2021

Office move from Battery Park City, Manhattan to Chicago

SIMON HOWARTH

Contents

Introduction

A start-up company based in Battery Park City, Manhattan has just been purchased by a larger company, part of the purchase terms requires the start up to move their office to Chicago. The location in Chicago can be decided by the start-up's leadership team, as long as the office is in a Chicago zip code.

The leadership team of the start-up company really like their current office location and realise the location is in an area the staff also enjoy. The leadership team of the start-up company are unfamiliar with Chicago but would like the new location to be like Battery Park City as this would potentially encourage their current staff to relocate with them rather than having to find new talent.

The audience for this project is the leadership team of the start-up and they would like to be presented with a handful of new locations in Chicago to relocate their office that are in a similar location to Battery Park City.

Data

Data for the Chicago Neighborhoods and zip codes will be scraped from https://www.dreamtown.com/maps/chicago-zipcode-map.

The latitude and longitude data for all US zip codes can be found at https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/ and exported as a csv.

Accessing the FourSquare API will grant access to information about venues within a radius of specific latitude and longitude – these latitude and longitude points will be gathered from opendatasoft.com as mentioned above.

Once the data from the three sources listed above is gathered, they will be placed into a dataframe (Table 1) and prepared to be used for K Means Clustering and comparison with Battery Park City.

Table 1

	Zip Code	Zip Code Latitude	Zip Code Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	60601	41.886456	-87.62325	sweetgreen	41.884964	-87.624728	Salad Place
1	60601	41.886456	-87.62325	Chicago Architecture Center	41.887720	-87.623650	Tour Provider
2	60601	41.886456	-87.62325	Roti Modern Mediterranean	41.886048	-87.624948	Mediterranean Restaurant
3	60601	41.886456	-87.62325	LondonHouse Chicago, Curio Collection by Hilton	41.887832	-87.625426	Hotel
4	60601	41.886456	-87.62325	Virgin Hotels Chicago	41.886065	-87.625853	Hotel

Location and venue data from FourSquare regarding Battery Park City has already been gathered from previous work and will be used to cluster the new data regarding Chicago Zip Codes and nearby venues.

Unfortunately the location data for each neighbourhood in Chicago was unable to be located so this project will focus on the zip code in Chicago rather than neighbourhood.

Modules, Packages, and Libraries used in the project include:

- Numpy
- Pandas
- Json
- Geopy
- Requests
- Matplotlib
- Sklearn
- Folium
- Bs4 (BeautifulSoup)

Methodology

After data collection was complete the locations of the Chicago zip codes were all plotted on a map (Figure 1) to get an idea of the area.

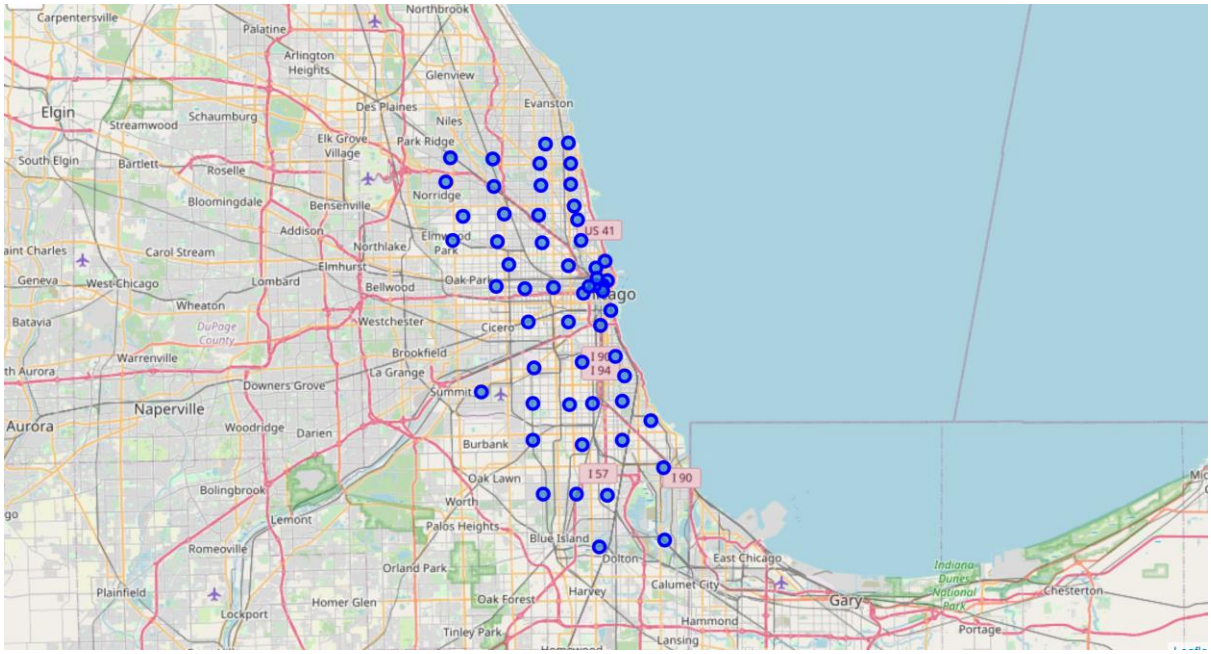

Figure 1

FourSquare data was accessed to find local venues within a walkable distance of the centre of each post code (Table 1).

A dataframe was created with each venue getting encoded to it's zip code and the mean of the frequency of occurrence in each category was calculated (Table 2).

Table 2

	Zip Code	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport	American Restaurant	Amphitheater	Animal Shelter	Antique Shop	Aquarium	Arcade	Argentinian Restaurant	Art Gallery	Art Museum
0	60601	0.000000	0.00	0.00	0.000000	0.000000	0.020000	0.01	0.00	0.000000	0.00	0.000000	0.00	0.010000	0.010000
1	60602	0.000000	0.00	0.00	0.000000	0.000000	0.010000	0.01	0.01	0.000000	0.00	0.000000	0.00	0.010000	0.010000
2	60603	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.01	0.00	0.000000	0.00	0.000000	0.00	0.010000	0.010000
3	60604	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.01	0.00	0.000000	0.00	0.000000	0.00	0.010000	0.010000
4	60605	0.000000	0.00	0.00	0.000000	0.000000	0.010000	0.01	0.00	0.000000	0.06	0.000000	0.00	0.000000	0.000000
5	60606	0.000000	0.00	0.00	0.000000	0.000000	0.010000	0.00	0.01	0.000000	0.00	0.000000	0.00	0.000000	0.010000
6	60607	0.000000	0.00	0.00	0.000000	0.000000	0.010000	0.00	0.00	0.000000	0.00	0.000000	0.00	0.000000	0.000000
7	60608	0.000000	0.00	0.00	0.000000	0.000000	0.000000	0.00	0.00	0.000000	0.00	0.000000	0.00	0.020000	0.000000
8	60609	0.000000	0.00	0.00	0.000000	0.000000	0.011905	0.00	0.00	0.011905	0.00	0.011905	0.00	0.047619	0.000000
9	60610	0.000000	0.00	0.00	0.000000	0.000000	0.030000	0.00	0.01	0.000000	0.00	0.000000	0.00	0.000000	0.000000

Venue frequency data had already been collated in a previous project and this was appended to Table 2. This resulted in a new dataframe with additional columns that weren't already included in Table 2, these additional columns resulted in null values in the original rows and new columns, these null values were all converted to 0 (Table 3).

Table 3

	Location	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport	American Restaurant	Amphitheater	Animal Shelter	Antique Shop	Aquarium	Arcade	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store
53	60660	0.0	0.0	0.0	0.01	0.0	0.000000	0.0	0.0	0.01	0.0	0.0	0.0	0.0	0.0	0.00
54	60661	0.0	0.0	0.0	0.00	0.0	0.010000	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.01
55	60707	0.0	0.0	0.0	0.00	0.0	0.020000	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.00
56	60827	0.0	0.0	0.0	0.00	0.0	0.000000	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.00
57	Battery Park City	0.0	0.0	0.0	0.00	0.0	0.012048	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.00

With this dataframe K Means Clustering could be implemented to group all the locations together, and importantly see which locations in Chicago were similar to Battery Park City.

The elbow method was implemented to find an optimal number of clusters. Looking at Figure 2 there are 2 potential elbows at k = 5 and k = 11.

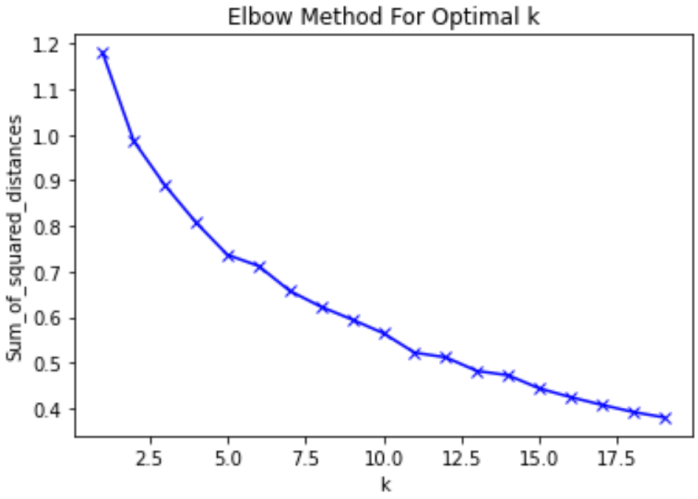

Figure 2

Using k= 5 would result in too many locations being returned in each cluster leaving too many options to consider. Running the K Means Cluster produced fewer clusters but there were still too many in cluster containing Battery Park City to make a recommendation of where to move the office. Increasing the k value again to k = 12 resulted in 9 zip codes with similar venues to Battery Park City.

Results

The 9 zip codes with similar venues to Batter Park City can be seen in Table 4.

Table 4

	Cluster	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	Latitude	Longitude
0	3	60601	Hotel	New American Restaurant	Bar	Steakhouse	Park	Italian Restaurant	Grocery Store	Café	Salad Place	Theater	41.886456	-87.623250
1	3	60602	Hotel	Steakhouse	Park	Theater	Italian Restaurant	Coffee Shop	Bar	Seafood Restaurant	Mediterranean Restaurant	Donut Shop	41.882937	-87.628740
2	3	60603	Hotel	Park	New American Restaurant	Theater	Coffee Shop	Bar	Steakhouse	Italian Restaurant	Pizza Place	Donut Shop	41.880446	-87.630140
3	3	60604	Hotel	Coffee Shop	Park	Theater	Steakhouse	Grocery Store	Italian Restaurant	Pizza Place	Salad Place	Gastropub	41.877589	-87.628180
5	3	60606	Hotel	New American Restaurant	Italian Restaurant	Coffee Shop	Bar	Steakhouse	Donut Shop	Mexican Restaurant	Mediterranean Restaurant	Grocery Store	41.882582	-87.637600
9	3	60610	Steakhouse	Hotel	Pizza Place	New American Restaurant	Italian Restaurant	Restaurant	Mexican Restaurant	Bar	Gym / Fitness Center	Seafood Restaurant	41.898582	-87.637100
10	3	60611	Hotel	American Restaurant	New American Restaurant	Café	Pizza Place	Steakhouse	Restaurant	Grocery Store	Gym / Fitness Center	Coffee Shop	41.904667	-87.625040
48	3	60654	Hotel	Italian Restaurant	Steakhouse	Bar	Mediterranean Restaurant	Coffee Shop	New American Restaurant	Mexican Restaurant	Seafood Restaurant	Burger Joint	41.888627	-87.635380
54	3	60661	Coffee Shop	Hotel	Italian Restaurant	Steakhouse	New American Restaurant	Donut Shop	Pizza Place	Restaurant	Sandwich Place	Burger Joint	41.882082	-87.644610
57	3	Battery Park City	Park	Coffee Shop	Hotel	Clothing Store	Memorial Site	Playground	Gym	Boat or Ferry	Plaza	Sandwich Place	40.711932	-74.016869

Figure 3 shows the 12 different clusters plotted on the map of Chicago.

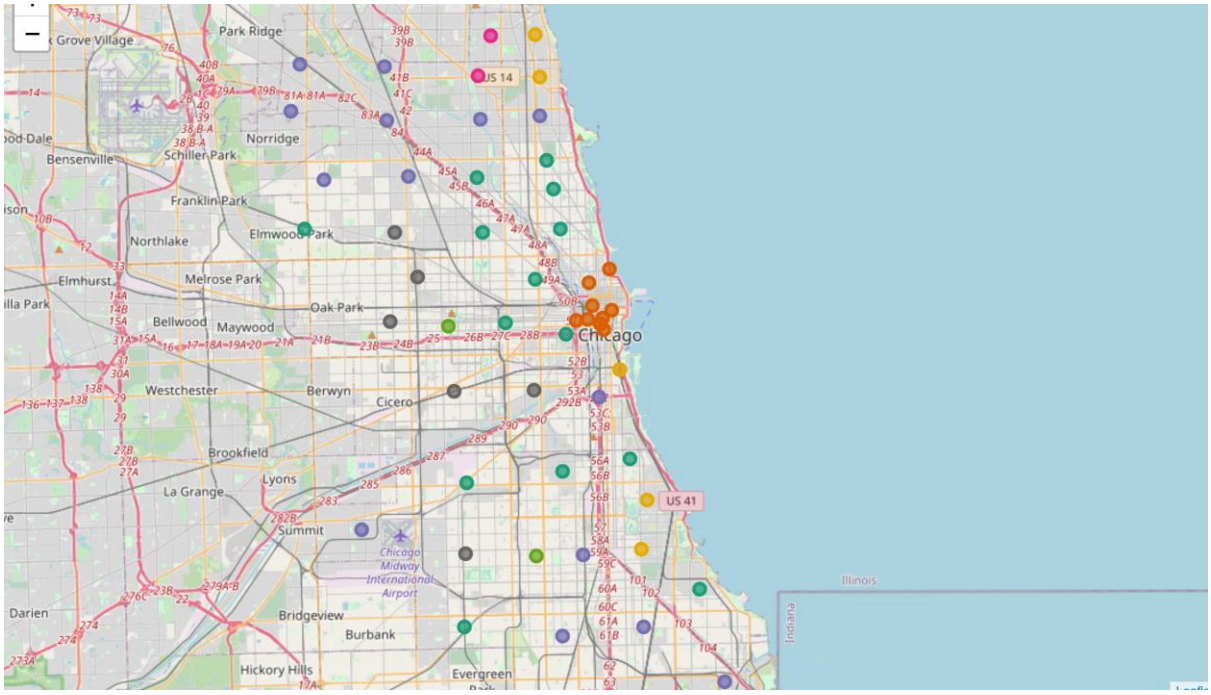

Figure 3

The cluster containing Battery Park City was cluster 3, Figure 4 shows the cluster locations zoomed in.

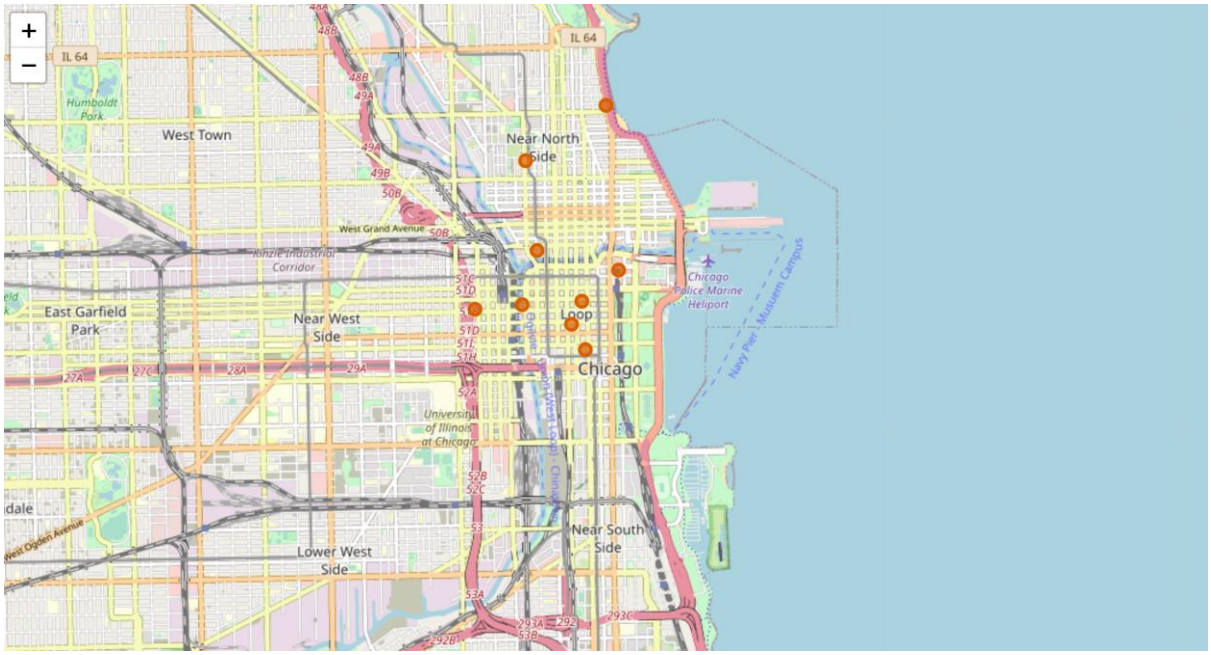

Figure 4

Discussion

There are 9 location in Chicago that K Means Clustering has discovered as being similar to Battery Park City based on FourSquare venue data.

The 9 locations in Chicago have the following zip codes:

- 60601
- 60602
- 60603
- 60604
- 60606
- 60610
- 60611
- 60654
- 60661

The locations are all within a generally similar area of Chicago (Figure 4) and the venues that appear to be the major grouping decisions is park, coffee shops, and hotels (Table 4) - but it is probably expected that the vicinity to good hotels will be less of a factor compared to parks and coffee shops when encouraging staff to move city.

The start up company looking to relocate to a similar area in Chicago should focus their search in the zip codes listed above, or more widely in The Loop and Near North Side.

Conclusion

The report has identified 9 zip codes (or 2 areas) of Chicago for the start-up company to focus on when looking to move office. To encourage their staff to move from New York to Chicago it might be a good idea to focus on the new office's proximity to good coffee shops and parks.

Looking at the cost of renting an office space was out of scope of this project but having narrowed down the search area in Chicago it would now be an easier to task to compare the cost of office space.