Opening a new restaurant in London

IBM DATA SCIENCE PROFSSIONAL CERTIFICATE – CAPSTONE PRESENTAION

JAZZ BHANGU, 3 JUL 2020

Intro / Business Problem

Our client owns a successful **vegetarian** restaurant in **West Hampstead, NW6** postcode

He wishes to open a second branch in London. But where to open it?

He has specific strategic criteria in mind regarding choice of location:

- **Similar area** to his current restaurant
- **Away from alternative dining options** – as customers will likely avoid taking a 'risk' with vegetarian if other options are available
- **Close to tube stations**
- **Ideally in same N, NW side of London**

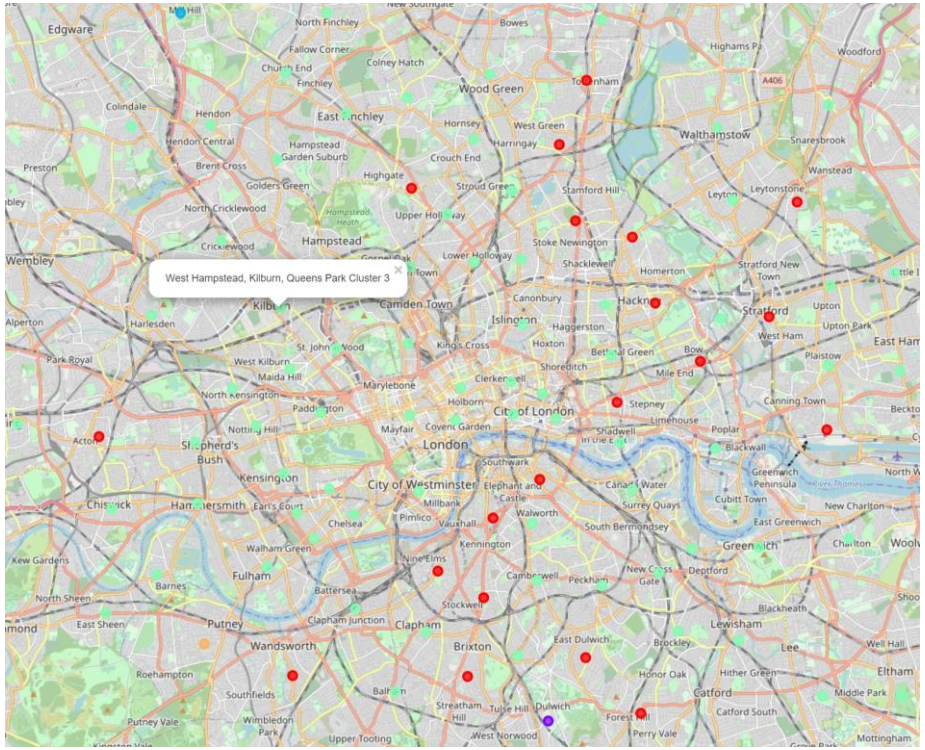
Data analysis / methodology

Initial data cleaning and merging into **pandas dataframes** was carried out.

Foursquare API was used to obtain location data and determine the venues in each area.

Basic exploration of this location data, using **histograms**, yielded an insight that neighbourhoods in London are quite different from each other.

So, **K-Means clustering** machine learning algorithm was used to find other similar areas in London, shown here in green circles.

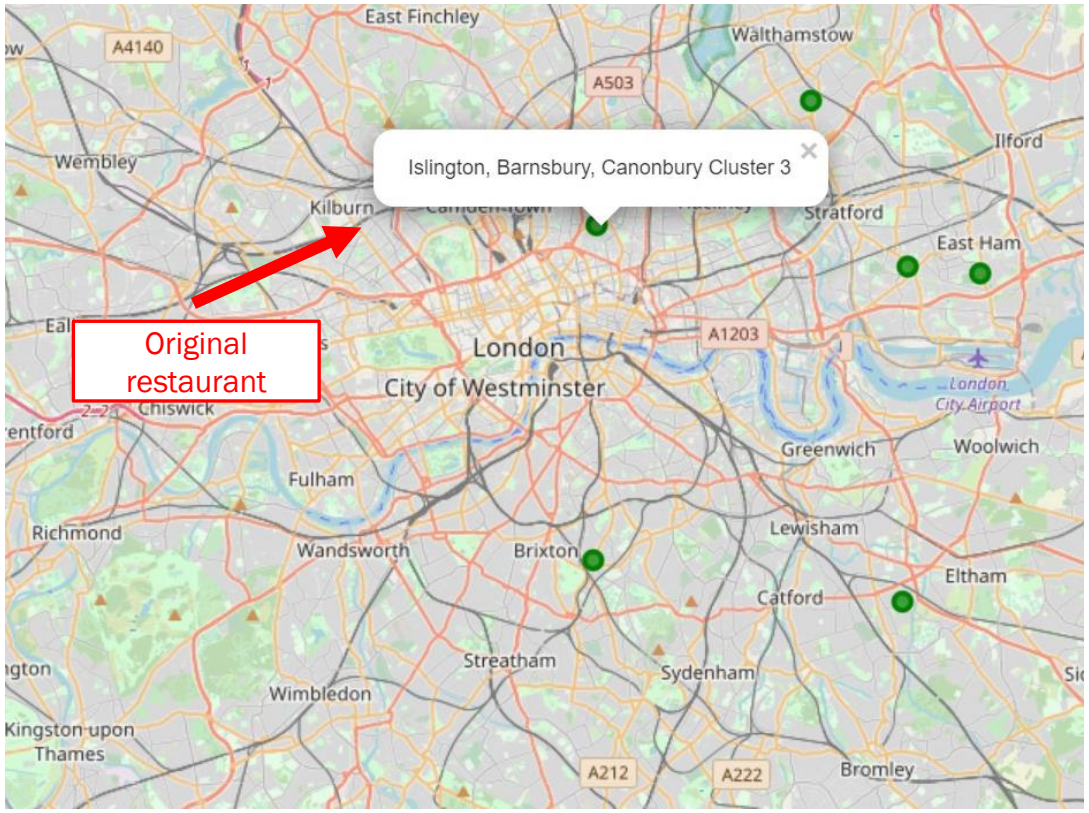


Results / discussion

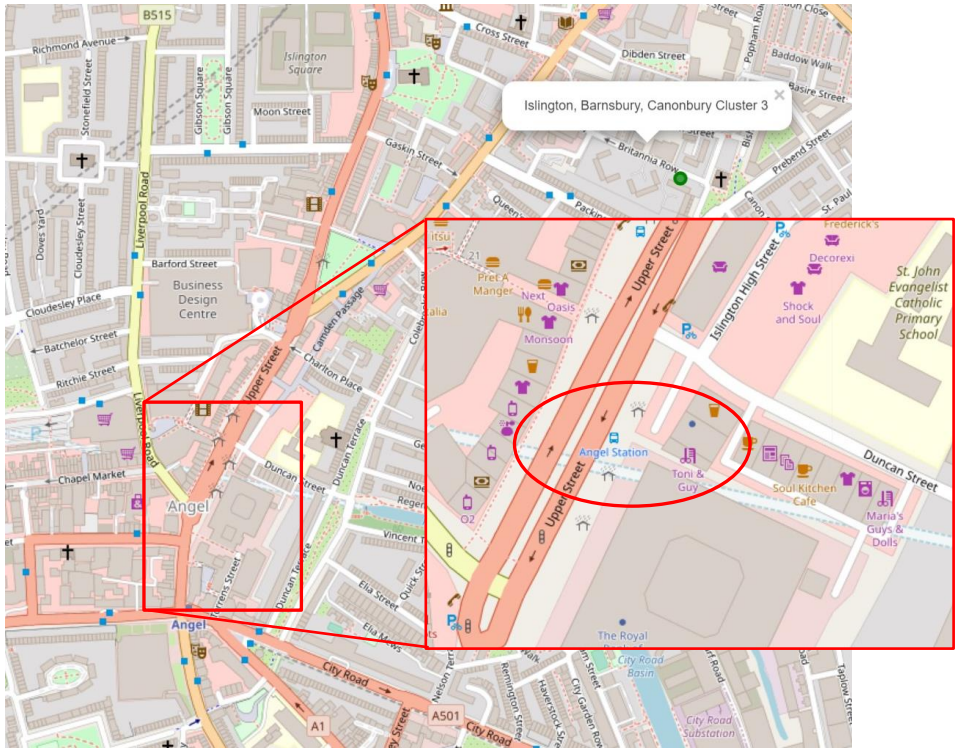
The locations in the same cluster as the original restaurant's neighbourhood ('West Hampstead') were then filtered to:

- Remove those that had Indian restaurants in the top venues (as Indian restaurants are a direct competition, with a sizeable vegetarian menu), and
- Further remove any where Restaurants featured in the top 10 venues count in that area. Fast-food outlets were not considered a direct competitor.

This then yielded 6 potential options for choice of location. Clearly, the only one in the same 'side of London', is **'Islington, Barnsbury, Canonbury'**, **NW6**.



Conclusion



The client should open his restaurant in the centre of **NW6 postcode area**. Nearby tube station – a key requirement – is confirmed by closer inspection; **'Angle'** tube.

Further research is recommended prior to final decision to confirm similarity of additional factors pertaining to neighbourhoods:

- Population demographics – average income etc.
- Type of building use – commercial vs residential mix.
- Peak pedestrian traffic flow-rates.