City of Boston: Analyzing Crime and Streetlamps

Nicholas Louie, Thomas Hsu - Boston University CS591 L1 Spring 2016
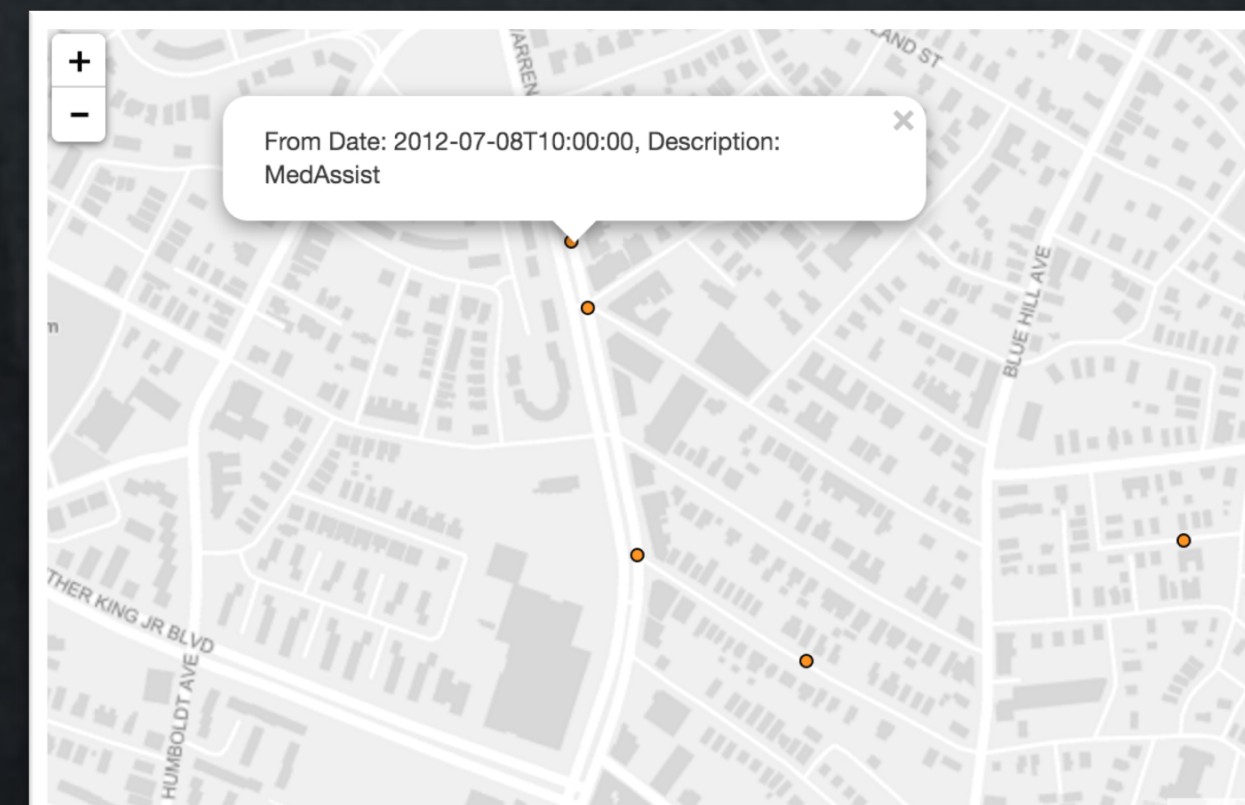
Questions

- What can we learn about crime in Boston given the data from 'data.cityofboston.gov' ?
- Do crimes happen further away from streetlights?
- Can we visualize crimes and streetlights on a map?

Possibilities

- It's possible that depending on the incidences of crime and lighting within a given area, a metric for 'danger' could be created to be used in further applications such as navigation as displayed in Google's controversial "Avoid ghetto" feature.
- By having a dataset of all the streetlamps, it's also possible to introduce this data to make 311 service requests easier to fulfill.

Visualizations

- Using the Leaflet Javascript API and Boston's GeoJSON data from Mapbox, we plotted each streetlight from Boston's Streetlamp Data, and a subset of crimes from Boston's Crime incidence Data.
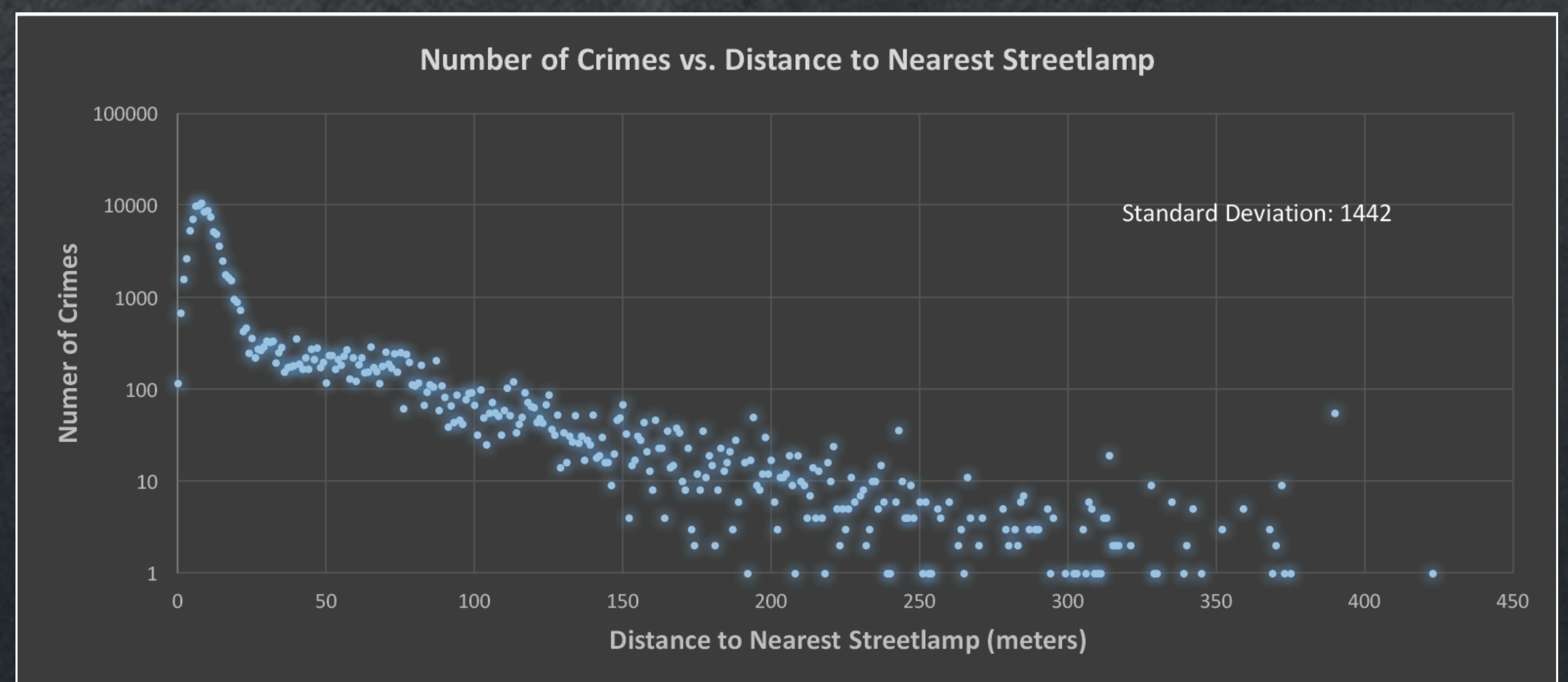

Take me for a spin at
http://nlouie.com/CS591

Implementation

The data files were retrieved manually through the browser, a lot simpler than going through the API. The files were transformed from the retrieved JSON to text file for usage in Hadoop. Originally the plan was to use HIVE to perform SQL-like queries, but in creating an intermediate dataset, HIVE stopped executing. So that step was relegated to a local machine, performed serially. The new intermediate set was then put back on the cluster. Data analysis was done on HIVE.

Limitations

There are quite a few limitations on the data and the analysis. We made a number of assumptions for simplicity. It would be too much work to try to geofence the entirety of indoor Boston, so we assumed that all crime occurs outdoors. We tried to filter only night crimes, night defined as from 18:00 to 6:00. There were a few nonsensical data points in the crimes set (Crimes occurring more than 8000 kilometers away from the nearest streetlamp), because not all reports in crime incidence were properly GPS tagged.

DataSets

https://data.cityofboston.gov/Facilities/Streetlight-Locations/7hu5-gg2y
https://data.cityofboston.gov/resource/7cdf-6fgx.json