

The battle of Neighborhoods

Cape Town, South Africa – Sushi Restaurants



Introduction

- ❖ When looking to open a restaurant in any city, it is very important to do research on the cities or neighborhoods. This gives you a good understanding of what kind of restaurants are in the area, which ones are doing well and why they are/ are not doing so well. The amount of restaurants, different cuisines and ratings of the restaurants are good indicators of how well a new one will fair in its chosen location.
- ❖ For the purpose of this project, we will assume that a person is looking to open a Sushi restaurant in the city of Cape Town, South Africa. However, Sushi is a very popular food in Cape Town and there is no shortage of Sushi restaurants.
- ❖ The person opening the restaurant wants to know:
 - How many Sushi restaurants are in each neighborhood
 - How many Sushi restaurants are in Cape Town in general
 - Which neighborhoods there are Sushi restaurants.
 - Which neighborhood is the best one to open their restaurant in.

Data description

The data I will be using will be as follows:

- ❖ **The Cape Town postal codes:** The table containing all of the postal codes and neighborhood names for the Cape Town area, which are found on the South African Postal Codes website (<http://sapostalcodes.info/queryPostal/Cape+Town>).
- ❖ **The Foursquare API data:** Obtaining the location data of each neighborhood, such as: venue names, geospatial data and categories.

Data description - Continued

❖ **Python Libraries/packages:** In order to create the machine learning algorithms, generate maps reflecting the data, encode the data, normalize the data and display the results.

❖ **The Python Libraries/packages:**

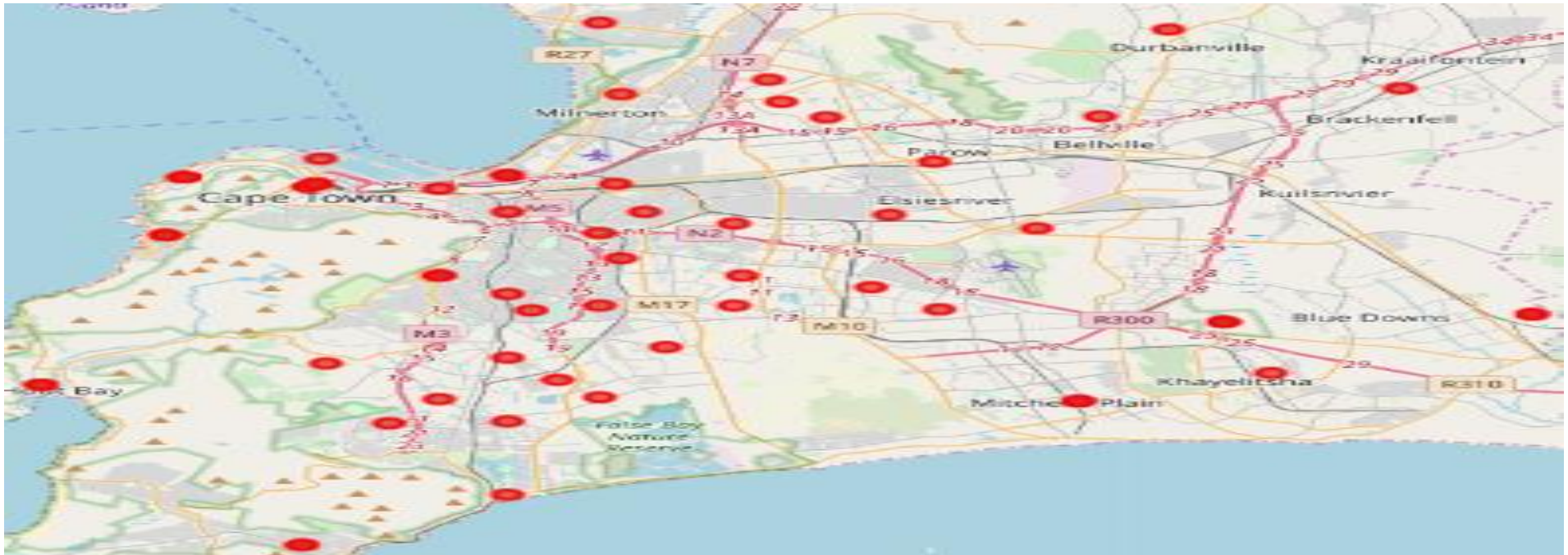
- ❖ Pandas
- ❖ Numpy
- ❖ Matplotlib
- ❖ JSON
- ❖ Sklearn
- ❖ Folium
- ❖ Requests
- ❖ Geopy

How the Data will be used to solve the problem

- ❖ Using the Cape Town postal codes as well as the Foursquare API data, create data frames in order to normalize the data.
- ❖ These data frames will contain the geospatial data and venue data.
- ❖ The data will be processed, encoded using “One Hot” encoding, run through a K-means clustering algorithm and mapped out visually.
- ❖ The results will be used to determine which neighborhood will be the best suited to open the new restaurant in.

Methodology: Identifying the Neighborhoods

- ❖ Identify the neighborhoods in Cape Town, we will use the Postal Codes and neighborhoods obtained using URL requests to (<http://sapostalcodes.info/queryPostal/Cape+Town>).
- ❖ Data scraped and read into Pandas dataframes,
- ❖ A URL request will be made to the IBM Watson data assets storage point to obtain the CSV file containing the coordinates for the postal codes.
- ❖ This is read into dataframes.
- ❖ Dataframes will be merged to form one dataframe containing the Neighborhood names, coordinates, postal codes and city name.
- ❖ Using the Folium package, this will be visually displayed on a map.



Methodology: Retrieving Location Data

- ❖ To retrieve location data for each of the neighborhoods in Cape Town, a request to the Foursquare API server will be made.
- ❖ After obtaining the data for the neighborhoods, request the Foursquare API server to return information about the Sushi venues in Cape Town.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	ATHLONE	-33.965260	18.501795	Hongs Sushi	-33.966223	18.492854	Asian Restaurant
1	ATHLONE	-33.965260	18.501795	Edo Sushi	-33.977238	18.493823	Sushi Restaurant
2	BELLMILLE	-33.881273	18.626469	Sushi Suki	-33.876608	18.628815	Asian Restaurant
3	BELLMILLE	-33.881273	18.626469	Yuz Sushi	-33.880080	18.635716	Sushi Restaurant
4	BELLMILLE	-33.881273	18.626469	Jasons Sushi	-33.873840	18.619392	Sushi Restaurant

Methodology: Processing the retrieved data

- ❖ Once the JSON file is retrieved from the Foursquare API server, read the results into dataframes.
- ❖ From there, normalize the data and then the data will be grouped by the Neighborhoods column and a count will be performed.
- ❖ Perform “One Hot” encoding on the resulting dataframe and re-group by neighborhoods.

	Neighborhood	Asian Restaurant	Chinese Restaurant	Dim Sum Restaurant	Food	Food & Drink Shop	Italian Restaurant	Japanese Restaurant	Korean Restaurant	Nightclub	Seafood Restaurant	Sushi Restaurant	Thai Restaurant
0	ATHLONE	0.500000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.500000	0.00
1	BELLEVILLE	0.200000	0.200000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.600000	0.00
2	BERGVLiet	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.750000	0.25
3	BO-KAAP	0.066667	0.000000	0.000000	0.0	0.066667	0.033333	0.000000	0.000000	0.00	0.066667	0.766667	0.00
4	BOTHASIG	0.500000	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.500000	0.00

Methodology – Unsupervised Machine Learning

- ❖ Select the top 10 Sushi restaurants from each neighborhood for statistical analysis.
- ❖ Perform K-Means Clustering on the restaurant data and merge the dataframes in order to solve the problem.
- ❖ K-means Clustering selected due to large amounts of unlabeled data.
- ❖ Using Folium, the cluster will be visually displayed on a map.
- ❖ The data clusters will also be examined in order to solve the problem.

Cluster 1

```
kmc_merged.loc[kmc_merged['Cluster Labels'] == 0, kmc_merged.columns[[1] + list(range(5, kmc_merged.shape[1]))]]
```

99	-33.921116	18.421201	Restaurant	Restaurant	Restaurant	Drink Shop	Restaurant	Restaurant	Restaurant	Nightclub	Restaurant	Restaurant	Food	0.0
100	-33.921118	18.412314	Asian Restaurant	Sushi Restaurant	Seafood Restaurant	Food & Drink Shop	Asian Restaurant	Italian Restaurant	Thai Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Food	0.0
128	-33.986004	18.472452	Sushi Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0
			Sushi	Sushi	Japanese	Dim Sum	Asian	Thai	Seafood		Korean	Italian	Food & Drink	

Results

- ❖ We can see that there are 310 Sushi restaurants in Cape Town.
- ❖ We can also see that there are 161 unique venues in 12 categories serving Sushi. The amount of venues per neighborhood can also be seen in the table below.
- ❖ The Neighborhoods where the restaurants are situated also listed in the below table.
- ❖ The table below is extracted from the dataframe “sushi_venues”.

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
ATHLONE	2	2	2	2	2	2
BELLVILLE	5	5	5	5	5	5
BERGVLIET	4	4	4	4	4	4
BO-KAAP	90	90	90	90	90	90
BOTHASIG	2	2	2	2	2	2
CAMPS BAY	2	2	2	2	2	2
CLAREMONT	9	9	9	9	9	9
CONSTANTIA	1	1	1	1	1	1
DURBANVILLE	4	4	4	4	4	4
EDGEMEAD	1	1	1	1	1	1
FISH HOEK	2	2	2	2	2	2
HOUT BAY	8	8	8	8	8	8
KENILWORTH	4	4	4	4	4	4

Results – Continued

LANSDOWNE	4	4	4	4	4	4
MILNERTON	2	2	2	2	2	2
MONTE VISTA	1	1	1	1	1	1
MOWBRAY	3	3	3	3	3	3
NEWLANDS	14	14	14	14	14	14
OBSERVATORY	7	7	7	7	7	7
OTTERY	1	1	1	1	1	1
PAROW	1	1	1	1	1	1
PLUMSTEAD	1	1	1	1	1	1
RETREAT	2	2	2	2	2	2
SEA POINT	18	18	18	18	18	18
SOUTHFIELD	1	1	1	1	1	1
TABLE VIEW	4	4	4	4	4	4
TOKAI	5	5	5	5	5	5
VICTORIA JUNCTION	90	90	90	90	90	90
WATERFRONT	17	17	17	17	17	17
WOODSTOCK	5	5	5	5	5	5

Results – Continued

❖ While the 1st most common restaurants serving Sushi were Sushi Restaurants, there were other common restaurants serving Sushi are as follows: Asian Restaurants, Thai Restaurants, Seafood Restaurants, Japanese Restaurants, Chinese Restaurants, etc.

	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	Cluster Labels
0	ATHLONE	Sushi Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	1
1	BELLVILLE	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	1
2	BERGVLIET	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	0
3	BO-KAAP	Sushi Restaurant	Seafood Restaurant	Food & Drink Shop	Asian Restaurant	Italian Restaurant	Thai Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Food	0
4	BOTHASIG	Sushi Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	1

Clusters – 1 & 2

Cluster 1

```
kmc_merged.loc[kmc_merged['Cluster Labels'] == 0, kmc_merged.columns[[1] + list(range(5, kmc_merged.shape[1]))]]
```

99	-33.921118	18.421201	Restaurant	Restaurant	Restaurant	Drink Shop	Restaurant	Restaurant	Restaurant	Nightclub	Restaurant	Restaurant	Food	0.0
100	-33.921118	18.412314	Asian Restaurant	Sushi Restaurant	Seafood Restaurant	Food & Drink Shop	Asian Restaurant	Italian Restaurant	Thai Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Food	0.0
128	-33.986004	18.472452	Sushi Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0
129	-33.986004	18.459243	Sushi Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0
130	-33.986004	18.463931	Sushi Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0
131	-33.986004	18.483823	Sushi Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0
132	-33.986004	18.488550	Sushi Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0
133	-33.986004	18.465246	Japanese Restaurant	Sushi Restaurant	Japanese Restaurant	Dim Sum Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Italian Restaurant	Food & Drink Shop	0.0

Cluster 2

```
kmc_merged.loc[kmc_merged['Cluster Labels'] == 1, kmc_merged.columns[[1] + list(range(5, kmc_merged.shape[1]))]]
```

	Neighborhood Latitude	Venue Longitude	Venue Category	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	Cluster Labels
0	-33.965260	18.492854	Asian Restaurant	Sushi Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	1.0
1	-33.965260	18.483823	Sushi Restaurant	Sushi Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	1.0
2	-33.881273	18.628815	Asian Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	1.0
3	-33.881273	18.635716	Sushi Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	1.0
4	-33.881273	18.619392	Sushi Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	1.0
5	-33.881273	18.633199	Sushi Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	1.0
6	-33.881273	18.633274	Chinese Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	1.0

Clusters – 3 & 4

Cluster 3

```
kmc_merged.loc[kmc_merged['Cluster Labels'] == 2, kmc_merged.columns[[1] + list(range(5, kmc_merged.shape[1]))]]
```

:

	Neighborhood Latitude	Venue Longitude	Venue Category	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	Cluster Labels
103	-33.951298	18.377194	Sushi Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	2.0
104	-33.951298	18.377194	Sushi Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	2.0
137	-34.026883	18.425341	Sushi Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	2.0
142	-33.872667	18.546940	Sushi Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	2.0
161	-33.868335	18.509757	Sushi Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	2.0
162	-33.868335	18.510070	Sushi Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	2.0

Cluster 4

```
kmc_merged.loc[kmc_merged['Cluster Labels'] == 3, kmc_merged.columns[[1] + list(range(5, kmc_merged.shape[1]))]]
```

:

	Neighborhood Latitude	Venue Longitude	Venue Category	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	Cluster Labels
143	-34.134088	18.430755	Food	Food	Thai Restaurant	Sushi Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Dim Sum Restaurant	3.0
144	-34.134088	18.430755	Food	Food	Thai Restaurant	Sushi Restaurant	Seafood Restaurant	Nightclub	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Dim Sum Restaurant	3.0

Clusters – 5

Cluster 5

```
kmc_merged.loc[kmc_merged['Cluster Labels'] == 4, kmc_merged.columns[[1] + list(range(5, kmc_merged.shape[1]))]]
```

:

	Neighborhood Latitude	Venue Longitude	Venue Category	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	Cluster Labels
188	-34.017221	18.506424	Nightclub	Nightclub	Thai Restaurant	Sushi Restaurant	Seafood Restaurant	Korean Restaurant	Japanese Restaurant	Italian Restaurant	Food & Drink Shop	Food	Dim Sum Restaurant	4.0

Discussion:

- ❖ Based on the results of the above table, observed the highest amounts of restaurants serving Sushi in the Bo-Kaap and Victoria Junction Neighborhoods.
- ❖ Each of the two has 90 Sushi Restaurants.
- ❖ These two neighborhoods seem to be abnormal.
- ❖ Both fall in the Cape Town City Centre
- ❖ Both have the same 90 restaurants surrounding them.
- ❖ Therefore they are joint first
- ❖ Followed by Sea Point with 18 in 2nd
- ❖ The Waterfront with 17 in 3rd.
- ❖ Based off of the Clustering, it appears as if most of the Sushi restaurants do not only sell Sushi, they usually have multiple cuisines.

Recommendation:

Based on the results and observations my recommendation is as follows:

- ❖ Due to the high volume of Sushi restaurants in and around the Bo-Kaap/Victoria junction neighborhoods, I would recommend that those neighborhoods be avoided as they are highly saturated.
- ❖ I would recommend that the person looks at opening the restaurant in the Waterfront or Newlands neighborhoods as these neighborhoods have already shown that there is a high demand for Sushi, while at the same time, not being overly saturated.
- ❖ I would also recommend that the person look at serving other cuisines in their restaurant as this seems to be something customers look for.

Conclusion:

- ❖ The goal of this project was to assist a potential restaurant owner by answering the Geospatial questions on how many Sushi restaurants are in each neighborhood, as well as Cape Town in general, and in which neighborhoods there are Sushi restaurants.
- ❖ They also requested a recommendation as to which neighborhood is the best one to open their restaurant in.
- ❖ Using the described data, I queried Foursquare API servers, used the results to perform clustering by the K-Means Machine Learning Algorithm, and made a recommendation that the person look at opening their restaurant in either Sea Point or the Waterfront neighborhood and that they look at serving an additional cuisine with their Sushi.



Thank you!