

Battle of the Neighborhoods

Cape Town, South Africa



Table of Contents

A description of the problem and a discussion of the background:	3
A description of the data and how it will be used to solve the problem:.....	3
Description of the data:	3
How the data will be used to solve the problem:.....	4
Methodology:.....	4
Results:.....	5
Discussion and Recommendation:.....	6
Discussion:	6
Recommendation:.....	6
Conclusion:.....	6



A description of the problem and a discussion of the background:

When looking to open a restaurant in any city, it is very important to do research on the city or neighborhood in which you are looking to open the restaurant. This will give 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. Researching the amount of restaurants, the different cuisines and the ratings of these restaurants will be the best indicators of how well the 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. Home to approximately 3.5 million people, Cape Town, South Africa is a very diverse City, with many restaurants serving food from all over the world. 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 as well as Cape Town in general and in which neighborhoods there are Sushi restaurants. From this information they would like a recommendation as to which neighborhood is the best one to open their restaurant in.

A description of the data and how it will be used to solve the problem:

Description of the data:

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.
- **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, I will 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 then be used to determine which neighborhood will be the best suited to open the new restaurant in.

Methodology:

Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, and what machine learnings were used and why.

Identifying the Neighborhoods in Cape Town:

To 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>). Once the data has been 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 will then be read into dataframes. Using Pandas, the 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.

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, we then request the Foursquare API server to return information about all of the Sushi venues in Cape Town.

Processing the retrieved data:

Once the JSON file is retrieved from the Foursquare API server, we will read the results into dataframes. From there we will normalize the data and then the data will be grouped by the Neighborhoods column and a count will be performed. We will then perform “One Hot” encoding on the resulting dataframe and re-group by neighborhoods. At this point the dataset will be ready to be used for machine learning and statistical analysis purposes.

Applying Unsupervised Machine Learning Techniques (K-Means Clustering) and statistical analysis:

Here we will select the top 10 Sushi restaurants from each neighborhood for statistical analysis. We will then perform K-Means Clustering on the restaurant data and merge the dataframes in order to solve the

problem. K-means Clustering was selected due to the large amount of unlabeled data. Using the Folium package, the cluster will be visually displayed on a map. The data clusters will also be examined in order to solve the problem.

Results:

Based off of the data, 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	Venue
ATHLONE	2
BELLVILLE	5
BERGVLIET	4
BO-KAAP	90
BOTHASIG	2
CAMPS BAY	2
CLAREMONT	9
CONSTANTIA	1
DURBANVILLE	4
EDGEMEAD	1
FISH HOEK	2
HOUT BAY	8
KENILWORTH	4
LANSDOWNE	4
MILNERTON	2
MONTE VISTA	1
MOWBRAY	3
NEWLANDS	14
OBSERVATORY	7
OTTERY	1
PAROW	1
PLUMSTEAD	1
RETREAT	2
SEA POINT	18
SOUTHFIELD	1
TABLE VIEW	4
TOKAI	5
VICTORIA JUNCTION	90
WATERFRONT	17
WOODSTOCK	5

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.

Discussion and Recommendation:

Discussion:

Based on the results of the above table I observed the highest amounts of restaurants serving Sushi in the Bo-Kaap and Victoria Junction Neighborhoods. Each of the two has 90 Sushi Restaurants. Looking at the rest of the data, these two neighborhoods seem to be abnormal. Both of these neighborhoods fall in the Cape Town City Centre and therefore both have the same 90 restaurants surrounding them. Therefore this would mean that they are joint first with the highest amount of Sushi Restaurants surrounding them, followed by Sea Point with 18 in 2nd and 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:

In 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.