Riyadh city gym & fitness center analysis: The best DISTRICTS to open a gym?

APPLIED DATA SCIENCE CAPSTONE

Abdelrahman Elsayed

Table of Contents

Introduction

The global health club industry has generated $94 billion over the past 12 months, up from $87.2 billion the previous year. With a global growth rate of 8.7%, we're seeing positive signals that the fitness industry will continue to grow. On this trajectory, the fitness market is set to hit $106 billion in 2020.

According to the report of *" Saudi Arabia Fitness Services Market Analysis 2012-2017 & Outlook 2018-2022"*:

Saudi Arabia youth population has increased at a steady pace over the review period (2012-2017), which has resulted into a significant increase in the fitness center going population. Arabia fitness services market has witnessed a tremendous growth in the past few years witnessing a strong double digit CAGR from 2012 to 2017. Rapidly rising obesity rate, heart problems, and increase in the consciousness about the personal health have propelled momentum to the market. The demographic profile of the country has also changed over the years, which has resulted as a supporting factor to this industry. Gyms membership has been the largest revenue generating segment in Saudi Arabia.

Riyadh has accounted for the maximum number of organized fitness centers in KSA, primarily for being the commercial capital of the country.

Here, we present some insights to help business owners who are interested in this industry and are looking to build a gym inside Riyadh to know which districts should they target.

Data:

Data related to Riyadh districts was tricky to find, however after going through several github links I found some large datasets in the formats of JSON and SQL.

The JSON link includes all boundaries of all the districts in Saudi Arabia, however we are only interested in the coordinates of Riyadh (Latitudes, Longitudes)

The links:

https://github.com/Faisal0sal/Saudi-Arabia-Regions-Cities-and-Districts

https://github.com/aalmangour/Saudi_GIS_Data/blob/master/sa_neighborhoods.sql

After analyzing the data, preprocessing the SQL file through Excel Power Query to clean it and process it to be in the desired format (shown here)

The Foursquare API is also used to find out popular venue details and categories in each district

GeoPy Python library is used as well for Maps visualizations

Disclaimer:

The overall data quality is rather poor due to the fact that places in Saudi Arabia doesn't provide Foursquare with the level of detail that is essential to this study. It's also worth mentioning that spelling of districts names makes it difficult to match the data required and thus that results in missing data points in some cases.

Methodology:

The methodology and approach are comprised into 4 steps:

I.   Data preparation including sourcing and cleansing.
II.   Visualization of Gyms in Riyadh to get an initial idea.
III.   Obtaining the most common venue categories in Riyadh neighborhoods.
IV.   Clustering the neighborhoods.

Simply, we source data for Riyadh districts and we learn how to convert addresses into their equivalent latitude and longitude values. Also, we use the Foursquare API to explore neighborhoods in Riyadh City. We use the explore function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. We use the k-means clustering algorithm to complete this task. Finally, we use the Folium library to visualize the neighborhoods in Riyadh City and their emerging clusters.-*just as we learned in the lab-*
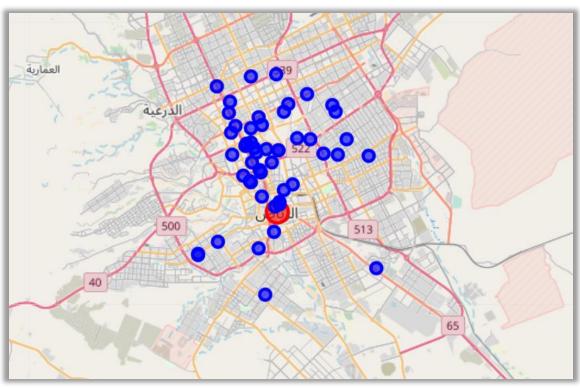
Here are snapshots of the Notebook created:

Riyadh Gym & Fitness centers analysis

Send the GET Request and examine the results

```
[29]: results = requests.get(url).json()
      results
```

```
[29]: {'meta': {'code': 200, 'requestId': '5ee35e709ca1e36a52e22bc9'},
       'response': {'venues': [{'id': '5b8ab80be96d0c00397ae123',
         'name': 'Gold's Gym',
         'location': {'lat': 24.665898,
          'lng': 46.681739,
          'labeledLatLngs': [{'label': 'display',
            'lat': 24.665898,
            'lng': 46.681739}],
          'distance': 5062,
          'postalCode': '12714',
          'cc': 'SA',
```

	id	name	categories	referralId	hasPerk	location.lat	location.lng
0	5b8ab80be96d0c00397ae123	Gold's Gym	[{'id': '4bf58dd8d48988d176941735', 'name': 'G...	v-1591959186	False	24.665898	46.681739
1	53b309c3498ee6ba6d20a978	Royal Guard Gym	[]	v-1591959186	False	24.648900	46.695501
2	5d2c85c4c386450030ee8e6e	Gold's Gym (جيم جولدز)	[{'id': '4bf58dd8d48988d175941735', 'name': 'G...	v-1591959186	False	24.666643	46.681526
3	52ada10311d2696d665f2f2c	Sheraton Gym	[{'id': '4bf58dd8d48988d176941735', 'name': 'G...	v-1591959186	False	24.638206	46.712072

Let's visualize the Gyms that are nearby

```
dataframe_filtered.name
```

```
0                                    Gold's Gym
1                              Royal Guard Gym
2                    Gold's Gym (جيم جولدز)
3                                 Sheraton Gym
4      Male Gym / Fitness Center - Armed Forces Offic...
5                                   Extreme Gym
6          Alfrusiya Gym (الفروسية - نادي الرياضي)
7                       Holiday Inn Izdihar - Gym
8                          Tulip Inn Fitness Gym
9                     Gym of knig salman Air Base
10                      Gym Kitchen (جيم كتشن)
11                                    Hilton Gym
12                        GYM - Executive Hotel
13     LAVA Fitness at KFMC | Ladies Gym (النادي الري...
14                Gym, Spa&Sauna. Akariya Compound
15                       Nwc's GYM (Life Fitness)
16                      Vitamin Gym (نادي فيتامين)
```

Data processing

```python
import pandas as pd
import numpy as np

#the data were obtained online through an SQL file found on GitHub that was scrabed an
#All the preprocessing was done in Power Query to turn the SQL file to a suitable data
Dist=pd.read_excel('https://srv-file10.gofile.io/download/Y9FcIf/RiyadhDistrictsData2.
Dist
```

	ID	DistrictAR	DistrictEN	City	Latitude	Longitude
0	763	حطين	Hitien	Riyadh	26.331606	44.883862
1	764	قرطبة	Qurtubah	Riyadh	26.327879	44.848296
2	765	الفيصلية	Al Faysaliyyah	Riyadh	26.293811	44.802280
3	766	العزيزية	Al Aziziyyah	Riyadh	26.307914	44.815816
4	767	القدس	Al Quds	Riyadh	26.302859	44.804438
5	768	الخالدية	Al Khalidiyyah	Riyadh	26.299103	44.825027
6	769	الروضة	Ar Rawdah	Riyadh	26.293558	44.825436
7	770	الربوة	Al Rabwah	Riyadh	26.321627	44.887540

Results:

The results are as shown below, illustrating the most common venue categories in Riyadh neighborhoods. Conclusions and insights are discussed in the following section.

	Neighborhood	BBQ Joint	Bakery	Burger Joint	Café	Candy Store	Castle	Clothing Store	Coffee Shop	Exhibit	...	Park
0	Al Aziziyyah	0.00	0.00	0.000000	0.000000	0.5	0.0	0.000000	0.000000	0.000000	...	0.0
1	Al Farouq	0.25	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	...	0.0
2	Al Faysaliyyah	0.00	0.00	0.090909	0.090909	0.0	0.0	0.090909	0.181818	0.000000	...	0.0
3	Al Khalidiyyah	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	...	0.0
4	Al Muntazah	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	...	1.0
5	Al Quds	0.00	0.25	0.000000	0.000000	0.0	0.0	0.000000	0.250000	0.000000	...	0.0
6	Al Yamamah	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	...	0.0
7	Al Yarmok	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	...	0.0
8	Alkah	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.333333	...	0.0
9	Ar Rawdah	0.00	0.00	0.000000	0.200000	0.0	0.2	0.000000	0.000000	0.000000	...	0.0
10	As Sanaiyyah	0.00	0.00	0.000000	0.500000	0.0	0.0	0.000000	0.000000	0.000000	...	0.0
11	As Sediq	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.200000	0.000000	...	0.0
12	Badr	0.00	0.00	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	...	0.0

```
         ---- Ar Rawdah----
                       venue   freq
0   Fast Food Restaurant    0.2
1                   Café    0.2
2            Supermarket    0.2
3                 Castle    0.2
4                   Farm    0.2


         ---- As Sanaiyyah----
               venue   freq
0               Café    0.5
1              River    0.5
2          BBQ Joint    0.0
3        Pizza Place    0.0
4             Market    0.0


         ---- As Sediq----
              venue   freq
0             Plaza    0.2
1        Restaurant    0.2
2               Gym    0.2
3        Coffee Shop   0.2
```

```python
# move neighborhood column to the first column
fixed_columns = [Riyadh_onehot.columns[-1]] + list(Riyadh_onehot.columns[:-1])
Riyadh_onehot = Riyadh_onehot[fixed_columns]

Riyadh_onehot.head()
```

	Neighborhood	BBQ Joint	Bakery	Burger Joint	Café	Candy Store	Castle	Clothing Store	Coffee Shop	Exhibit	...	Park	Pizza Place	Plaza
0	Al Faysaliyyah	0	0	0	0	0	0	0	0	0	...	0	0	0
1	Al Faysaliyyah	0	0	0	0	0	0	0	1	0	...	0	0	0
2	Al Faysaliyyah	0	0	0	0	0	0	0	0	0	...	0	0	0
3	Al Faysaliyyah	0	0	0	0	0	0	1	0	0	...	0	0	0
4	Al Faysaliyyah	0	0	0	1	0	0	0	0	0	...	0	0	0

5 rows × 34 columns

	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
0	Al Aziziyyah	Candy Store	Flower Shop	Windmill	Gym	Bakery	Burger Joint	Café
1	Al Farouq	BBQ Joint	Shawarma Place	Fast Food Restaurant	Falafel Restaurant	Coffee Shop	Flower Shop	Flea Market
2	Al Faysaliyyah	Coffee Shop	Gym / Fitness Center	Clothing Store	Flea Market	Fast Food Restaurant	Mobile Phone Shop	Grocery Store
3	Al Khalidiyyah	Fast Food Restaurant	Market	Middle Eastern Restaurant	Windmill	Exhibit	Flea Market	Farm
4	Al Muntazah	Park	Windmill	Flower Shop	Flea Market	Fast Food Restaurant	Farm	Falafel Restaurant
5	Al Quds	Coffee Shop	Falafel Restaurant	Turkish Restaurant	Bakery	Burger Joint	Café	Candy Store
6	Al Yamamah	Fast Food Restaurant	Windmill	Gym	Bakery	Burger Joint	Café	Candy Store
7	Al Yarmok	Farm	Ice Cream Shop	Optical Shop	Windmill	Exhibit	Flea Market	Fast Food Restaurant

Conclusion:

Based on the results shown and illustrated in the notebook as well as the previous section, we developed a criterion to categorize each of the venues into three categories;
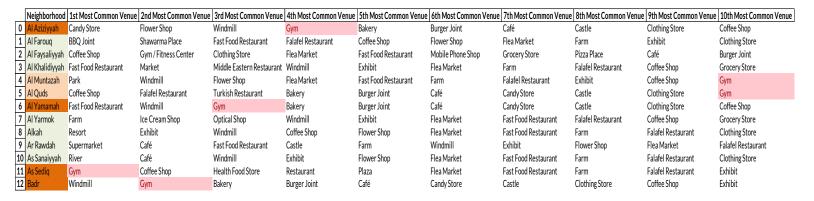
Safe Zone

(the recommended neighborhoods to open a new Gym) since currently it doesn't include any Gyms in the top 10 venues.

Moderate-risk Zone

To open a Gym here, a decision maker must take further steps in analyzing the neighborhoods, and the needs and Characteristics of people living in it. (This neighborhood includes Gyms in the top Venues in the rankings from 6th till 10th)

Red Zone

(the recommended neighborhoods to open a new Gym) since currently it doesn't include any Gyms in the top 10 venues.

	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
0	Al Aziziyyah	Candy Store	Flower Shop	Windmill	Gym	Bakery	Burger Joint	Café	Castle	Clothing Store	Coffee Shop
1	Al Farouq	BBQ Joint	Shawarma Place	Fast Food Restaurant	Falafel Restaurant	Coffee Shop	Flower Shop	Flea Market	Farm	Exhibit	Clothing Store
2	Al Faysaliyyah	Coffee Shop	Gym / Fitness Center	Clothing Store	Flea Market	Fast Food Restaurant	Mobile Phone Shop	Grocery Store	Pizza Place	Café	Burger Joint
3	Al Khalidiyyah	Fast Food Restaurant	Market	Middle Eastern Restaurant	Windmill	Exhibit	Flea Market	Farm	Falafel Restaurant	Coffee Shop	Grocery Store
4	Al Muntazah	Park	Windmill	Flower Shop	Flea Market	Fast Food Restaurant	Farm	Falafel Restaurant	Exhibit	Coffee Shop	Gym
5	Al Quds	Coffee Shop	Falafel Restaurant	Turkish Restaurant	Bakery	Burger Joint	Café	Candy Store	Castle	Clothing Store	Gym
6	Al Yamamah	Fast Food Restaurant	Windmill	Gym	Bakery	Burger Joint	Café	Candy Store	Castle	Clothing Store	Coffee Shop
7	Al Yarmok	Farm	Ice Cream Shop	Optical Shop	Windmill	Exhibit	Flea Market	Fast Food Restaurant	Falafel Restaurant	Coffee Shop	Grocery Store
8	Alkah	Resort	Exhibit	Windmill	Coffee Shop	Flower Shop	Flea Market	Fast Food Restaurant	Farm	Falafel Restaurant	Clothing Store
9	Ar Rawdah	Supermarket	Café	Fast Food Restaurant	Castle	Farm	Windmill	Exhibit	Flower Shop	Flea Market	Falafel Restaurant
10	As Sanaiyyah	River	Café	Windmill	Exhibit	Flower Shop	Flea Market	Fast Food Restaurant	Farm	Falafel Restaurant	Clothing Store
11	As Sediq	Gym	Coffee Shop	Health Food Store	Restaurant	Plaza	Flea Market	Fast Food Restaurant	Farm	Falafel Restaurant	Exhibit
12	Badr	Windmill	Gym	Bakery	Burger Joint	Café	Candy Store	Castle	Clothing Store	Coffee Shop	Exhibit

The results are shown in the table:

Neighborhood
Al Aziziyyah
Al Farouq
Al Faysaliyyah
Al Khalidiyyah
Al Muntazah
Al Quds
Al Yamamah
Al Yarmok
Alkah
Ar Rawdah
As Sanaiyyah
As Sediq
Badr

*Color indicates the categories each neighborhood belongs to.

Clustering

```python
# import k-means from clustering stage
from sklearn.cluster import KMeans

# set number of clusters
kclusters = 10

Riyadh_grouped_clustering = Riyadh_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Riyadh_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[:]
```

```
array([5, 0, 2, 9, 1, 2, 3, 8, 6, 2, 7, 2, 4])
```

Discussion:

The results show a great potential for opening Gyms, giving that there are currently 7 districts which doesn't have Gyms or fitness centers in their top 10 venues. There are also 2 locations which could be considered after further analysis. However, we recommend staying away from the 3 locations depicted in the previous section in the "Red Zone" category.

Finally, this work is just an example of how wonderful this field could be, and how it is integrated in our daily lives and in business.

There are minor comments described in the Data section, however there is no limit to what could be built based on this work as a foundation.

The aspect considered here is from a "naïve" point of view, and it shouldn't be the only criteria considered, it is just to illustrate what could be done with the tools given and the powerful Foursquare API.

Thank you for taking the time to read this.