

Battle of Neighborhood - Chicago Neighborhood Analysis

1) Introduction/Business Problem

The basis of this study is to help a small group of investors planning to open a coding Franchise in Chicago. They wanted to choose top 10 communities in Chicago. The plan is to teach young kids who are in the age range 5 to 9. In order, for them to choose a location the community should have a high combined score based on the criteria listed below

Criteria

- High income area
- More Kids in the age group of 5 to 9
- Good male and female population in the age group of 25 to 45.

2) Data

I got all the data for this problem by searching for Chicago data. I found the following data in the CSV format in government sites

1. Per_Capita_Income.csv - This is formatted file that has income information for each of the community in the Chicago City
2. Chicago_Neighborhood_Zip.csv - Neighborhoods and the zip code
3. Chicago_Neighborhood_stats.csv - Census information of the Chicago community
4. Neighborhood_community.csv - Neighborhoods in each community

Foursquare API - To collect information on other venues/competitors in the Chicago community

3) Methodology

- Load the data
- Transform the data
- Find the top 10 communities
- Analyze the venue in these neighborhoods using the Foursquare API
- Do a onehot coding to cluster the data
- Find the optimum value using the Elbow curve
- Find the clusters and plot them