The Battle of Neighborhoods | Finding a place in Boston, MA

Business Problem

Boston is a great place and there are plenty of people there to fill a new establishment. For a venue to succeed we need to be able to have our venue in the ideal location. In a big city, that can be hard because of the possibility of opening in a dead area. So, the problem is around where we should open this venue. We also need to know what kind of venue we should open.

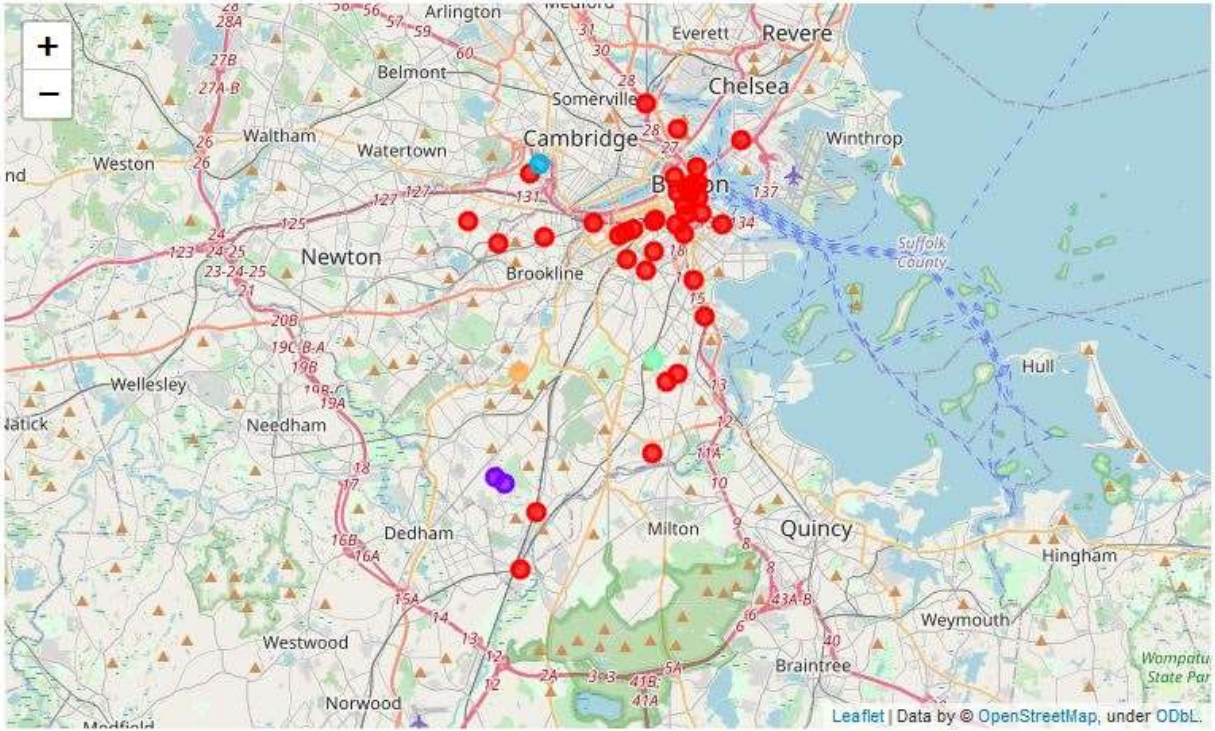
Data

- We need the name of the neighborhoods in Boston.
- Next we need the zip code of these areas.
- With the zip code we can derive the Longitude and Latitude of these neighborhoods.
- From this information we can get the Venue information that contains the venue name, category, longitude, latitude, and the neighborhood that this venue is found in.

Methodology

- We get the neighborhood names and zip codes by scraping the web. We can put this information into a data frame.
- From here we can get the longitude and latitude.
- Using the Foursquare API we can get the most popular venues in these neighborhoods.
- We can use the k-means clustering algorithm to cluster this data.
- From the clustered data we can derive results to our problem.

Clusters map



Results

- In the end, we can see that the best area to open a venue is in the area of cluster one.
  - Cluster one has the most neighborhoods. That means it has the most people.
- We can also understand that the best type of venue that we should open is a coffee shop or an Italian restaurant because those are the most popular in cluster one.