

Open a coffee shop in Wellington, New Zealand

Naoto Sai, Story data scientist based in Munich

September 7, 2019

1. Introduction

In this IBM capstone project on Coursera, I set a problem from my interest and future vision. In the near future, I am going to live in New Zealand over 1 year and am wondering which city would have interesting things to offer to me. In addition to which city I go to, I add another issue that I has to be taken care of. Let us say that I am coffee addicted and have to open a new coffee shop in order to survive at a city in which I am living.

Here there are two problems to be solved: The first one is to decide to which city I am going and the other one is to find out a suburb where I will open a coffee shop in the city. This means that at first I have to define my criteria of a city where I want to move into, and then need to analyze characteristics of several cities in New Zealand and see what they look like. Afterwards it is necessary for me to see each suburb in the city. However, it would be still complicated to me that I, as a beginner data scientist, would get and scrape various data such as weather information in different cities from different sources and compare them. Therefore, I would go for a simple premise: I am going to live in “Wellington”. This could actually happen in my future for various reasons that I would not mention here and where actually IBM headquarters in New Zealand is located. In addition, ICT and Degitech industry are growing in Wellington and this sounds fascinating to me as well.

The city where I am going was chosen, so the first problem is solved. But, in which part of Wellington would it make sense for me to open a coffee shop? For that answer, I am going to check venues and inhabitants in each suburb and find out a suburb having comparatively many inhabitants with less competitors, coffee shops or cafes.

2. Data acquisition and cleaning

These data as shown below can be obtained without any cost:

- Suburbs of Wellington city including data of Longitude and Latitude (*1)
- Population in each suburb (*1)
- Geopy library, folium for mapping
- Foursquare API, which is used to get common venue data of given suburbs in Wellington.

Table 1 shows the suburbs of Wellington city, population of each suburb, Latitude and Longitude respectively. There are 51 suburbs in Wellington and this table is already sorted by population. It is found out that the maximum number of population among 51 suburbs is 14,736, in Karori.

	Suburb	Population	Latitude	Longitude
16	Karori	14736	-41.284855	174.736795
43	Tawa, New Zealand	13764	-41.171567	174.824786
44	Te Aro	11313	-41.294000	174.777000
13	Johnsonville, New Zealand	10239	-41.223540	174.807240
25	Miramar, New Zealand	8898	-41.316667	174.816667
18	Khandallah	8478	-41.233333	174.783333
30	Newtown, New Zealand	8418	-41.313700	174.779800
29	Newlands, New Zealand	7278	-41.224000	174.823000
12	Island Bay, New Zealand	6861	-41.337730	174.772310
19	Kilbirnie, New Zealand	6429	-41.320400	174.794100

Table 1. Suburbs of Wellington and numbers in each suburb

In order to decide where to open a coffee shop, as written above, I will take venues data in each suburb into account. These data can be obtained by using Foursquare API and k-means clustering would be applied for looking at correlation in each suburb.

3. Exploratory data analysis

3.1 Create of a map of Wellington with suburbs superimposed on top

Table 2 shows the whole data on suburbs in Wellington. Figure 1 shows the map of Wellington with suburbs superimposed on top. In order to create this map I imported geopy and folium library, and used the data showed on the table 1. In this report, explanations about the details of the code are not necessary, so you can check my codes on my GitHub account page(*2) as a reference.

	Suburb	Population	Latitude	Longitude				
0	Aro Valley	3621	-41.295278	174.767500	23	Maupua	1503	-41.298413 174.823105
1	Berhampore, New Zealand	3609	-41.316667	174.783333	24	Melrose, New Zealand	1215	-41.325468 174.788161
2	Breaker Bay	0	-41.329877	174.832294	25	Miramar, New Zealand	8898	-41.316667 174.816667
3	Brooklyn, New Zealand	4074	-41.306111	174.763056	26	Mornington, Wellington	2208	-41.317896 174.764115
4	Churton Park	5862	-41.208355	174.808190	27	Mount Cook, Wellington	5112	-41.301764 174.774031
5	Crofton Downs	1508	-41.257000	174.764000	28	Mount Victoria, Wellington	5400	-41.296155 174.794479
6	Glenside, New Zealand	345	-41.199355	174.819190	29	Newlands, New Zealand	7278	-41.224000 174.823000
7	Grenada North	336	-41.191056	174.845438	30	Newtown, New Zealand	8418	-41.313700 174.779800
8	Grenada, New Zealand	1248	-41.204824	174.825544	31	Ngaio, New Zealand	5502	-41.233333 174.766667
9	Hataitai	4563	-41.300000	174.800000	32	Ngauranga	39	-41.250000 174.816667
10	Horokiri	180	-41.207000	174.849000	33	Northland, Wellington	3357	-41.281389 174.758889
11	Houghton Bay	0	-41.350000	174.783333	34	Ohariu, New Zealand	783	-41.223611 174.743889
12	Island Bay, New Zealand	6861	-41.337730	174.772310	35	Oriental Bay	1056	-41.291667 174.794444
13	Johnsonville, New Zealand	10239	-41.223540	174.807240	36	Owhiro Bay	1743	-41.344000 174.759000
14	Kaiwharawhara	144	-41.260001	174.789466	37	Paparangi	2841	-41.216643 174.818866
15	Karaka Bays	0	-41.303142	174.830419	38	Rongotai	1104	-41.326186 174.806131
16	Karori	14736	-41.284855	174.736795	39	Roseneath, Wellington	1728	-41.292297 174.801257
17	Kelburn, New Zealand	3615	-41.288771	174.763298	40	Seatoun	2235	-41.316667 174.833333
18	Khandallah	8478	-41.233333	174.783333	41	Strathmore Park	3708	-41.333300 174.820000
19	Kilbirnie, New Zealand	6429	-41.320400	174.794100	42	Takapu Valley	306	-41.161192 174.872087
20	Kingston, Wellington	2208	-41.324318	174.763107	43	Tawa, New Zealand	13764	-41.171567 174.824786
21	Lyall Bay	2607	-41.329993	174.793505	44	Te Aro	11313	-41.294000 174.777000
22	Makara, New Zealand	759	-41.220009	174.713791	45	Thorndon, New Zealand	4125	-41.275000 174.777900
					46	Vogeltown, Wellington	2004	-41.313000 174.769000
					47	Wadestown, New Zealand	3516	-41.261820 174.773570
					48	Wellington Central, Wellington	3913	-41.284945 174.775958
					49	Wilton, New Zealand	2052	-41.268160 174.756944
					50	Woodridge, New Zealand	1329	-41.212849 174.834145

Table 2. The whole data of suburbs in Wellington and numbers in each suburb

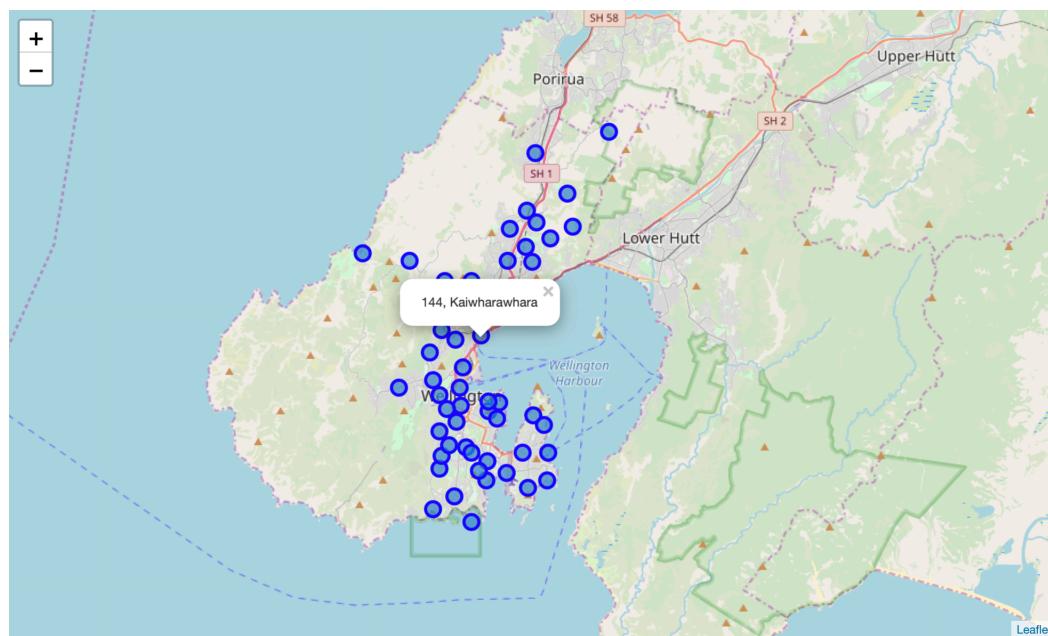


Fig. 1. Map of Wellington with suburbs superimposed on top

3.2 Acquire venues data in each suburb

As a next step, I got data of venues in each suburb by using Foursquare API and found out that the data from only 43 suburbs were available. As for Foursquare API, I defined max. 100 venues as a limit and 500 m radius from the centroid of each suburb. Figure 2 shows the number of venues in each suburb. These suburbs are already sorted by the number of venues. Since there are over 50 venues both in Te Aro and Wellington Central, I assume that these suburbs must be prospering and the places where many inhabitants in Wellington work.

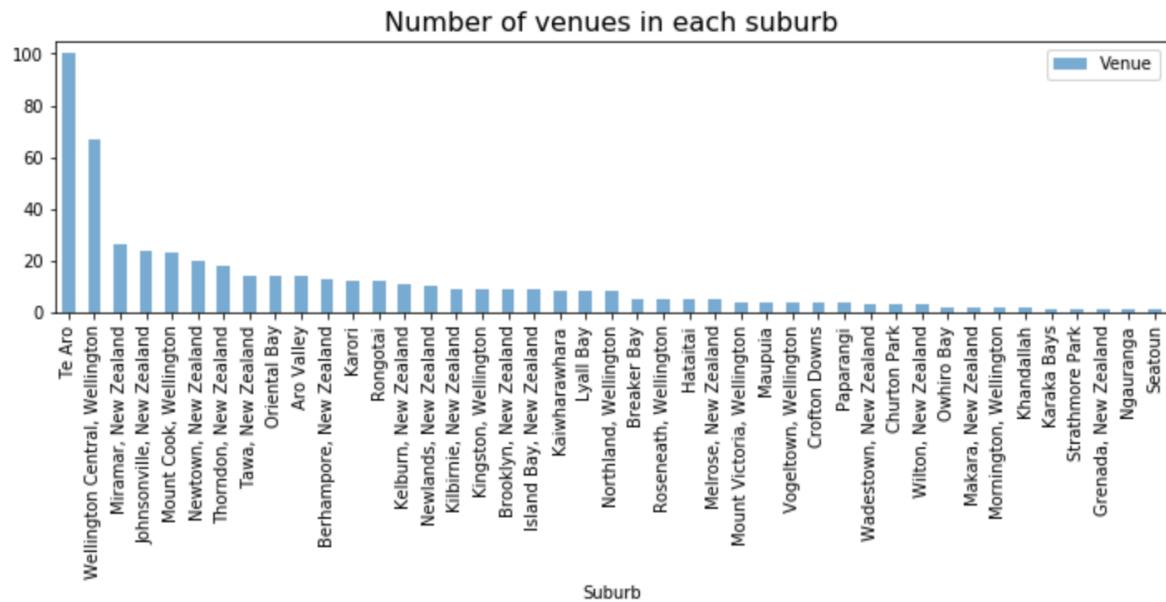


Fig. 2. Number of venues in each suburb

3.3 Print each suburb along with the top 10 most common venues

In order to find out which suburb has the most coffee shops, I need to dig down the acquired data above and will print out each suburb along with the top 10 most common venues. Therefore, one hot encoding process is applied for the given data and there is a data frame showing a suburb and count of venue categories as shown on the table 3.

Suburb	Airport Terminal	Art Gallery	Arts & Crafts Store	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bar	Bay	...	Tunnel	Turkish Restaurant	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Waterfront
0 Aro Valley	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
1 Aro Valley	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
2 Aro Valley	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
3 Aro Valley	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
4 Aro Valley	0	0	0	0	0	0	1	0	0	0	...	0	0	0	0	0	0

5 rows × 150 columns

Table 3. Count of venue category in each suburb in the first 5 rows

Afterwards, I made columns according to the number of top venues in each suburb and created a new data frame shown as Table 4.

	Suburb	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	Aro Valley	Bakery	Liquor Store	Beer Bar	Farmers Market	Coffee Shop	Fish Market	Café	Brewery	Hotel	Restaurant
1	Berhampore, New Zealand	Coffee Shop	Fish & Chips Shop	Ramen Restaurant	Zoo	Dive Bar	Italian Restaurant	Bus Station	Park	Gourmet Shop	Thai Restaurant
2	Breaker Bay	Scenic Lookout	Nudist Beach	Bed & Breakfast	Beach	Dive Bar	Electronics Store	Emergency Room	Exhibit	Farmers Market	Fast Food Restaurant
3	Brooklyn, New Zealand	Deli / Bodega	Indie Movie Theater	Café	Pie Shop	Fast Food Restaurant	Gastropub	Burger Joint	Video Store	Convenience Store	Electronics Store
4	Churton Park	Garden Center	Gym	Café	Zoo Exhibit	Fish Market	Gastropub	Garden	Furniture / Home Store	French Restaurant	Fish & Chips Shop

Table 4. Each suburb along with the top 10 most common venues

4. Methods

K-means clustering is one of the simplest and popular unsupervised machine learning algorithms and this clustering method was chosen in this case to find an optimal suburb for opening a new coffee shop in Wellington.

For this clustering, population of each suburb is also one of the key numbers and should be involved. Therefore, I calculated normalized numbers of population in each suburb and these numbers were also included in the calculation.

It is essential to find out which k is the most promising for the given data in K-means clustering. I applied the elbow method to find a promising k which will be used in clustering shown later. Figure 3 shows the number of clusters and its distortion respectively. From this figure, I set 2 as an optimal k in the clustering.

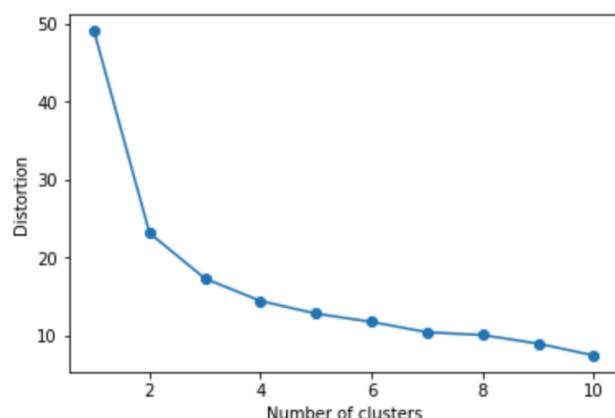


Fig. 3. Distortion and number of clusters

5. Results

5.1 Visualization of the clustering

As you can see on Figure 4., this is the map of Wellington with clustered suburbs superimposed on top. Cluster 0 is illustrated in red and cluster 1 is shown in purple, respectively.

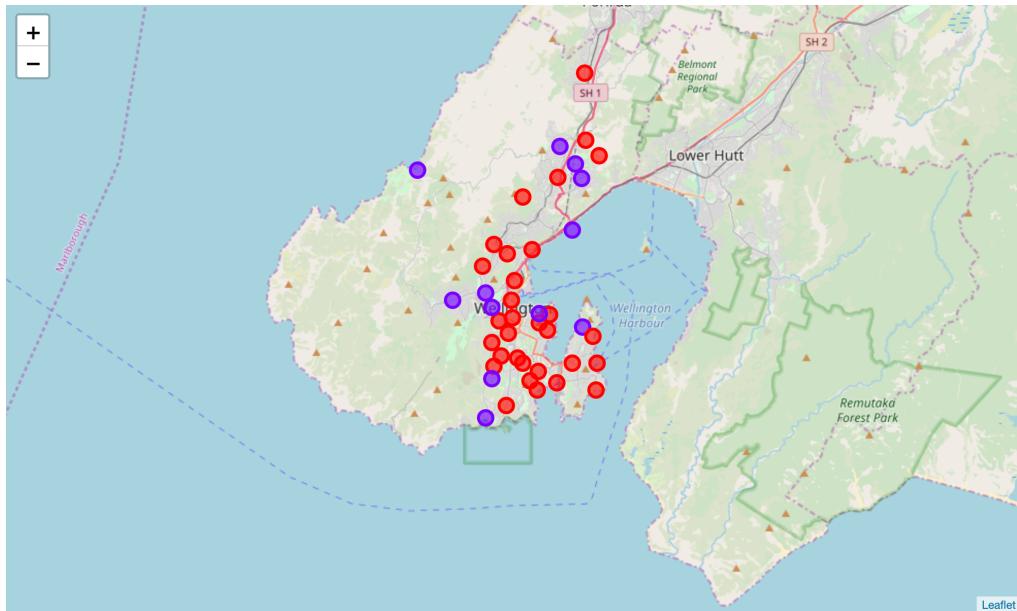


Fig. 4. Map of Wellington with the clustered suburbs superimposed on top

Examples of clustering 0 and clustering 1 are shown in the table 5 and table 6 as shown below. From these tables, it is obvious that population does have a big impact on this clustering.

	Suburb	Latitude	Longitude	Population	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
35	Tawa, New Zealand	-41.171567	174.824786	2.859002	Bakery	Café	Sandwich Place	Supermarket	Pharmacy	Photography Lab	Train Station	Fish & Chips Shop
36	Te Aro	-41.294000	174.777000	2.163839	Coffee Shop	Restaurant	Café	Vietnamese Restaurant	Bar	Burger Joint	Pizza Place	Asian Restaurant
9	Johnsonville, New Zealand	-41.223540	174.807240	1.859226	Supermarket	Fast Food Restaurant	Pizza Place	Bus Stop	Burger Joint	Bookstore	Coffee Shop	Hotel
21	Miramar, New Zealand	-41.316667	174.816667	1.478885	Bakery	Gastropub	Pizza Place	Thai Restaurant	Café	Sandwich Place	Fish & Chips Shop	Fish Market
14	Khandallah	-41.233333	174.783333	1.359763	Scenic Lookout	Mountain	Zoo Exhibit	Electronics Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
26	Newtown, New Zealand	-41.313700	174.779800	1.342745	Bus Station	Coffee Shop	Bakery	Bar	Mexican Restaurant	Bus Line	Café	Fish & Chips Shop
8	Island Bay, New Zealand	-41.337730	174.772310	0.901142	Café	Supermarket	Fast Food Restaurant	Video Store	Pizza Place	Movie Theater	Pie Shop	Beer Garden
15	Kilbirnie, New Zealand	-41.320400	174.794100	0.778616	Supermarket	Grocery Store	Bookstore	Paper / Office Supplies Store	Fast Food Restaurant	Bakery	Café	Zoo Exhibit

Table 5. Data frame of cluster 0

	Suburb	Latitude	Longitude	Population	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
12	Karori	-41.284855	174.736795	3.134686	Supermarket	Kebab Restaurant	Bakery	Pool	Deli / Bodega	Sandwich Place	Fish & Chips Shop	Multiplex	Pub
25	Newlands, New Zealand	-41.224000	174.823000	1.019413	Indian Restaurant	Grocery Store	Fast Food Restaurant	Bus Line	Coffee Shop	Pizza Place	Park	Supermarket	Pharmacy
4	Churton Park	-41.208355	174.808190	0.617801	Cosmetics Shop	Café	Garden Center	Italian Restaurant	Emergency Room	Fish Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market
13	Kelburn, New Zealand	-41.288771	174.763298	-0.019503	Café	Pub	Bus Stop	Indian Restaurant	Pizza Place	Malay Restaurant	Bakery	Coffee Shop	Fish Chip Shop
28	Northland, Wellington	-41.281389	174.758889	-0.092679	Pizza Place	Convenience Store	Burger Joint	Tunnel	Café	Shopping Mall	Zoo Exhibit	Electronics Store	Fast Food Restaurant
31	Paparangi	-41.216643	174.818866	-0.239029	Photography Studio	Other Repair Shop	Grocery Store	Zoo Exhibit	Fish Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Exhibit
16	Kingston, Wellington	-41.324318	174.763107	-0.418563	Nature Preserve	Convenience Store	Mini Golf	Bus Station	Fast Food Restaurant	Soccer Field	Beach	Trail	Electronic Store
30	Owhiro Bay	-41.344000	174.759000	-0.550449	Country Dance Club	Neighborhood	Shoe Store	Tree	Electronics Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market

Table 6. Data frame of cluster 1

5.2 The top 5 suburbs for opening a coffee shop

Since my goal is to find out which suburb has the most inhabitants while having less competitors at the same time, I eliminated suburbs that have coffee shops and cafes in the 1st, 2nd, and 3rd most common Venue. Table 7 shows the best 5 suburbs to open a new coffee shop from my research after the process. Figure 5 shows the map of Wellington with the top 5 suburbs superimposed on top.

	Suburb	Latitude	Longitude	Population	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	Karori	-41.284855	174.736795	3.134686	Supermarket	Kebab Restaurant	Bakery	Pool	Deli / Bodega	Sandwich Place	Fish & Chips Shop	Multiplex	Pub
1	Johnsonville, New Zealand	-41.223540	174.807240	1.859226	Supermarket	Fast Food Restaurant	Pizza Place	Bus Stop	Burger Joint	Bookstore	Coffee Shop	Hotel	Gym
2	Miramar, New Zealand	-41.316667	174.816667	1.478885	Bakery	Gastropub	Pizza Place	Thai Restaurant	Café	Sandwich Place	Fish & Chips Shop	Fish Market	French Restaurant
3	Khandallah	-41.233333	174.783333	1.359763	Scenic Lookout	Mountain	Zoo Exhibit	Electronics Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Exhibit
4	Newlands, New Zealand	-41.224000	174.823000	1.019413	Indian Restaurant	Grocery Store	Fast Food Restaurant	Bus Line	Coffee Shop	Pizza Place	Park	Supermarket	Pharmacy

Table 7. Top 5 suburbs

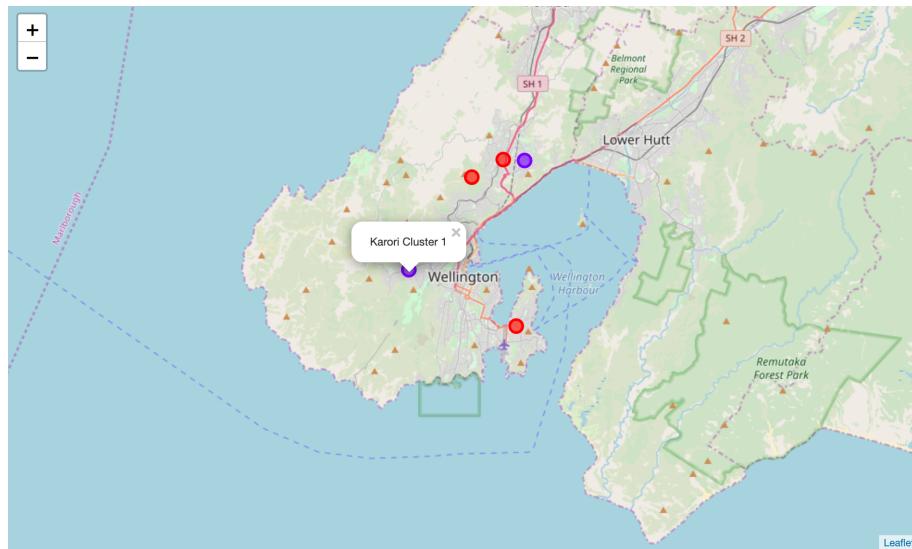


Fig 5. Map of Wellington with the top 5 suburbs superimposed on top

5.3 Number of population and venues in the top 5 suburbs

Table 8 and Figure 6 show the number of population and venues in the top 5 suburbs in Wellington. From these numbers, it is found out that Karori is the biggest city by population and has relatively low venues.

If I would open a coffee shop in Karori, although it would be of course depending on the quality of coffee that I offer, the probability that my coffee shop is successful could be high. Johnsonville and Miramar are the 2 suburbs that have both high population and numbers of venues at the same time. I assume that these 2 suburbs are prospering and have a lot of potential, too. In addition, there are less competitors in the suburbs, I could also recommend myself to open a coffee shop in the suburbs, for example.

	Suburb	Population	Venue
4	Karori	14736	11
0	Johnsonville, New Zealand	10239	24
1	Miramar, New Zealand	8898	26
2	Khandallah	8478	2
3	Newlands, New Zealand	7278	10

Table 8. Number of population and venues in the top 5 suburbs

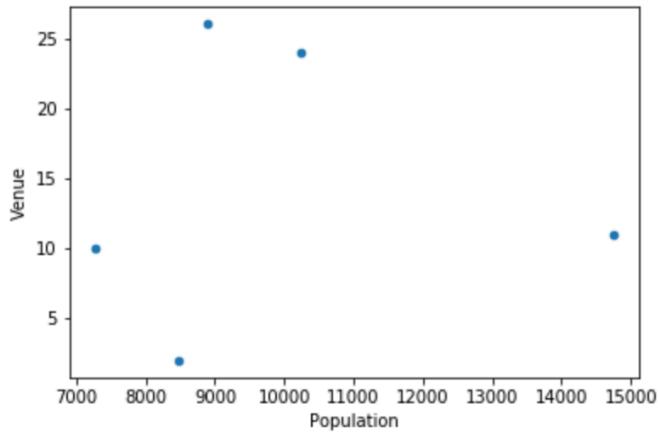


Fig 6. Number of population and venues in the top 5 suburbs

6. Discussion

In this report, I basically figured out the best suburb and other top 4 suburbs in Wellington that could be suitable to open a new coffee shop based on the data of population and venues in each suburb. It would be possible for you to grasp a correlation between the 2 clusters from Figure 4, but this clustering might not be necessary in this simple case.

For further analysis, there are some information that I would like to involve such as the weather information, number of tourists in each suburb, price of the land and so on. Then this analysis could be more interesting and reliable.

7. Conclusion

In case that I would go to Wellington and have to open a coffee shop in the near future due to my addiction to coffee, I would recommend myself to open one in these 5 suburbs: Karori, Johnsonville, Miramar, Khandallah, and Newslands. These suburbs have relatively many inhabitants with less coffee shops and cafes. Therefore, it could be possible that my coffee shop would be unique to the inhabitants since there are not much competitors there, although there are other aspects to be considered such as the quality of coffee, of course.

As you can see on Figure 5, Karori is a bit far away from the central of Wellington and surrounded by mountains, nature. So it is expected that people or tourists visit the area for hiking or mountain trips and my future coffee shop could be exclusively unique to those explorers.

As for further analysis, as mentioned in the discussion section, I would need some more information such as weather information, number of tourists, price of land etc. With these data this research

could be more realistic and meaningful. I would like to finish my report at this point and am wondering which name I give a coffee shop that I might open in the future. Maybe “Sai Sai Coffee”.

References

(*1) Wikipedia on Willington City: https://en.wikipedia.org/wiki/Category:Suburbs_of_Wellington_City

(*2) My GitHub Page: [https://github.com/Naoto31/Coursera_Capstone/blob/master/The%20Battle%20of%20Neighborhoods%20\(Final\).ipynb](https://github.com/Naoto31/Coursera_Capstone/blob/master/The%20Battle%20of%20Neighborhoods%20(Final).ipynb)