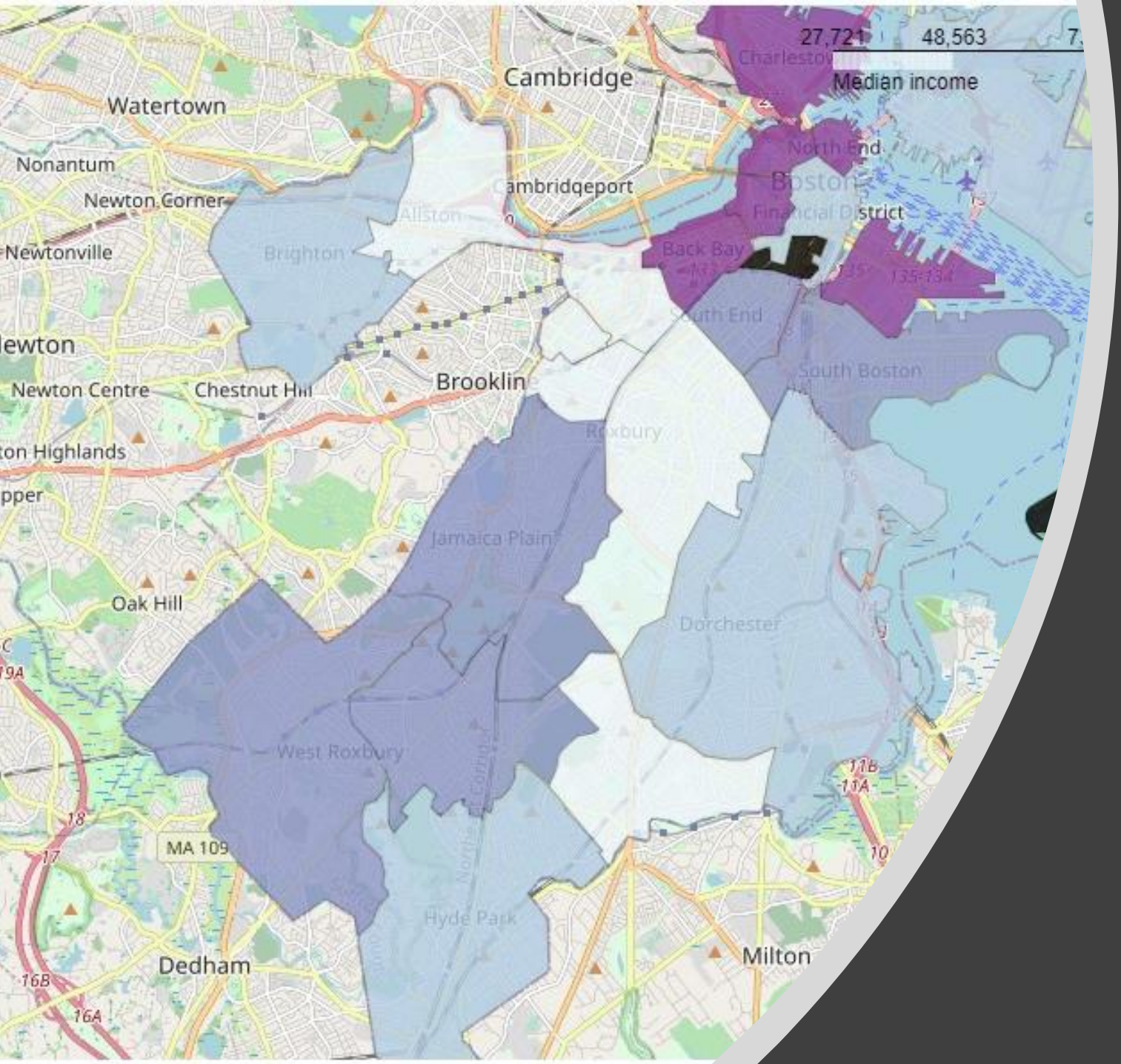


The Battle of Neighborhoods Part 2

Opening a restaurant in Boston

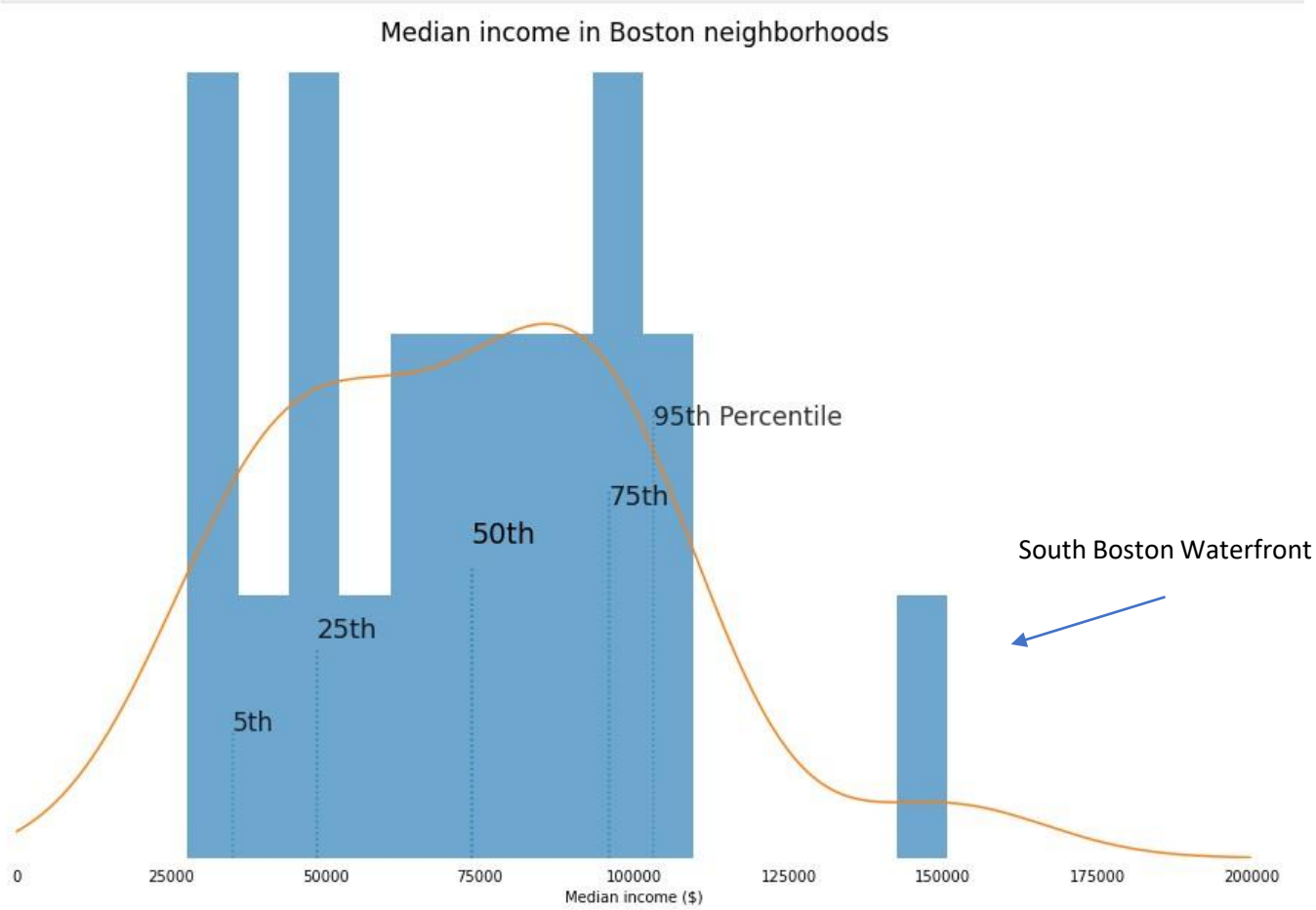
About Boston

- Boston, MA, USA
- The city itself is home to 645 thousand people, with the suburbs it forms Greater Boston and the population exceeds 4.5 million people
- High concentration of hi-tech companies, world's top universities, financial center, and center of medical and biotech developments
- Attracts over 20 million of tourists annually

Data acquisition

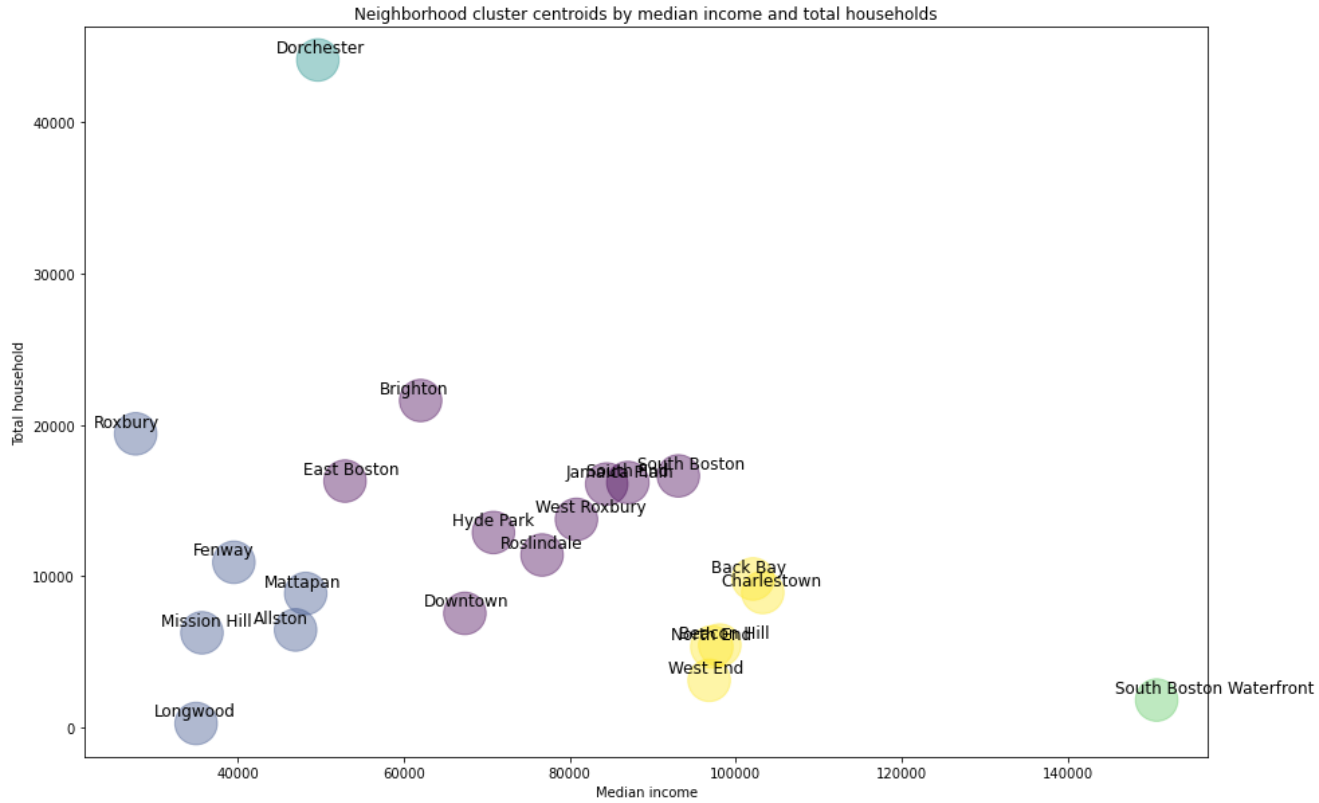
- The neighborhoods households and income data
  - Analyze Boston, the City of Boston's open data hub
  - Published by Department of Innovation and Technology
  - https://data.boston.gov/dataset/neighborhood-demographics
- The Neighborhood boundaries data
  - Analyze Boston, the City of Boston's open data hub
  - Published by Boston Maps
  - https://data.boston.gov/dataset/boston-neighborhoods
- Venues retrieved from Foursquare

Distribution of median income



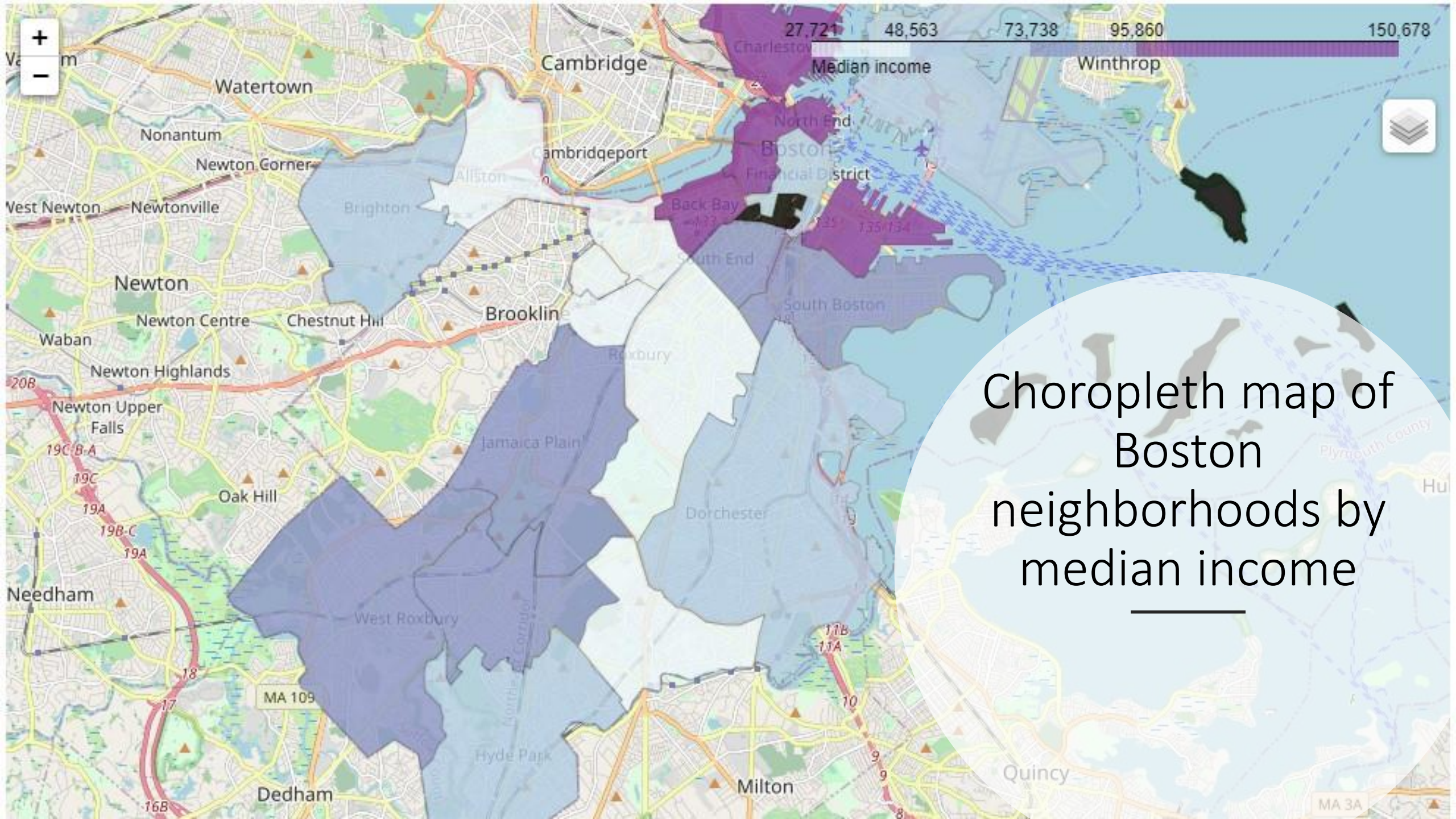
- Median income of South Boston Waterfront is significantly higher than other neighborhoods, thus is an outlier. Also, this neighborhood has one of the lowest number households

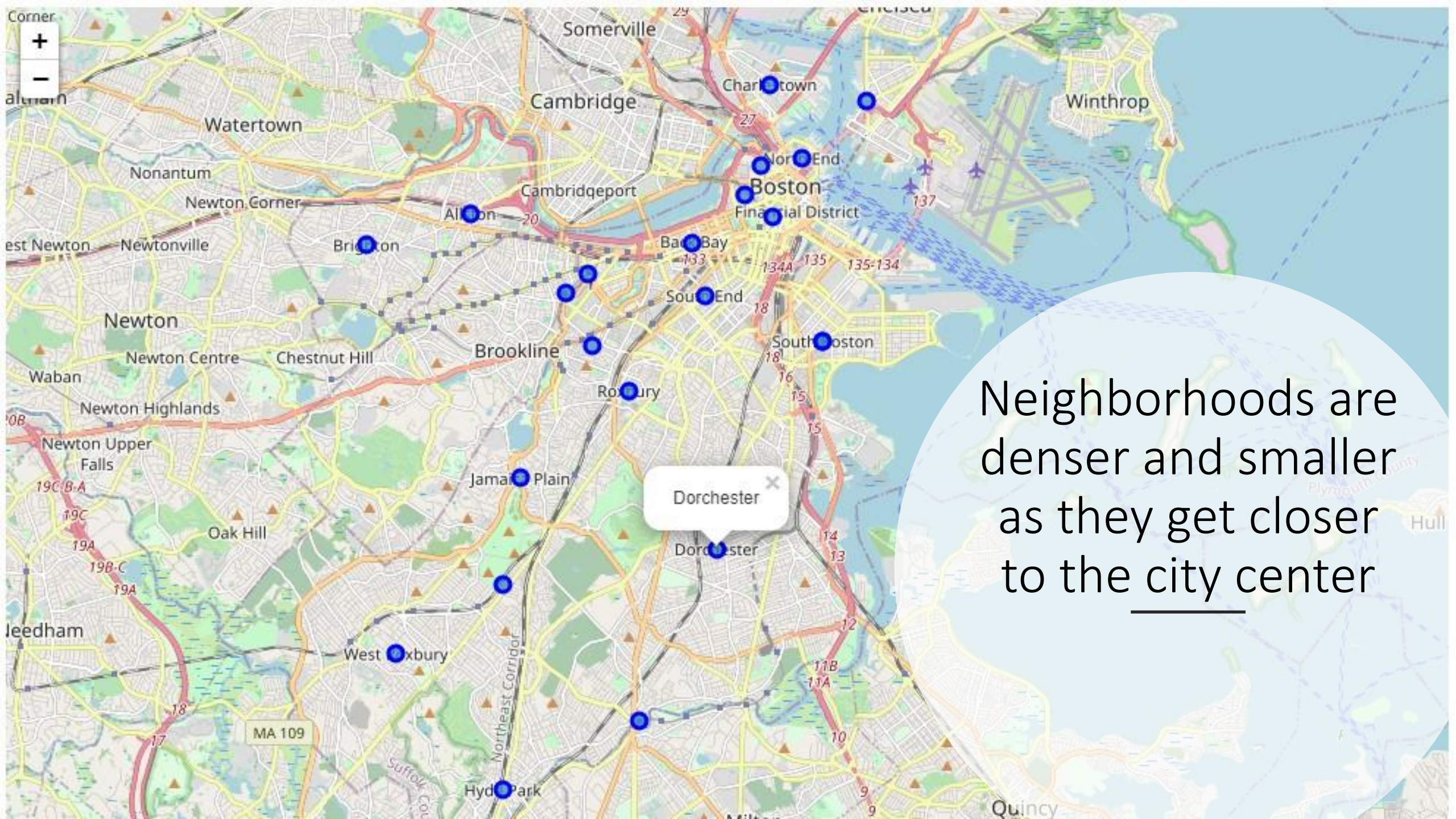
Neighborhood clusters by income: k-means algorithm



The neighborhoods in each cluster are similar to each other in terms of the features included in the dataset, i.e. median income and total households.
Example profiles for each group, considering the common characteristics of each cluster:

- 0: "Higher tier"
- 1: "Lower tier"
- 2: "Inbetweeners"
- 3: "Mid tier"
- 4: "The rich few"

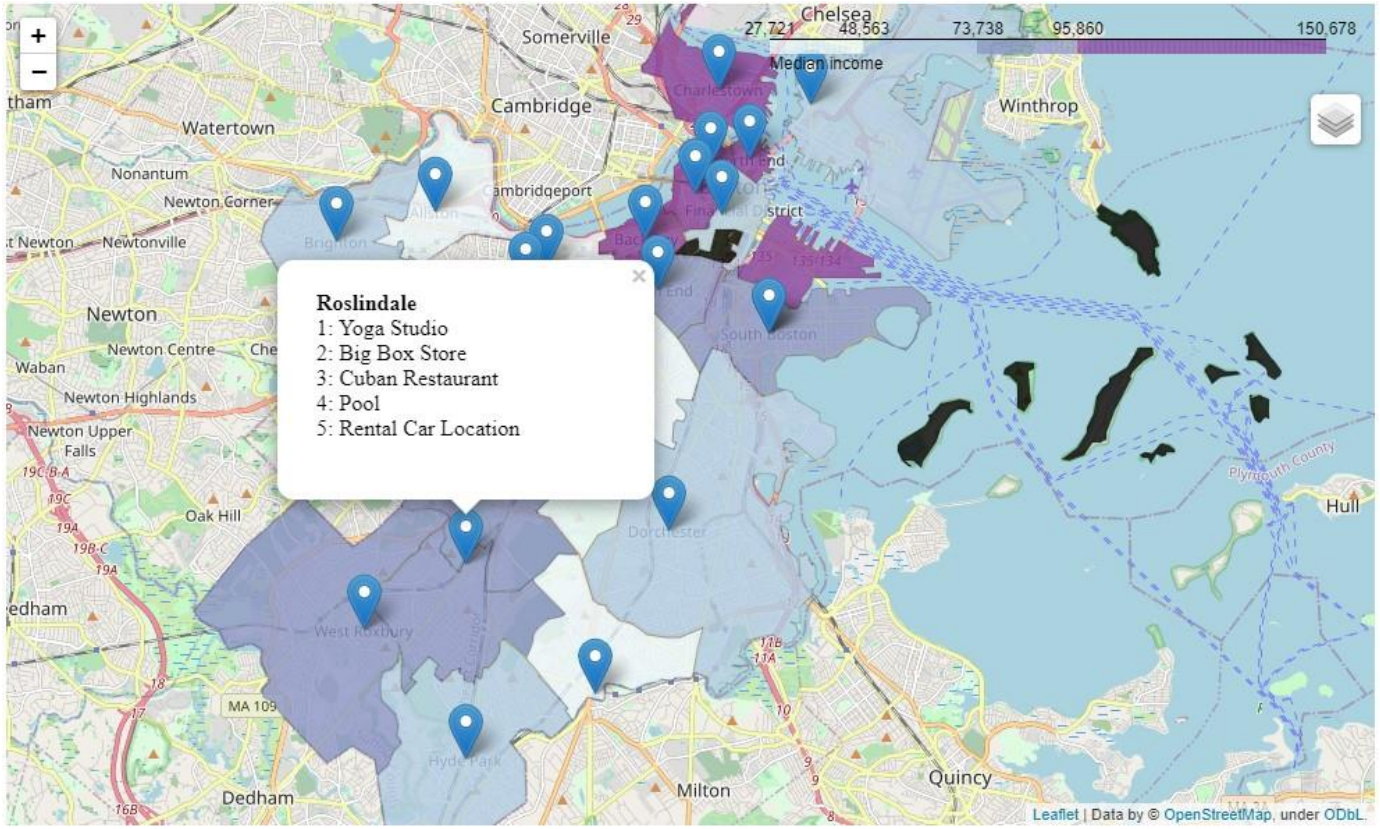

27,721
48,563
73,738
95,860
150,678
Median income
Cambridge
Watertown
Nonantum
Newton Corner
West Newton
Newtonville
Newton
Newton Centre
Chestnut Hill
Waban
Newton Highlands
Newton Upper Falls
Oak Hill
Brooklin
Brighton
Allston
Back Bay
North End
Winthrop
Jamaica Plain
Roxbury
Dorchester
South Boston
West Roxbury
Hyde Park
Needham
Dedham
Milton
Quincy
MA 109
MA 3A
Choropleth map of Boston neighborhoods by median income


Somerville
Cambridge
Watertown
Nonantum
Newton Corner
Cambridgeport
Boston
Winthrop
Newton
Newton Centre
Chestnut Hill
Brookline
Waban
Newton Highlands
Newton Upper Falls
Oak Hill
Dorchester
Jamaica Plain
West Roxbury
Hyde Park
Quincy
MA 109
Northeast Corridor
Neighborhoods are denser and smaller as they get closer to the city center

Venues from Foursquare

- Overall, there are **184** unique categories retrieved within **1000** meters radius from the coordinates of each of the neighborhoods.
- For example, within **1000** meters radius in Dorchester there are **21** venues across **19** unique categories.
- In total, for **22** neighborhoods in our data frame Foursquare returned **882** venues.

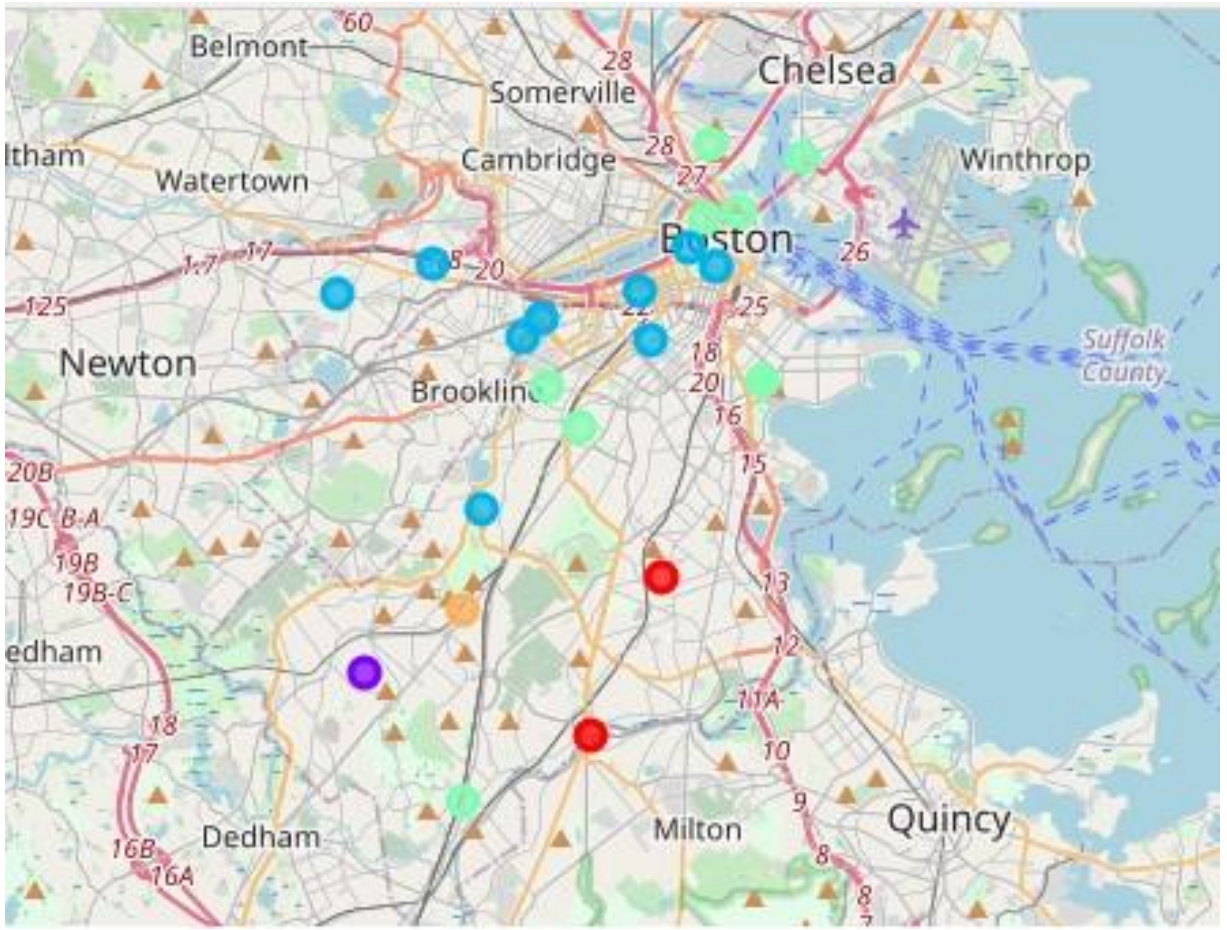
Top 5 venue categories for each neighborhood



Folium package allows neighborhoods to be explored interactively by visually displaying the boundaries and data from Foursquare.

In this snapshot top 5 most common venue types shown for Roslindale.

Neighborhood clusters by venue types : k-means algorithm



As seen before, Roslindale appears to be distinct from others and form its own cluster. This correlates also with income data – despite being away from downtown Boston, median income is higher that some neighborhoods that are closer to city center.

Neighborhood clusters by venue types

- According to the list of top 5 venue categories, **Cluster 0** and **Cluster 1** are not rich for restaurants. Restaurants in these neighborhoods are 5th most common categories, but diverse (Caribbean, Moroccan, and Southern/Soul).

- **Cluster 2** is the most diverse in its restaurants. Among top 5 venue categories there are American, Italian, Korean, French, Chinese, Asian, Thai cousins as well as Seafood, Falafel, and fast-food restaurants.

- **Cluster 3** seems to be popular with Italian cuisine judging by the number of pizza places these neighborhoods have. This may induce competitors or, on the contrary, drive crowd of pizza lovers to these neighborhoods. Therefore, further analysis needed if one would like to open here a restaurant with Italian cuisine.

- **Cluster 4** has only neighborhood. Roslindale neighborhood is distinct from others with its Yoga Studios as the most common categories, followed by Big box stores and Cuban restaurants.

Conclusion and way forward

- Since we have not specified the type of a restaurant or targeted customer segment, we limited this analysis only to describe the neighborhoods and their similarity by applying clustering algorithms.
- We identified clusters by income level and by venue type in each neighborhood with 1km radius
- This can help to narrow down the search for location for a restaurant
- Some additional ideas on how one can continue to refine recommendations:
  - Add population density data
  - Add demographic data such as age and occupation
  - Distance between venues
  - Reviews of competitor venues in the chosen location
  - Properties available to rent/buy, prices, taxes, etc.