The Battle of Neighborhoods

Introduction

The purpose of this Project is to help people exploring better facilities around their place of choice. It will help people making smart and efficient decision on selecting great place out of a big number of other places in New Zealand.

I been planning to move to New Zealand and will be great if I have lots of information about which places have the facilities that best fit my family lifestyle. For ease of accessing to Cafe, School, Super market, medical shops, grocery shops, mall, theatre, hospital, like minded people, etc.

This Project aim to create an analysis of features for a people (like me) migrating to New Zealand to search the best fitted place as a comparative analysis between populated places. The features include median housing price and the most common facilities around in that place.

It will help people to get awareness of the area and place before moving to a new city, state, country or place for their work or to start a new fresh life

Foursquare API

This project will use Four-square API as its prime data gathering source because it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

Work Flow

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

Clustering Approach

We will cluster the 20ths most populated places based on theirs urban development and average house prices.

Libraries used to develop the Project

Pandas: For creating and manipulating dataframes.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

Scikit Learn: For importing k-means clustering.

JSON: Library to handle JSON files.

XML: To separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.

Matplotlib: Python Plotting Module.

Data Description

We will use a dataset that locates the populted places in New Zealand. Populated places are defined here as settled areas, either urban or rural where densitys of around 20 persons per hectare exist, and something is able to be seen from the air.

Data Link: https://koordinates.com/layer/3657-nz-populated-places-points/

Additionally, we also will be using data from a page that contains average house prices in New Zealand

Data Link: https://www.enz.org/new-zealand-house-prices.html

Foursquare API Data

We will need data about different venues in different places. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of places, we then connect to the Foursquare API to gather information about venues inside each and every place. For each place, we have chosen the radius to be 500 meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

Place Name
Place Latitude
Place Longitude
Venue
Name of the venue e.g. the name of a store or restaurant
Venue Latitude
Venue Longitude
Venue Category

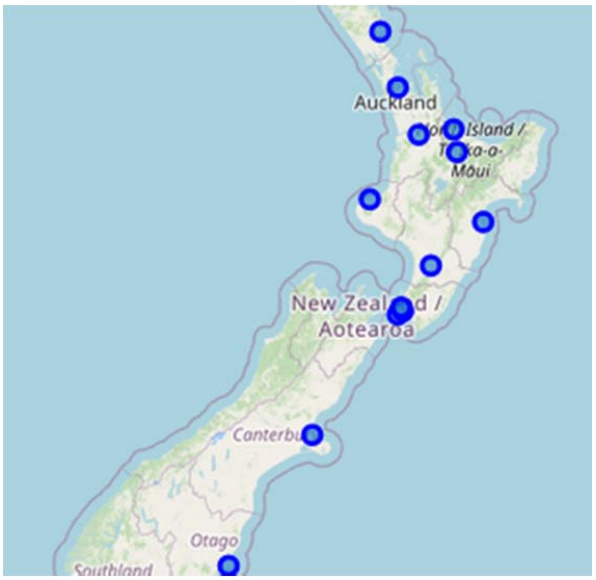
Methodology

In this section we discuss the data analysis that I conducted in order to arrive at the final result

Scraping data

A dataframe *df_2* is generated that consists of the union of the location data ("GEOMETRY_X" and "GEOMETRY_Y") and the amount of population together with the average values of houses in that area ("HAP"). Only the places where the average value of the houses was found were taken.

	name	desc_code	pop	region	tla	class	GEOMETRY_X	GEOMETRY_Y	HAP
0	Auckland	METR	1137141	Auckland Region	Auckland City	2	174.765989	-36.847013	800000
1	Christchurch	METR	333204	Canterbury Region	Christchurch City	2	172.636656	-43.532049	450000
2	Wellington	METR	164508	Wellington Region	Wellington City	2	174.776890	-41.291179	600000
3	Hamilton	TOWN	128988	Waikato Region	Hamilton City	2	175.281742	-37.785863	575000
4	Dunedin	METR	87384	Otago Region	Dunedin City	2	170.503575	-45.874123	460000
5	Lower Hutt	USAT	76845	Wellington Region	Lower Hutt City	2	174.904349	-41.209652	657100
6	Palmerston North	TOWN	68439	Manawatu-Wanganui Region	Palmerston North City	2	175.610996	-40.356124	460000
7	Tauranga	TOWN	64986	Bay of Plenty Region	Tauranga City	2	176.169876	-37.684330	670100
8	Napier	TOWN	52302	Hawke's Bay Region	Napier City	2	176.918314	-39.490174	555000
9	Rotorua	TOWN	49533	Bay of Plenty Region	Rotorua District	2	176.251523	-38.137847	455000
10	Porirua	USAT	46023	Wellington Region	Porirua City	2	174.842505	-41.138316	600000
11	Whangarei	TOWN	45087	Northland Region	Whangarei District	2	174.322041	-35.726603	515000
12	Invercargill	TOWN	42129	Southland Region	Invercargill City	2	168.347374	-46.412985	318000
13	New Plymouth	TOWN	41925	Taranaki Region	New Plymouth District	2	174.076095	-39.057195	475000



Next, each location was explored with a radius of 1,000 meters and a limit of 100 locations.

The response of the Foursquare Explore endpoint was JSON collection which required to be wrangled in order to extract meaningful information.

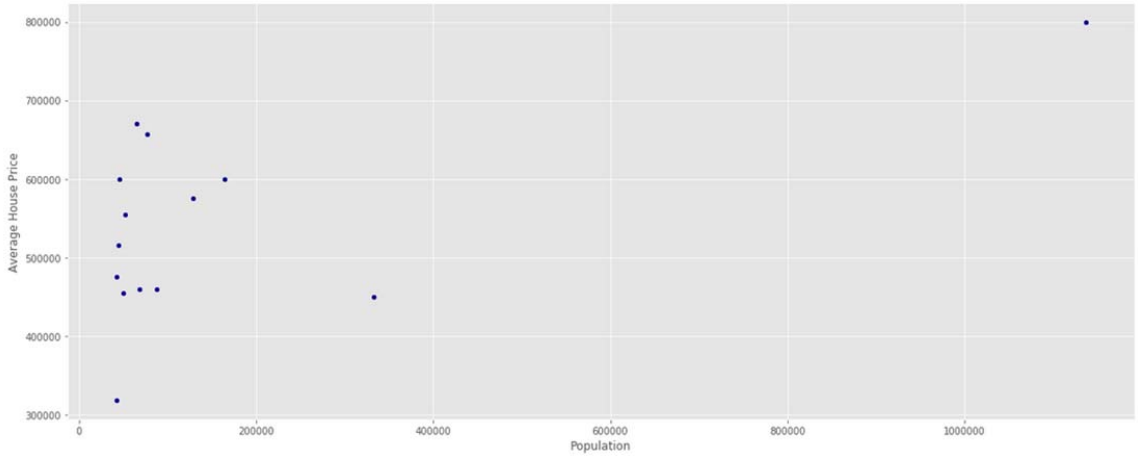
The resulted dataframe was then sorted in descending order on the basis of the population count and the top 10 columns (corresponding to the top 10 venues) were stored.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Auckland	Café	Japanese Restaurant	Pizza Place	Plaza	Korean Restaurant	Steakhouse	Mexican Restaurant	Bar	Indian Restaurant	Pub
1	Christchurch	Café	Hotel	Bar	Japanese Restaurant	Thai Restaurant	Coffee Shop	Restaurant	Gastropub	Supermarket	Italian Restaurant
2	Dunedin	Café	Coffee Shop	Bar	Asian Restaurant	Hotel	Shopping Mall	Brewery	Theater	Restaurant	Pub
3	Hamilton	Café	Restaurant	Bar	Fast Food Restaurant	Hotel	Coffee Shop	Department Store	Mexican Restaurant	Performing Arts Venue	Steakhouse
4	Invercargill	Café	Burger Joint	Department Store	Bar	Food	Food & Drink Shop	Fast Food Restaurant	Supermarket	Movie Theater	Flower Shop
5	Lower Hutt	Café	Supermarket	Fast Food Restaurant	Indian Restaurant	Electronics Store	Burger Joint	Shoe Store	Shopping Mall	Portuguese Restaurant	Japanese Restaurant
6	Napier	Café	Restaurant	Fast Food Restaurant	Electronics Store	Thai Restaurant	Hotel	Beach	Organic Grocery	Paper / Office Supplies Store	Pub
7	New Plymouth	Restaurant	Café	Coffee Shop	American Restaurant	Hotel	Thai Restaurant	Electronics Store	Bar	Bakery	Burger Joint
8	Palmerston North	Café	Fast Food Restaurant	Asian Restaurant	Coffee Shop	Grocery Store	Shopping Mall	Japanese Restaurant	Thai Restaurant	Brewery	Burger Joint
9	Porirua	Furniture / Home Store	Grocery Store	Café	Paper / Office Supplies Store	Skate Park	Pizza Place	Burger Joint	Pet Store	Fast Food Restaurant	Supermarket
10	Rotorua	Café	Hotel	Pizza Place	Hostel	Motel	Pedestrian Plaza	Pool	Steakhouse	Furniture / Home Store	Middle Eastern Restaurant
11	Tauranga	Café	Coffee Shop	Hotel	Pub	Japanese Restaurant	Middle Eastern Restaurant	Gastropub	Park	Seafood Restaurant	Fast Food Restaurant
12	Wellington	Café	Coffee Shop	Bar	Restaurant	Vietnamese Restaurant	Italian Restaurant	Pizza Place	Asian Restaurant	Burger Joint	Chinese Restaurant
13	Whangarei	Café	Supermarket	Fast Food Restaurant	Gastropub	Middle Eastern Restaurant	Coffee Shop	Bakery	Restaurant	Electronics Store	Filipino Restaurant

Next, a Dataframe was generated only with the population values and average house value, with the intention of generating a k-cluster algorithm.

	name	pop	HAP
0	Auckland	1137141	800000
1	Christchurch	333204	450000
2	Wellington	164508	600000
3	Hamilton	128988	575000
4	Dunedin	87384	460000
5	Lower Hutt	76845	657100
6	Palmerston North	68439	460000
7	Tauranga	64986	670100
8	Napier	52302	555000
9	Rotorua	49533	455000
10	Porirua	46023	600000
11	Whangarei	45087	515000
12	Invercargill	42129	318000
13	New Plymouth	41925	475000

A scatter plot was made to observe the points.



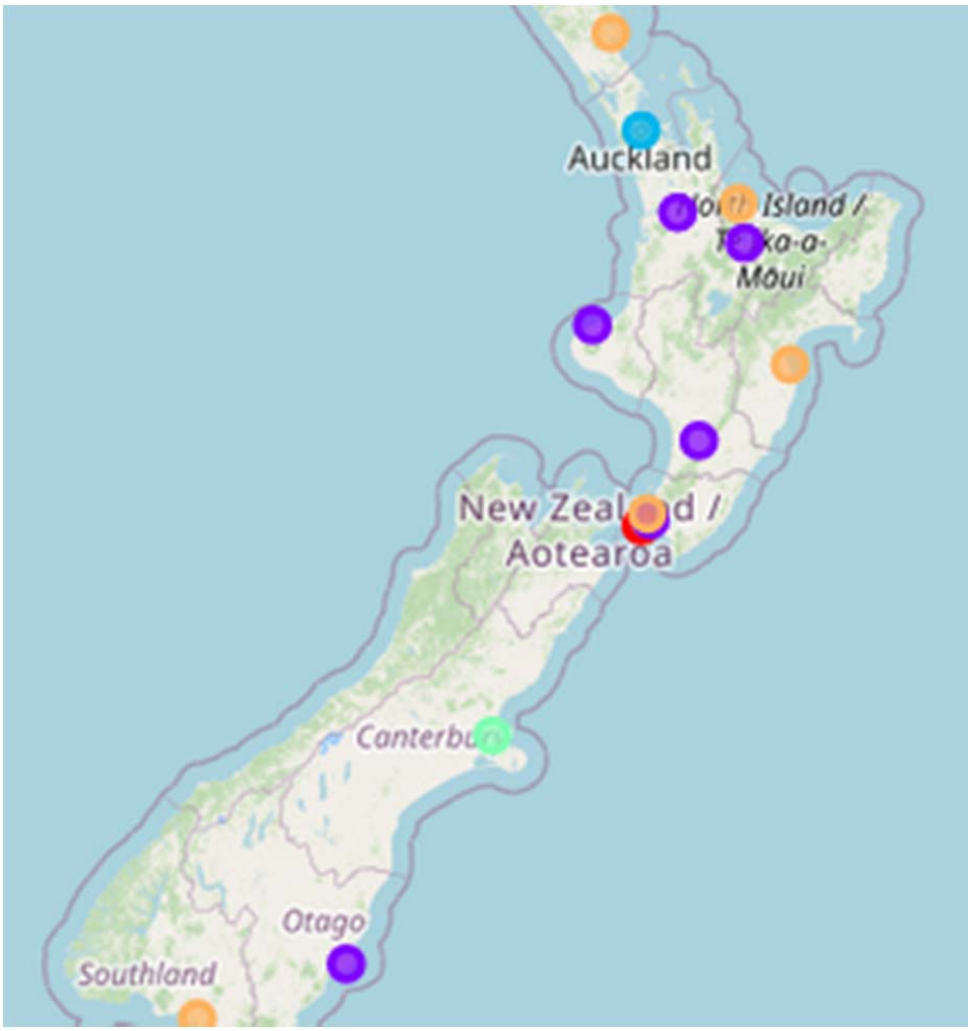
Next, the StandardScaler was used to normalize the dataset.

```
array([[ 3.47270088,  2.21867512],
       [ 0.59483645, -0.79298982],
       [-0.00904696,  0.49772373],
       [-0.1361984 ,  0.2826048 ],
       [-0.28512881, -0.70694225],
       [-0.32285542,  0.98905535],
       [-0.35294649, -0.70694225],
       [-0.36530724,  1.10091719],
       [-0.41071233,  0.11050966],
       [-0.42062456, -0.74996603],
       [-0.43318936,  0.49772373],
       [-0.43653997, -0.23368061],
       [-0.44712876, -1.92881774],
       [-0.44785903, -0.57787089]])
```

Next, a k-means algorithm was generated taking as k = 5. With this, a Dataframe was created with the generated labels for each place.

	name	desc_code	pop	region	tla	class	GEOMETRY_X	GEOMETRY_Y	HAP	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Auckland	METR	1137141	Auckland Region	Auckland City	2	174.765989	-36.847013	800000	2	Café	Japanese Restaurant	Pizza Place	Plaza	Korean Restaurant	Steakhouse	Mexican Restaurant	Bar	Indian Restaurant
1	Christchurch	METR	333204	Canterbury Region	Christchurch City	2	172.636656	-43.532049	450000	3	Café	Hotel	Bar	Japanese Restaurant	Thai Restaurant	Coffee Shop	Restaurant	Gastropub	Supermarket
2	Wellington	METR	164508	Wellington Region	Wellington City	2	174.776890	-41.291179	600000	0	Café	Coffee Shop	Bar	Restaurant	Vietnamese Restaurant	Italian Restaurant	Pizza Place	Asian Restaurant	Burger Joint
3	Hamilton	TOWN	128988	Waikato Region	Hamilton City	2	175.281742	-37.785863	575000	1	Café	Restaurant	Bar	Fast Food Restaurant	Hotel	Coffee Shop	Department Store	Mexican Restaurant	Performing Arts Venue
4	Dunedin	METR	87384	Otago Region	Dunedin City	2	170.503575	-45.874123	460000	1	Café	Coffee Shop	Bar	Asian Restaurant	Hotel	Shopping Mall	Brewery	Theater	Restaurant
5	Lower Hutt	USAT	76845	Wellington Region	Lower Hutt City	2	174.904349	-41.209652	657100	1	Café	Supermarket	Fast Food Restaurant	Indian Restaurant	Electronics Store	Burger Joint	Shoe Store	Shopping Mall	Portuguese Restaurant
6	Palmerston North	TOWN	68439	Manawatu-Wanganui Region	Palmerston North City	2	175.610996	-40.356124	460000	1	Café	Fast Food Restaurant	Asian Restaurant	Coffee Shop	Grocery Store	Shopping Mall	Japanese Restaurant	Thai Restaurant	Brewery
7	Tauranga	TOWN	64986	Bay of Plenty Region	Tauranga City	2	176.169876	-37.684330	670100	4	Café	Coffee Shop	Hotel	Pub	Japanese Restaurant	Middle Eastern Restaurant	Gastropub	Park	Seafood Restaurant
8	Napier	TOWN	52302	Hawke's Bay Region	Napier City	2	176.918314	-39.490174	555000	4	Café	Restaurant	Fast Food Restaurant	Electronics Store	Thai Restaurant	Hotel	Beach	Organic Grocery	Paper / Office Supplies Store
9	Rotorua	TOWN	49533	Bay of Plenty Region	Rotorua District	2	176.251523	-38.137847	455000	1	Café	Hotel	Pizza Place	Hostel	Motel	Pedestrian Plaza	Pool	Steakhouse	Furniture / Home Store
10	Porirua	USAT	46023	Wellington Region	Porirua City	2	174.842505	-41.138316	600000	4	Furniture / Home Store	Grocery Store	Café	Paper / Office Supplies Store	Skate Park	Pizza Place	Burger Joint	Pet Store	Fast Food Restaurant
11	Whangarei	TOWN	45087	Northland Region	Whangarei District	2	174.322041	-35.726603	515000	4	Café	Supermarket	Fast Food Restaurant	Gastropub	Middle Eastern Restaurant	Coffee Shop	Bakery	Restaurant	Electronics Store
12	Invercargill	TOWN	42129	Southland Region	Invercargill City	2	168.347374	-46.412985	318000	4	Café	Burger Joint	Department Store	Bar	Food	Food & Drink Shop	Fast Food Restaurant	Supermarket	Movie Theater
13	New Plymouth	TOWN	41925	Taranaki Region	New Plymouth District	2	174.076095	-39.057195	475000	1	Restaurant	Café	Coffee Shop	American Restaurant	Hotel	Thai Restaurant	Electronics Store	Bar	Bakery

A map was generated to show the classification



Results

After conducting the study and checking a variety of details, the following results are the results:

- 3 of the 5 clusters consist only one city, that demonstrates that this cities are very particular
- The most common venue in almost every cluster is “Café”

Discussion

Based on the data and the charts plotted, we can infer the following:

- Cluster 1, which consist of the city of Wellington, has very high population and very high house prices.

- Cluster 2, which consists of the cities of Hamilton, Dunedin, Lower Hutt, Palmerston North, Rotorua and New Plymouth, has low population and low house prices.
- Cluster 3, which consists of the city of Auckland, has very highest population of the country and the highest house values of the study.
- Cluster 4, which consists of the city of Christchurch, has high population and the lowest house values of the study.
- Cluster 5, which consists of the cities of Tauranga, Napier, Porirua, Whangarei, Invercargill, has the lowest population of the study but high house prices.

Conclusion

As a part of this project and study, we demonstrated that with the right dataset and the right skills, plausible relationships can be established between features which can lead to deeper insights. We classified the 14 most populated cities of New Zealand. These studies can be vital for people looking to migrate to New Zealand, helping then to select the suited city for theirs needs.

This concludes the report for this Capstone project. Thank you.