Battle of the Neighborhoods: Top businesses in a city

Sudha N

Jan 20 2020

1. Introduction

EasyBiz is a New Age Marketing company that seeks out local businesses in major cities to sell Social Media PR services. They manage all facets of Public Relations for their clients by working with several third-party companies in implementing web, social presence and by various types of promotions benefiting their clients.

In the past, EasyBiz had worked with companies in Los Angeles and helped them become profitable by enhancing Social Media presence, by adding the client's location as a Poke stop [to target millennials who play Pokemon Go], by bringing foot-traffic to the business by implementing offers such as BOGO [Buy One Get One] and several other PR services.

Based on the success in Los Angeles by putting together portfolios targeted to businesses, EasyBiz's unique business model is profitable if they manage many of the same type of businesses. To accomplish this bulk sales model, EasyBiz prepares targeted portfolios to present to all potential clients in the same category, as it is easier for potential clients to visualize and relate to their services if the examples are in their domain.

Based on previous experiences, their main leads are businesses in urban neighborhoods, whose clientele is mainly urban, who live in and around downtown areas of major cities, whose mode of commute is walk.

Identifying the top business categories, that are commutable in each major city is the problem we would like to solve here.

1.1 Business problem

Identify the top 10 business categories within a one-mile radius of a major city.

2. Data

We will primarily be using FourSquare Locations data for the cities of New York and Toronto and obtain all the businesses within a one-mile radius of the two cities. We would then determine the Categories that are most frequent in the cities.

With this list of business categories, EasyBiz can put together a portfolio using examples from the same business category that they had already implemented in Los Angeles. This portfolio will help EasyBiz reach many businesses in the same category with the same portfolio, reducing their preparation time with increased chance of landing multiple clientele in the same business.

For example, if "Deli" is one of the top categories in New York or Toronto, then EasyBiz would compile a case study using a Deli they had already worked with, listing out specifics like the following:

- Performance of the Deli before they became an EasyBiz client
- Metrics for top items sold
- EasyBiz implemented marketing campaigns
- Social media presence or boost in presence if they already had a presence
- Types of promotions that were implemented [Buy one, Get one etc.]
- Customer reviews and feedback process, documenting reviews in a popular platform
- Increased foot traffic during the campaigns
- Foot traffic after one month, two months up to a year after the campaigns were implemented
- Performance of the Deli one year after they became an EasyBiz client

2.1 Data Sources

- New York City data containing neighborhoods, boroughs, latitudes and longitudes are obtained from https://cocl.us/new_york_dataset
- List of postal codes of Canada are obtained from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

2.2 Data Wrangling

We use Json library to handle Json files as raw data for New York City is in this format. Using the Pandas library, this raw data is transformed into a pandas dataframe, by first creating an empty dataframe and then looping through the data and filling the dataframe, one row at a time. Then using the geopy library, latitude and longitude values are obtained for all the entries in the the New York City dataframe. Using this data, FourSquare service is called to return venues that are within 500 ft of each of the neighborhoods.

3. Analysis

The analysis starts with initial data with the Neighborhood, Borough, Latitude and Longitude transformed into a pandas dataframe. The dataset comprises of data for all 5 boroughs and 306 neighborhoods in New York City.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Using the latitude and longitude of all the neighborhoods in New York city, FourSquare service is called which returns all the associated details for places of interest that are within a one mile radius of the center of the neighborhood.

After having obtained all the details from all the neighborhoods in a city, we find that the Category for each of the venues is stored in the “Categories” tag. This value is extracted and saved for each of the venues that are returned. If there is more than one category returned, then we simply use the first value. This way, a new dataset, which is also a pandas dataframe is created with the name of the venue, the venue’s category and its latitude and longitude.

	name	categories	lat	lng
0	Bella Abzug Park	Park	40.755580	-74.000344
1	505W37	Residential Building (Apartment / Condo)	40.756909	-73.998007
2	Legacy Records	American Restaurant	40.757565	-73.998111
3	Sean Kelly Gallery	Art Gallery	40.756022	-73.998306
4	Manhattan Aquariums	Pet Store	40.757029	-73.999085

We need to look at the top venues or businesses in each of the neighborhoods in a city that are in a walkable distance from the center of the city. This is to attract mainly urban population whose primary mode of commute is by walk. In this example, we look at the venues or businesses in New York City that are within a mile from each of the neighborhoods. Continuing the process or getting this information from FourSquare for all the neighborhoods in New York city results in a dataframe with a row for each business or venue for each of the neighborhoods with exactly one Venue Category for each of the venues.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Shell	40.894187	-73.845862	Gas Station
2	Wakefield	40.894705	-73.847201	Shepherd ENG Heating Cooling And Refrigeration	40.895539	-73.848241	Construction & Landscaping
3	Wakefield	40.894705	-73.847201	Pitman Deli	40.894149	-73.845748	Food
4	Co-op City	40.874294	-73.829939	Modell's Sporting Goods	40.872584	-73.829532	Sporting Goods Shop

Once the venues and the associated venue categories within a mile are determined for each of the neighborhoods in the city, we then determine how many venues/businesses are returned for each of the neighborhoods. This is to ensure that there was atleast one business returned for each of the neighborhoods.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Allerton	11	11	11	11	11	11
Annadale	1	1	1	1	1	1
Arden Heights	1	1	1	1	1	1
Arlington	2	2	2	2	2	2
Arrochar	2	2	2	2	2	2
Astoria	3	3	3	3	3	3
Astoria Heights	5	5	5	5	5	5
Auburndale	3	3	3	3	3	3
Bath Beach	4	4	4	4	4	4
Battery Park City	28	28	28	28	28	28

The next step is to determine the relative frequency of each of the venues/businesses within each neighborhood. In this example, we see that in Bayside, Spa is a top business whereas in Hillcrest, Ice Cream Shop may be a top business. After this step, we then take the top businesses in each of the neighborhoods and then find the most common businesses from this list.

```
----Bayside----
                       venue  freq
0                        Spa  0.12
1                Pizza Place  0.12
2      Vietnamese Restaurant  0.12
3           Greek Restaurant  0.12
4           Sushi Restaurant  0.12
5             Cosmetics Shop  0.12
6           Asian Restaurant  0.12
7   Mediterranean Restaurant  0.12
8                 Non-Profit  0.00
9       Other Great Outdoors  0.00
10              Optical Shop  0.00
```

```
----Hillcrest----
                              venue  freq
0                  Ice Cream Shop  0.12
1                     Coffee Shop  0.12
2                       Food Truck  0.12
3                       Food Court  0.12
4                       Donut Shop  0.12
5                      Pizza Place  0.12
6       College Academic Building  0.12
7                      Deli / Bodega  0.12
8   Vegetarian / Vegan Restaurant  0.00
9           Outdoors & Recreation  0.00
10              Outdoor Sculpture  0.00
```

```
----Upper East Side----
                              venue  freq
0                          Boutique  0.13
1                    Clothing Store  0.07
2                    Chocolate Shop  0.07
3                           Bakery  0.07
4                         Gift Shop  0.07
5                Miscellaneous Shop  0.07
6                     Grocery Store  0.07
7                      Burger Joint  0.07
8                         Juice Bar  0.07
9                       Art Gallery  0.07
10  Vegetarian / Vegan Restaurant  0.07
```

The final step in this process is to consolidate the most common venues in each of the neighborhoods. This step identifies which are the top businesses that have the most relative frequency in most neighborhoods. In this case, Deli (sandwich shops) and Bodega (kind of a wine bar) are top categories in 26 of the 306 NYC neighborhoods. Armed with this data, EasyBiz can put together a marketing strategy targeted to these business categories when they meet with businesses in the neighborhoods.

	1st Most Common Venue	Neighborhood
0	Deli / Bodega	26
1	Pizza Place	16
2	Playground	11
3	Park	8
4	Bus Station	8
5	Italian Restaurant	8
6	Bar	7
7	Grocery Store	7
8	Chinese Restaurant	7
9	Bank	5

4. Conclusion

With the knowledge of which business is the most frequent in which neighborhood, EasyBiz can put together a marketing strategy targeted to just these business categories when they meet with the said businesses in the neighborhoods. In this case, the presentation that EasyBiz uses will be used in 26 neighborhoods at each of the businesses, which maximizes the probability that some of the business presentations end with ah positive outcome where the business signs up with Easy Biz to become their social media partner.

Ensuring that the strategy presentation is tailored to each of the businesses is key to this exercise since EasyBiz has determined that it’s easier to convince potential clients to become clients when the business owner is able to relate to the information presented and the marketing strategy proposed. This concludes this exercise, where the final list of most common businesses for each neighborhood is the desired outcome.