Capstone Project - Open an Italian restaurant in London

Introduction

This project is to find an optimal location for investors interested in opening an Italian restaurant in London, U.K., preferring areas with no Italian restaurants in vicinity.

Data description

The problem description leads to the following three factors:

- 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

The required information is acquired as below:

- centers of candidate areas will be generated and their approximate addresses will be obtained using Azure Maps API reverse geocoding
- number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API

Data gathering

The first step of creating latitude and longitude coordinates for centroids of candidate neighborhoods is to get the coordinates of London city center, which is acquired by Azure Maps geocoding API.

```
print(query, ": ", center)
```

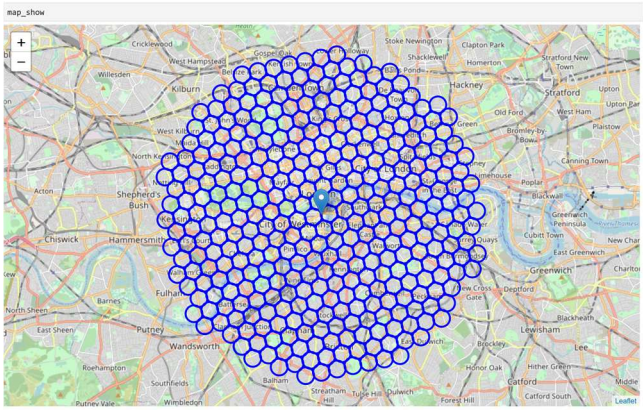
```
London, UK :  [51.50015, -0.12624]
```

A grid of area candidates which are equally spaced and centered around city center is created. The neighborhoods are defined as circular areas with a radius of 300 meters. The neighborhood centers are 600 meters apart.

Accurate distance calculation requires the grid creation of locations in Cartesian 2D coordinate system which allows calculating distances in meters instead of latitude/longitude degrees. The coordinates will be projected back to latitude/longitude degrees for Folium map display. The functions are created to convert between WGS84 spherical coordinate system (latitude/longitude degrees) and UTM Cartesian coordinate system (X/Y coordinates in meters).

```
London center longitude=-0.12624, latitude=51.50015
London center UTM X=-547082.9460837919, Y=5814745.518286032
London center longitude=-0.12623999999999394, latitude=51.50014999999999
```

The coordinates of centers of neighborhoods/areas are ready to be evaluated, visualized as below:



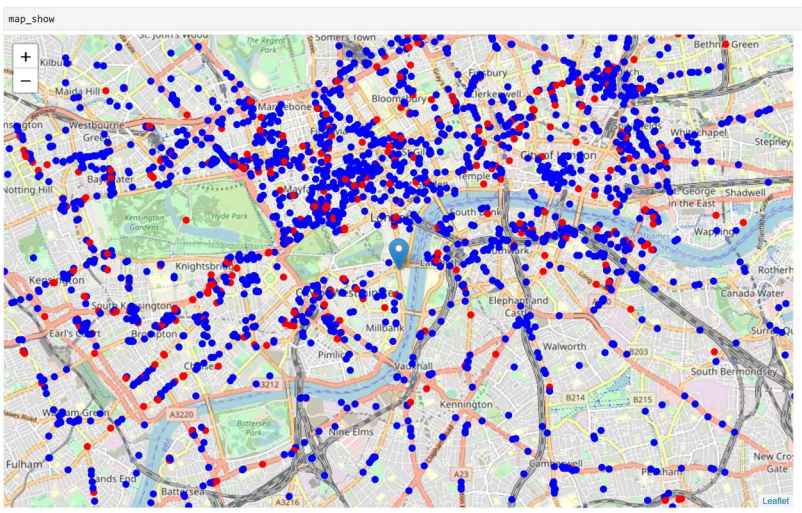
The approximate addresses of those locations are acquired via Google Maps API and stored as pandas data frame.

```
df_locations.head(10)
```

	Address	Latitude	Longitude	X	Y	Distance from center
0	10 Scrutton Close, London, SW12 0AW	51.447233	-0.134452	-548882.946084	5.809030e+06	5992.495307
1	London, SW2 5	51.448335	-0.126118	-548282.946084	5.809030e+06	5840.376700
2	68 Elm Park, London, SW2 2TZ	51.449435	-0.117783	-547682.946084	5.809030e+06	5747.173218
3	23 Brockwell Park Row, London, SW2 2YJ	51.450536	-0.109447	-547082.946084	5.809030e+06	5715.767665
4	39 Norwood Road, London, SE24 9QG	51.451635	-0.101111	-546482.946084	5.809030e+06	5747.173218
5	90 Delawyk Crescent, London, SE24 9JD	51.452734	-0.092774	-545882.946084	5.809030e+06	5840.376700
6	5 Great Spilmans, London, SE22 8SZ	51.453833	-0.084437	-545282.946084	5.809030e+06	5992.495307
7	8 Clarence Mews, London, SW12 9SR	51.450090	-0.148481	-549782.946084	5.809549e+06	5855.766389
8	Oaklands Estate, London, SW4 8AH	51.451193	-0.140146	-549182.946084	5.809549e+06	5604.462508
9	147 Clarence Crescent, London, SW4 8LN	51.452294	-0.131811	-548582.946084	5.809549e+06	5408.326913

The information on restaurants in each neighborhood is acquired via Foursquare API, only including data with venues in the categories of ‘food’, ‘restaurant’ (blue) and ‘Italian restaurant’ (red).

```
Total number of restaurants: 2814
Total number of Italian restaurants: 412
Percentage of Italian restaurants: 14.64%
Average number of restaurants in neighborhood: 6.851648351648351
```



Data analysis

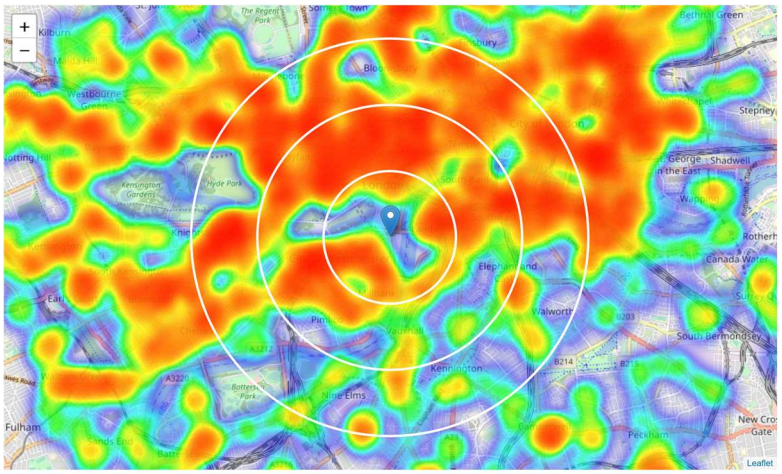
The calculation and exploration of restaurant density uses heatmaps to identify a few promising areas close to center with low number of restaurants in general and no Italian restaurants in vicinity.

The area candidate analysis consists of counting the number of restaurants in the area and calculating the distance to nearest Italian restaurant from every area candidate center.

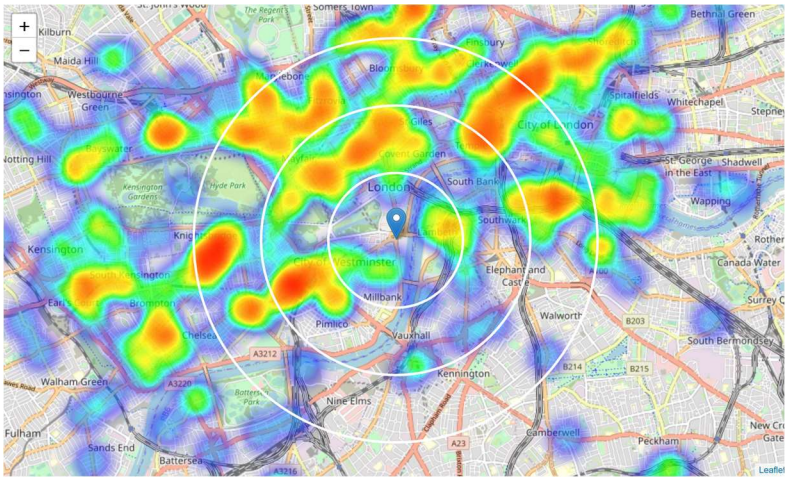
```
df_locations.head(20)
```

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Italian restaurant
0	10 Scrutton Close, London, SW12 0AW	51.447233	-0.134452	-548882.946084	5.809030e+06	5992.495307	0	1023.394283
1	London, SW2 5	51.448335	-0.126118	-548282.946084	5.809030e+06	5840.376700	3	1615.934228
2	68 Elm Park, London, SW2 2TZ	51.449435	-0.117783	-547682.946084	5.809030e+06	5747.173218	1	1528.942918
3	23 Brockwell Park Row, London, SW2 2YJ	51.450536	-0.109447	-547082.946084	5.809030e+06	5715.767665	0	1355.775841
4	39 Norwood Road, London, SE24 9QG	51.451635	-0.101111	-546482.946084	5.809030e+06	5747.173218	3	1077.761891
5	90 Delawyk Crescent, London, SE24 9JD	51.452734	-0.092774	-545882.946084	5.809030e+06	5840.376700	1	528.184522
6	5 Great Spilmans, London, SE22 8SZ	51.453833	-0.084437	-545282.946084	5.809030e+06	5992.495307	0	341.155514
7	8 Clarence Mews, London, SW12 9SR	51.450090	-0.148481	-549782.946084	5.809549e+06	5855.766389	3	333.957673
8	Oaklands Estate, London, SW4 8AH	51.451193	-0.140146	-549182.946084	5.809549e+06	5604.462508	0	771.644541
9	147 Clarence Crescent, London, SW4 8LN	51.452294	-0.131811	-548582.946084	5.809549e+06	5408.326913	2	1190.348424

A heatmap is created to visualize the density of restaurants.

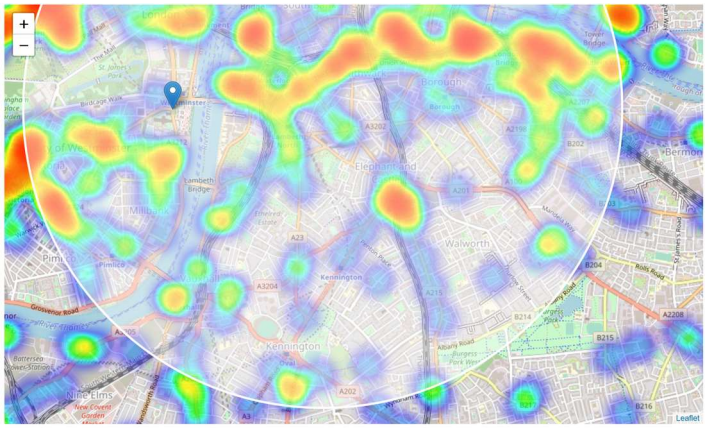


Another heatmap is created to visualize the density of Italian restaurants.

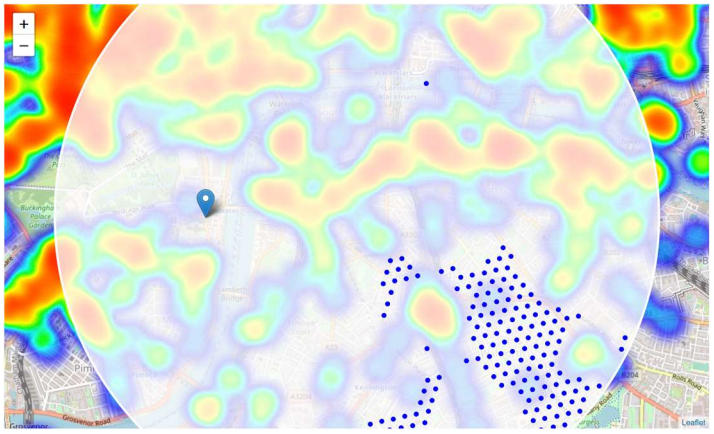


Two maps indicate higher density of existing Italian restaurants directly north and west from London center, with closest pockets of low Italian restaurant density positioned east, south-east and south from city center. A new map is created to visualize narrower region of interest, closer to Elephant and Castle, Vauxhall, Walworth and Kennington, which will include low-restaurant-count parts. However, those areas are more residential than tourists-populous, so the most promising candidates shouldn't

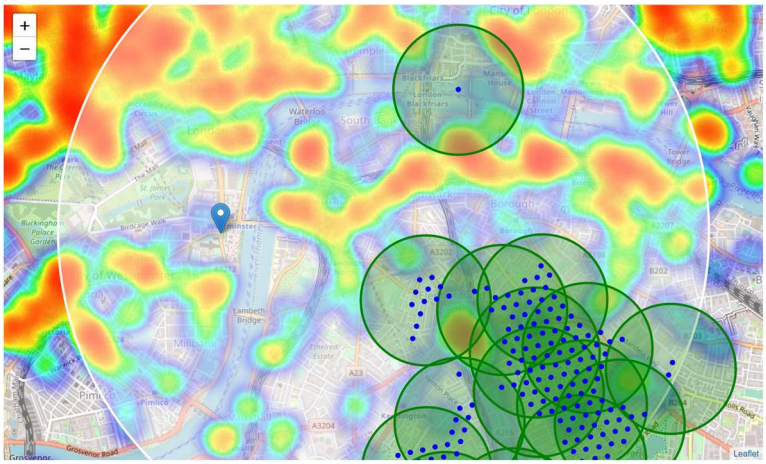
be too distant from the hottest areas.



Candidate locations are focused on those with no more than two restaurants in radius of 250 meters, and no Italian restaurants in radius of 400 meters.



Locations are clustered by K-mean to create centers of zones containing good locations.



Candidate area centers close to Lambeth are reverse geocoded to get the addresses for further consideration.

```
============================================================
Addresses of centers of areas recommended for further analysis
============================================================

Deverell Street, London, SE1 4AA                  => 2.6km from Lambeth
126 Bethwin Road, London, SE5 0YY                 => 2.8km from Lambeth
Townley Street, London, SE17 1DZ                  => 2.8km from Lambeth
37 Saint George's Road, London, SE1 6ER           => 1.7km from Lambeth
London, SE17 2                                    => 3.5km from Lambeth
23 Townsend Street, London, SE17 1HY              => 3.0km from Lambeth
27 Forsyth Gardens, London, SE17 3NE              => 2.5km from Lambeth
Westminster to Greenwich Thames River Services, London => 2.2km from Lambeth
(7 - 5) Pasley Close, London, SE17 3JY            => 2.4km from Lambeth
23 Thurlow Street, London, SE17 2FY               => 3.1km from Lambeth
93 Balfour Street, London, SE17 1PB               => 2.6km from Lambeth
4 Addington Square, London, SE5 7JZ               => 3.2km from Lambeth
2 Villa Street, London, SE17 2DF                  => 3.3km from Lambeth
Dunton Road, London, SE1 5TW                      => 3.7km from Lambeth
County Street, London, SE1 6AH                    => 2.3km from Lambeth
```

Result and Discussion

Highest concentration of restaurants was detected north and west from Westminster, so the attention is focused to areas south, south-east and east, corresponding to Elephant and Castle, Vauxhall, Walworth and Kennington, and finally on Lambeth 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 the attention to this narrower area, a dense grid of location candidates is created; those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Italian 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 for more detailed local analysis based on other factors.

The result contains potential new restaurant locations based on number of and distance to existing venues, both restaurants in general and Italian restaurants particularly, 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.

Conclusion

The purpose of this project was to identify London areas close to center with low number of restaurants in order to narrow down the search for optimal location for a new Italian restaurant. Final decision on optimal restaurant location will be based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness, proximity to major roads, real estate availability, prices, social and economic dynamics.