Coursera Capstone

Introduction

This analysis is trying to answer the questions:

1) Where to open an Indian restaurant in Cambridge 2) Whether to open an Indian restaurant in an area of historically high house prices or whether to open in Central Cambridge. While Central Cambridge has high house prices, there are other areas that have similarly high house prices.

This analysis can hopefully be replicated for a) Other restaurant types b) Other cities.

Data

The factors that will impact our analysis are :

- The number of existing restaurants in a mircrolocation
- The number and distance of India restaurants in the neighborhood
- The distance of the neighborhood from the city center.

I used a regularly space location grid centered around a key location in the city center.
The following data source will be needed to generate the information and analysis required:
Data Sources :
Venue Data

- Foursquare.com #Location and Borough Data Geonames : Cambridge Borough Data Set and GPS data Downloaded from Geonames (in the UK_full zip file) http://download.geonames.org/export/zip/

Housing Price Data

- Downloaded for 2019 and 2020 from HMRC https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads

Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using *geocoder
number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API
coordinate of Cambridge center will be obtained using Google Maps API geocoding of well known Cambrdige location (King's College)

Methodology

We have collected the location and type data for every restaurant around Cambridge city center.

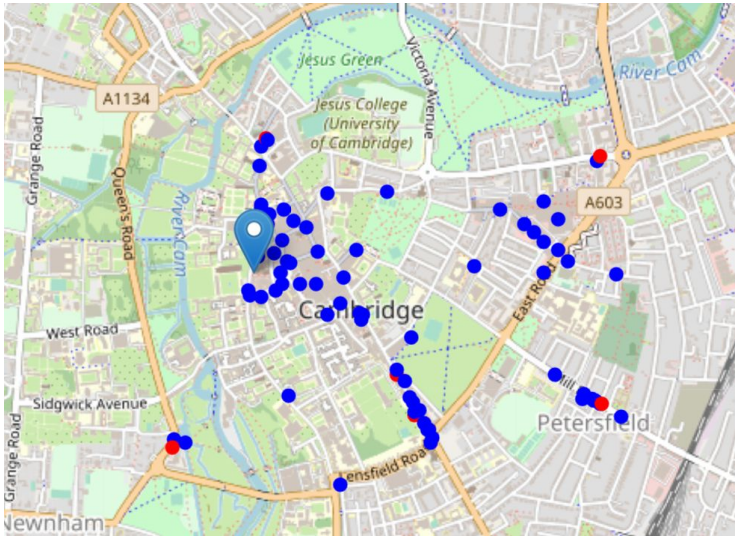
We have also identified the indian restaurants around cambridge city center.

We will first look at restaurant density across different parts of cambridge city center. We will further use heatmaps to determine key areas near the city center with 1) Low density of restaurants 2) Low density of Indian restaurants.

We will then focus on promising areas and create clusters that meet basic requirements including no more than 2 restaurants within 50 meters and no indian restaurants within 500 meters.

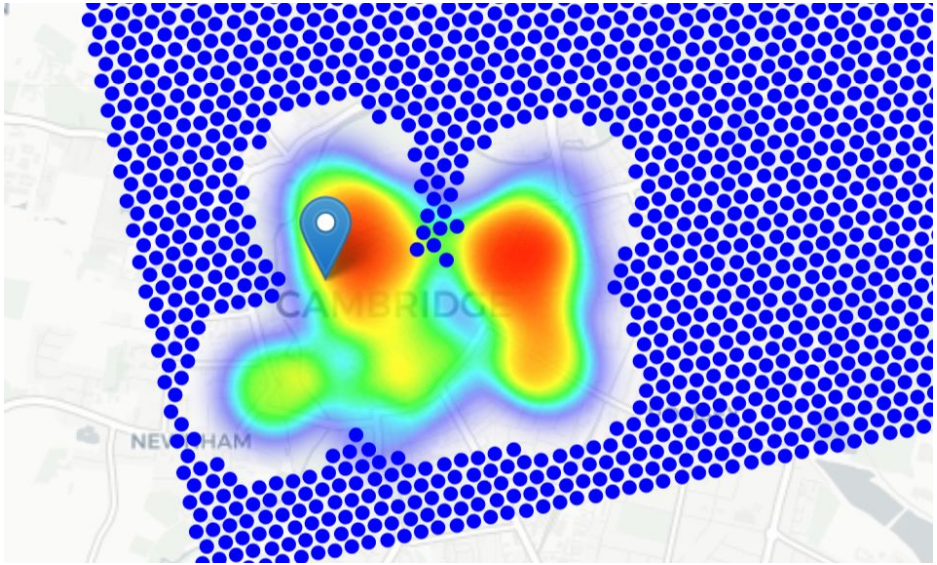
We will then present a map of all such locations and create clusters using k-means clustering of these locations to identify general zones/neighborhoods which should be a starting point for street level exploration by local stakeholders.

Analysis



Red indicates Indian restaurants

We now know which restaurants are within a few kilometers of the city center. We also know which ones are indian restaurants and under which microneighborhood they reside under.



Red indicates Indian restaurants

We now know which restaurants are within a few kilometers of the city center. We also know which ones are indian restaurants and under which microneighborhood they reside under.

Results and Discussion

Our results indicate that town center is indeed a crowded area to open an indian restaurant, both due to the total number of restaurants nearby and the number of Indian restaurants.

Using K-Means, it can be seen that there are clusters near town center that may be feasible to open a restaurant. These include :

Around Grange Road
Further down Hills Road
Around Milton Road

Conclusion

The conclusion notes that there are a few good places to open up a restaurant that is not in very central cambridge. Namely,

Around Grange Road
Further down Hills Road
Around Milton Road
These three areas provide a data driven opportunity, particularly for owners that want to open up an Indian restaurant.

The results do validate the hypothesis that central cambridge is overcrowded in terms of restaurants, however, it also shows that there are few Indian restaurants in central Cambridge, with many of them located further away.