

Target the best neighborhoods to open a new night club in Boston

Introduction/Business Problem:

There are hundreds of night clubs opened in Boston nowadays. In recent years, as the population and economy are still growing, business man sees potential opportunities to open new night clubs and make profits.

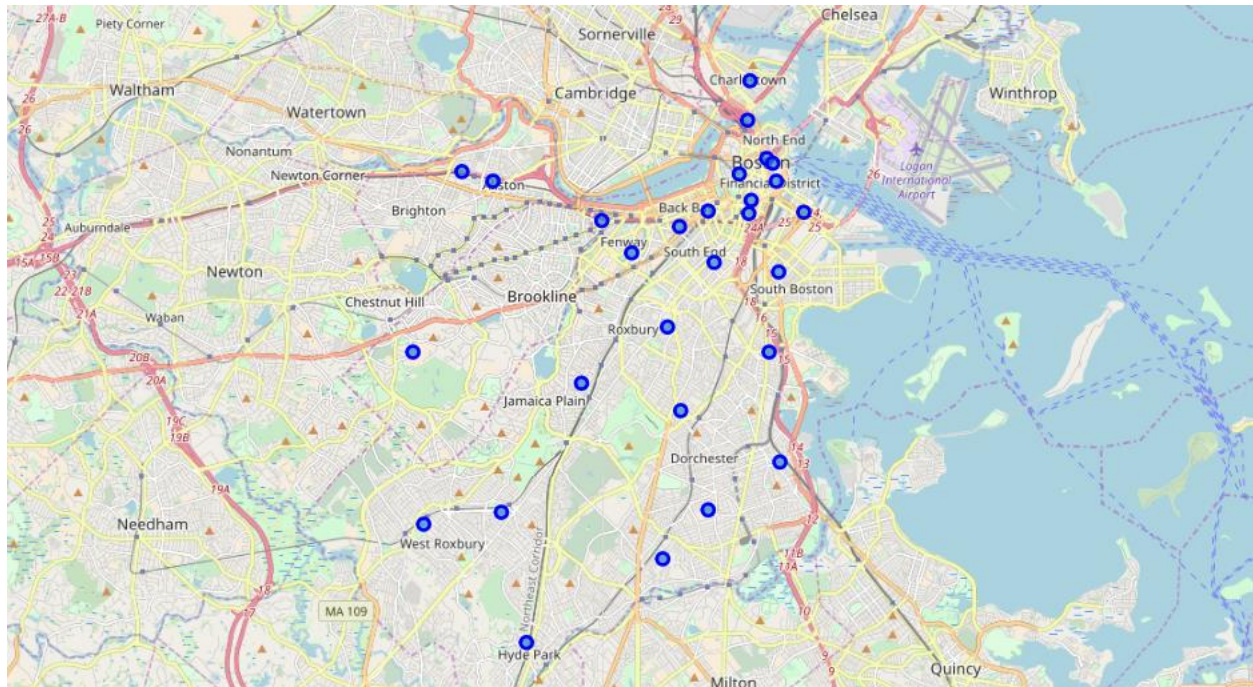
However, there are nearly 30 neighborhoods in Boston area. Each of them has different situations, such as different population, number of existing clubs, house price and income levels. Therefore, the question is that what the best area or neighborhood to open the new night club and how to make the decisions.

In this project, we would like to use the data we can collect from internet for free to build our simple model and help business owner to support their decision-making process. The good choice of the new night club should put multiple factors into considerations.

Data:

To build our model, we have to use different types of data. In general, we will need the followings:

- Neighborhoods: the data to define the neighborhoods of Boston area. For each neighborhood, we may also need the zip code that we can use to link other datasets later. We scrape the data from Boston's wiki page where defines the neighborhoods
- Geographic coordinate: for each zip code or neighborhood, we use geocoder to figure out the latitude and longitude which we can use to collect details from Foursquare.
- Number of existing night clubs: we use the coordinates for each neighborhood to get the number of night clubs by using Foursquare
- Census data: such as population, household size, income per each neighborhood. It's also originally from the Boston's wiki page. We need to do some preprocessing in order to get the data we like to use



ZipCode	Neighborhood	Venue	Area	Per capita income	Median household income	Median family income	Population	Number of households
0 02134	Allston	20.0	Allston	\$25,319	\$37,638	\$49,355	20,478	8,916
1 02163	Allston-Harvard Business School	28.0	Allston-Harvard Business School	\$21,915	\$43,889	\$91,190	1,842	562
2 02116	Back Bay/Bay Village	46.0	Back Bay/Bay Village	\$81,458	\$87,630	\$134,875	21,318	10,938
3 02115	Back Bay/Fenway-Kenmore	48.0	Back Bay/Fenway-Kenmore	\$21,654	\$23,677	\$50,303	29,178	9,958
4 02108	Beacon Hill/Financial District	43.0	Beacon Hill/Financial District	\$78,569	\$95,753	\$153,618	4,155	2,337
5 02114	Beacon Hill/West End	41.0	Beacon Hill/West End	\$65,865	\$79,734	\$169,107	11,933	6,752
6 02135	Brighton	18.0	Brighton	\$31,773	\$50,291	\$62,602	38,839	18,336
7 02129	Charlestown	41.0	Charlestown	\$56,267	\$89,105	\$98,445	17,052	8,083
8 02467	Chestnut Hill	5.0	Chestnut Hill	\$53,382	\$113,952	\$148,396	22,796	6,351
9 02111	Chinatown/Financial District/Leather District	40.0	Chinatown/Financial District/Leather District	\$56,716	\$44,758	\$88,333	7,616	3,390
10 02124	Dorchester-Codman Square-Ashmont	5.0	Dorchester-Codman Square-Ashmont	\$23,115	\$48,329	\$55,031	49,867	17,275
11 02122	Dorchester-Fields Corner	6.0	Dorchester-Fields Corner	\$23,432	\$51,798	\$50,246	25,437	8,216
12 02121	Dorchester-Mount Bowdoin	18.0	Dorchester-Mount Bowdoin	\$18,226	\$30,419	\$35,439	26,801	9,739
13 02125	Dorchester-Uphams Corner-Savin Hill	35.0	Dorchester-Uphams Corner-Savin Hill	\$22,158	\$42,298	\$44,397	31,996	11,481
14 02128	East Boston	41.0	East Boston	\$23,450	\$49,549	\$49,470	41,680	14,965

Methodology:

When we want to open a new business, we usually ask the following questions:

1. Are residents rich enough to spend money?
It's for entertainment purpose. We know for people who make higher income will tentatively have more potential to bring profits. Therefore, we like to see how rich each neighborhood is.
2. Do residents have such life style to enjoy the new business?
We know relatively young adults like to visit night clubs. We first thought about using age to do the ranking. However, using the age to make the decision is not enough since a younger neighborhood does not necessary mean it has more young adults since it could also cause by families with young kids who are not consumers for nigh club business. Therefore, we found family size could be a good indicator. For the smaller family size, it usually means singles or

young couples. For larger family size, it usually means the family have kids so most likely the family won't have too much time for nigh clubs.

3. Existing business: Boston is a fully developed metro area. It has had a lot of clubs opened already. If a neighborhood has a lot of nigh clubs, it may not be a good idea to open a new one due to the competition. However, we cannot just consider the venue numbers since areas may have different populations

To address the three questions above, we introduced a "score". The score is equal to:

Score = Per Capita Income Ranking + Household Size Ranking + Venue Per Population Ranking

Per Capita Income Ranking: sort by Per capita income in ascending order. It will indicate the wealth of the residents

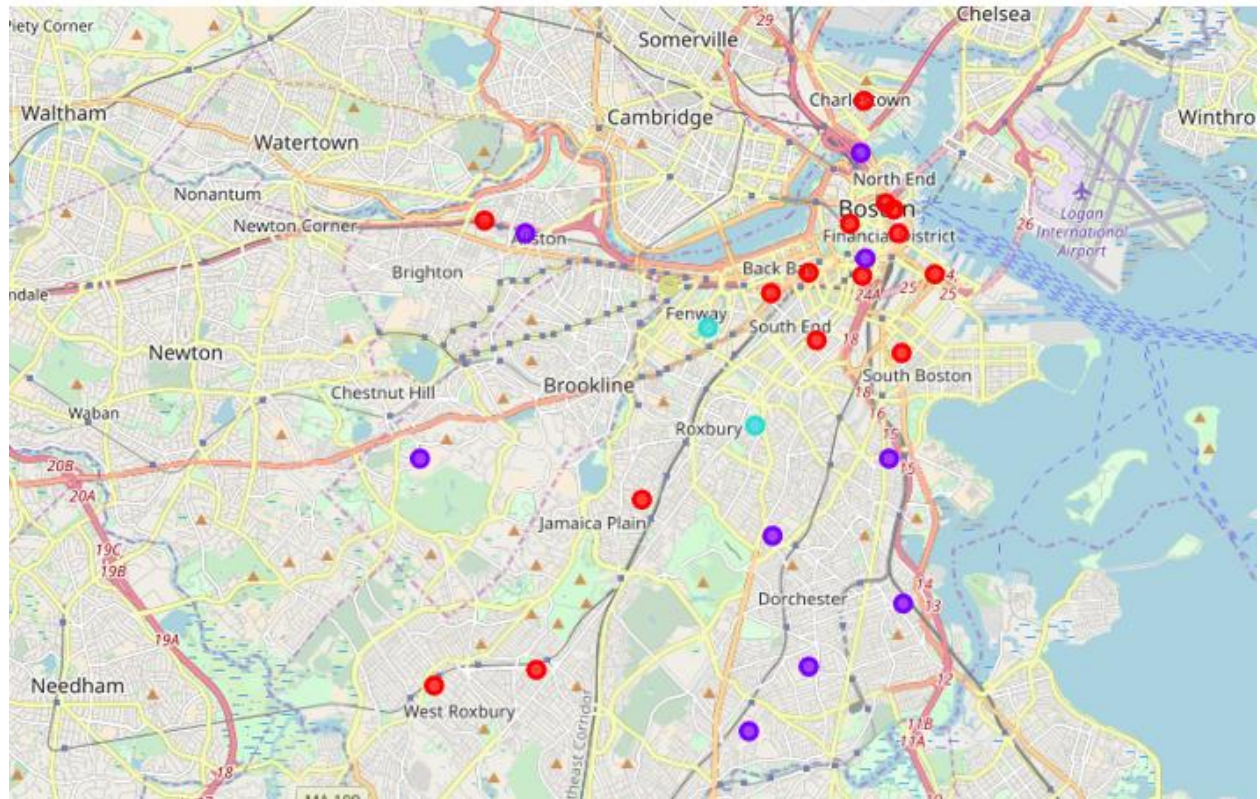
Household Size Ranking: total population divided by number of family in each neighborhood. It's sorted by descending order. The higher, the more young adults

Venue Per Population Ranking: number of night clubs divided by population in descending order. It will indicate how full the market is

In this model, the high the score is the more we favor to open a new nigh club. Once we get the scores, we will do the segmentation by give each neighborhood a label. We will use the label to indicate which area(s) are the best to open the new business

Results :

The downtown area plus the surrounding neighbors are better area to open the new clubs



Discussion:

Our model is a simple linear model based on some assumptions. In general, it's good enough to show the difference. However, the more accurate way is to figure out the similarities of each neighborhood and bring more factors into considerations

The Foresquare API can only return up to 100 venues for each query. So if the area has more than 100 clubs, the way we calculate the total number of venue may not accurate.

Conclusion:

In this project, we use all public information we get from open sources tools and wiki page. By using our heatmap, the business owner can decide which area to go.