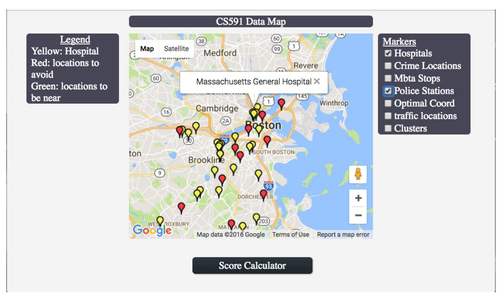
1) Close to transportation stops for reliable access.
2) Close to, but not surrounded by, crime areas for quicker response times.
3) Distanced from high traffic areas.
4) Distanced from police stations to avoid two emergency vehicles in one location

Techniques

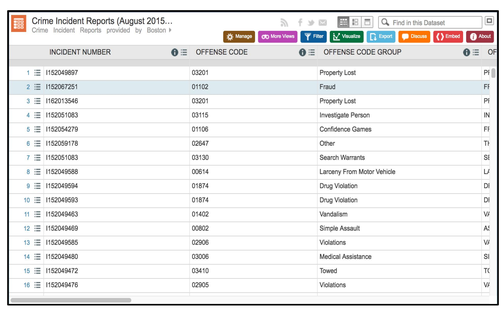
- We used k-means to cluster our data sets into two different cluster groups, locations we wanted to keep far and ones we wanted to keep close.
- Our optimization algorithm tried to minimize the Euclidean distance to all close proximity clusters, and maximize the distance to all far proximity clusters.

Data

The dataset we used is accessible through the city of Boston at cityofboston.gov, where a lot of public Boston data is hosted. The data that we chose to work with included current hospital locations, MBTA stops, crime incident reports, police station locations, and traffic. These datasets were transformed so that only names/identifiers (if applicable) and their lattitude longitude were the only columns that were saved into our database. Then we clustered the data by categorizing them with labels "F" and "C" which represent whether we want to be near them or far from them.



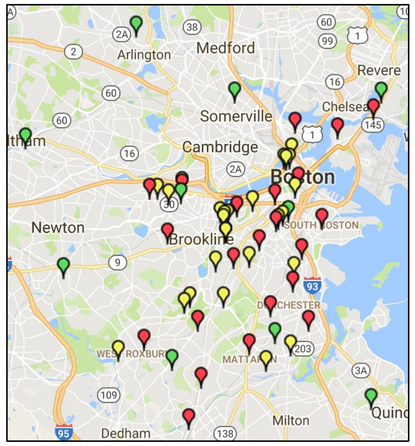
Web Interface created to show our data points



Data interface available through cityofboston.gov

Analysis

- Our rankings suffered because the latitude and longitude was not bounded to the abnormal Boston polygon shape. The algorithm suggested that points outside of Boston were ranked higher, but this is due the lack of cluster data outside of the city.
- Despite this, how our scores compared to Google results was not too far off. We found that on average, the two different rankings, once sorted, were off by only around 7 placements over a list of 26.
- Our recommendations followed in the same way and were located further from the Boston area, but this could be improved by bounding the points to the Boston polygon.



Google Ratings (out of 5)

Massachusetts General Hospital	4.3
Beth Israel Deaconess Medical Center West Campus	4.1
Boston Medical Center	4
Franciscan Children's Hospital	3.9
New England Medical Center	3.7
Beth Israel Deaconess Medical Center East Campus	3.6
Kindred Hospital	3.4

Our Ratings (not normalized)

Carney Hospital	0.004737505267096494
Kindred Hospital	0.002084855056278357
Boston Medical Center	0.0005538733174830423
Franciscan Children's Hospital	-0.004174204470076006
New England Medical Center	-0.006835661400495226
St Elizabeth's Hospital	-0.009744925356388302
Massachusetts Eye & Ear Infirmary	-0.012101165898738787

Conclusion

Moving forward with this project, more research should be done into picking more factors. Two other factors we would have liked to include were the hospital locations and population dense areas. Ideally hospitals won't be close to one another but nearby areas with dense groups of people. We weren't able to find data on the population density of Boston that we could easily use, and we wanted to keep the clusters with the same number of factors in each so we chose to exclude the hospital locations. For a basic implementation we feel that this produced results that are worth looking more into. With the improvements mentioned in the extensions section, this algorithm could be used to determine whether or not hospitals are placed well throughout Boston and extended to other essential city buildings as well.