Mexican Restaurant possibilities in London

Mike Gessner

A Report for Chef Gordon Gonzalez

Introduction and Background

Our client is a world-renown chef and restauranteur and he's looking to expand to new cities and continents. He has hired us to examine potential locations for his new venue in London, United Kingdom. Chef Gordon Gonzalez owns several successful Mexican restaurants in several United States cities. He believes Londoners are craving authentic Mexican food and tacos in particular, and has decided that his potential London eatery will be taco themed. He needs our help in analyzing the market and the best place to put the restaurant in London.

Business Problem

Clearly, London has many restaurants. Therefore, we will need to find locations that are not already crowded. That said, we need to also find areas where there are very few or no Mexican restaurants in the vicinity of a potential location. Location is quite important to Chef Gonzalez (as it is for any restaurant). He would like to have a restaurant with high foot traffic every day of the week. He states that he wants to make sure that he is near the city center, if possible.

Using our expertise in data science and analysis, we will provide some potentially promising locations. We will describe the advantages and disadvantages of each location so that our client can make an informed decision.

Data

Based on the business problem presented above, here are the factors that influence the potential location:

1. The number of existing restaurants in a neighborhood/area (i.e., the density)
2. The number of and distance to Mexican restaurants in the neighborhood/area (our direct competition)
3. The distance of the neighborhood/area from the city center

Rather than use pre-defined neighborhoods/boroughs in London, we use regularly spaced locations that are centered around the city center to define those areas. We chose Trafalgar Square as our city center for this analysis.

We will rely on the following data sources to help us make the recommendation:

- The potential areas will be generated using an algorithm and the approximate addresses of those generated areas will be obtained via the TomTom API for reverse geocoding
- The number of restaurants and their types (e.g., Italian, Mexican) and their location in each candidate area will be obtained via the Foursquare API
- The city center coordinates of London will be obtained via the TomTom API for geocoding for Trafalgar Square

Foursquare data

Foursquare categorizes venues (e.g., Arts & Entertainment, Food). For our problem, we're interested in the broad "Food" category that is identified with a unique alphanumeric code. The "Food" category contains many sub-categories. Of the sub-categories, we focus on "Mexican Restaurant" that is also identified with a unique alphanumeric code. Additionally, there are sub-categories to "Mexican Restaurant" like "Taco Place." We use those unique alphanumeric codes to help us identify potential direct competitors. The categories and their alphanumeric codes are found at: https://developer.foursquare.com/docs/resources/categories.

In addition to the category and sub-categories of a venue, Foursquare also stores the address of the venue. We will use the address from Foursquare and then obtain the coordinates of the venue from TomTom using the reverse geocoding API.

Combining the data sources

As mentioned, TomTom's APIs allow us to map locations in London. We start with a London center, Trafalgar Square, and consider an area 6km in every direction from that center. This defines our boundaries for the analysis. We then define equal-sized "areas" or "neighborhoods" within that boundary.

The Foursquare data allows us to map the restaurants within the defined boundaries. We obtain a total count of restaurants within our defined area (and each neighborhood), as well as the number of Mexican restaurants. So within each equal-sized neighborhood, we can calculate the density of restaurants and the percentage of those restaurants that are our direct competitors.

Methodology

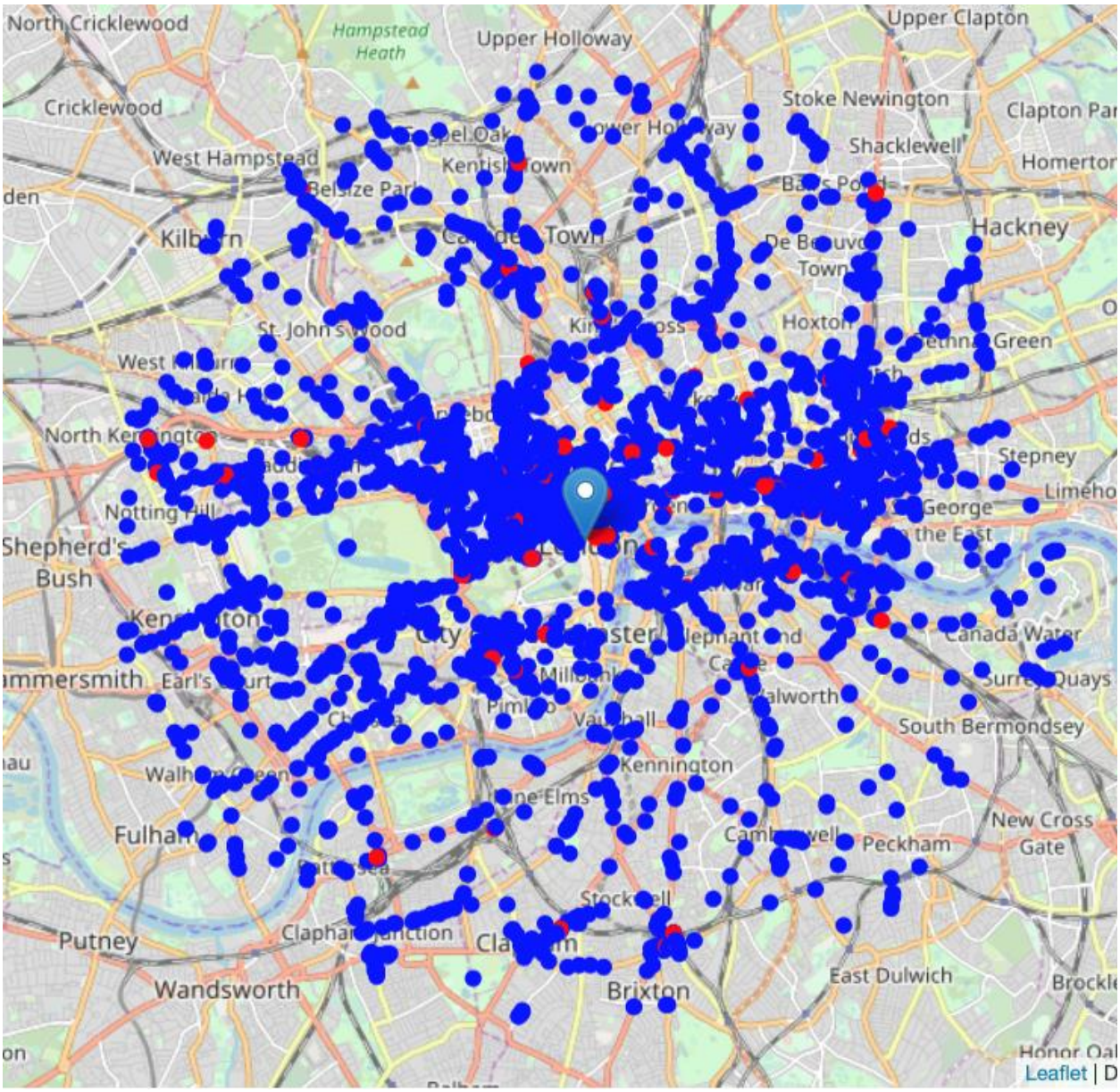
As mentioned above, we start with a city center location in London. For the purpose of this report, we use the coordinates of Trafalgar Square as the center of London. In order to get the coordinates of Trafalgar Square, we use the TomTom geocoder API. We store those coordinates and then do the following:

- We create a few functions to help convert xy coordinates to latitude and longitude and vice versa, and one to calculate the distance between two points on a map
- Next, we define neighborhoods with a maximum of 6 kilometers from the city center that creates equal sized neighborhoods from the city center and map them
- Then we create a function that gets the long address from latitude and longitude that uses TomTom's reverse geocoding API
- Using the reverse geocoding, we obtain approximate addresses from each of the defined neighborhoods and create a dataframe that contains the street address, the latitude and longitude, the xy coordinates, and the distance from the city center
- With the addresses stored, we then rely on the Foursquare API to obtain restaurants near each of those addresses
- With that accomplished, we then map those restaurants and identify the Mexican restaurants
- We will create a heatmap of restaurants as well as one for the Mexican restaurants, this will help us identify low density restaurant areas in London for further investigation
- From the refocusing, we will identify potentially good locations using 2 criteria:
    - The number of restaurants within a radius of 250m must be 2 or fewer
    - The distance to the nearest Mexican restaurant must be greater than 399m
- The criteria will identify potential locations matching those two criteria
- Using KMeans, we cluster those potential locations (limited to 15 clusters)
- Obtain the rough address of the center of each cluster and map those as well as create a dataframe for those 15 clusters that represent the recommended areas

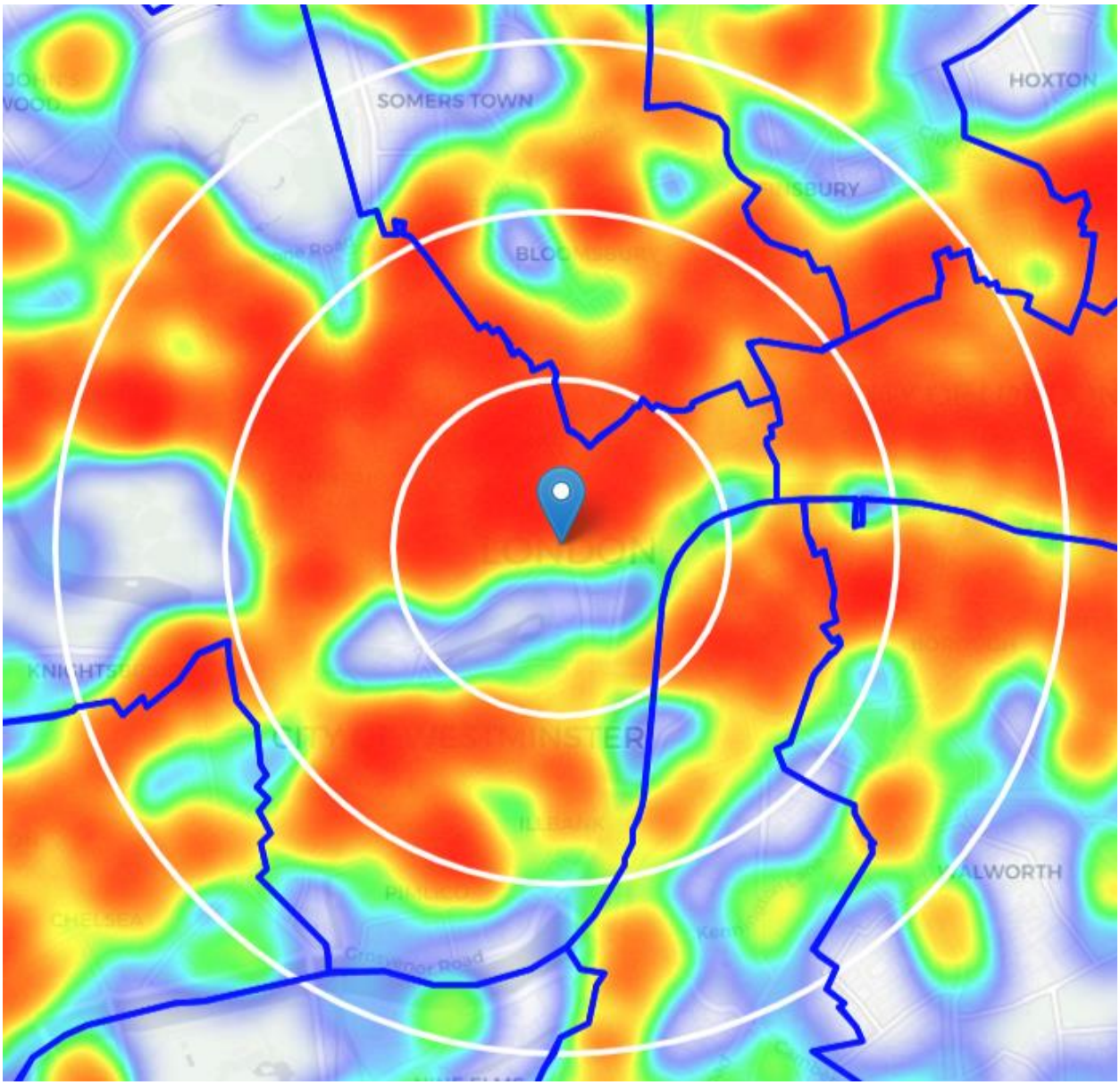
Analysis

Our analysis of the data obtained from Foursquare and TomTom identified 3,064 restaurants within our 6km radius from Trafalgar Square. Of those restaurants, 93 of them are Mexican—representing just over 3% of the restaurants in our defined area.

Below is a map of the restaurants within the 6km radius of Trafalgar Square:



As the map shows, restaurants are marked with a blue circle. Mexican restaurants are red circles. The marker represents Trafalgar Square. The data and map suggests that there is not a lot of competition in the Mexican restaurant category—and London may be ripe for a taco invasion. That said, we need to find potential locations. Let's use a heat map to help us out.



The heatmap of all restaurants shows a lot of heat! There might be some potential (or pockets of relatively low density) south, southwest, and southeast. There might also be some potential west of Trafalgar Square. But what about our main competition?

That's a lot better than all restaurants! But that's expected, given that 3% of the restaurants are Mexican. We can see that there is relatively high intensity north and east of Trafalgar Square—essentially SOHO and Covent Garden areas in Westminster. That said, we still have areas of interest within Westminster (south, southwest, and west), as well as Lambeth and Southwark (across the Thames River).

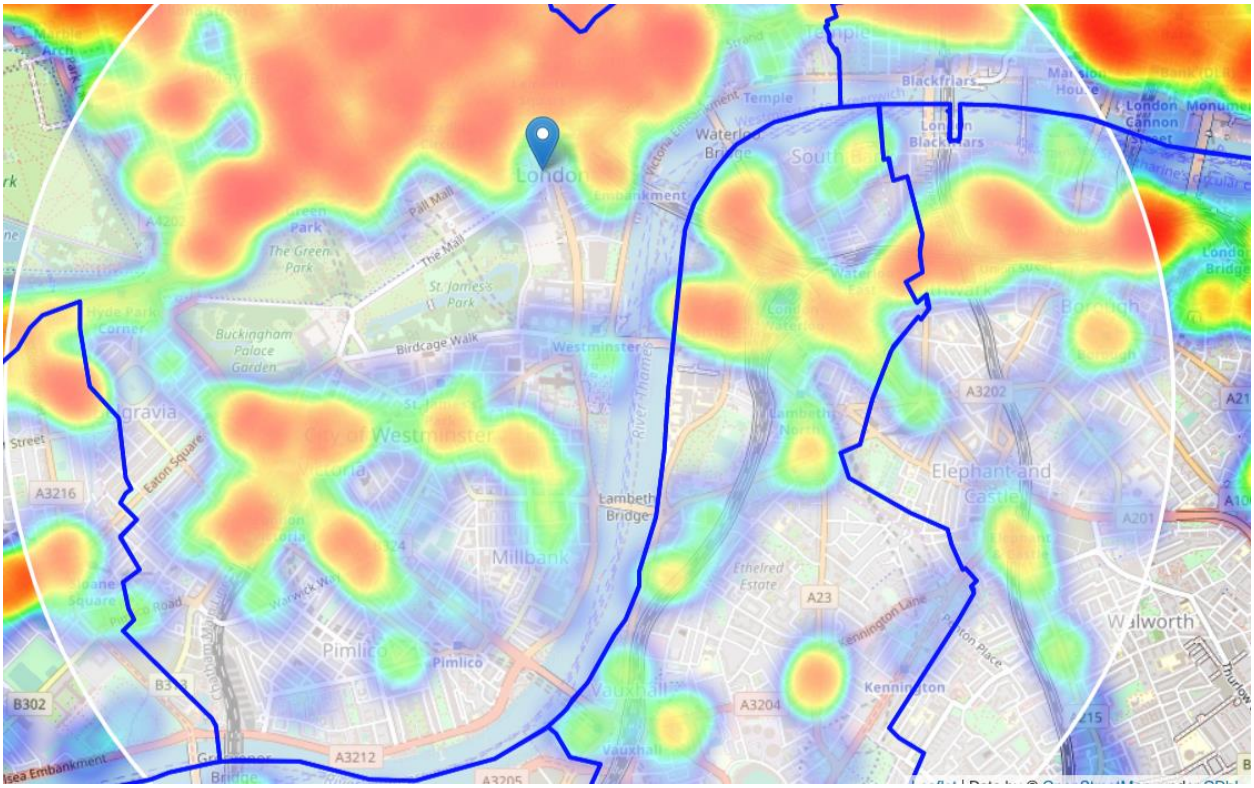
Within Westminster, the Mayfair district to the west has potential. The district is ritzy and upscale and is bounded by Oxford Street to the north (lots of shopping and foot traffic, popular with tourists and locals). Across the Thames River, Southwark is also interesting. It has landmarks like the Tate Modern Museum and Borough Market and is also popular with hipsters (who love tacos). Lambeth may also have potential as directly across the river from Trafalgar Square (at least somewhat close) is the London Eye and is being revitalized.

This quick analysis suggests that we should alter the focus a bit from our original 6km around Trafalgar Square.

Refocus and drill in!

The data shows that the average distance from an area center to a Mexican restaurant is nearly 900 meters. That's not too close and not too far from each center. That's probably expected as there aren't a lot of Mexican restaurants within our 6 km borders from Trafalgar Square.

We'll reduce the focus to 2.5 km from that center.

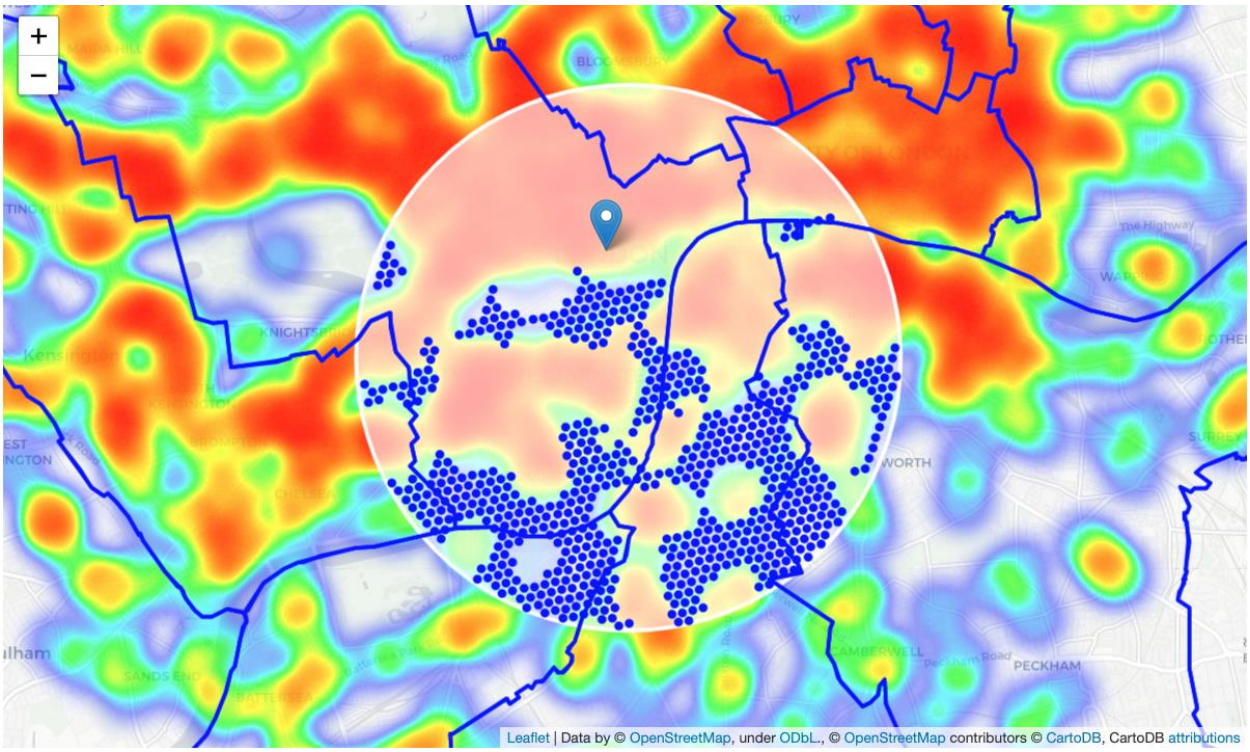


So that new focus does cover many of the pockets of low density—especially in Lambeth and Southwark, as well as Mayfair and southern Westminster.

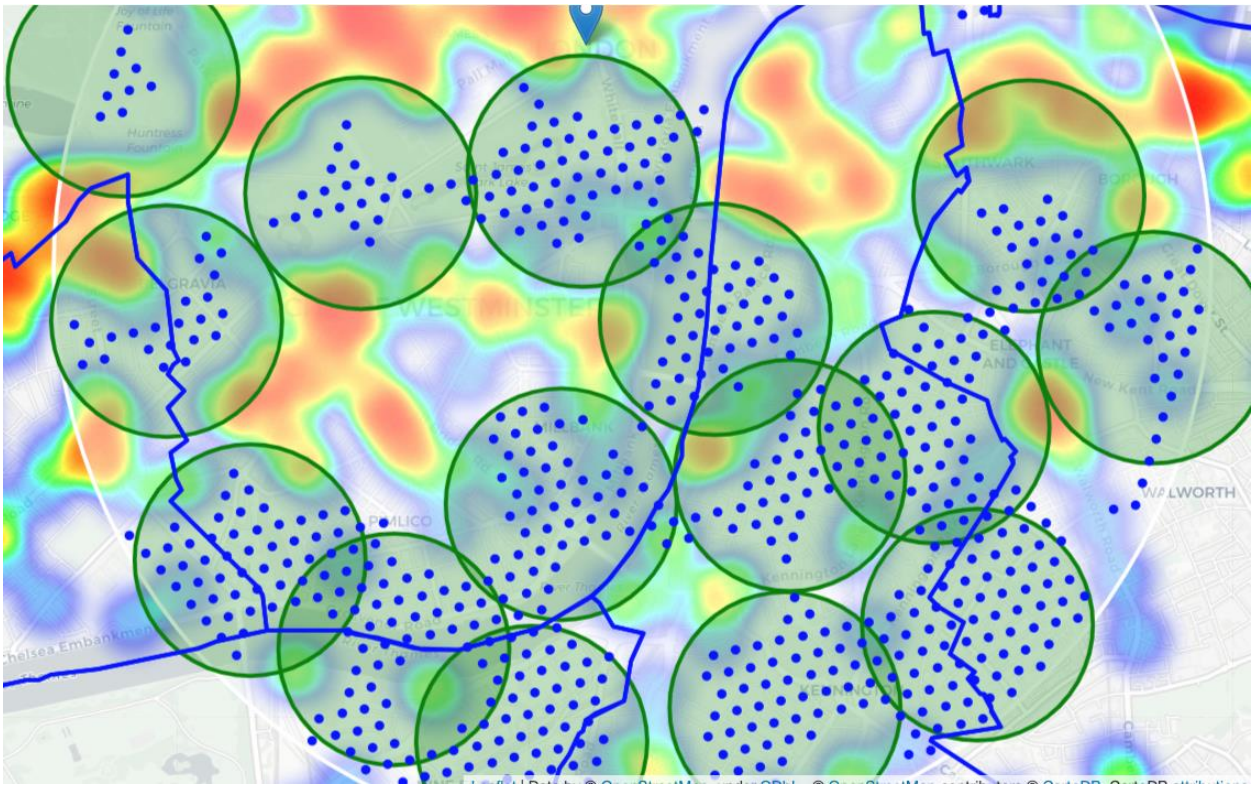
Let's also create a new, more dense grid of location candidates restricted to our new region of interest (roi) and potential locations 100m apart. Doing so identifies 2,261 candidate neighborhood centers. But we have to narrow that down some more as well.

To do so, we use two criteria. First, the number of restaurants within a radius of 250m from the center must be 2 or fewer. Second, the distance to the closest Mexican restaurant must be greater than 399m. Applying those filters to the data finds that 673 locations meet both criteria. That’s much better than 2,261!

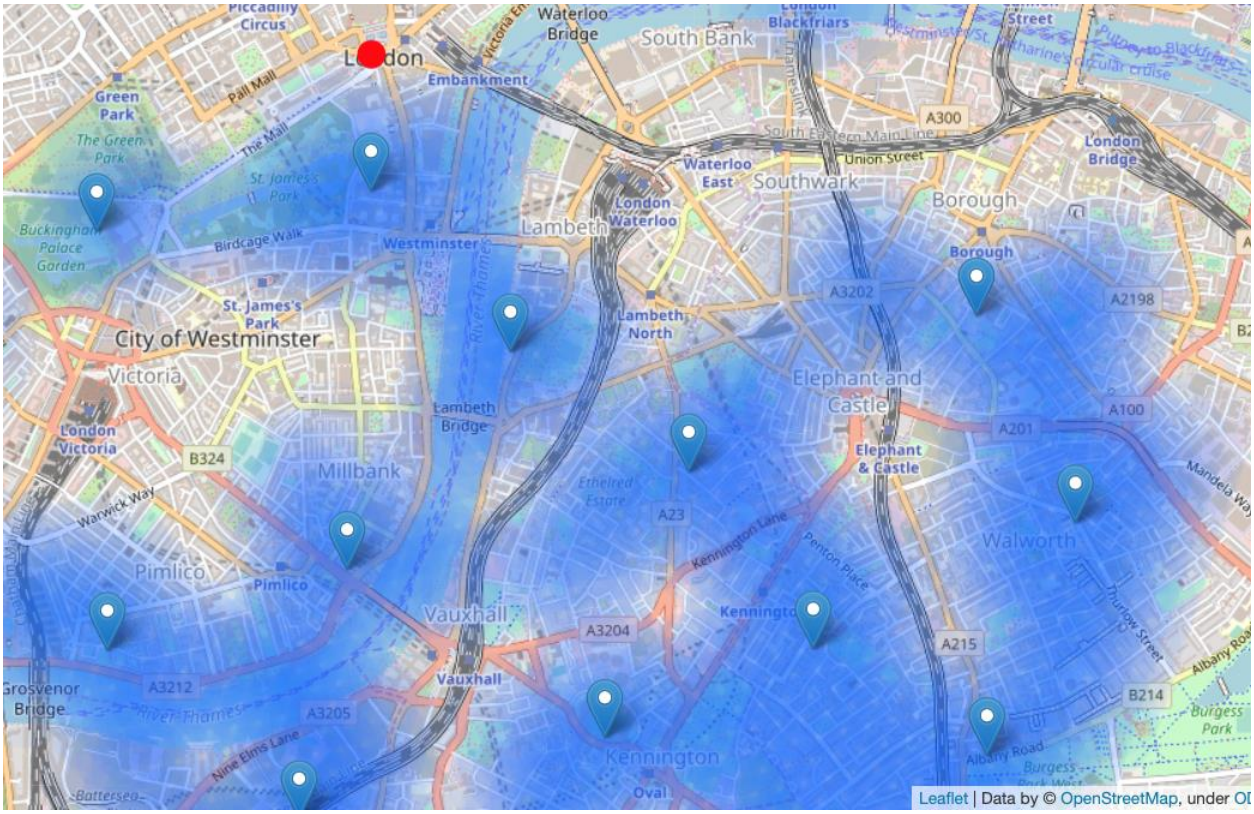
Now we can map those centers:



The blue circles represent our 673 locations. We can cluster those locations to create centers of zones that represent potentially good locations. The center of those clusters will generate approximate addresses and help Chef Gonzalez and his team make an informed decision on the location of his next restaurant!



Now that we have the clusters, we can then map markers with the approximate addresses to help Chef's team take the analysis to the next step.



Results and Discussion

Our analysis shows that although there is a great number of restaurants in London (more than 3000 in our initial area of interest which was 12x12km around Trafalgar Square), there are pockets of low restaurant density fairly close to city center. The highest concentration of restaurants was detected north and west from Trafalgar Square, so we focused our attention to areas south, south-east and east, corresponding to boroughs of Lambeth and Southwark and south-west portion of Westminster. Attention was focused on which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics and a number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest (covering approx. 5x5km from Trafalgar Square) we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Mexican 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 a more detailed local analysis based on other factors.

The result of all this is 15 zones containing the largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Mexican restaurants in particular. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! The purpose of this analysis was to only provide information on areas close to London center (Trafalgar Square) but not crowded with existing restaurants (particularly Mexican) - 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 the lack of competition in that area. The 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.

Conclusion

The purpose of this project was to identify London areas close to its center with a low number of restaurants (particularly Mexican restaurants) in order to aid Chef Gordon Gonzalez and his staff in narrowing down the search for an optimal location for a new Mexican restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis (Westminster, Lambeth, and Southwark), 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.

The final decision on optimal restaurant location will be made by Chef Gonzalez and his staff, 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, access to Underground or other public transit, real estate availability, prices, social and economic dynamics of every neighborhood, etc.