IBM DATA SCIENCE CAPSTONE PROJECT

AUCKLAND BOOKSTORE

By: Rodrigo Castro
rodrigo20acc@gmail.com

June 2020

Introduction

The objective of this Data Science project is to determine an optimal location for a new bookstore in the city of Auckland, New Zealand.

Background

Even if nowadays the e-commerce of electronic books and books sent by delivery is showing a remarkable growth, it's still very important for booklovers to find a place to take their time and interact with the objects of their affection. Bookstores are appealing places in which readers can have unexpected encounters and be welcomed with enjoyable surprises.

Business Understanding

Auckland, which is the most important metropolitan city in New Zealand and is home to its largest population, provides an attractive environment in which to open a new bookstore. Thus, is reasonable to select this city as the subject of study for this project.

Business Understanding

What is curious about bookstores is that they can benefit from sharing the same environment with their competitors as they complement each other. In different bookstores we can find the same title not only at a different price, but also with a different editorial work behind it. By consequence, having bookstores located near each other enrich the experience of the costumer and ensure that all of them are eventually visited. In this project, we will focus only in the areas where there already are other bookstores.

Data

Neighbourhoods

- Name of each neighbourhood: https://en.wikipedia.org/wiki/Category:Suburbs_of_Auckland

Latitude and Longitude Coordinates

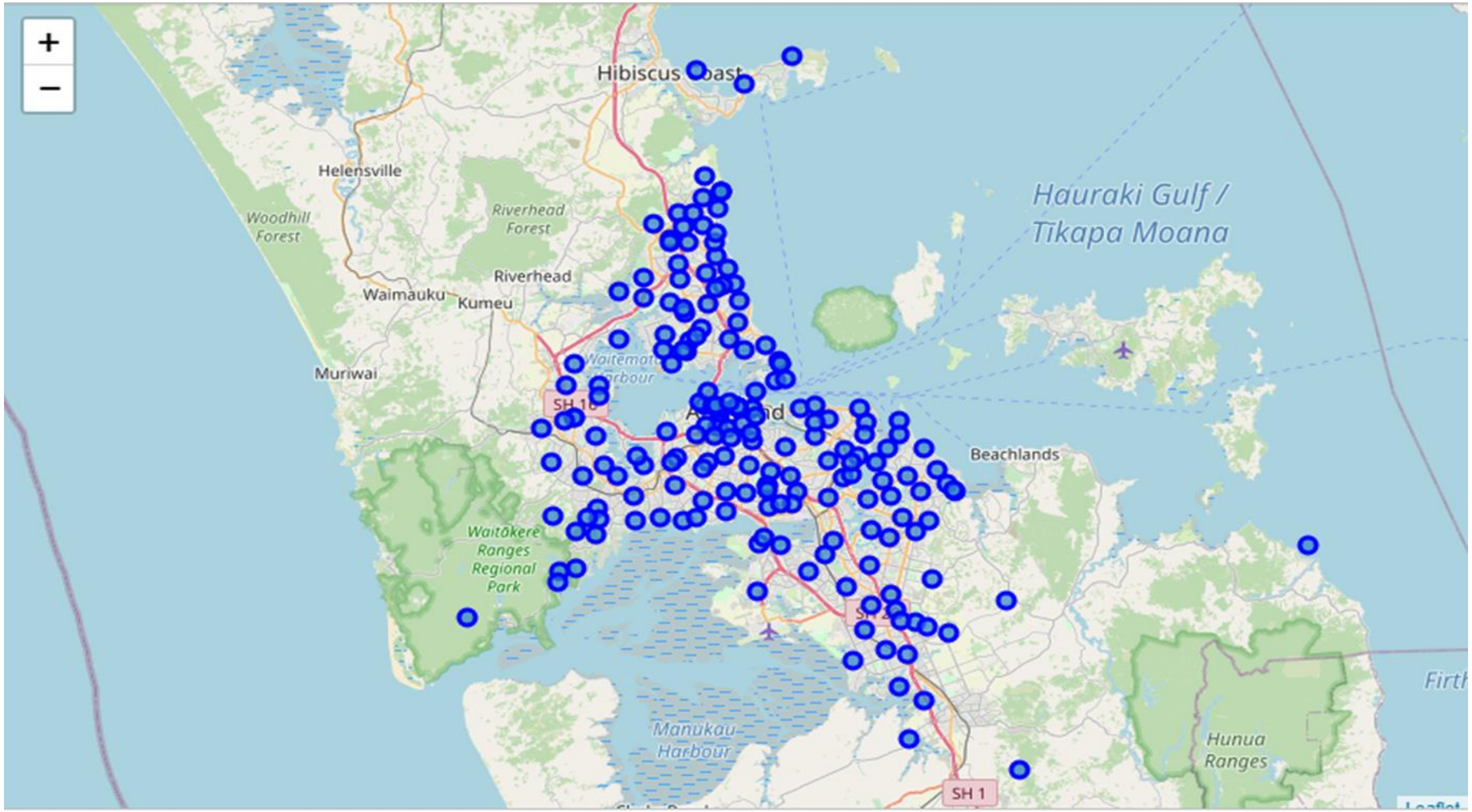
- Geocoder Package of Python

Venue Data

- Characteristics of venues: Foursquare API

Data Visualization

Auckland's map along with its neighbourhoods and marked coordinates.



Foursquare API

The location data was extracted from the Foursquare server considering the following parameters:

- Radius = 3 km
- Max. number of venues = 100

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Airport Oaks	-36.72916	174.70893	K-Mart	-36.728644	174.709722	Department Store
1	Airport Oaks	-36.72916	174.70893	QBE Stadium	-36.726937	174.702059	Stadium
2	Airport Oaks	-36.72916	174.70893	The Merchant Bar & Kitchen	-36.727810	174.709088	Bar
3	Airport Oaks	-36.72916	174.70893	Event Cinemas	-36.728335	174.708643	Movie Theater
4	Airport Oaks	-36.72916	174.70893	Albany Mega Centre	-36.731220	174.706719	Shopping Mall

One-hot Encoding

Is the process of transforming the categorical variables into numeric features. The unique items in the Venue Category column of the dataframe were one-hot encoded. Afterwards, the new dataframe was filtered considering only the neighbourhoods that contain bookstores, since they are the only ones that are of our interest.

	Neighbourhood	African Restaurant	Airport	Airport Lounge	Airport Service	American Restaurant	Aquarium	Arcade	Argentinian Restaurant	Art Gallery	...	Vietnamese Restaurant
9	Avondale, Auckland	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	...	0.000000
19	Brookby	0.0	0.0	0.0	0.0	0.010000	0.000000	0.000000	0.0	0.000000	...	0.000000
25	Chatswood, New Zealand	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	...	0.000000
30	Cockle Bay, New Zealand	0.0	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.000000	...	0.000000
33	Dannemora, New Zealand	0.0	0.0	0.0	0.0	0.000000	0.012195	0.000000	0.0	0.000000	...	0.000000
34	Devonport, New Zealand	0.0	0.0	0.0	0.0	0.000000	0.024390	0.000000	0.0	0.000000	...	0.000000
40	Ellerslie, New Zealand	0.0	0.0	0.0	0.0	0.010000	0.000000	0.000000	0.0	0.010000	...	0.010000

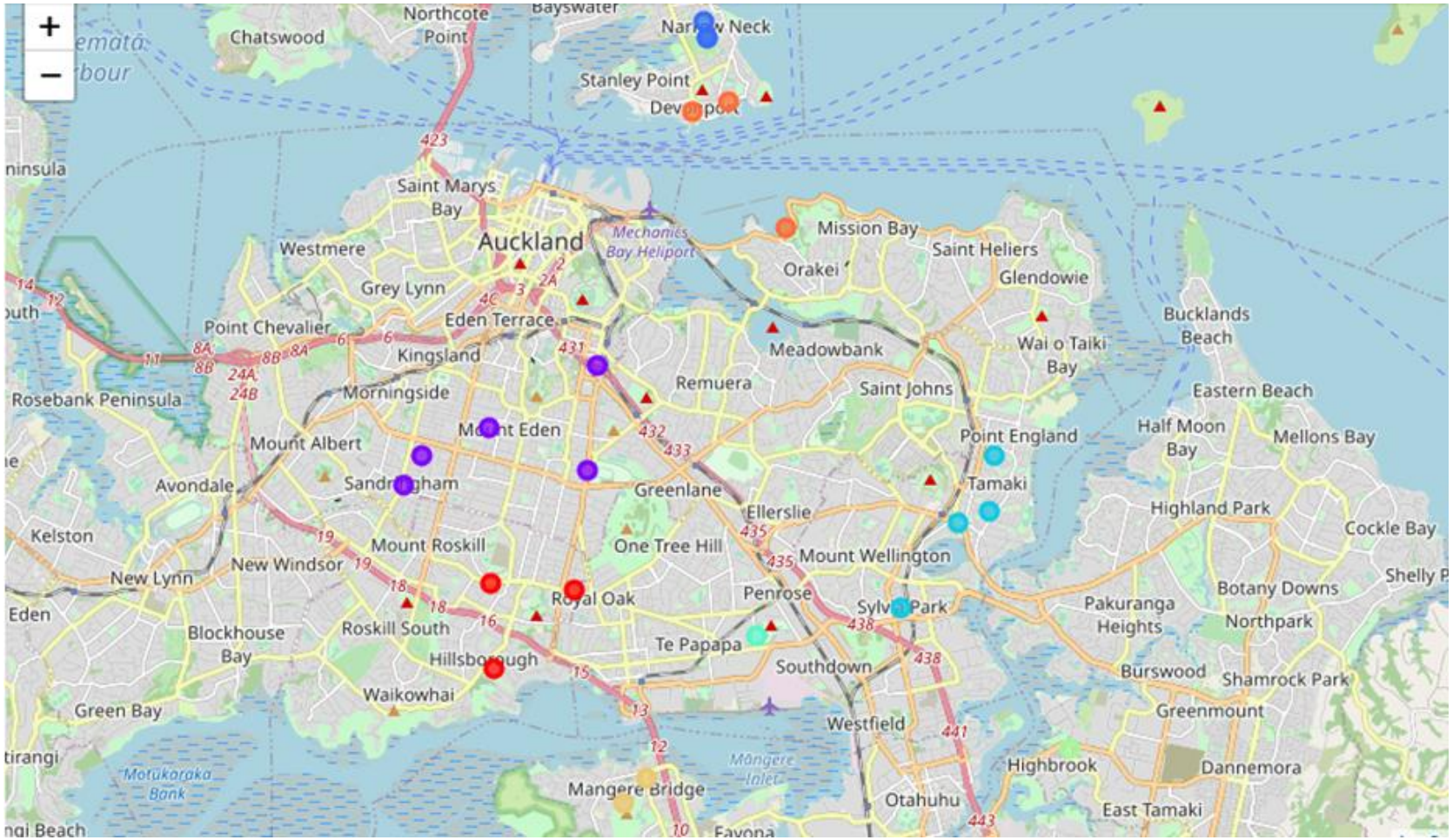
Top 10 Venues

For further understanding, for each of the selected neighbourhoods the 10 most popular venues were retrieved.

	Neighbourhood	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
9	Avondale, Auckland	Café	Chinese Restaurant	Bar	Coffee Shop	Bakery	Gym	Park	Indian Restaurant	Burger Joint	Asian Restaurant
19	Brookby	Café	Fast Food Restaurant	Beach	Coffee Shop	Neighborhood	Burger Joint	Japanese Restaurant	Supermarket	Indian Restaurant	Bar
25	Chatswood, New Zealand	Café	Scenic Lookout	Beach	Bar	History Museum	Bakery	Performing Arts Venue	Comfort Food Restaurant	Candy Store	Burger Joint
30	Cockle Bay, New Zealand	Café	Park	Clothing Store	Grocery Store	Fast Food Restaurant	Farm	Rental Car Location	Juice Bar	Sporting Goods Shop	Motel
33	Dannemora, New Zealand	Café	Hotel	Bar	Beach	Harbor / Marina	Restaurant	Coffee Shop	Pizza Place	Turkish Restaurant	Japanese Restaurant

Clustering

In this case, the K-mean algorithm was used since it is computationally faster when working with huge datasets. By running different clustering to the dataset it was concluded that 8 is the ideal number of clusters.



Results

We obtained the 8 desired clusters of neighbourhoods. A suitable name was given to each of them in a way that it summarizes their characteristics.

- Cluster 1: Diverse Restaurants and Parks
- Cluster 2: Asian Cuisine Restaurants and Bars
- Cluster 3: Scenic Lookout and Cultural Places
- Cluster 4: Fast Food Restaurants and Diverse Stores
- Cluster 5: Fast Food Restaurant Mostly
- Cluster 6: Asian Cuisine Restaurants Mostly
- Cluster 7: Parks and Diverse Stores
- Cluster 8: Beach and Harbor

Discussion

After conducting the K-mean clustering in Auckland's neighbourhoods and analysing each cluster, we can observe that Cluster 3 seems to be the most appropriate. The reason being that these neighbourhoods seem to be located in an attractive place to the kind of costumers who frequent bookstores given that it contains places of cultural interest and scenic lookouts where they may read. These elements enrich the experience of any booklover willing to walk through an entire neighbourhood visiting different bookstores since they complement each other.

	Neighbourhood	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
117	Murrays Bay	Café	Bar	Beach	Scenic Lookout	Movie Theater	Fruit & Vegetable Store	Sushi Restaurant	Supermarket	Bookstore	History Museum
26	Chatswood, New Zealand	Café	Scenic Lookout	Beach	Bar	History Museum	Bakery	Performing Arts Venue	Comfort Food Restaurant	Candy Store	Burger Joint
192	Unsworth Heights	Café	Scenic Lookout	Beach	Bar	History Museum	Bakery	Performing Arts Venue	Comfort Food Restaurant	Candy Store	Burger Joint

Cluster 3: Scenic Lookout and Cultural Places

Conclusions

- The task of identifying an optimal location to open a bookstore in the city of Auckland has required that we used multiple Data Science Skills. From data collection and data preparation to modeling and evaluation, we went through some of the Data Science Methodology steps.

- This project limits the scope of its analysis by deciding an optimal location based solely on the venues of a neighbourhood. Other factors such as socioeconomic statistics, land price or supply chain analysis, just to name a few, could have been used to produce a more in-depth investigation. Nevertheless, the findings of this study represent a significant step towards investigating the bookstore business in the city of Auckland.

рахмат
danke
謝謝
ngiyabonga
Баярлалаа
spas
teşekkür ederim
mersi
kia ora
barka
welalin
tack
mahalo
tapadh leat
спасибо
faafetai lava
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
thank you
gracias
хвала
kiitos
dankie
nanni
nandri
dhanyavad
asante
manana
maururu
koszonom
dankon
aciu
obrigada
tenki
bedankt
hvala
gracie
bayarlalaa
enkosi
djiere dieuf
tau
mochchakkeram
murakoze
akun
дякую
mamnun
dziękuję
chnorakaloutioun
gratias ago
gracies
sulpay
taiku
go raibh maith agat
chokrane
sobodi
dėkuji
sagolun
sukriya
kop khun krap
arigatô
takk
dakujem
trugarez
obrigado
mèsi
didi madloba
ありがとう
najis tuke
tanemirt
grazie
kam sah hamnida
rahmat
terima kasih
rahmet
diolch
dhanyavadagalu
shukriya
merce
мерси
ευχαριστώ
xiexie
merci
감사합니다