



CAPSTONE PROJECT

IBM DATA SCIENCE PROFESSIONAL CERTIFICATE

OPENING AN ITALIAN RESTAURANT IN BIRMINGHAM, UK

BUSINESS PROBLEM

We want to take opportunities in retail, in particular, by opening an italian restaurant

Goal: Location is a key point. For this reason, we want to analyse the area to find the best match

DATA

Data used:

List of Birmingham areas (https://en.wikipedia.org/wiki/B_postcode_area)

Latitude and longitude (pgeocoder,geocoder)

Info about venues (Foursquare API)

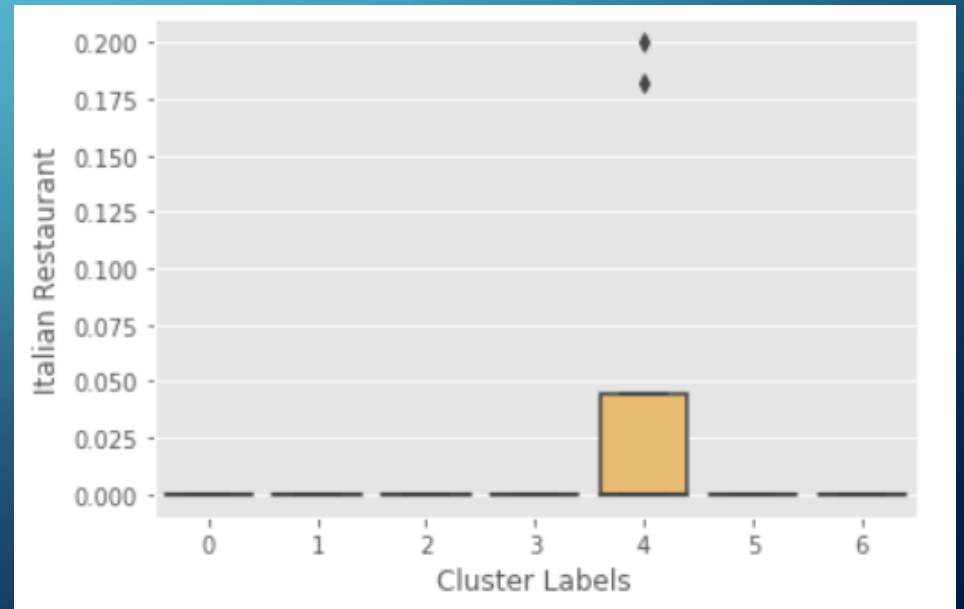
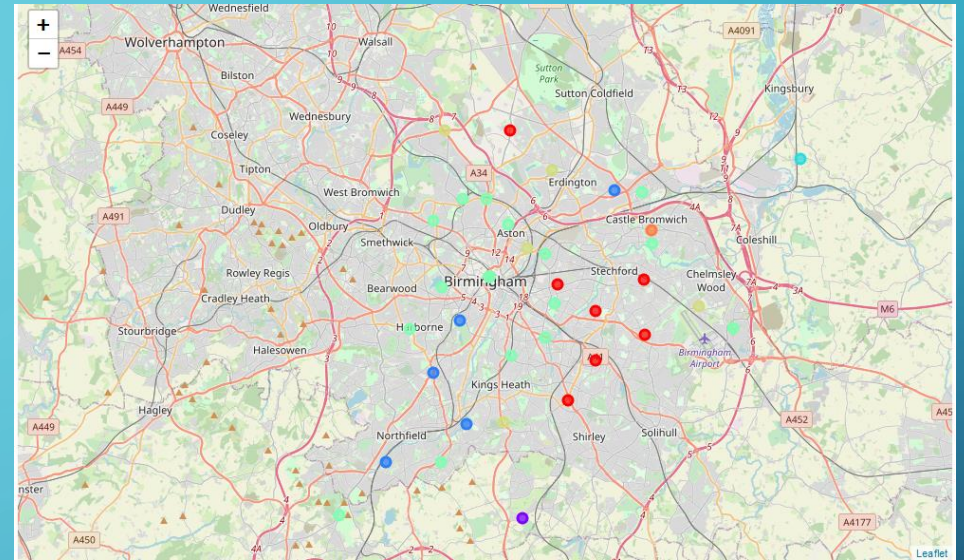
METHODOLOGY

- Webscraping (neighbourhoods list)
- Fetching latitude and longitude (pgeocoder,geocoder)
- Fetching (Foursquare API)
- One hot encoding of the venues (qualitative data→quantitative data)
- K-Means clustering

RESULTS

- Out of 7 clusters, one of interest was identified.

It is the only one where there are already italian restaurants while in the others there none.



DISCUSSION

- Since the complete absence of italian restaurants in the other cluster seems to suggest there is not much interest or that the clusters don't have the right features
- The presence is unbalanced also whitin the cluster of interest
- Recommendations:
- Open the business in the center
- Alterantively, in the other neighbourhoods of the cluster

