

What is a community?

Question

Given different types of city establishments that can either be categorized as indicators of "stronger" or "weaker community", can we find other correlations with other indicators of inequality or "instability" in Boston and address common problems in terms of this social aspect?

In other words, if we were to add a new "community hotspot" or take away an existing "tourist establishment" in order to maximize the benefits to Bostonians across a variety of metrics - economic, social, infrastructural, etc. - where would we do so?

Datasets

- Crime Incident Reports (July 2012 - August 2015)
- Property Assessment
- Public Access Fishing Locations
- Active Food Establishment Licenses
- Entertainment Licenses
- Community Supported Agriculture Pickups
- Boston Parking Lots (scraped from Google)
- Boston Libraries (scraped from Google)

Methodology

The public fishing locations, community supported agriculture pickups, and library locations were grouped and taken to be "community indicators".

The entertainment, active food establishment license, and parking lot locations were taken to be "anti-community indicators" — with the rationale that entertainment is a form of escapism from where one currently is (with large parking lots/garages designated as private, doing no public good), and that food establishments similarly exist to give people a break from eating around their community (and often being tourist spots as well).

Visualizations

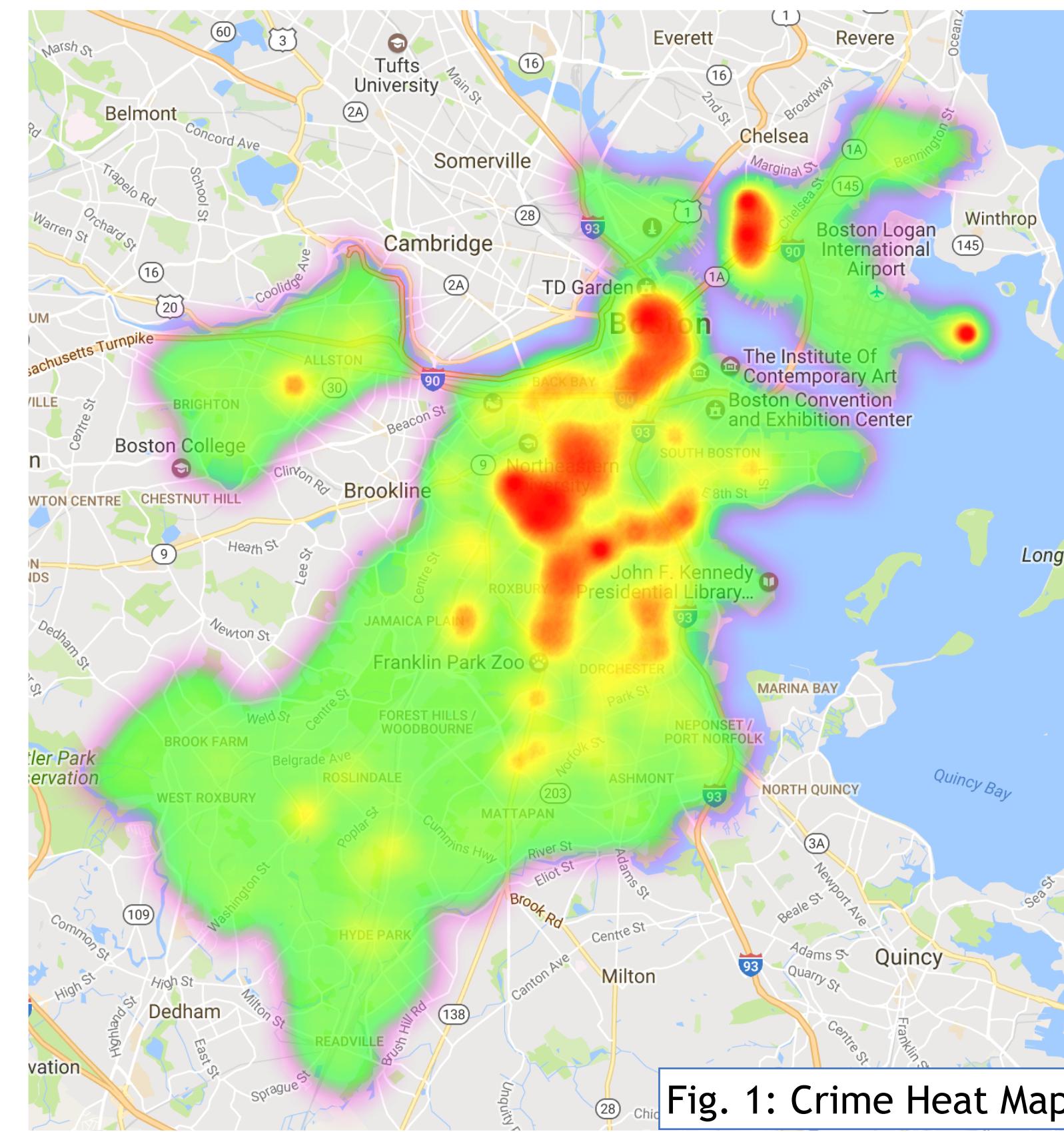


Fig. 1: Crime Heat Map

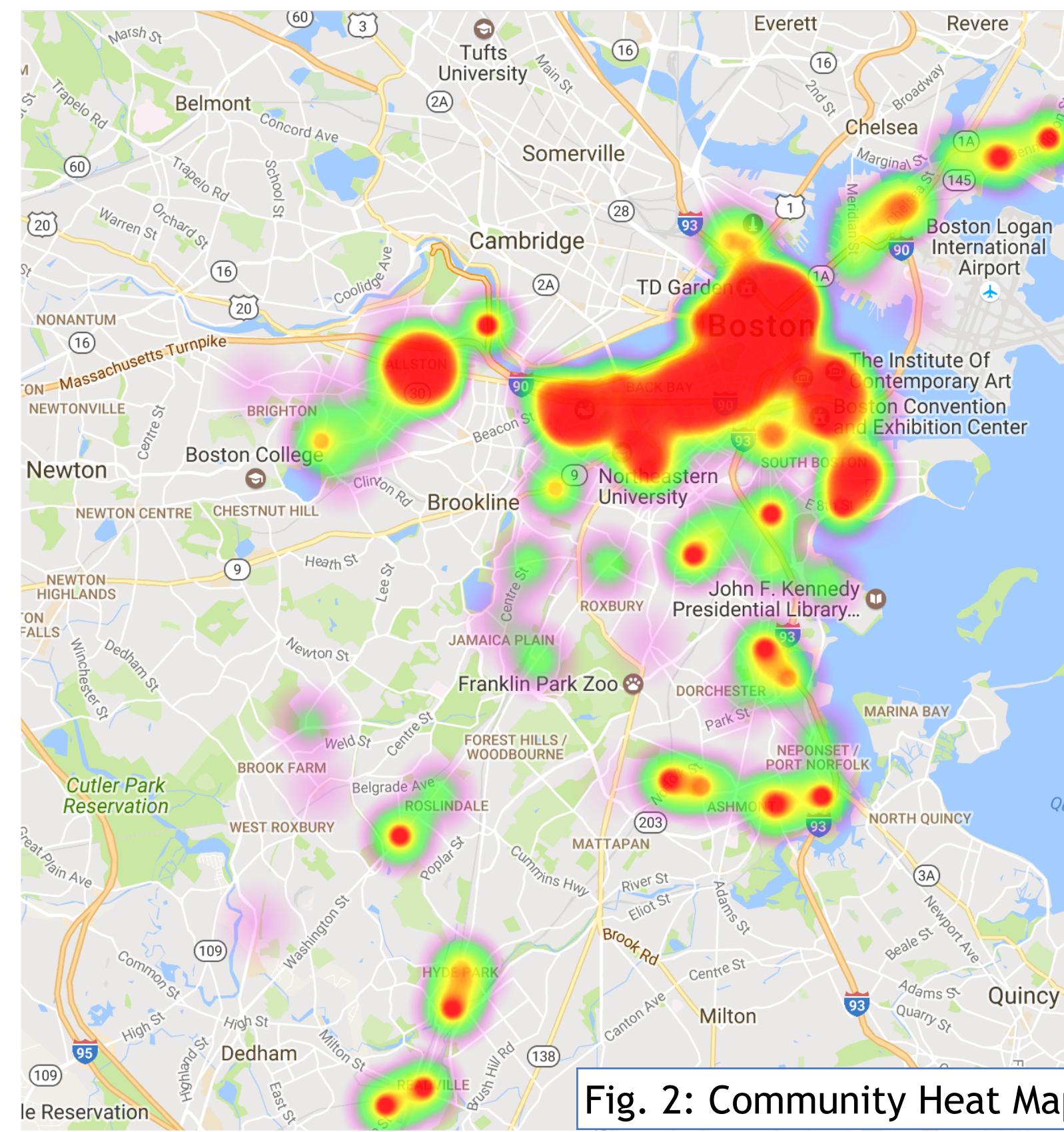


Fig. 2: Community Heat Map

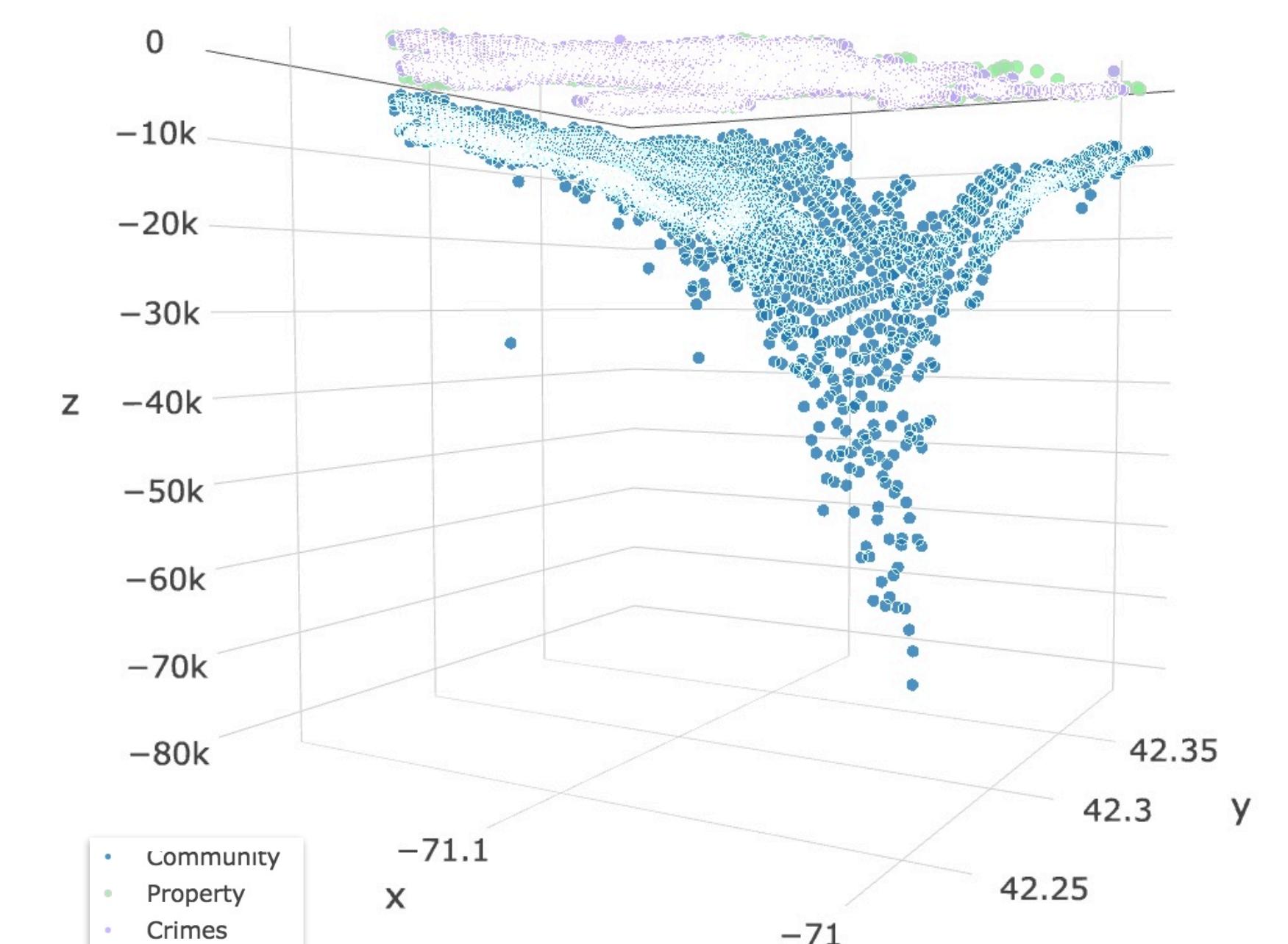
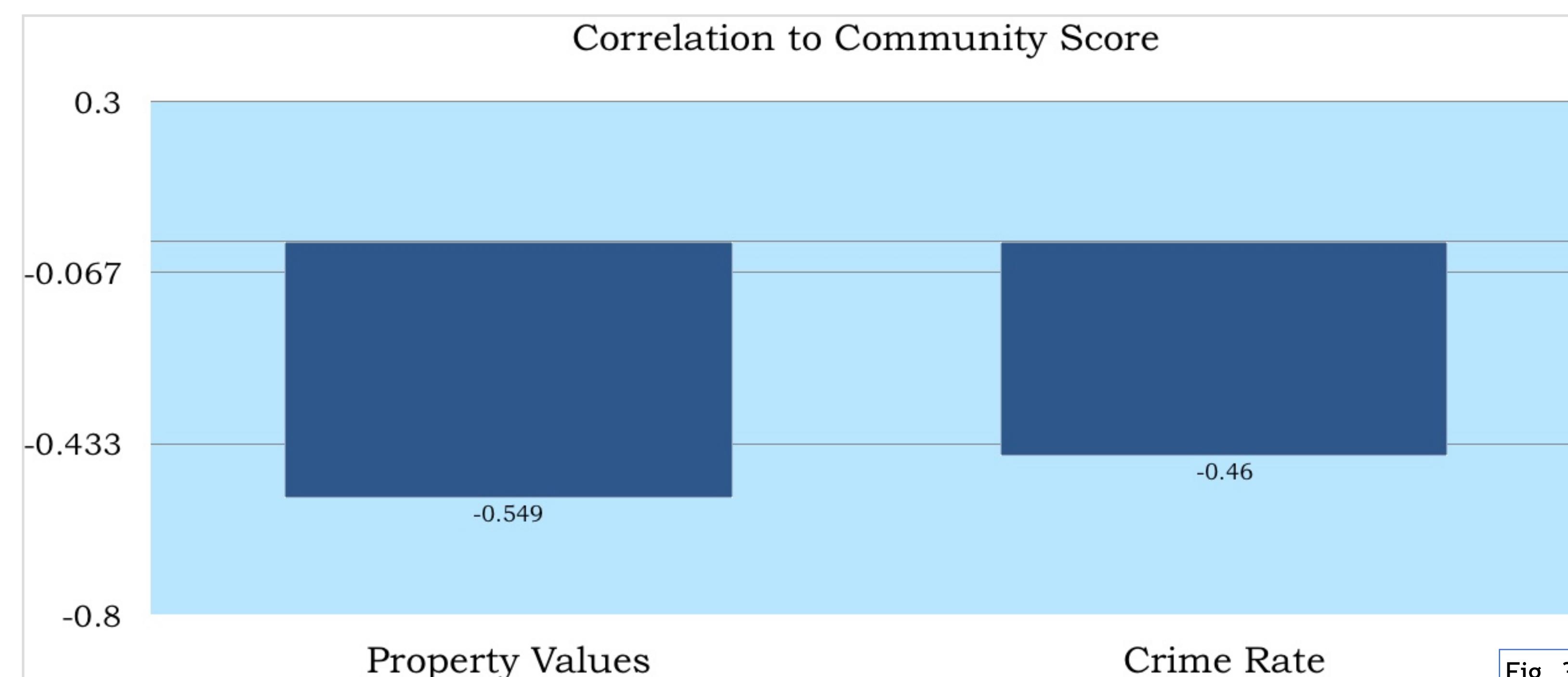


Fig. 4: Community Score vs. Crime vs. Property



Observations

Given the correlations, we found that there was a negative correlation between the community score and both crime and property values, respectively. Property values' correlation coefficient with community (as we defined it) appeared to be more negative than community's correlation coefficient with crime, which means that property value is more negatively correlated with community.

To extend on this, we may choose to extend on the idea of these correlations with community further. This would be interesting because the correlation coefficient may be a result, or even a victim of the ways that we obtained certain data sets. More specifically, the way that we created the community dataset may have introduced significant bias that we would have to overcome. To remedy this, we may find a point of comparison by calculating the p-value and considering every permutation of the data.

Analysis Techniques

- Fig. 1 — We looked at the crime data and the output from Boston Grid cell GPS Centers (1000x1000-ft cells), and kept a count of how many crimes occurred in that area.
- Fig. 2 — To make this, we take 1000x1000-ft cells in a grid of cells across Boston, and compute the indicator of how "strong" or "weak" community at that location is. The higher community areas have a redder coloring.
- Fig. 3 — This bar chart is made from a computation of the correlation coefficient of the "Crime Score" and "Community Scores" from Fig 1 and 2. Each value represents the correlation between the entries of each community dataset and the crime dataset.
- Fig. 4 — This shows the community score, crime count, and property score for each 1000x1000-ft. cell, all plotted against each other. This may help to better illustrate why the correlations are the way they are.

Sources

- <https://data.cityofboston.gov>