

Project title: Best neighbourhood to open an Indian restaurant in Leeds

Description of problem and background

A relative wants to open an Indian restaurant in Leeds, England. More specifically a South Indian restaurant. The relative believes there to be a niche in the market for this particular cuisine in Leeds. Leeds is a diverse city with restaurants of many cuisines, but perhaps a gap in the market for South Indian. The relative is also not very familiar with all the neighbourhoods of Leeds.

Thus, the relative wants to understand which neighbourhoods of Leeds are most appropriate for setting up a South Indian restaurant. This may be done in several ways, but the relative would like to see patterns in the types of competing restaurants in each neighbourhood area. The target audience is my relative or more broadly entrepreneurs seeking to develop a restaurant in Leeds in the most appropriate place for maximised profits.

Description of data and how it will be used to solve problem

There is an existing online wikipedia article which lists different places in Leeds as well as their postcode

(https://en.wikipedia.org/wiki/List_of_places_in_Leeds). The data from this will be accessed and manipulated using BeautifulSoup (python package).

Geocoder (python library) will be used to determine the geographical coordinates of these places in Leeds. Foursquare API will be used to get the data of the venues at each of these places. Notable venue types of interest will be the different types of restaurants, particularly (South) Indian restaurants.

Machine learning in the form of k-means clustering combined with visualisation using folium will allow the analysis of the data. Thus, we will be able to draw conclusions to address the problem statement of this project.