'Coffee in Cape Town'

IBM Data Science Professional Certificate
Capstone Project

Laura Tissing
January 2020

Introduction

A client is looking for a location to open a Coffee Shop in one of the suburbs of Cape Town.

This presentation covers the first phase of this project.

Initial wishes are: *to find a location*

- ✔ *with a balanced mixture of citizens and tourists*
- ✔ *within reasonable distance from the city centre*
- ✔ *with as little competition as possible*

Factors included in this *first* phase of the study:

- Location of competitors
- Surrounding venues, businesses and attractions
- Distance to city centre
- Attractriveness to national/international tourists and citizens

Data

The following data is used to address the task description:

– Website on South Africa's postal codes

– Geographical coordinates from Geolocator and Google Maps

– Foursquare API with a variety of data on venues

– Wikipedia

– Websites on suburbs of Cape Town

Methodology

Part 1 – Data collection and visualization

Results in

– a visualization of the different Street Codes of Cape Town (Street Codes contain one or more suburb(s))

|	STR-CODE	SUBURB	Latitude	Longitude
0	7100	EERSTE RIVER, EERSTERIVIER	-33.987809	18.667149
1	7405	MAITLAND, PAARDEN ISLAND, PAARDENEILAND, PINEL...	-33.926604	18.498313
2	7441	BOTHASIG, EDGEMEAD, MELKBOS, MELKBOSSTRAND, MI...	-33.844994	18.522595
3	7455	LANGA, LANGA ZONE 1, LANGA SONE 1, LANGA ZONE ...	-33.944727	18.532317
4	7460	MONTE VISTA	-33.908873	18.550382
5	7490	MATROOSFONTEIN	-33.955570	18.587988
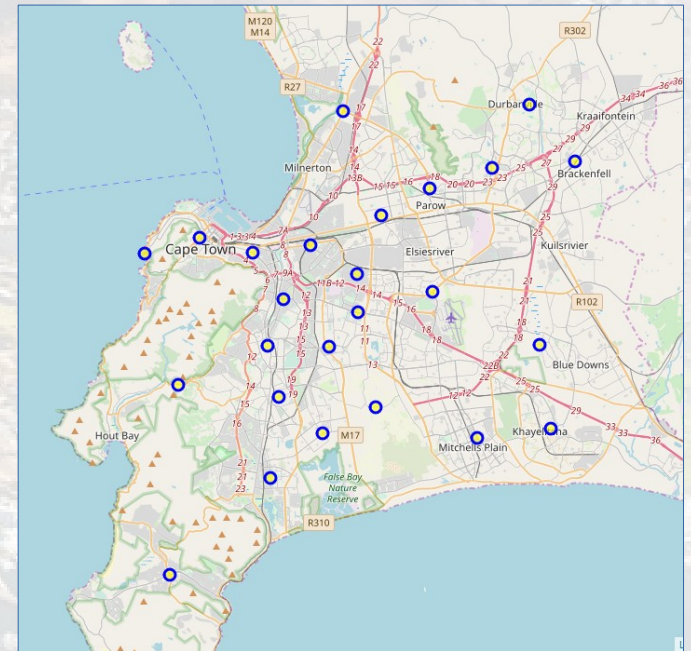
Methodology

Part 2 – Using Foursquare API to retrieve venue information

Results in:

– specific venue information on every Street Code

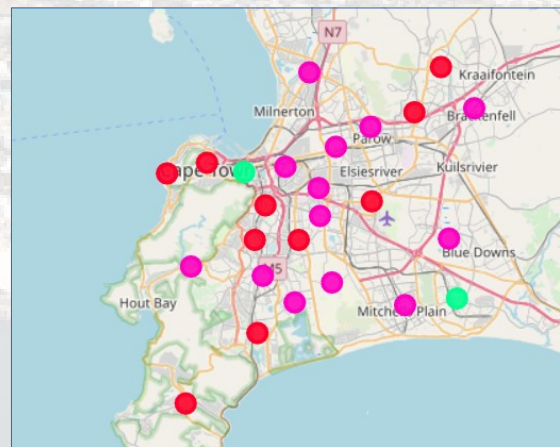|	Street Code	Street Code lat	Street Code lng	Venue	Venue lat	Venue lng	Category
0	7100	-33.987809	18.667149	Airport Mall	-33.994398	18.668121	Arts & Crafts Store
1	7100	-33.987809	18.667149	Engen Hindle Road Service Station	-33.979858	18.657894	Gas Station
2	7100	-33.987809	18.667149	Shoprite	-33.986895	18.680509	Supermarket
3	7100	-33.987809	18.667149	Allrich Trading (Pty) Ltd	-33.984661	18.654280	Print Shop
4	7100	-33.987809	18.667149	Wimpy	-33.979030	18.651890	Burger Joint
5	7100	-33.987809	18.667149	Shoprite Liquor Shop	-33.995441	18.650288	Liquor Store
6	7405	-33.926604	18.498313	Merrypak & Print	-33.928847	18.500087	Arts & Crafts Store
7	7405	-33.926604	18.498313	Magica Roma	-33.940149	18.497787	Italian Restaurant

– total number of venues in the Street Codes
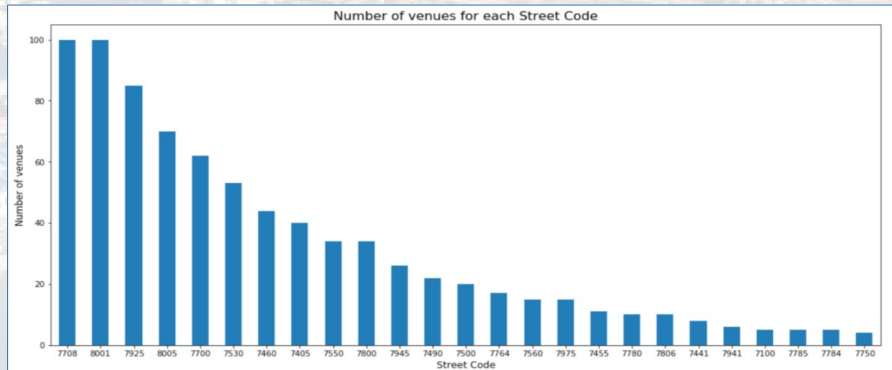
Methodology

Part 3 – One-hot encoding & K-means clustering

Results in:

- A dataframe with the mean frequency occurence of Coffee Shops in each Street Code

	Street Code	Coffee Shop
0	7100	0.000000
1	7405	0.000000
2	7441	0.000000
3	7455	0.000000
4	7460	0.000000
5	7490	0.083333
6	7500	0.000000
7	7530	0.074074

- Visualization of the clustering on occurence of Coffee Shops in each Street Code

Methodology

Part 4 – Recommendation

Results in:

- Target Street Codes that are
  - in Cluster 0 (with no or little competition)
  - located within a 15km radius of the city centre

	Street Code	Cluster Labels	Coffee Shop	Latitude	Longitude	Distance(km)
0	7405	0	0.0	-33.926604	18.498313	7.542683
1	7441	0	0.0	-33.844994	18.522595	13.030715
2	7455	0	0.0	-33.944727	18.532317	10.949512
3	7460	0	0.0	-33.908873	18.550382	12.427999
4	7764	0	0.0	-33.968128	18.533450	11.906517
5	7800	0	0.0	-34.019579	18.474911	12.060983
6	7806	0	0.0	-34.012013	18.400462	10.072664
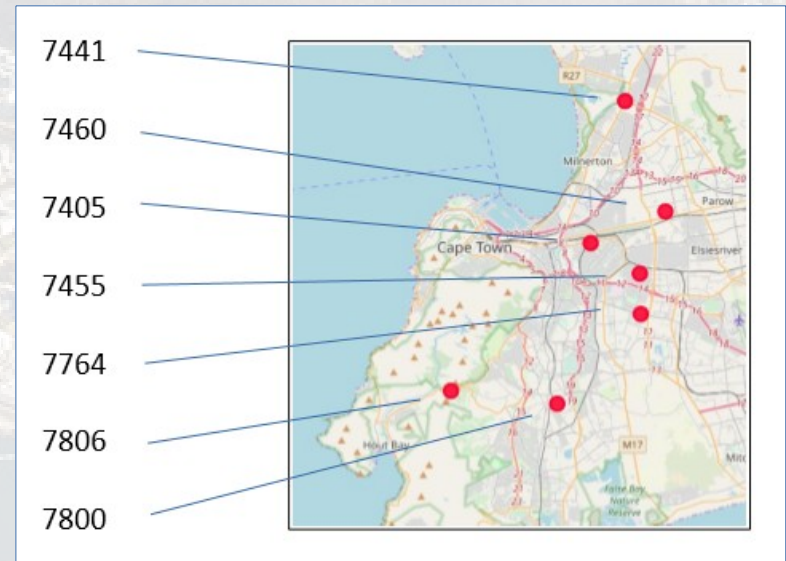
A dataframe with the 10 most common venues expands insight in the target Street Codes:

THE FINAL DATAFRAME WITH STREET CODES THAT ARE SUITABLE FOR RECOMMENDATION (TARGET STREET CODES):

	Street Code	Latitude	Longitude	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	7405	-33.926604	18.498313	Furniture / Home Store	Grocery Store	Gym	Train Station	Burger Joint	Café	Fast Food Restaurant	Seafood Restaurant	Pharmacy	Playground
1	7441	-33.844994	18.522595	Pizza Place	Gay Bar	Shopping Mall	Convenience Store	Stadium	Garden Center	Chinese Restaurant	Fast Food Restaurant	Dance Studio	Warehouse Store
2	7455	-33.944727	18.532317	Convenience Store	Gas Station	Seafood Restaurant	Fast Food Restaurant	Farmers Market	Steakhouse	Business Service	Shopping Mall	Pizza Place	Design Studio
3	7460	-33.908873	18.550382	Fast Food Restaurant	Portuguese Restaurant	Gas Station	Steakhouse	Casino	Bowling Alley	Diner	Hotel	Seafood Restaurant	Skating Rink
4	7764	-33.968128	18.533450	Fast Food Restaurant	Pizza Place	Nightclub	Convenience Store	Burger Joint	Bakery	Stadium	Steakhouse	Indian Restaurant	Seafood Restaurant
5	7800	-34.019579	18.474911	Grocery Store	Fast Food Restaurant	Golf Course	Pharmacy	Pizza Place	Café	Thai Restaurant	Bakery	Performing Arts Venue	Portuguese Restaurant
6	7806	-34.012013	18.400462	Vineyard	Café	Flower Shop	Playground	Trail	Restaurant	French Restaurant	Tapas Restaurant	Wine Bar	Event Space

...as do WordClouds of *every 1st suburb* in the target Street Codes; they reveal interesting information

Conclusion & Discussion

- Most of the target Street Codes seem to have good opportunities to start a business like a Coffee Shop

- especially Street Codes 7441, 7800 and 7806 are showing potential in terms of characteristics and types of venues that can be found there

- Further research is essential to be able to recommend more specific locations

- It might be interesting to see if the outcome returns even better options if the Street Codes in Cluster 1 (with moderate competition) are included in the research

Examples of factors that need to be studied in the next phase:

Rental prices/m² prices

Demographic structures

Accessability to and from city centre