CS591 Data Mechanics

Optimal Trash Collection Unit Installation Sites

Jacquelyn Andrade, Joseph Cho

Introduction:

Our goal is to gather information on the sanitation-related trends of Boston neighborhoods. Essentially, we want to be able to analyze this data to propose the optimal location to install trash collection units in order to help reduce sanitary code violations, improve overall cleanliness, and prevent trash overfill. The goal of this application is to create an interactive application for the City of Boston. The user will be able to specify the number of trash sites the city wishes to install as well as the radius a single site is expected to cover. The application will display the optimal locations as markers on a map.

Datasets:

All the following Datasets were obtained from the City of Boston.

- **Big Belly Alerts 2014**
- **Master Address List**
- **Code Enforcement - Building and Property Violations**
- **Food Establishments Inspections**
- **Mayor's 24 Hour Hotline (Cases created last 90 days**

Algorithm:

Analysis:

We look at the geolocations of sanitary violations and requests, the weights of those locations, and we also look at the locations of Big Belly's and their average fullness. In addition, the algorithm takes into consideration the geolocations of areas that already have heavy trash collection schedules and attempts to focus more on problematic areas where trash is collected less regularly. Based upon these datasets we find the optimal placement of trash collection units to reduce the overall amount of incidents associated with excess waste. We find these placements using weight metrics and KD trees. Our algorithm has been adjusted to accept number of units and radius parameters and to return a set of optimization coordinates.

Weighting:

In order to find the optimal location for new collection units, we calculate the weights of areas that a new collection unit would cover within its expected radius. Weights are scaled based upon the association and severity of the violation, request, average fullness, or pickup frequency.

Type	Scale
Improper Storage Trash	0.75
Illegal dumping	0.75
Overfilling of barrel/dumpster	1
Storage of Garbage & Rubbish	0.75
Insects Rodents Animals	0.3
Trash Illegally Dump Container	0.75

KD Trees:

"KD-trees" is a technique used to find coordinates within a specific radius given a central coordinate point. Using the points within these radii we calculate the total weight of the region associated with that coordinate. The algorithm will find the location coordinates of the regions that have maximal weight out of all randomly selected coordinates.

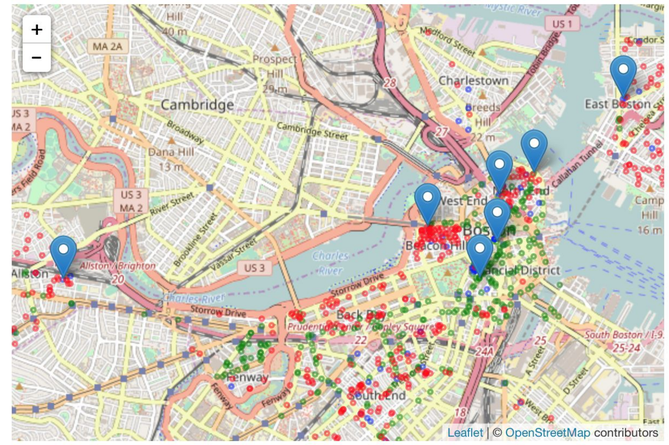
Key: ● = Violations, ● = Service Requests, ● = Big Belly Collection Sites

Results:

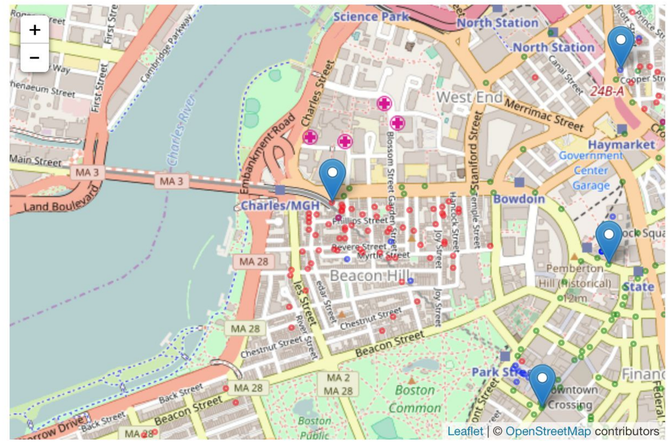
In our analysis, we found that the algorithm initially tends to favor the areas of Boston that are more metropolitan or have high levels of activity, but still tends to favor areas outside of densely trafficked areas when additional trash units are specified. Also, in multiple iterations of running the application, the same exact coordinates are not always returned, however they are within the same relative areas.

Limitations: Weight Derivation/Actual Severity of Incident

Key: ● = Violations, ● = Service Requests, ● = Big Belly Collection Sites



Key: ● = Violations, ● = Service Requests, ● = Big Belly Collection Sites



Future Work:

In future iterations of this project we hope to be able to also optimize trash schedules for the City of Boston. We would like to apply the same weighing techniques to pinpoint locations where trash could be collected more often and locations where trash is not a severe problem and funding to these locations can be allocated somewhere else.