Best Location for the Coffee Producing Firm in Johannesburg

Sizwe Mhlongo

April 01, 2019

1. Introduction

Ababa Coffee is a medium enterprise coffee producing firm based in Addis Ababa, Ethiopia which is located in the Eastern region of Africa. The leadership team of Ababa Coffee are planning to expand the business footprint by entering the South African coffee market. Ababa Coffee management team require advice regarding the best location to establish the business offices and warehouse in South Africa.

The leadership team of Ababa Coffee is aware that Johannesburg is the largest city and the economic hub of South Africa and they have decided that setting up the business in Johannesburg will provide it with the best chance to succeed. This is because Johannesburg has the highest number of working class citizens which are the highest consumers of coffee and as a result, Johannesburg has the highest density of potential customers for Ababa Coffee which include establishments such as: Coffee Shops, Restaurants, Diners, Fast Foods Outlets, Grocery Shops and Supermarkets.

Johannesburg is however a big city, and the management team's main problem is deciding where exactly in Johannesburg must they establish the business. To help the Ababa Coffee management team to make an informed decision they require services of a Data Scientist to analyse the geographical locations data of Johannesburg including all its Boroughs/ Regions and Neighbourhoods/ Suburbs. This data will also include GPS coordinates of all Suburbs in Johannesburg as well as locations of all potential Ababa Coffee customers mentioned above. The decision regarding the best location to setup the business will be based on the neighbourhood with the highest density of these customers. This will ensure that Ababa Coffee' location is strategically

situated closer to its customers which will reduce their marketing and logistics costs and reduce their delivery timelines which will increase their customer's satisfaction.

2. Data

To help Ababa Coffee in making an informed decision about the best location to setup the business offices in Johannesburg, the Data Scientist we will need to access following data:

1) The Boroughs/Regions and Neighbourhoods/Suburbs of Johannesburg from Wikipedia: https://en.wikipedia.org/wiki/Johannesburg.

   We will make us of the BeautifulSoup Python package to scrape this Wikipedia website and transform the data about the City of Johannesburg including all its Boroughs/Regions and Neighbourhoods/Suburbs into a Pandas Dataframe.

2) The GPS coordinates (latitude, longitude) of these Boroughs/Regions and Neighbourhoods/Suburbs of Johannesburg.

   We will use the Geocoder Python package to get these coordinates.

3) From Foursquare API we require the following venues data:
   - the Coffee Shops venues of the Johannesburg Neighbourhoods
   - the Restaurants venues of the Johannesburg Neighbourhoods
   - the Diners venues of the Johannesburg Neighbourhoods
   - the Fast Foods Outlets venues of the Johannesburg Places
   - the Grocery Shop venues of the Johannesburg Neighbourhoods

- the Supermarkets venues of the Johannesburg Neighbourhoods

We will then leverage the above data in order to determine which Johannesburg Neighbourhood is the most appropriate in order to locate the offices for Ababa Coffee in Johannesburg, South Africa.

3. Methodology and Discussion

After using BeautifulSoup Python package to scrape the Johannesburg Neighbourhoods data from the Wikipedia website, data was converted into the Pandas Dataframe for the ease of manipulation and computation:

```
In [45]: df = pd.DataFrame()
         df['Places'] = Places
         df
```

Out[45]:

	Places
0	Alexandra, Gauteng
1	None
2	None
3	None
4	Dainfern
5	None
6	Diepsloot
7	None
8	Drie Ziek
9	None
10	Ebony Park
11	None
12	Ennerdale, Gauteng
13	None
14	Farmall
15	None
16	Itsoseng
17	None

Data was then cleaned to remove all the empty fields:

```
In [49]: df
```

Out[49]:

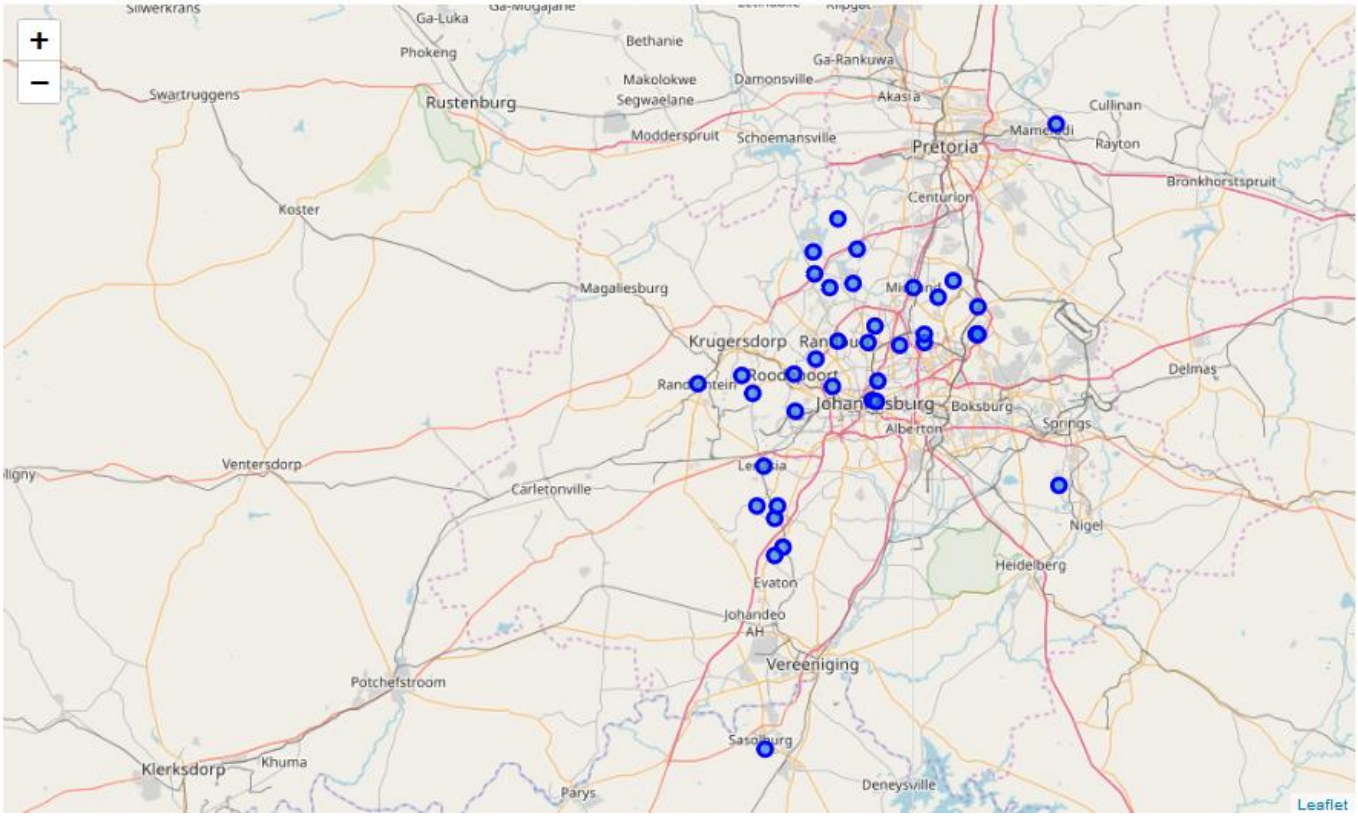
	Places
0	Alexandra
1	Dainfern
2	Diepsloot
3	Drie Ziek
4	Ebony Park
5	Ennerdale
6	Farmall
7	Itsoseng
8	Ivory Park
9	Johannesburg
10	Kaalfontein
11	Kagiso
12	Kanana Park
13	Lakeside
14	Lanseria

Using the Geocoder Python package the respective coordinates for the Neighbourhoods in Johannesburg were retrieved into a DataFrame. This and the previous DataFrame were merged to form the source data for further analysis:

Out[50]:

	Code	Places	Latitude	Longitude
0	798014	Alexandra	-26.104444	28.098889
1	798011	Chartwell	-26.167513	28.052286
2	798002	City of Johannesburg (non-urban)	-26.176730	27.963530
3	798012	Dainfern	-25.993149	28.004047
4	798003	Diepsloot	-25.930770	28.012252
5	798035	Drie Ziek	-26.202270	28.043630
6	798007	Ebony Park	-26.128640	27.930013
7	798033	Ennerdale	-26.414938	27.846326
8	798017	Farmall	-25.999010	27.957900
9	798021	Itsoseng	-26.083330	28.250000
10	798006	Ivory Park	-25.987500	28.203889
11	798015	Johannesburg	-26.205000	28.049722
12	798005	Kaalfontein	-26.035497	28.254675
13	798024	Kagiso	-26.157778	27.781389
14	798039	Kanana Park	-26.083330	28.250000
15	798037	Lakeside	-26.098792	28.148274
16	798019	Lanseria	-25.934926	27.925819
17	798038	Lawley	-26.391667	27.812778
18	798029	Lehae	-26.097945	28.034029
19	798028	Lenasia	-26.319631	27.824432
20	798032	Lenasia South	-26.391944	27.853056
21	798020	Lucky 7	-25.974132	27.927873
22	798001	Malatjie	-26.083330	28.254675
23	798009	Mayibuye	-25.705717	28.410015
24	798004	Midrand	-25.999262	28.125912
25	798018	Millgate Farm	-26.083330	28.148274
26	798034	Orange Farm	-26.466111	27.865278
27	798040	Poortjie	-26.456849	-26.456849
28	798008	Rabie Ridge	-26.017222	28.174722
29	798016	Randburg	-26.096111	27.975000

In order to better visualise this data, a Folium Map of Johannesburg and its Neighbourhoods was created:



Foursquare API was then used to retrieve the different categories of venues within Johannesburg:

In [55]:
```
# Use category id 4d4b7105d754a06374d81259 to only get the food venues
joburg_venues_food = getNearbyVenues(names=joburg_places['Places'], latitudes=joburg_places['Latitude'], longitudes=joburg_places['Longitude'], radius=5000, LIMIT=500, categoryIds='4d4b7105d754a06374d81259')
joburg_venues_food.head()
```

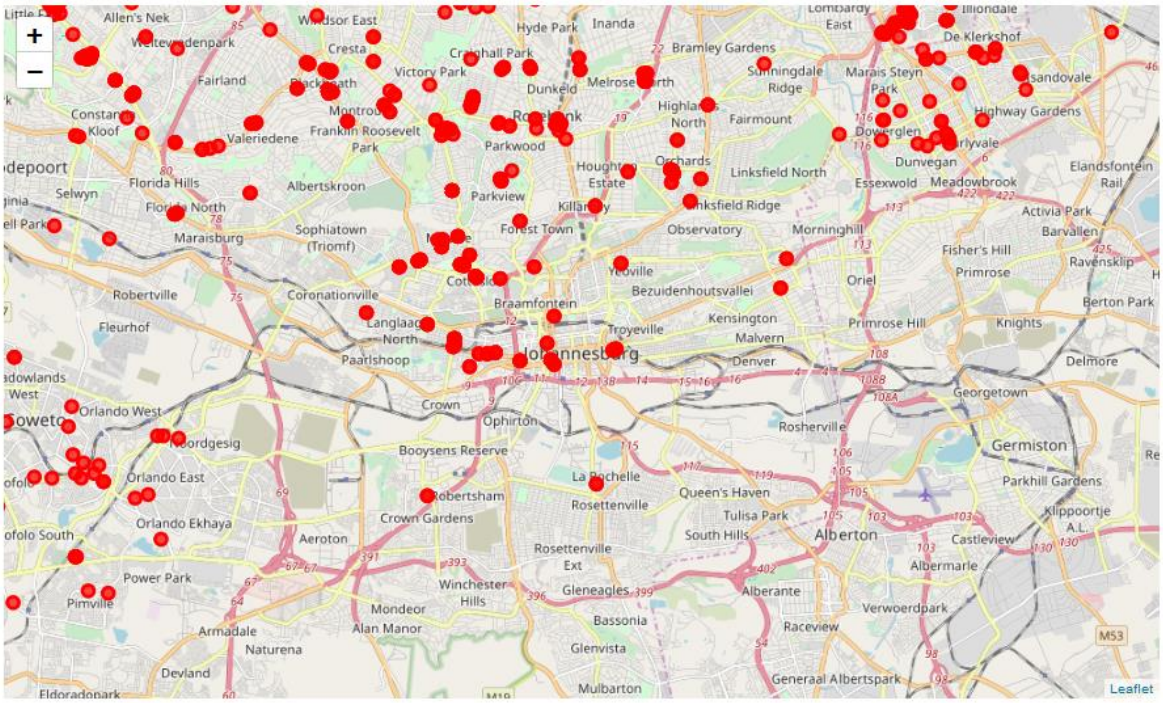
Out[55]:

	Places	Places Latitude	Places Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Alexandra	-26.104444	28.098889	Nu	-26.106396	28.049878	Juice Bar
1	Alexandra	-26.104444	28.098889	Trump's Grill	-26.107484	28.054812	BBQ Joint
2	Alexandra	-26.104444	28.098889	Moyo	-26.133844	28.068203	African Restaurant
3	Alexandra	-26.104444	28.098889	Tashas	-26.131583	28.068550	Café
4	Alexandra	-26.104444	28.098889	Walnut Grove	-26.108800	28.052731	Restaurant

In [56]:
```
joburg_venues_food.shape
```

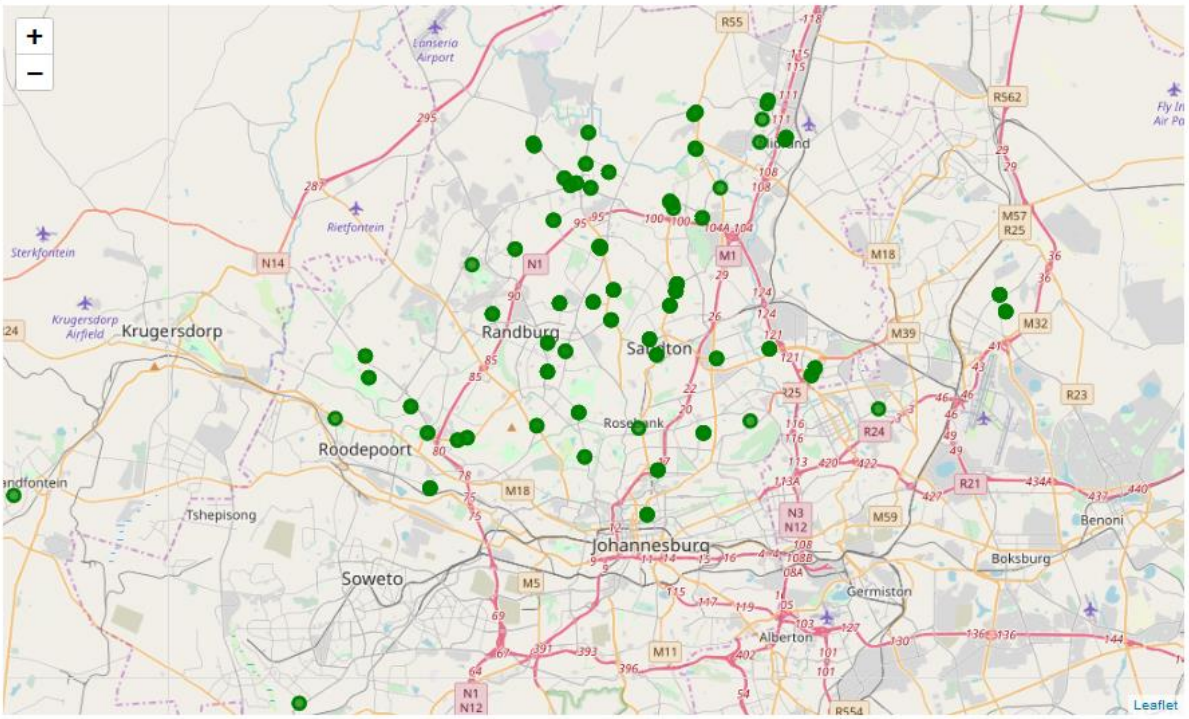
i. Folium Map of Food Venues in Johannesburg

```
map_joburg_food = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(joburg_venues_food, 'red', map_joburg_food)
map_joburg_food
```



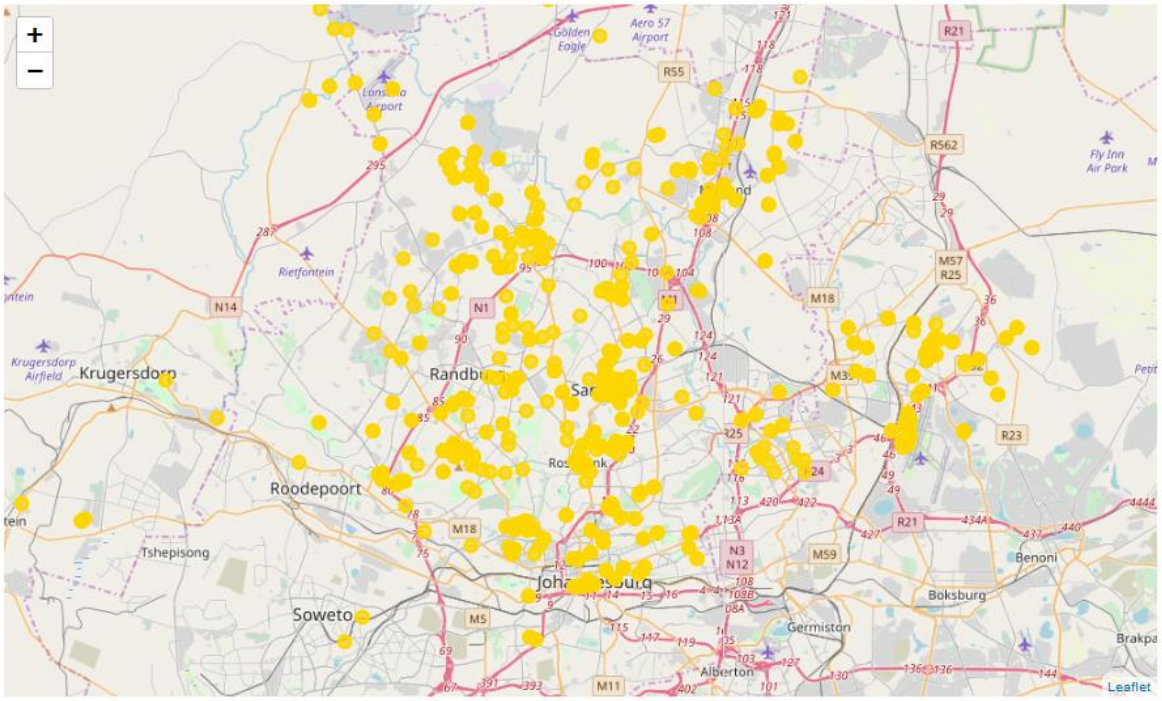
ii. Folium Map of Supermarkets Venues in Johannesburg

```
map_joburg_supermarket = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(joburg_venues_supermarket, 'green', map_joburg_supermarket)
map_joburg_supermarket
```



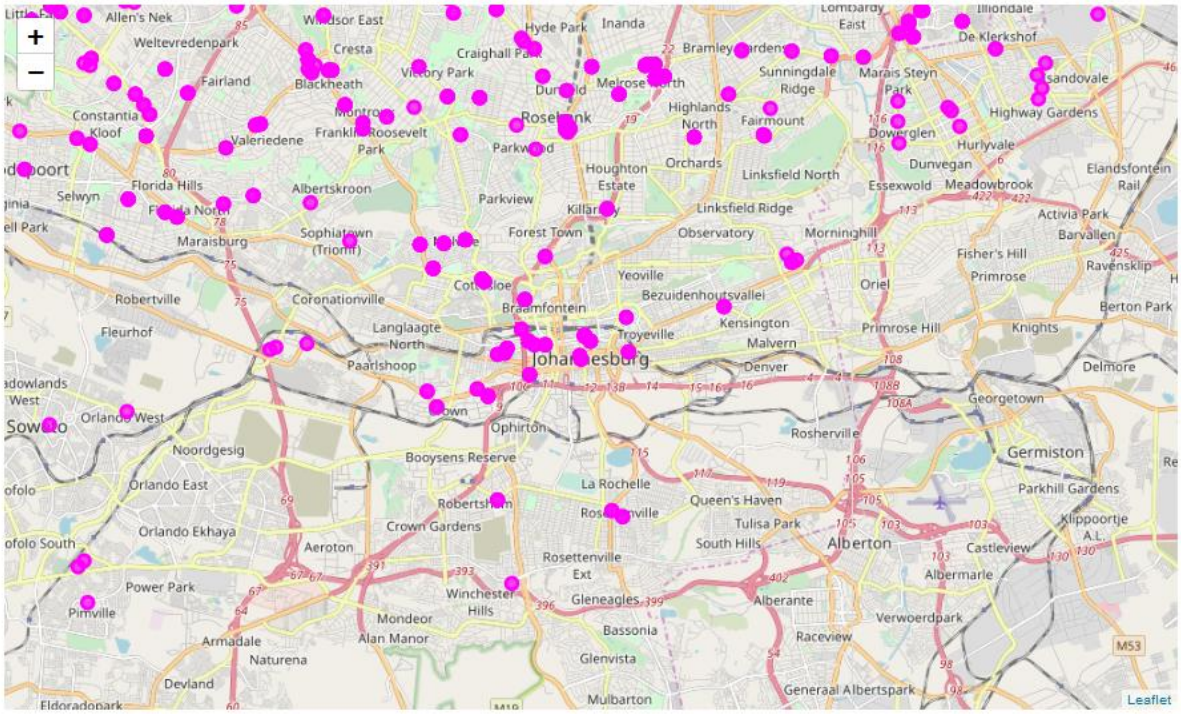
iii. Folium Map of Hotels Venues in Johannesburg

```
map_joburg_hotel = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(joburg_venues_hotel, 'gold', map_joburg_hotel)
map_joburg_hotel
```



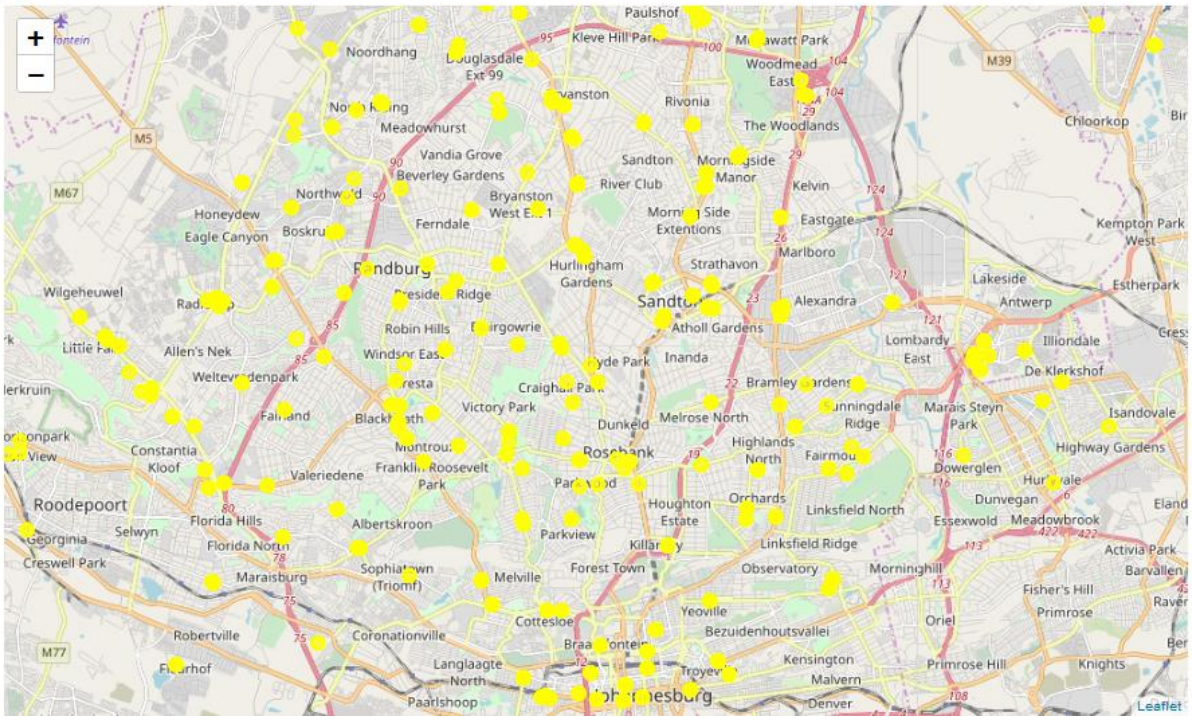
iv. Folium Map of Shopping Malls Venues in Johannesburg

```
map_joburg_mall = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(joburg_venues_mall, 'fuchsia', map_joburg_mall)
map_joburg_mall
```



v. Folium Map of Grocery Stores Venues in Johannesburg

```
map_joburg_grocery = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(joburg_venues_grocery, 'yellow', map_joburg_grocery)
map_joburg_grocery
```



We then computed a Statistic Weighted Average of each Venue based on its strategic importance to Ababa Coffee.

```
# positive weight, because Food places like Restaurants are big customers for Ababa Coffee due to high consumption rates
weight_food = 2

# positive weight, because consumers will purchase the Ababa Coffee products from their nearest Supermarkets
weight_supermarket = 1.5

# positive weight, because Hotels are big customers for Ababa Coffee due to high consumption rates
weight_hotel = 2

# positive weight, because consumers can purchase the Ababa Coffee products from their nearest Shopping Malls
weight_mall = 1

# positive weight, because consumers will purchase the Ababa Coffee products from their nearest Grocery Stores
weight_grocery = 1.5
```

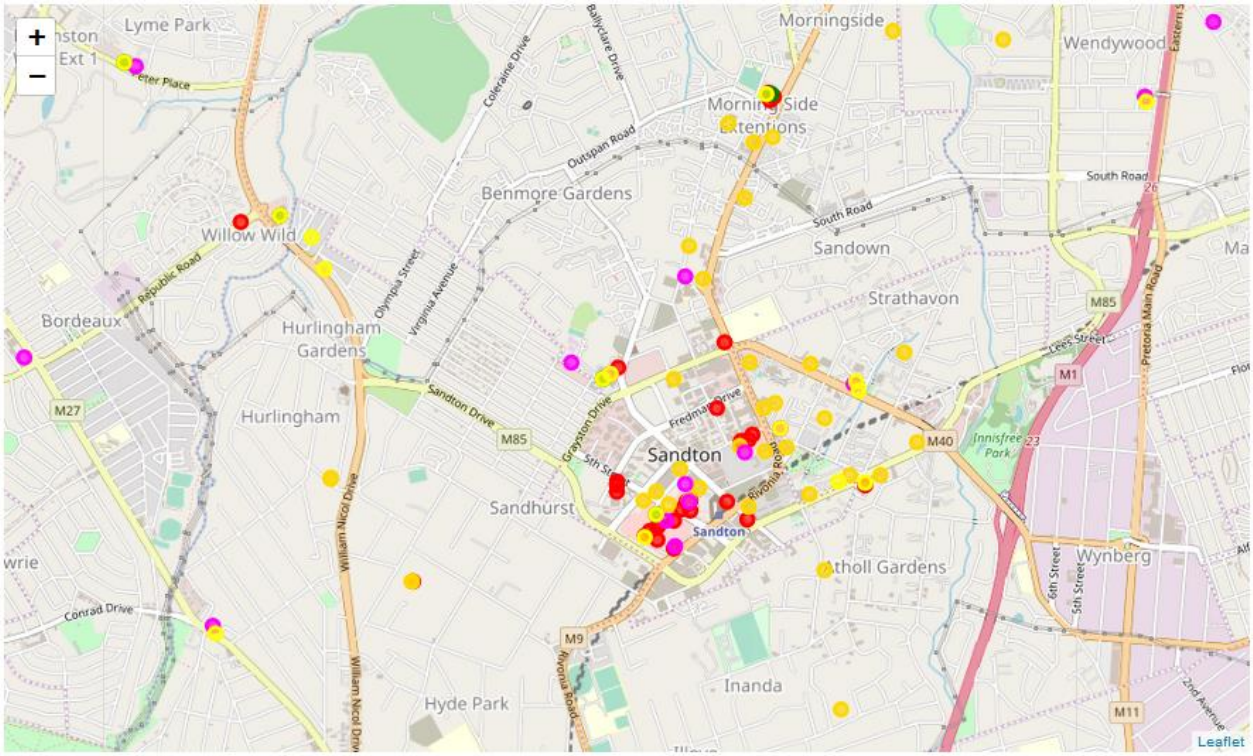
```
df_weighted = joburg_places[['Places']].copy()
```

```
df_weighted['Score'] = joburg_places['Foods'] * we
t + joburg_places['Hotels'] * weight_hotel + jobur
cery Stores'] * weight_grocery
df_weighted = df_weighted.sort_values(by=['Score']
df_weighted
```

	Places	Score
33	Sandton	352.5
18	Lehae	350.5
29	Randburg	343.0
1	Chartwell	328.0
0	Alexandra	320.0
6	Ebony Park	311.5
11	Johannesburg	293.0
5	Drie Ziek	284.5
3	Dainfern	280.0
2	City of Johannesburg (non-urban)	279.5
8	Farmall	272.5
24	Midrand	268.5
9	Itsoseng	236.5
22	Malatjie	236.5
14	Kanana Park	236.5
32	Roodepoort	200.0
15	Lakeside	192.5
25	Millgate Farm	159.0
12	Kaalfontein	133.5
28	Rabie Ridge	129.0
19	Lenasia	123.5
21	Lucky 7	109.5

Sandton achieved the highest accumulated score as shown above.

Lastly, we took a closer look into the distribution of the different Venue categories around Sandton.



4. Conclusion

Sandton is the best Place/ Neighbourhood for Ababa Coffee to setup their business within the City of Johannesburg, in South Africa. This is due to the highest concentration of major customers that Ababa Coffee is targeting.