



IBM

Coursera Capstone

A p p l i e d D a t a S c i e n c e

Opening a New Bakery in Brisbane, Australia

I N D E X

Coursera Capstone
Applied Data Science
Capstone _ Final Submission

01. Business Problem

02. Data

03. Methodology

04. Results

05. Discussion and Conclusion

- Main objective:
To analyse and select the best locations in the city of Brisbane, Australia to open a new bakery
- This project is timely as the city is currently suffering from oversupply of bakeries
- Main question
If a investor is looking to open a new bakery, where would you recommend that they open it?

- Essential Data:
 - List of neighborhoods in Brisbane
 - Latitude and longitude information of the neighborhoods
 - Address data of existing bakeries
- Sources of data
 - Neighborhoods data – Wikipedia website
([https://en.wikipedia.org/wiki/Category:Suburbs of Brisbane](https://en.wikipedia.org/wiki/Category:Suburbs_of_Brisbane))
 - Geocoder package for latitude and longitude
 - Foursquare API for venue data

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

04. Results

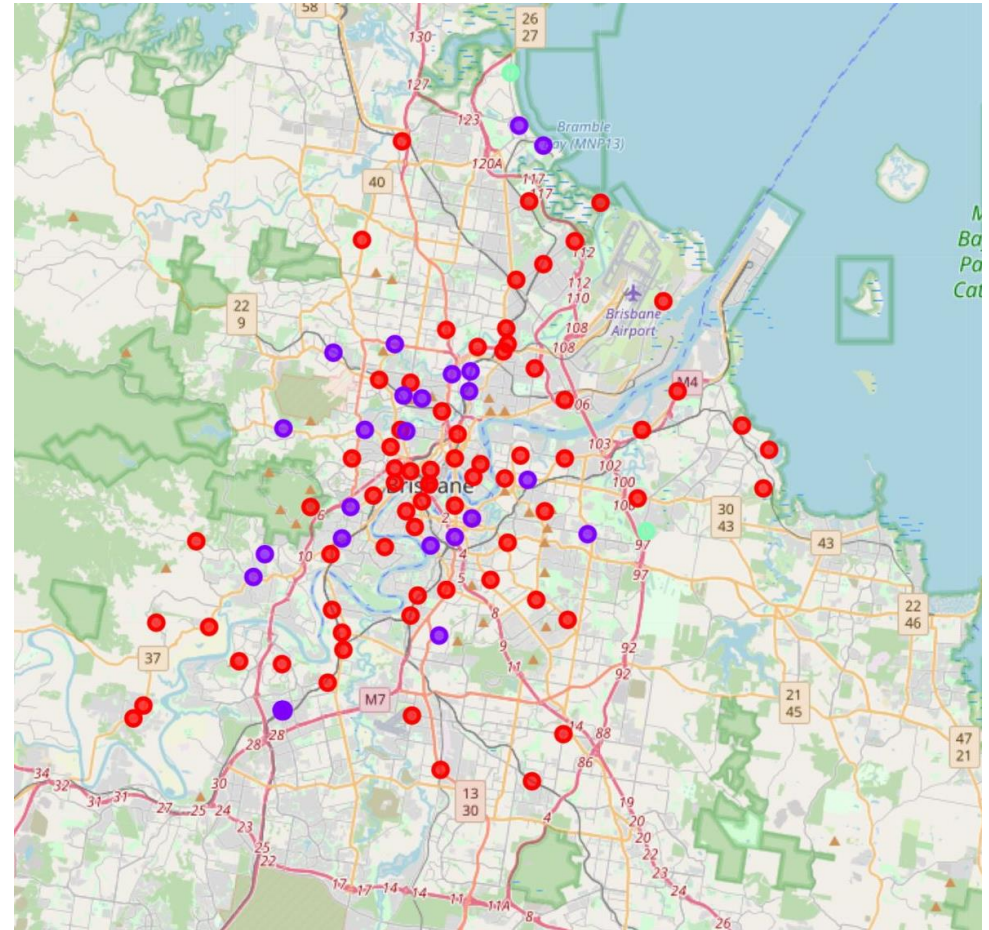


- Categorized the neighborhoods into 3 clusters:

Cluster 0: Neighborhoods with low number to no existence of bakeries

Cluster 1: Neighborhoods with moderate number of bakeries

Cluster 2: Neighborhoods with high concentration of bakeries



- Highest number in Cluster 2 and moderate number in cluster 1
Oversupply of bakeries mostly happened in the Asian town
Avoid cluster 2, already high concentration of bakeries
- Cluster 0 has very low number to no bakeries in the near place
It is recommended to open new bakeries on cluster 0 with now competition
- Answer to business question
In terms of facility density, Cluster 0 are the most preferred locations
Additional research are needed

The background of the slide features a grayscale photograph of a modern building with large windows and a balcony, partially obscured by a large, gnarled tree. A semi-transparent red overlay covers the left and bottom portions of the image, creating a diagonal split effect. The text "Thank you" is centered within the red area.

Thank you