Capstone Project

Ideal Location to Open a Chinese Restaurant in Dublin

- Dhruv Panghal

Table of Contents

Table of Figures

1. Introduction

Background

As the capital city of Ireland, Dublin is a city which is popular among the tourists. On the other hand, there are many people from different countries since most European IT and high-tech companies such as Google and Facebook are headquartered here. It should be pointed out that Dublin is full of potentials for the future development and enlargement due to the booming of tourism and immigration of the employees in the giant organizations. However, the number of restaurants in Dublin is relatively low especially Chinese restaurants considering its enlarging trend in the near future. Therefore, this project is to analyze the status of restaurants particularly Chinese restaurants to identify the optimal locations to open a new Chinese restaurant.

Business Problem

In this project, we aim to find the optimal location to open Chinese restaurant in Dublin.

The question of business interest would be:

"Would you recommend a location in Dublin to open a new Chinese restaurant?"

The number of Chinese restaurants in Dublin is relatively low. There might be potential benefits if it opens. Therefore, this report will be targeted to stakeholders interested in opening a Chinese restaurant in Dublin, Ireland. Since there are lots of restaurants in Dublin, we will try to detect locations that are not already crowded with restaurants. We are also particularly interested in areas with no Chinese restaurants in vicinity. We would also prefer locations as close to city center as possible, assuming that first two conditions are met. In this project, the selection of the locations is assumed to be highly dependent on its nearby environment instead of the economic factors such as rental prince. The main task would be to select the 5 possible locations to open the cinema and explain them to stakeholders.

2. Data Manipulation

Data Criteria

Based on definition of our problem, factors that will influence our decision are:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to Italian restaurants in the neighborhood, if any
- distance of neighborhood from city center

We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Data Sources

Following data sources will be needed to extract/generate the required information:

- centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using Google Maps API reverse geocoding;

• number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API;

• coordinate of Dublin center will be obtained using Google Maps API geocoding of well-known Dublin location.

3. Methodology

In this project we will focus our efforts on detecting areas of Dublin that have low restaurant density, particularly those with low number of Chinese restaurants. We will limit our analysis to area 5km around city center. The main steps are shown as follows:

1) At first, we have collected the required data: location and type (category) of every restaurant within 5km from Dublin center (N Earl St). We have also identified Chinese restaurants (according to Foursquare categorization).
2) Secondly, our analysis will be calculation and exploration of restaurant density across different areas of Dublin - we will use heatmaps to identify a few promising areas close to center with low number of restaurants in general (and no Chinese restaurants in vicinity) and focus our attention on those areas.
3) Finally, we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than two restaurants in radius of 250 meters, and we want locations without Chinese restaurants in radius of 400 meters. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

4. Dublin City Center

Let's create latitude & longitude coordinates for centroids of our candidate neighborhoods. We will create a grid of cells covering our area of interest which is approximately 12x12 kilometers centered around Dublin city center.

Let's first find the latitude & longitude of Dublin city center, using specific, well-known address and Google Maps geocoding API.

Out[39]:

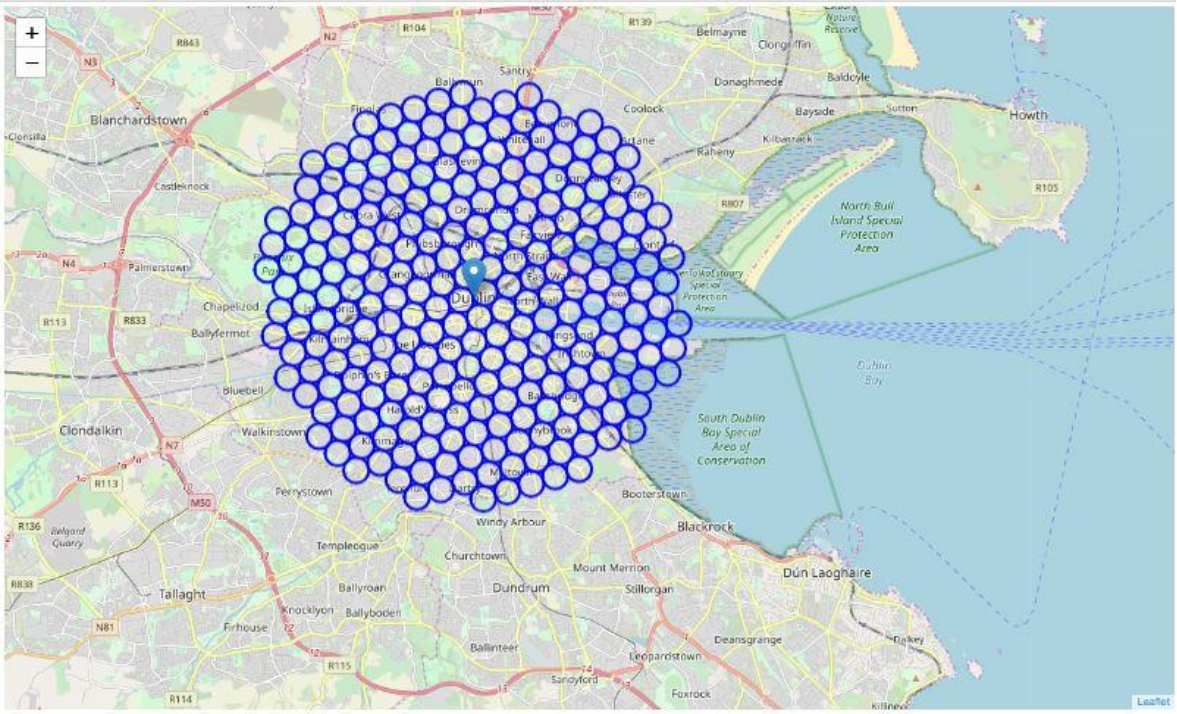


Figure 1. Grid of cells covering area of interest

Foursquare

Let's use Foursquare API to get info on restaurants in each neighborhood.

1. We are only interested in 'food' category, but only those that are proper restaurants;

2. The other Non-specified restaurants are ignored in this project.

3. Make sure to detect and include all the subcategories of specific 'Chinese restaurant' Category.

Out[43]:

	Address	Latitude	Longitude	X	Y	Distance from center
0	NO ADDRESS	53.307489	-6.256669	-906422.228865	6.119873e+06	4828.043082
1	NO ADDRESS	53.309056	-6.248277	-905822.228865	6.119873e+06	4714.870094
2	NO ADDRESS	53.310623	-6.239885	-905222.228865	6.119873e+06	4676.537180
3	NO ADDRESS	53.312190	-6.231491	-904622.228865	6.119873e+06	4714.870094
4	NO ADDRESS	53.313756	-6.223097	-904022.228865	6.119873e+06	4828.043082
5	NO ADDRESS	53.307919	-6.279915	-907922.228865	6.120393e+06	4956.813493
6	NO ADDRESS	53.309488	-6.271524	-907322.228865	6.120393e+06	4657.252409
7	NO ADDRESS	53.311057	-6.263133	-906722.228865	6.120393e+06	4419.276959
8	NO ADDRESS	53.312625	-6.254741	-906122.228865	6.120393e+06	4253.234064
9	NO ADDRESS	53.314193	-6.246348	-905522.228865	6.120393e+06	4167.733197

Figure 2. Details of Each Location

5. Restaurants

After obtaining the longitudes and latitudes of each area, it is necessary to go over these neighborhood locations and get nearby restaurants. In addition, we also need maintain a dictionary of all-found restaurants and all-found Chinese restaurants. The overview of the number of restaurants is shown as follows:

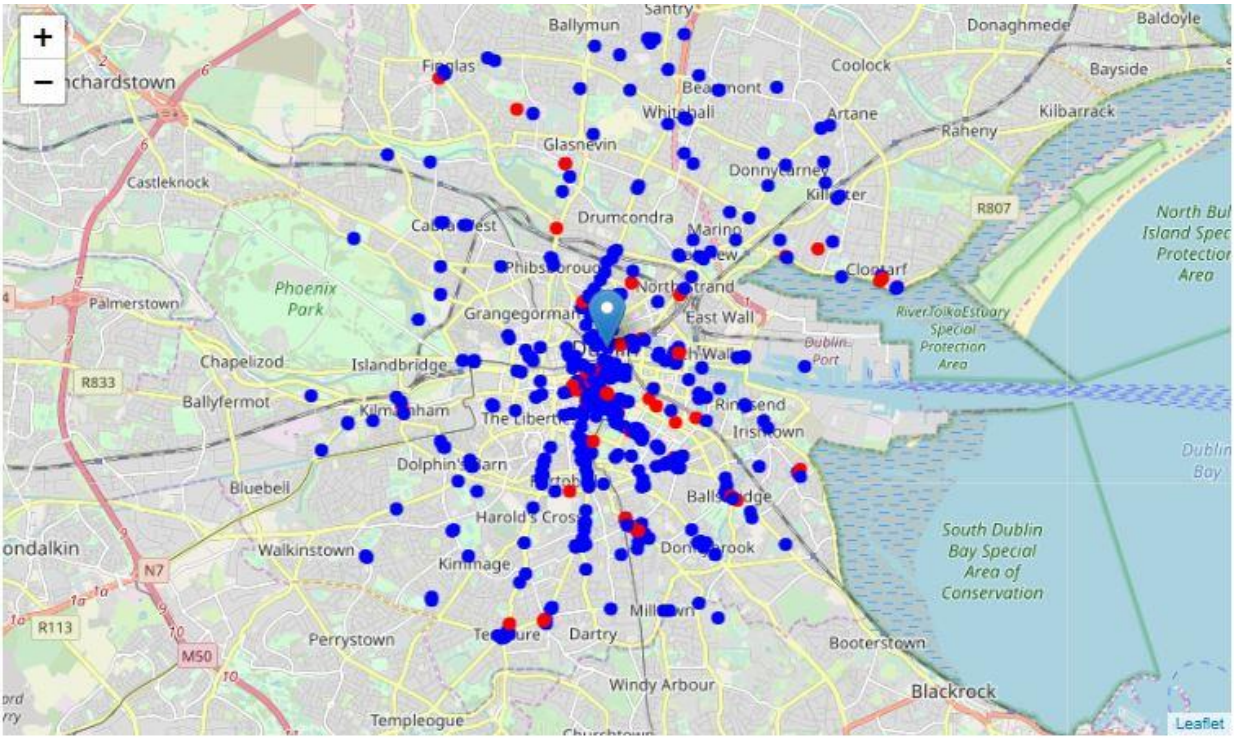
• Total number of restaurants: 494

• Total number of Chinese restaurants: 54

• Percentage of Chinese restaurants: 10.93%

• Average number of restaurants in neighborhood: 1.7992125984251968

In order to have a clear view of all the collected restaurants in our area of interest on map, Chinese restaurants are shown in the map below with a different color.



By now, we have all the restaurants in area within few kilometers from N Earl St (Dublin City Center), and we know which ones are Chinese restaurants. We also know which restaurants exactly are in vicinity of every neighborhood candidate center. The next step is to use this data for analysis to produce the report on optimal locations for a new Chinese restaurant!

6. Optimal Locations

The Chinese restaurants in each area are shown as below.

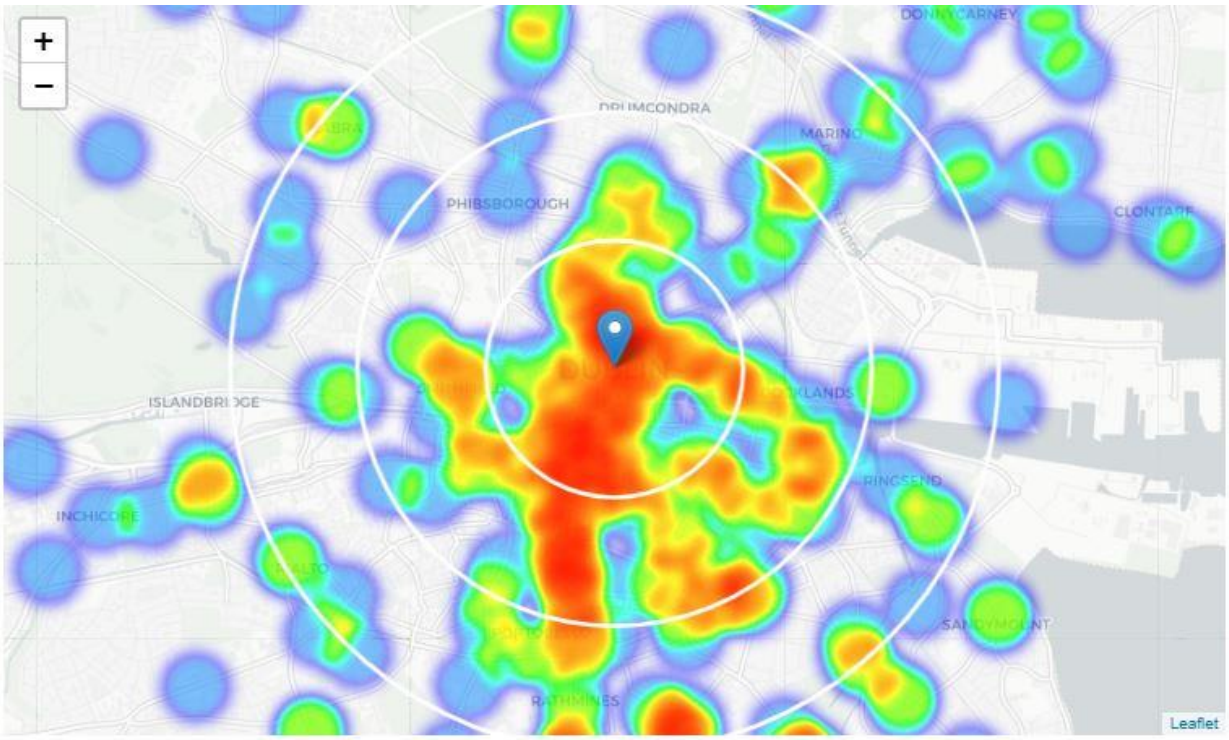
	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	NO ADDRESS	53.307489	-6.256669	-906422.228865	6.119873e+06	4828.043082	0
1	NO ADDRESS	53.309056	-6.248277	-905822.228865	6.119873e+06	4714.870094	0
2	NO ADDRESS	53.310623	-6.239885	-905222.228865	6.119873e+06	4676.537180	0
3	NO ADDRESS	53.312190	-6.231491	-904622.228865	6.119873e+06	4714.870094	1
4	NO ADDRESS	53.313756	-6.223097	-904022.228865	6.119873e+06	4828.043082	0
5	NO ADDRESS	53.307919	-6.279915	-907922.228865	6.120393e+06	4956.813493	0
6	NO ADDRESS	53.309488	-6.271524	-907322.228865	6.120393e+06	4657.252409	1
7	NO ADDRESS	53.311057	-6.263133	-906722.228865	6.120393e+06	4419.275959	0
8	NO ADDRESS	53.312625	-6.254741	-906122.228865	6.120393e+06	4253.234064	0
9	NO ADDRESS	53.314193	-6.246348	-905522.228865	6.120393e+06	4167.733197	2

The next step is to calculate the distance to nearest Chinese restaurant from every area candidate center (not only those within 300m - we want distance to closest one, regardless of how distant it is).

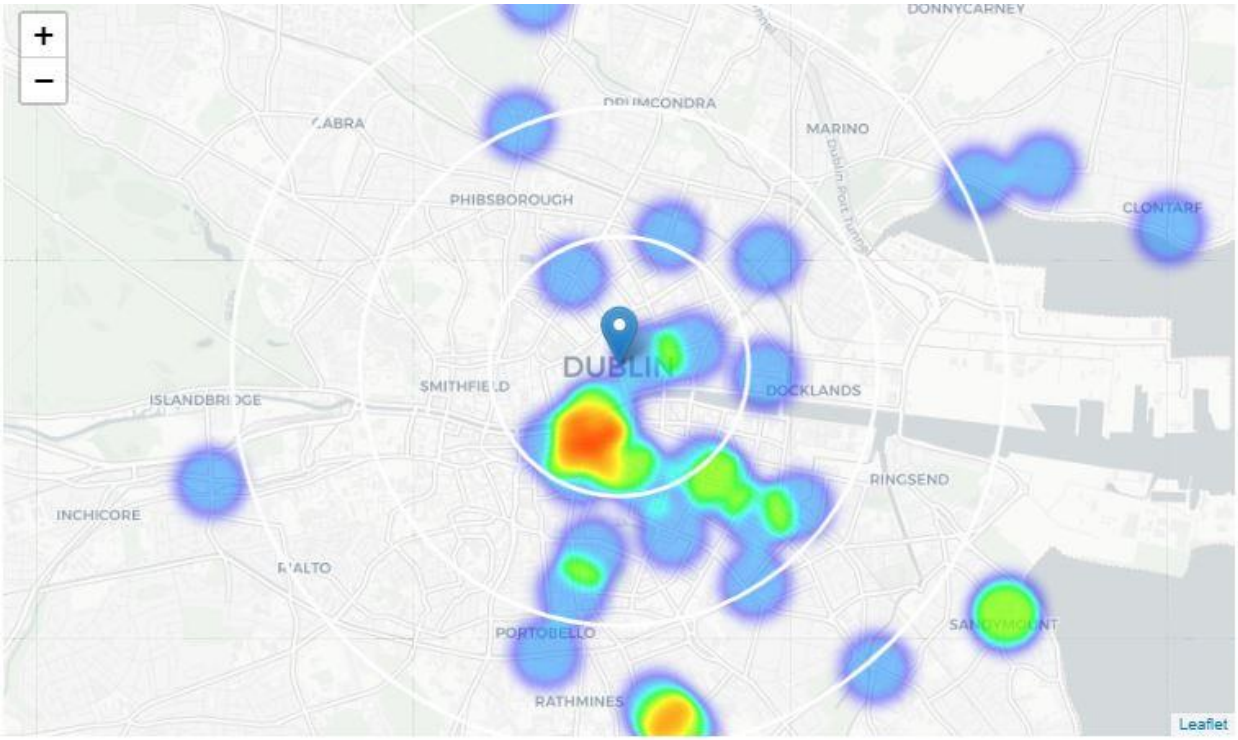
	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Chinese restaurant
0	NO ADDRESS	53.307489	-6.256669	-906422.228865	6.119873e+06	4828.043082	0	1332.307314
1	NO ADDRESS	53.309056	-6.248277	-905822.228865	6.119873e+06	4714.870094	0	1764.046953
2	NO ADDRESS	53.310623	-6.239885	-905222.228865	6.119873e+06	4676.537180	0	1788.930479
3	NO ADDRESS	53.312190	-6.231491	-904622.228865	6.119873e+06	4714.870094	1	1878.376693
4	NO ADDRESS	53.313756	-6.223097	-904022.228865	6.119873e+06	4828.043082	0	1750.526976
5	NO ADDRESS	53.307919	-6.279915	-907922.228865	6.120393e+06	4956.813493	0	325.603992
6	NO ADDRESS	53.309488	-6.271524	-907322.228865	6.120393e+06	4657.252409	1	358.774626
7	NO ADDRESS	53.311057	-6.263133	-906722.228865	6.120393e+06	4419.275959	0	792.302325
8	NO ADDRESS	53.312625	-6.254741	-906122.228865	6.120393e+06	4253.234064	0	1337.742762
9	NO ADDRESS	53.314193	-6.246348	-905522.228865	6.120393e+06	4167.733197	2	1232.057751

Average distance to closest Chinese restaurant for each area center is 1054.755.

It is necessary to create a map showing heatmap / density of restaurants and try to extract some meaningful information from that. A few circles are plotted to indicate distance of 1km, 2km and 3km from N Earl St.



Another heatmap showing density of Chinese Restaurants only:

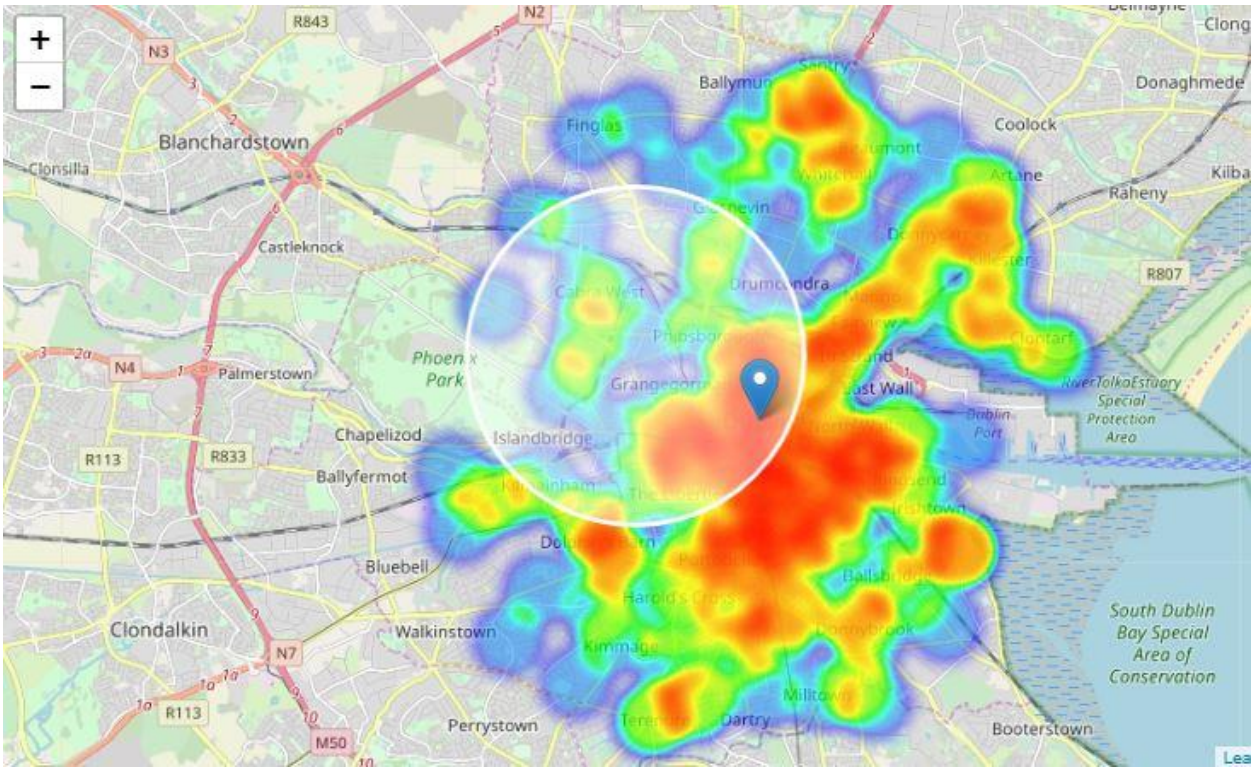


As you can see from the map above:

• It is not so 'hot' which means Chinese restaurants represent a subset of approximately 10% of all restaurants in Dublin

• It also indicates higher density of existing Chinese restaurants directly south from N Earl St.

Based on this we will now focus our analysis on areas south-west, south-east and west from Dublin center - we will move the center of our area of interest and reduce its size to have a radius of 2.5km. Let's define new, narrow region of interest, which will include low-restaurant-count parts.



Let's calculate two most important things for each location candidate:

• Number of restaurants in vicinity (using radius of **250 meters**)

• Distance to closest Chinese restaurant.

	Latitude	Longitude	X	Y	Restaurants nearby	Distance to Chinese restaurant
0	53.337369	-6.277623	-906772.228865	6.123550e+06	0	895.459525
1	53.337631	-6.276224	-906672.228865	6.123550e+06	0	823.453016
2	53.336655	-6.285698	-907322.228865	6.123636e+06	0	1392.416460
3	53.336917	-6.284299	-907222.228865	6.123636e+06	0	1310.323347
4	53.337178	-6.282900	-907122.228865	6.123636e+06	0	1226.040122
5	53.337440	-6.281500	-907022.228865	6.123636e+06	0	1144.290823
6	53.337702	-6.280101	-906922.228865	6.123636e+06	0	1065.658761
7	53.337964	-6.278702	-906822.228865	6.123636e+06	0	990.886322
8	53.338225	-6.277302	-906722.228865	6.123636e+06	0	920.914117
9	53.338487	-6.275903	-906622.228865	6.123636e+06	0	856.918852

Based on the filtering criteria:

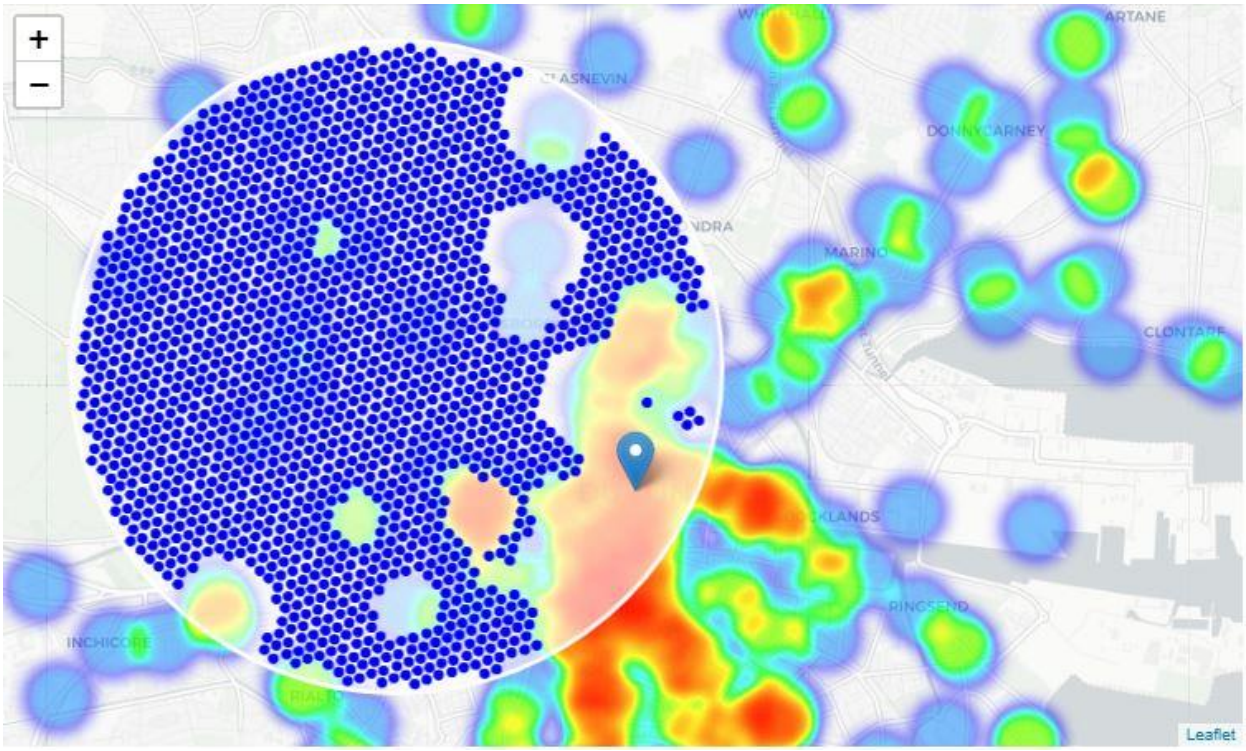
• locations with no more than two restaurants in radius of 250 meters,

• no Chinese restaurants in radius of 400 meters. The results are shown as follows:

• Locations with no more than two restaurants nearby: 1879

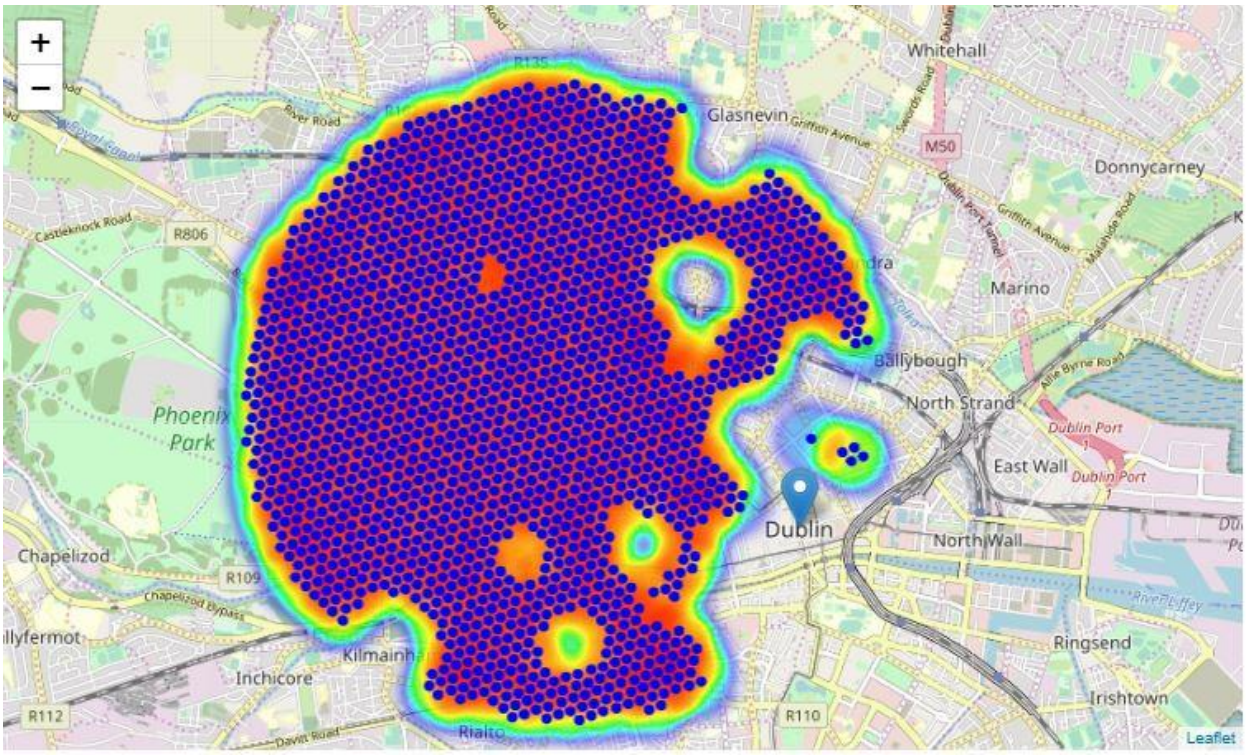
• Locations with no Chinese restaurants within 400m: 1869

• Locations with both conditions met: 1715



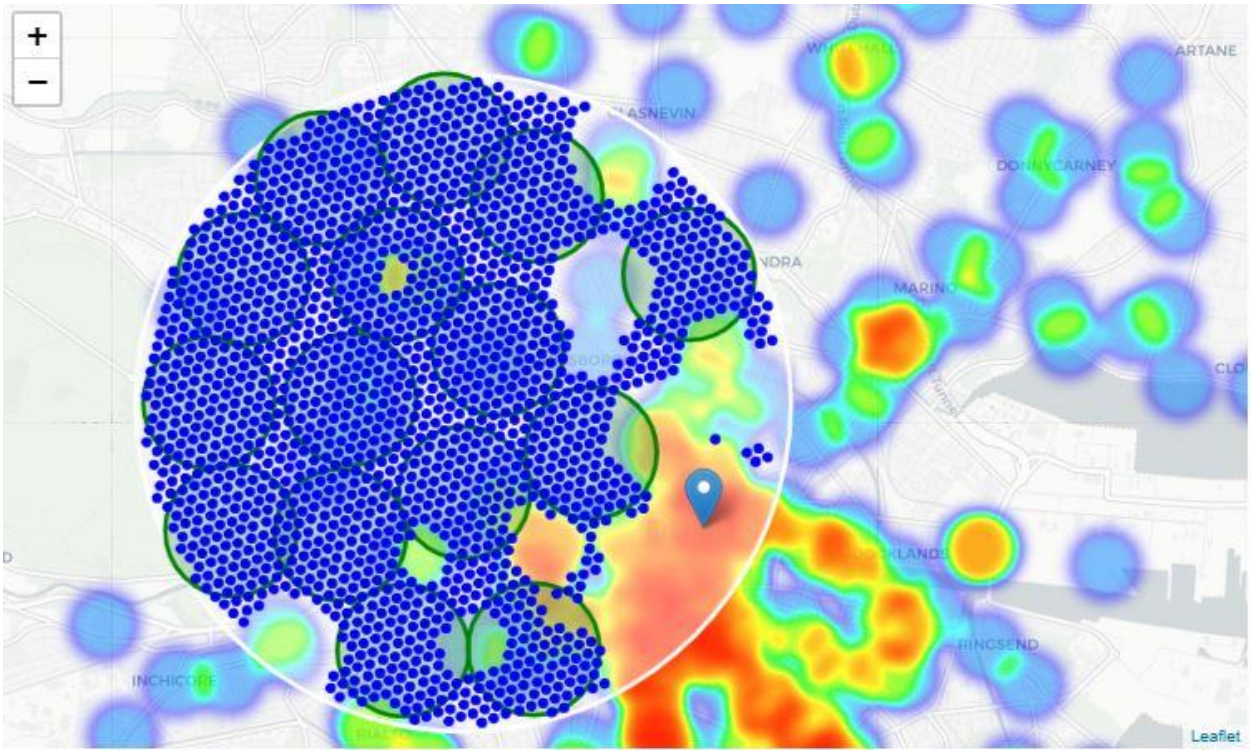
By now, a bunch of locations fairly close to N Earl St has been identified, and we know that each of those locations has no more than two restaurants in radius of 250m, and no Chinese restaurant closer than 400m. Any of those locations is a potential candidate for a new Chinese restaurant, at least based on nearby competition.

Let's now show those good locations in a form of heatmap:



7. Clustering

The final step is to cluster those locations to create centers of zones containing good locations. Those zones, their centers and addresses will be the final result of our analysis



These clusters represent groupings of most of the candidate locations and cluster centers are placed nicely in the middle of the zones 'rich' with location candidates. Addresses of those cluster centers will be a good starting point for exploring the neighborhoods to find the best possible location based on neighborhood specifics.

8. Results

This concludes our analysis. We have created 15 addresses representing centers of zones containing locations with low number of restaurants and no Chinese restaurants nearby, all zones being fairly close to city center (all less than 4km from N Earl St, and about half of those less than 2km from N Earl St). Although zones are shown on map with a radius of ~500 meters (green circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Our analysis shows that although there is a great number of restaurants in Dublin, there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected south and east from N Earl St, so we focused our attention to areas west and north-west. After directing our attention to this more narrow area of interest, we first created a dense grid of location candidates (spaced 100m appart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Chinese restaurant closer than 400m were removed. Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors. Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both

restaurants in general and Chinese restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Dublin center but not crowded with existing restaurants (particularly Chinese) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

9. Conclusion

The main objective of this project was to identify Dublin areas close to center with low number of restaurants (particularly Chinese restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Chinese restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders. Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.