IBM Coursera Capstone Project

A Bird's view of London

By: Ankush Babbar

Business Problem:

- The London city is the financial capital of the country and is diverse in culture as well as the city have many visitors every year.

- To analyze the interests of the people living in the city and determine the localities that are popular for some specific venue categories.

- While many immigrants come and search for a house to live, we would look into what localities and neighborhoods differ from other.

- Those who are planning or interested in visiting the city might be interested what are good localities and local people interest overall

Data:

- Data Required:
  - List of Boroughs and Neighbourhoods of London.
  - Latitude and Longitude of areas in London.
  - Venue data particularly for comparing and clustering.

- Source of Data:
  - Wikipedia page for neighbourhoods , https://en.wikipedia.org/wiki/List_of_areas_of_London
  - Geocoder Module for Latitude and Longitude of Area.
  - Four Square API for exploring venues.

Methodology:

- Web Scrapping Wikipedia page for neighbourhoods list.

- Get Latitude and Longitude coordinates using geocoder.

- Use Foursquare API to get venue data

- Group data by neighbourhood , filtering and cleaning Data.

- Perform Clustering on the data by using K-Means clustering.

- Visualize the clusters in a map using Folium.

Results:

- Categorized neighborhoods into 5 Clusters:
  - ➤ Cluster 1 : Neighborhoods with Restaurants, Clubs, Pubs, Bars, Cafes.
  - ➤ Cluster 2: Neighborhoods with Best Grocery Stores
  - ➤ Cluster 3: Neighborhoods with good Hardware Stores.
  - ➤ Cluster 4: Neighborhoods with good public Transits.
  - ➤ Cluster 5: Neighborhoods with Famous Parks.

Discussion:

- If someone is deciding to buy a house in London we can see the neighborhood clusters venues and based on venues decide where he should buy the house basically analyzing the locality.

- London is a big city with a high population density in a narrow area. The total number of measurements and population densities of the all districts in total can vary. As there is such a complexity, very different approaches can be tried in clustering and classification studies.

Conclusion:

- As a result, people are turning to big cities to start a business or work. For this reason, people can achieve better outcomes through their access to the platforms where such information is provided.

- Not only for investors but also city managers can manage the city more regularly by using similar data analysis types or platforms.

Thank you..!!!