Coursera Capstone

IBM Applied Data Science

Capstone

opening a new bar in the city of London& looking for a happy place to live.

By:Mohammad Aslam

July 2020

Business Problem

- 1)The objective of this capstone project is to analyse and select the best locations in the city of London to start a family or living or moving in from other places
-
- 2).The best place in the city of london to start a “Bar”. Using the techniques learnt throughout the course I aim to solve the problem.

Data

- Data required
  - List of neighbourhoods in London
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to Bar.
  - Data on happiness index.

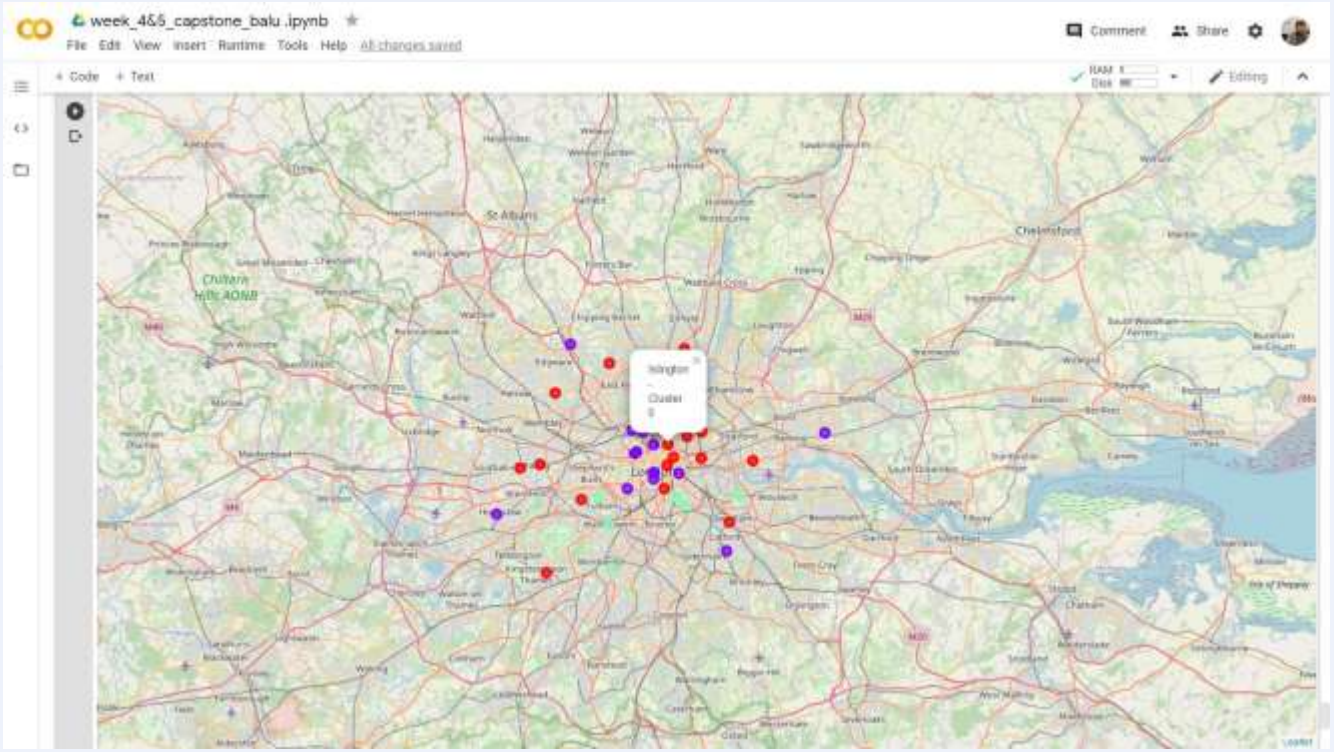
- Sources of data
  - london data store:burrough profiles.
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

Methodology

- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using Geocoder
- map and clustter the happiness index data
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Bar
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

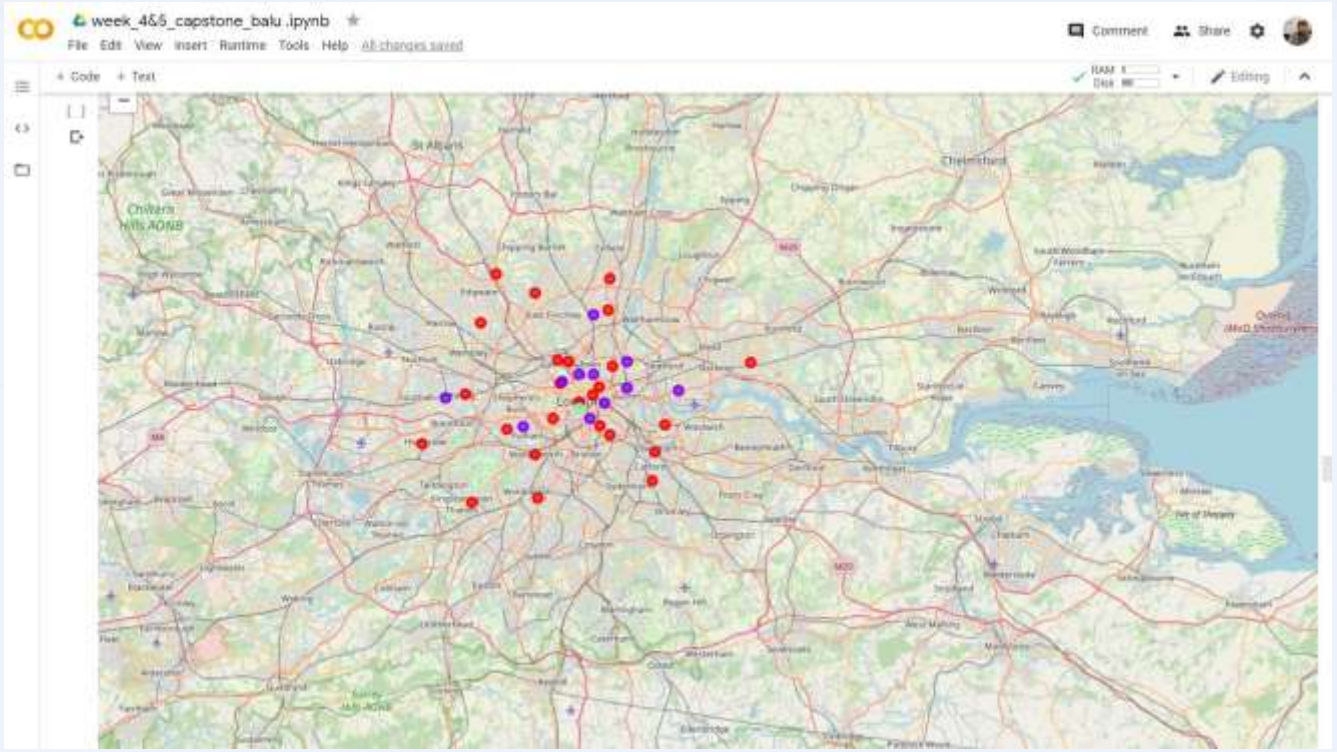
Results

- Categorized the neighbourhoods into 3 clusters :
  - Cluster 0: Neighbourhoods with highno of bars.
  - Cluster 1: Neighbourhoods with moderate number to no of bars
  - Cluster 2: Neighbourhoods with low concentration of shopping malls



Results

- All the areas in the city of London ,England have been clustered into three regions as shown above in the map .They have been clustered based on the happiness index,worthwhileness index ,life satisfaction indices that give the overall living index of these areas.



Discussion

- The results from the k-means clustering of bars show that we can categorize the neighbourhoods into 3 clusters based on the frequency of occurrence for “bars”
- Cluster 0: Neighbourhoods with high number of bars
- Cluster 1: Neighbourhoods with moderate number to no existence of “bars”
- Cluster 2: Neighbourhoods with low concentration of bara

Recommendation

- Open new shopping malls in neighbourhoods in cluster 1 with little to no competition
- Can also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition
- Avoid neighbourhoods in cluster 2, already high concentration of shopping malls and intense competition

Conclusion

- Answer to business question: The neighbourhoods in cluster 1 are the most preferred locations to open a new shopping mall
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new shopping mall

Thank you!

thank you
danke
謝謝
ngiyabonga
teşekkür ederim
рахмат
Баярлалаа
спасибо
faafetai lava
vinaka
blagodaram
dank je
misaotra
matondo
paldies
grazzi
gracias
tapadh leat
хвала
asante
manana
obrigada
kiitos
dankie
dhanyavad
bedankt
hvala
maururu
enkosi
dziękuję
mochchakkeram
go raibh maith agat
chnorakaloutioun
gratias ago
sukriya
kop khun krap
taiku
obrigado
sagolun
dekuji
arigatô
takk
dakujem
trugarez
ありがとう
grazie
terima kasih
najis tuke
tanemirt
rahmet
dhanyavadagalu
shukriya
мерси
merci
감사합니다
xiexie
ευχαριστώ
diolch
kam sah hamnida