

Al Khobar Rentability Study

2020

By Bader Ababtain

Abstract: When you try to build a real estate project, one of the biggest questions is the location and whether you will be able to rent spaces in that location or not. Location selection could be one of the biggest challenges in a real estate business as depending on that location would end up with a rentable place or not. Given that location selection is an important aspect of any real estate project, we shall study the rentability of a store given its location in the small city of Al Khobar in Saudi Arabia.

Contents

Introduction _____	1
Data _____	1
Data Selection.....	1
Data Quality	2
Data Processing.....	2
Methodology _____	3
Data Exploration	3
Models	15
Results _____	16
Discussion _____	19
Conclusion _____	20
Contact Information _____	20
Information _____	Error! Bookmark not defined.

Introduction

In a real estate project, location selection can be challenging. As one wouldn't have an idea whether one will be able to rent the place one is trying to build. Depending on that location and quality of build, the price of meter is set. We will explore the price in the supply data analysis but the demand data for the pricing is not available. Since those factors data are not readily available, we will focus in this report on rentability of a place based on their location. More specifically, we will study the rentability of a store location in Al Khobar city in Saudi Arabia. Real estate development companies in Al Khobar would benefit from such analysis. Al Khobar is a small city in Saudi Arabia with a total population of 219,679 and a total area of 571 square kilometers. A city with forty-one neighborhoods only. Within those forty-one neighborhoods, we will study the rentability of a store in a given location.

Data

Data Selection

Obtaining the data was not an easy task as most of the needed data was not readily available. I had to find the location data for each neighborhood in my city which was not an easy search but I finally found it [here](#). Using the foursquare data as rented stores, I had to also find the data for the unrented or available stores which I scrapped from one of the most famous sites in Saudi Arabia for renting a place [Aqar](#).

Data Quality

The quality of the data is mediocre at best as the only usable piece of information is the location data which was obtained after a heavy processing and some approximation of the Aqar data especially. Aqar data lacked the location (latitude and longitude) and the foursquare data lacked the price of the rent and the square meter area.

Data Processing

After a long search, I found the coordinates for neighborhoods in Saudi Arabia but it was in SQL files format for cities and neighborhoods which made it really hard to process and open. After downloading a collection of software to open the file, I found a website that opens SQL files but it refused to run the create table query in the file so I had to copy it manually to a text file and turn it into a CSV file. I also had to manipulate it through Excel in order to make it proper to join the cities table with the neighborhood table. Then, I filtered the data in order to obtain the neighborhood data for Al Khobar. For the Aqar data, the only location data that was available is the unofficial address in Arabic. So, I needed to do a fuzzy match between the two data sets in order to obtain the coordinates of the neighborhood the spaces lie in. After that I had to randomize the location through a normal distribution with a mean as the center of the neighborhood and standard deviation as the third of distance between the center and the edges of the neighborhood in order to obtain a unique location for each data point. The points are randomized based on the neighborhood they are located in. Then, I added for each data set a target value of one for rented stores (the foursquare data) and a value of zero for unrented stores (the Aqar data).

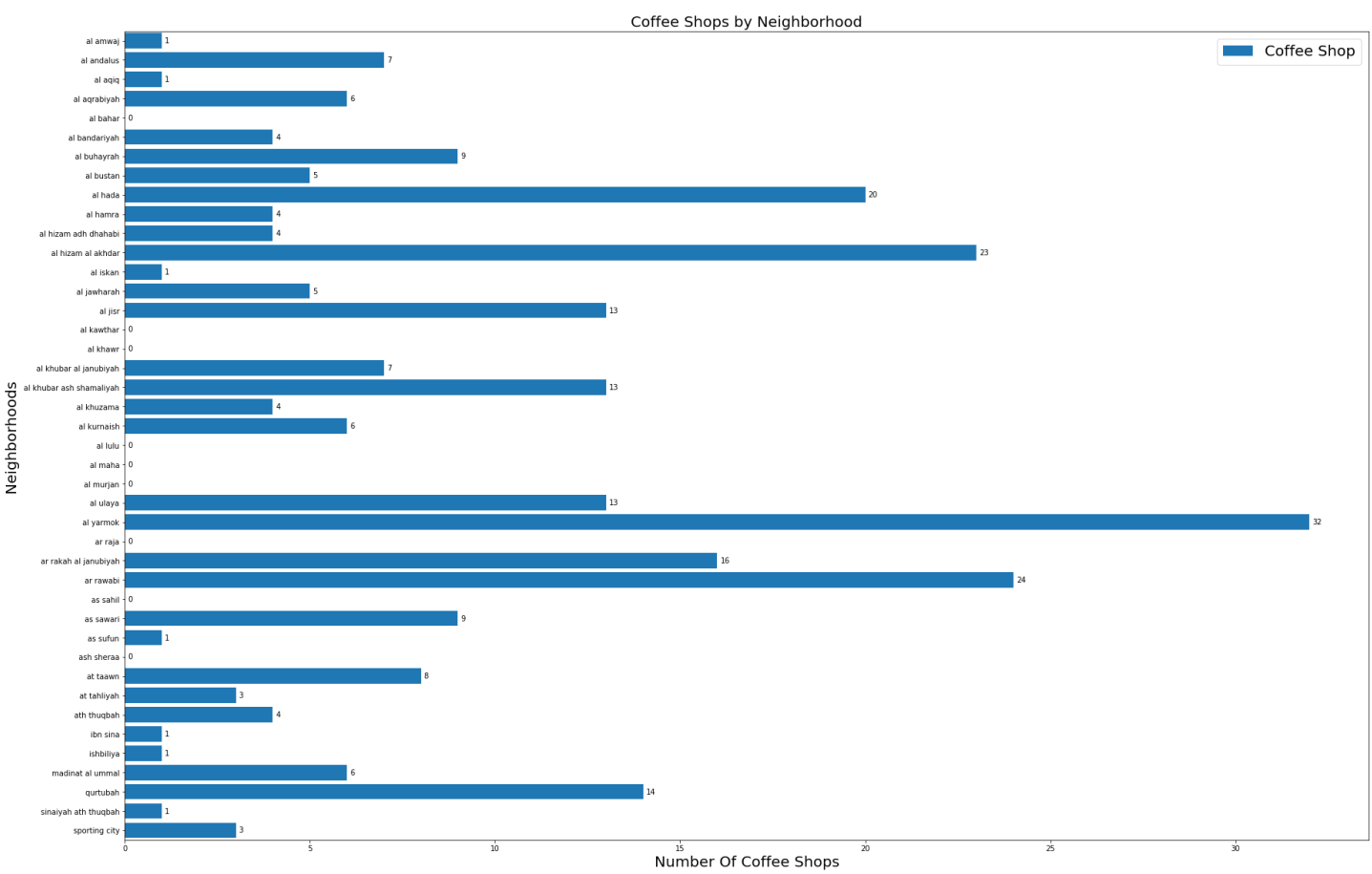
Methodology

Data Exploration

Here we will explore the data for the demand and supply. We will begin with the demand. Here is a pivot table of the neighborhoods by venue category:

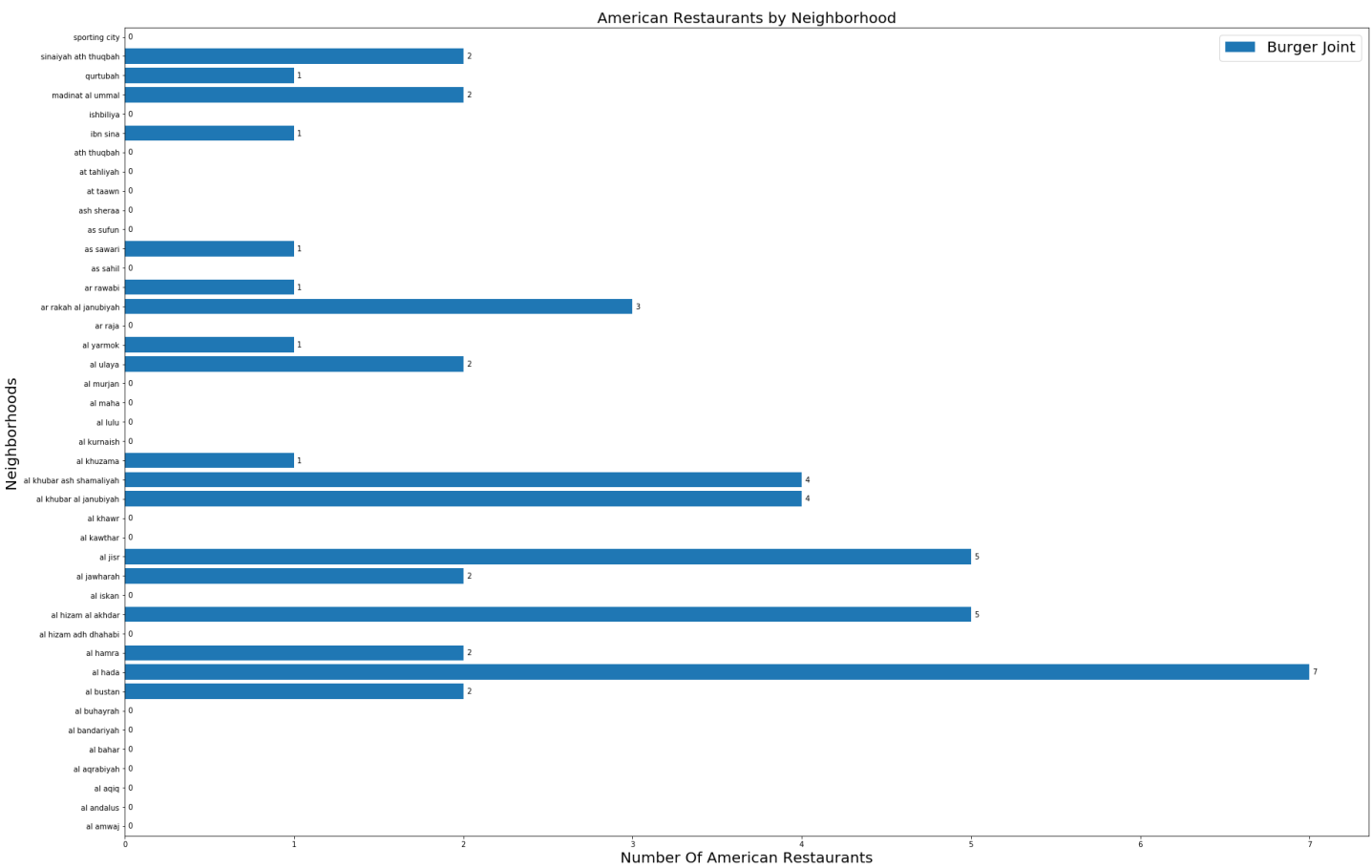
Neighborhood	Yoga Studio	Waterfront	Watch Shop	Video Game Store	Veterinarian	Vegetarian / Vegan Restaurant	Turkish Restaurant	Truck Stop
al amwaj	0	0	0	0	0	0	0	0
al andalus	0	0	0	0	0	1	1	0
al aqiq	0	0	0	0	0	0	0	0
al aqrabiyah	0	0	0	0	0	0	1	0
al bahar	0	0	0	0	0	0	0	0
al bandariyah	0	0	0	0	0	0	0	0
al buhayrah	0	0	0	0	0	0	0	0
al bustan	0	0	0	0	0	0	0	0
al hada	0	0	0	0	0	1	0	1
al hamra	0	0	0	1	0	0	1	0
al hizam adh dhahabi	0	0	0	0	0	0	0	0
al hizam al akhdar	1	0	0	0	0	0	0	0
al iskan	0	0	0	0	0	0	0	0
al jawharah	0	0	0	0	0	0	0	0
al jisr	0	0	0	0	0	0	0	0
al kawthar	0	0	0	0	0	0	0	0
al khawr	0	0	0	0	0	0	0	0
al khubar al janubiyah	0	0	0	0	0	0	0	0
al khubar ash shamaliyah	0	0	0	0	0	0	1	0
al khuzama	0	0	0	0	0	0	0	0
al kurnaish	0	0	0	0	0	0	0	0
al lulu	0	0	0	0	0	0	0	0
al maha	0	0	0	0	0	0	0	0
al murjan	0	0	0	0	0	0	0	0
al ulaya	0	0	0	0	0	0	0	0
al yarmok	2	1	1	0	0	0	0	0

As there is a lot of zeros here. Let's visualize parts of it. Let's check out how many coffee shops are in each neighborhood:



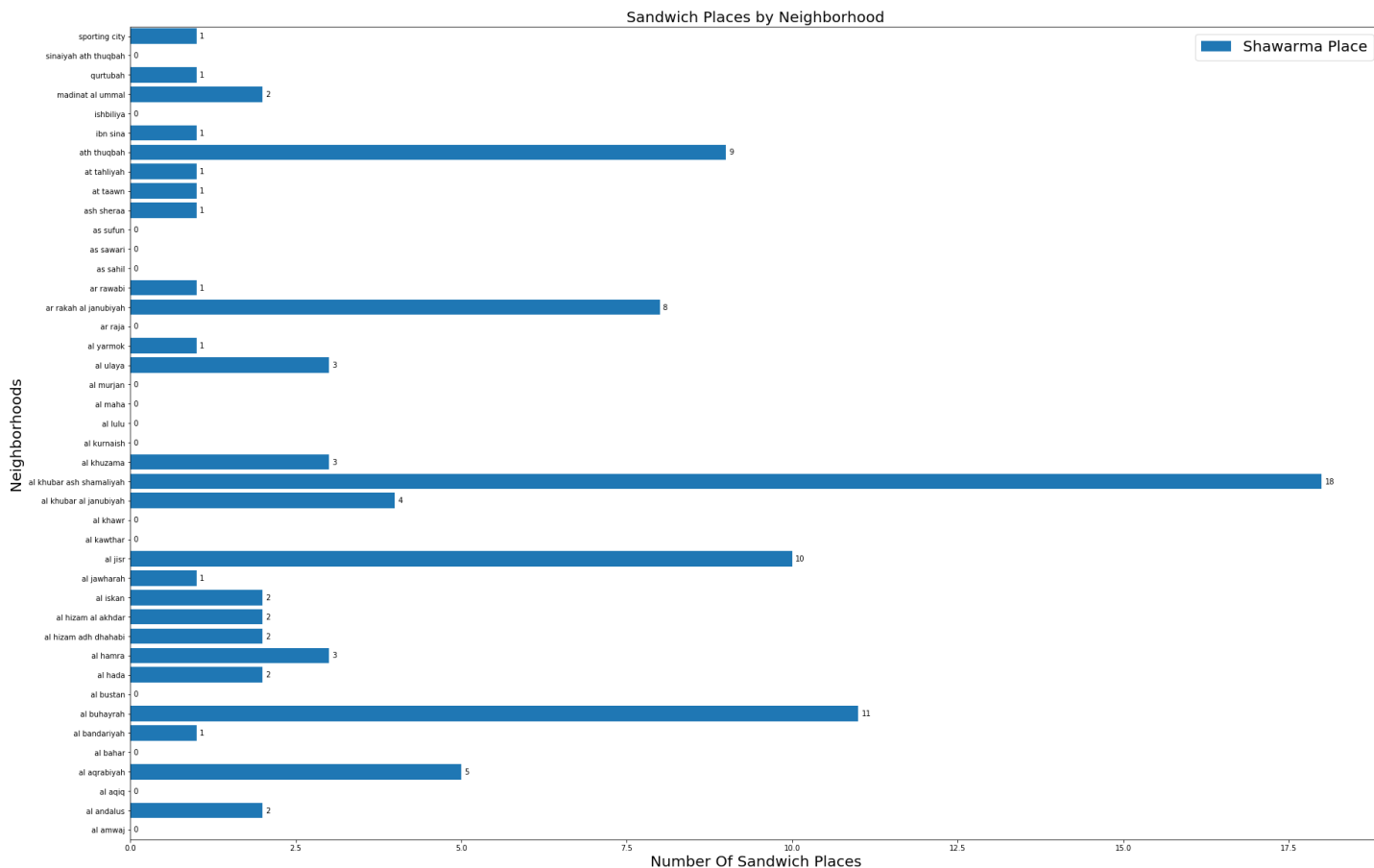
There seems to be coffee shops everywhere. The highest concentration of coffee shops seems to be in Al Yarmok. Amongst least is Sinaiyah Ath Thoqbah which makes sense since that is area where garages exist (places to repair cars). Overall, it seems that coffee shops are very popular here and the data support it.

Now let us take a look at the burger joints in Al Khobar



As you can see, American restaurants are popular in Al Khobar. The most concentration of American restaurants are in Al Hada neighborhood with a total number of seven restaurants between burger joints and American restaurants. There is a collection of neighborhoods which lack American restaurants and burger joints. They are either totally residential such as Al Hizam Al Thahabi and with very few commercial lands or of low-income class such as Ath Thuqbah.

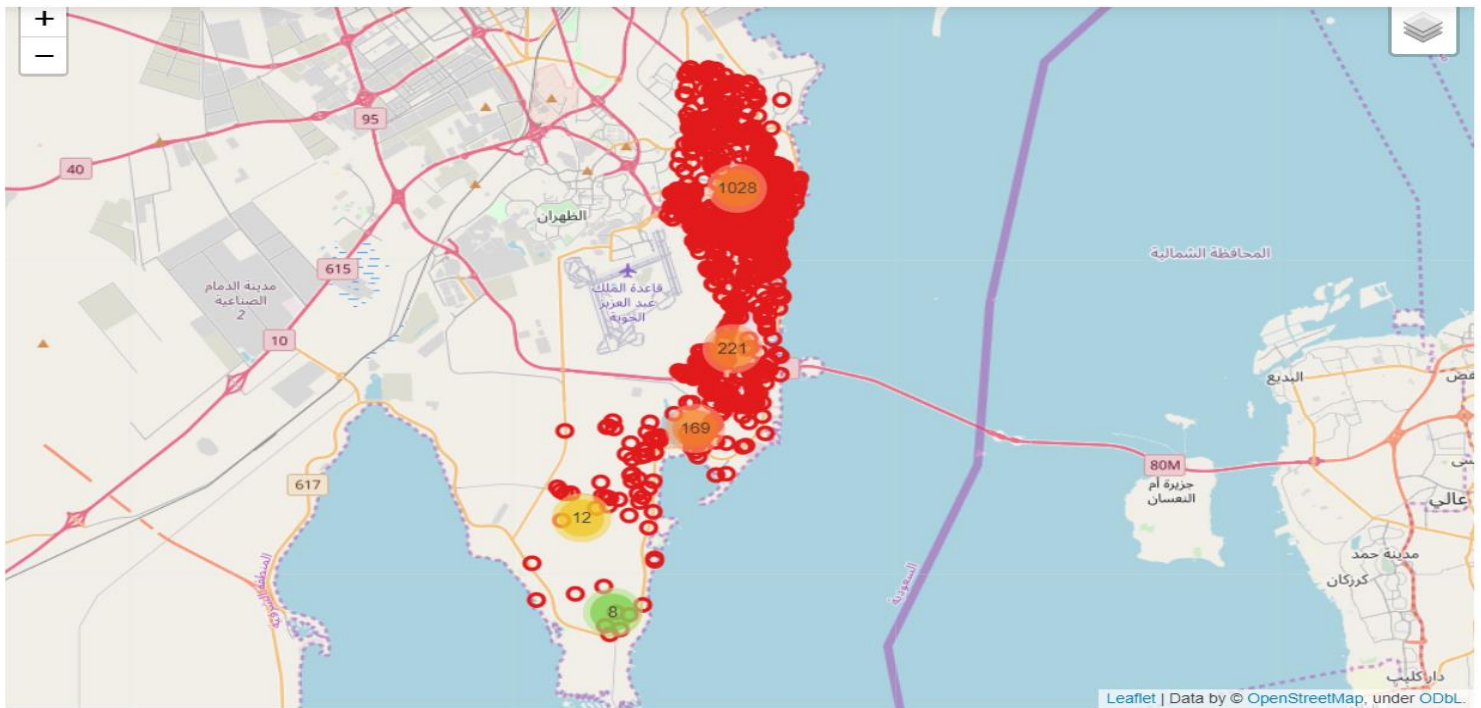
Now let us explore everybody's restaurant. Sandwiches are very popular here especially shawarma places.



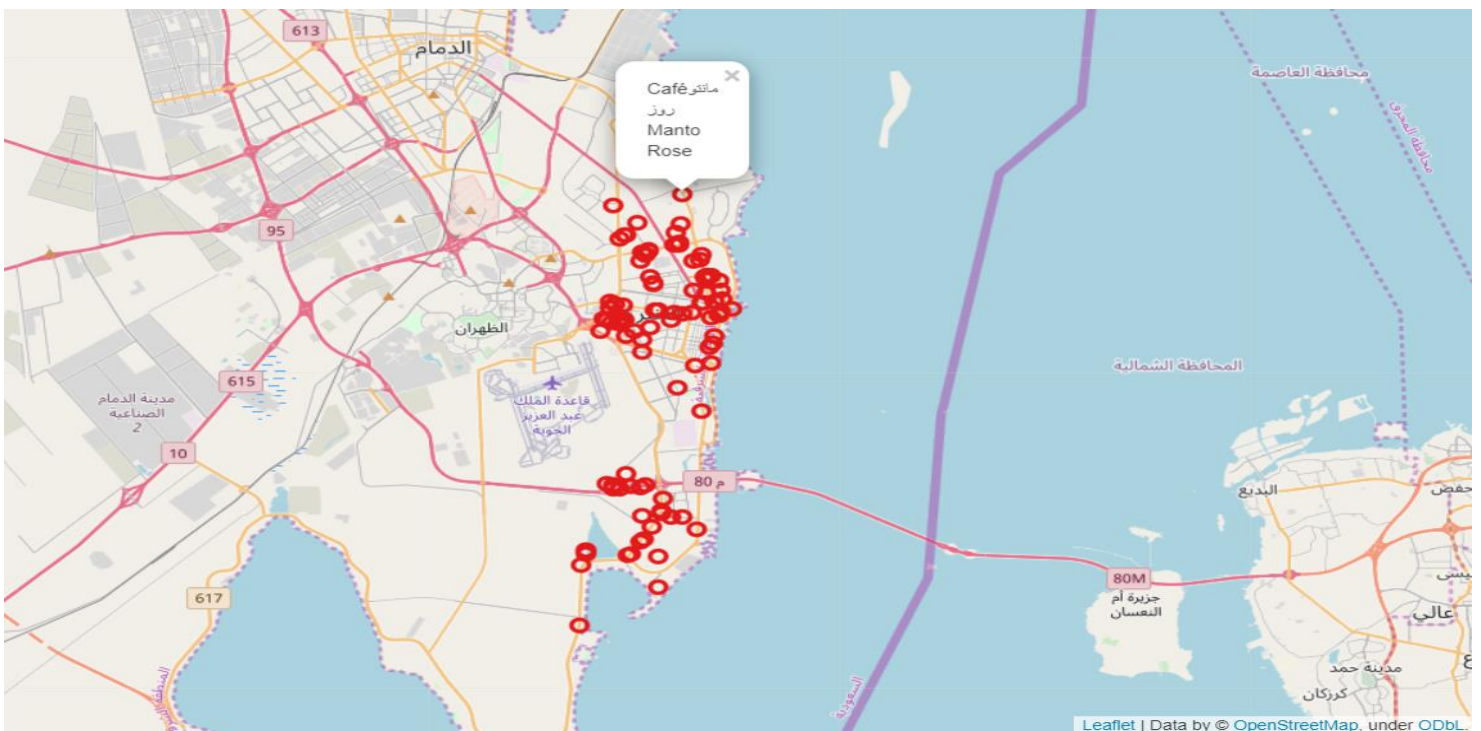
As you can see, sandwiches are very popular here. With Al Khubar Ash Shamaliyah having the biggest piece of the cake with around eleven shawarma places and a total number of sandwich places of eighteen.

Methodology

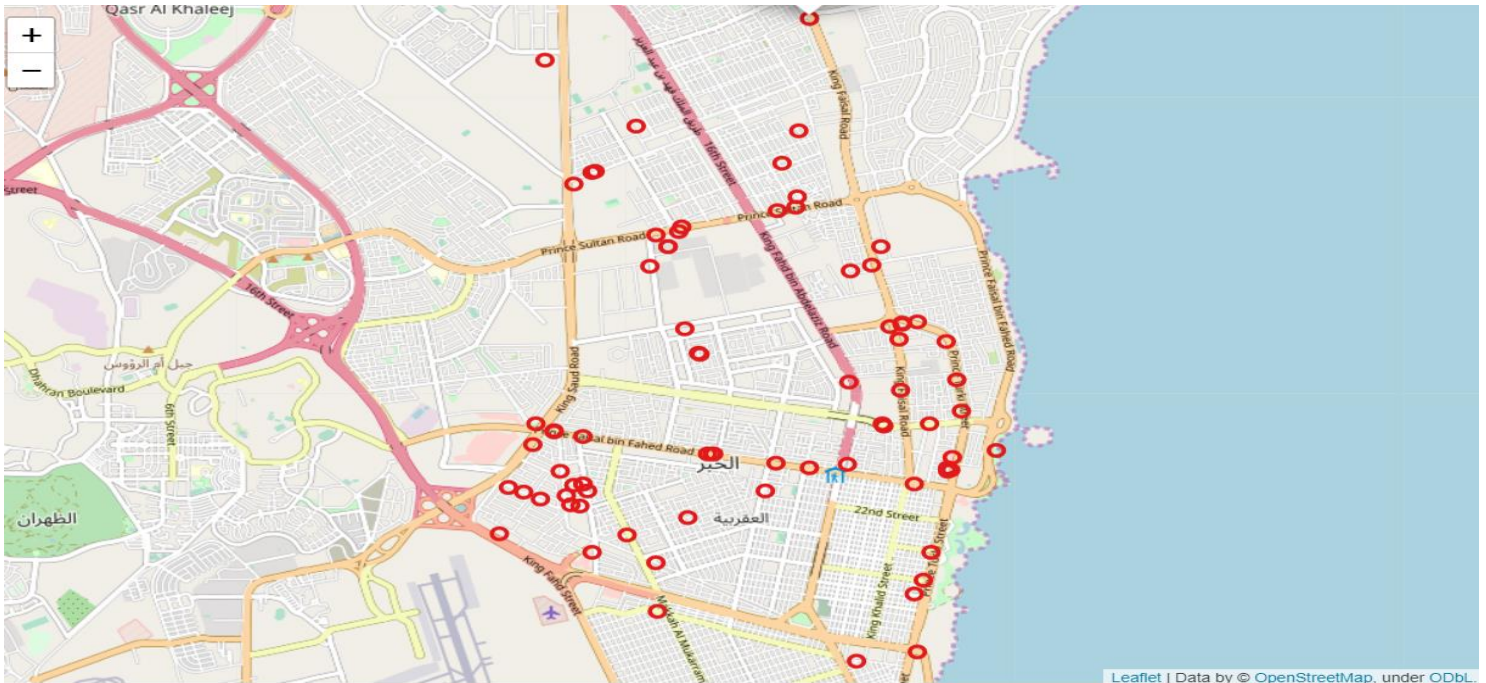
Now let us plot them all together on a map.



The map looks a bit messy. Let us filter out some places like coffee shops and plot those on the map.

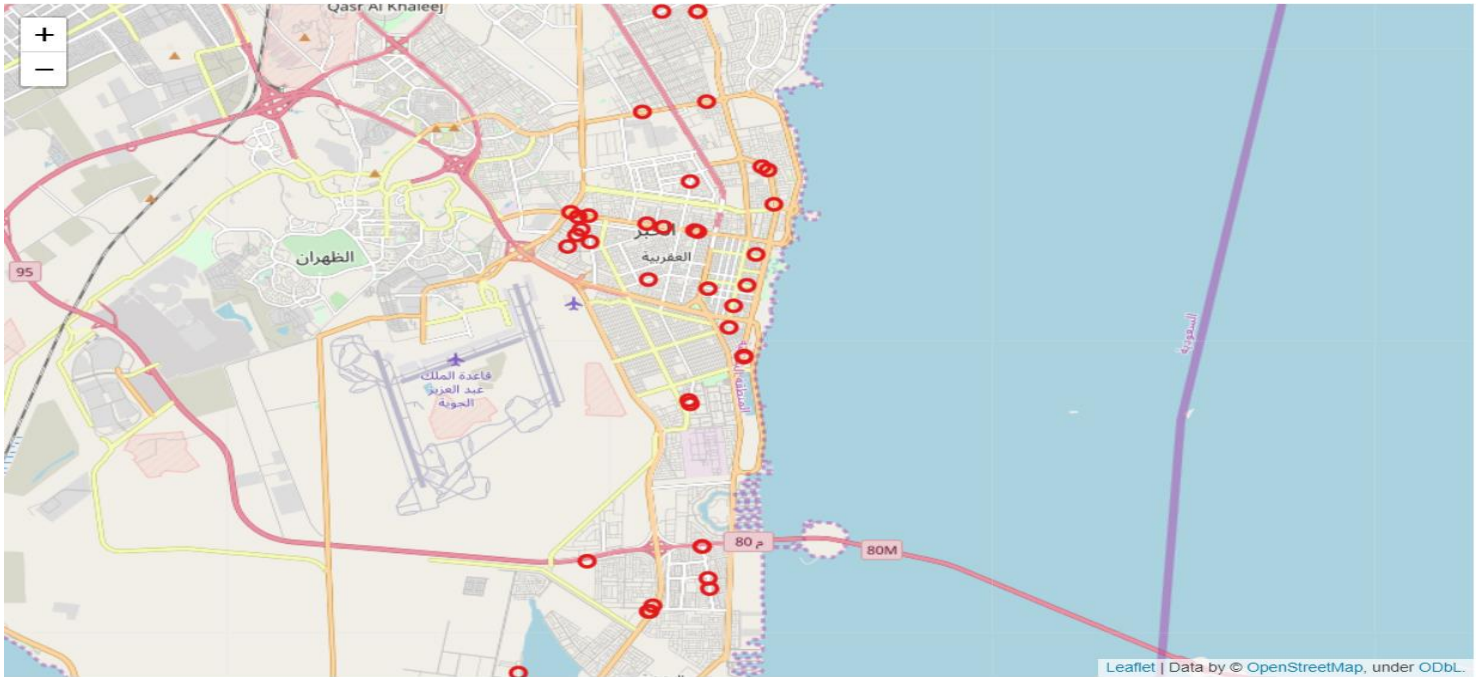


There seems to be a heavy concentration of coffee shops in the north of the city. Let us explore that.



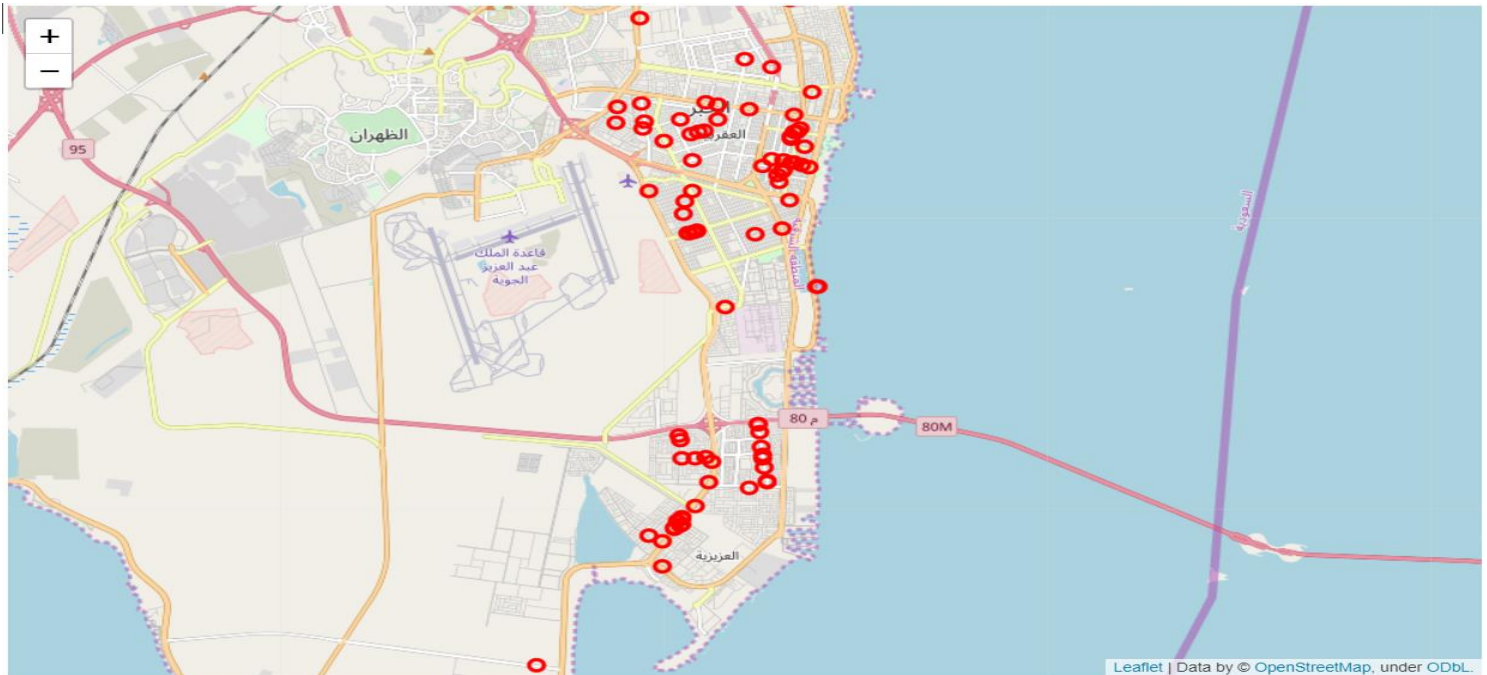
Coffee shops seem to be everywhere. It would not be a good idea to invest in one of these unless one can differentiate themselves from others.

Now let us explore the burger joints in the city.



As you can see, it is spread out across the city with a heavy concentration in the heavily populated areas near Al Hada and Al Hizam Al Akhdar. Ath Thogba seems to have no burger joint at all. A cheap burger joint there might be a good idea. Most of the burger joints are around two streets Prince Faisal Bin Fahd Road and Prince Turkey Street as those are the main two streets in Al Khobar.

Now let us explore the sandwich places in Al Khobar.



As you can see, sandwich places are very popular. Sandwich places are everywhere as they are cheap and very tasty.

Now let us look at the most popular venues in each neighborhood.

Methodology

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 amwaj	Beach	Coffee Shop	Supermarket	Fast Food Restaurant	Food Truck	ATM	Pharmacy	Outdoor Sculpture	Outlet Store
1	al andalus	Coffee Shop	Gym / Fitness Center	Trail	Home Service	Museum	Steakhouse	Border Crossing	Bowling Alley	Breakfast Spot
2	al aqiq	Beach		Coffee Shop	ATM	Pharmacy	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant
3	al aqrabiyah	Café	Bakery	Indian Restaurant	Health & Beauty Service	Lounge	Clothing Store	Coffee Shop	Shawarma Place	Hotel
4	al bahar	Bay	Outdoor Sculpture	ATM	Nail Salon	Optical Shop	Organic Grocery	Outlet Store	Pakistani Restaurant	Park
5	al bandariyah	Coffee Shop	Dessert Shop	Outlet Store	Middle Eastern Restaurant	Pizza Place	Food Truck	Hookah Bar	Gift Shop	Frozen Yogurt Shop
6	al buhayrah	Middle Eastern Restaurant	Coffee Shop	Café	Breakfast Spot	Fast Food Restaurant	Pastry Shop	Shawarma Place	Bakery	Dessert Shop
7	al bustan	Coffee Shop	Café	Burger Joint	Juice Bar	Bakery	Seafood Restaurant	Sushi Restaurant	Steakhouse	Market
8	al hada	Food Truck	Café	Coffee Shop	Burger Joint	Hotel	Arcade	Boutique	Flower Shop	Juice Bar
9	al hamra	Coffee Shop	Fast Food Restaurant	Burger Joint	Lounge	Bakery	Tourist Information Center	Grocery Store	Pizza Place	Shopping Mall
10	al hizam adh dhahabi	Coffee Shop	Café	Dessert Shop	Asian Restaurant	Restaurant	Shawarma Place	Park	Tailor Shop	Middle Eastern Restaurant
11	al hizam al akhdar	Coffee Shop	Dessert Shop	Café	Burger Joint	Italian Restaurant	Bakery	Gym / Fitness Center	Juice Bar	Ice Cream Shop
12	al iskan	Auto Workshop	Beach	Juice Bar	Middle Eastern Restaurant	Playground	Smoke Shop	Hookah Bar	Auto Garage	Fried Chicken Joint
13	al jawharah	Café	Coffee Shop	Pastry Shop	Food Truck	Gym	Garden	Health & Beauty Service	Health Food Store	Supermarket
14	al jisr	Coffee Shop	Café	Burger Joint	Donut Shop	Italian Restaurant	Bakery	Dessert Shop	Falafel Restaurant	Middle Eastern Restaurant
15	al kawthar	Farm	Campground	Park	French Restaurant	Casino	ATM	Outlet Store	Pakistani Restaurant	Pastry Shop
16	al khawr	Beach	River	ATM	Pharmacy	Optical Shop	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant
17	al khubar al janubiyah	Hotel	Bakery	Coffee Shop	Ice Cream Shop	Donut Shop	Burger Joint	Middle Eastern Restaurant	Gym	Dessert Shop
18	al khubar ash shamaliyah	Middle Eastern Restaurant	Coffee Shop	Breakfast Spot	Clothing Store	Furniture / Home Store	Fried Chicken Joint	Café	Falafel Restaurant	Indian Restaurant
19	al khuzama	Juice Bar	Café	Pizza Place	Shawarma Place	Pharmacy	Restaurant	Burger Joint	Coffee Shop	Cupcake Shop
20	al kurnaish	Coffee Shop	Café	Tea Room	Hookah Bar	French Restaurant	Hotel	Chocolate Shop	Mexican Restaurant	Rest Area
21	al lulu	Hookah Bar	ATM	Pharmacy	Optical Shop	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant	Park
22	al maha	Hookah Bar	Farm	Stables	Food Truck	Resort	Print Shop	Pop-Up Shop	Rental Car Location	Organic Grocery
23	al murjan	Beach	Farm	ATM	Pharmacy	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant	Park

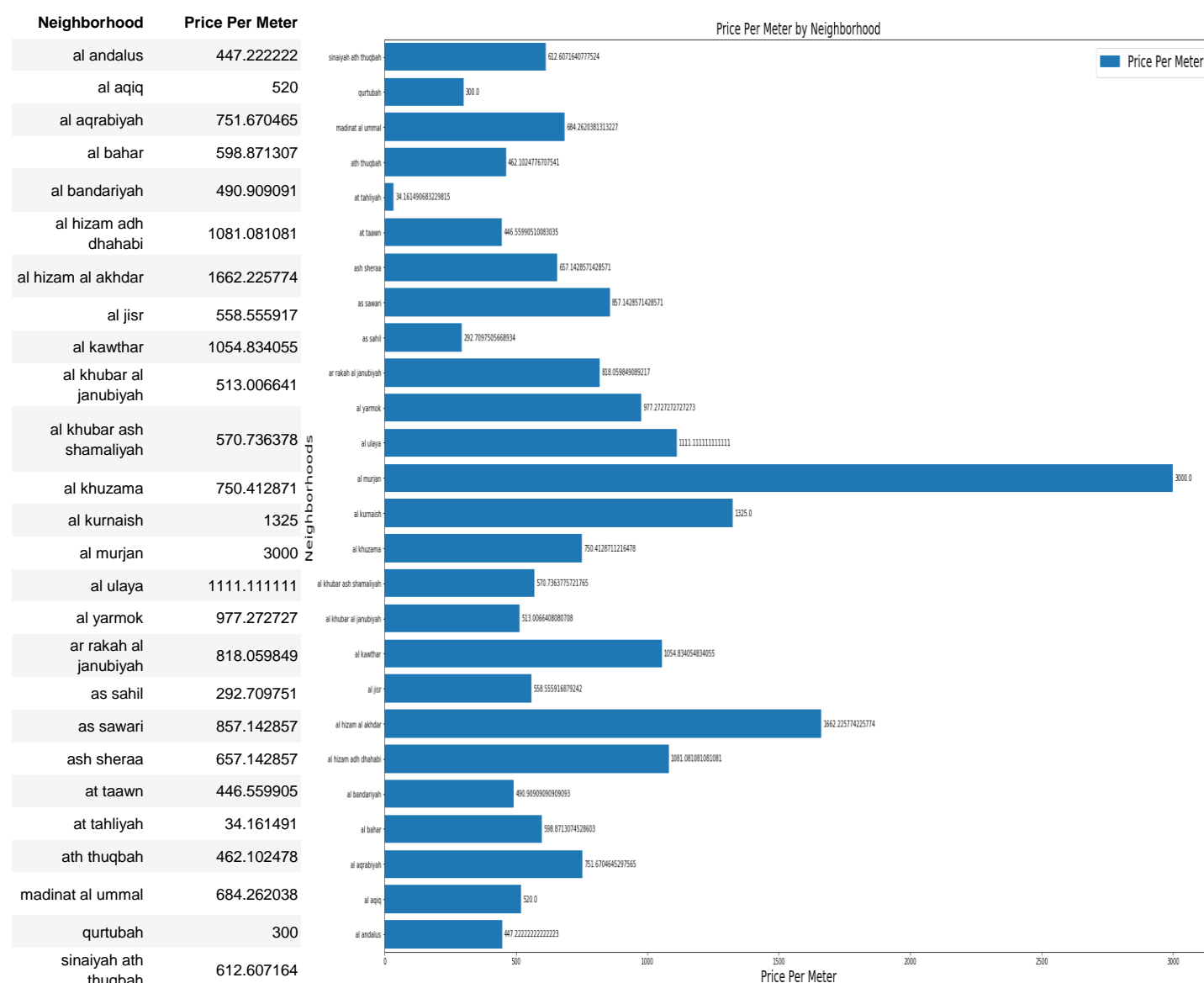
Methodology

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
24	al ulaya	Café	Coffee Shop	Food Truck	Jewelry Store	Health Food Store	Hotel	Furniture / Home Store	Shopping Mall	Donut Shop
25	al yarmok	Coffee Shop	Café	Food Truck	Dessert Shop	Hotel	Jewelry Store	Pharmacy	Donut Shop	Flower Shop
26	ar raja	Lounge	ATM	Pharmacy	Optical Shop	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant	Park
27	ar rakah al janubiyah	Coffee Shop	Hookah Bar	Middle Eastern Restaurant	Café	Bakery	Furniture / Home Store	Supermarket	Fried Chicken Joint	Burger Joint
28	ar rawabi	Coffee Shop	Café	Furniture / Home Store	Hotel	Fast Food Restaurant	Breakfast Spot	Indian Restaurant	Restaurant	Gym
29	as sahil	Restaurant	Vegetarian / Vegan Restaurant	Pharmacy	Optical Shop	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant	Park
30	as sawari	Café	Coffee Shop	Soccer Field	Seafood Restaurant	Hookah Bar	Beach	Athletics & Sports	Candy Store	Campground
31	as sufun	Boat or Ferry	Beach	Café	ATM	Noodle House	Organic Grocery	Outdoor Sculpture	Outlet Store	Pakistani Restaurant
32	ash sheraa	Hookah Bar	Resort	Soccer Field	Botanical Garden	Gym	Breakfast Spot	Falafel Restaurant	Summer Camp	Beach
33	at taawn	Hookah Bar	Café	Soccer Field	Restaurant	Rental Car Location	Rental Service	Athletics & Sports	Pet Store	Spa
34	at tahliyah	Café	Resort	Intersection	Beach	Turkish Restaurant	Coffee Shop	Pharmacy	Outdoor Sculpture	Outlet Store
35	ath thuqbah	Middle Eastern Restaurant	Market	Ice Cream Shop	Hotel	Coffee Shop	Pizza Place	Restaurant	Breakfast Spot	Fried Chicken Joint
36	ibn sina	Auto Workshop	Hookah Bar	Auto Garage	Restaurant	Track	Campground	Smoke Shop	Hotel	Shipping Store
37	ishbiliya	Creperie	Fried Chicken Joint	Motorcycle Shop	Surf Spot	Massage Studio	Café	ATM	Pharmacy	Outlet Store
38	madinat al ummal	Coffee Shop	Bakery	Seafood Restaurant	Gift Shop	Breakfast Spot	Burger Joint	Middle Eastern Restaurant	Gym / Fitness Center	Juice Bar
39	qurtubah	Café	Coffee Shop	Bakery	Gym	Spa	Soccer Stadium	Hookah Bar	Middle Eastern Restaurant	Insurance Office
40	sinaiyah ath thuqbah	Burger Joint	Soccer Field	Juice Bar	Astrologer	Café	Supermarket	Beach	Bakery	Tourist Information Center
41	sporting city	Hotel	Lounge	Pharmacy	Market	Coffee Shop	Lebanese Restaurant	Bar	Motorcycle Shop	Chocolate Shop

Interesting, Cafés and Coffee shops seems to be in almost every neighborhood among the top ten places.

Methodology

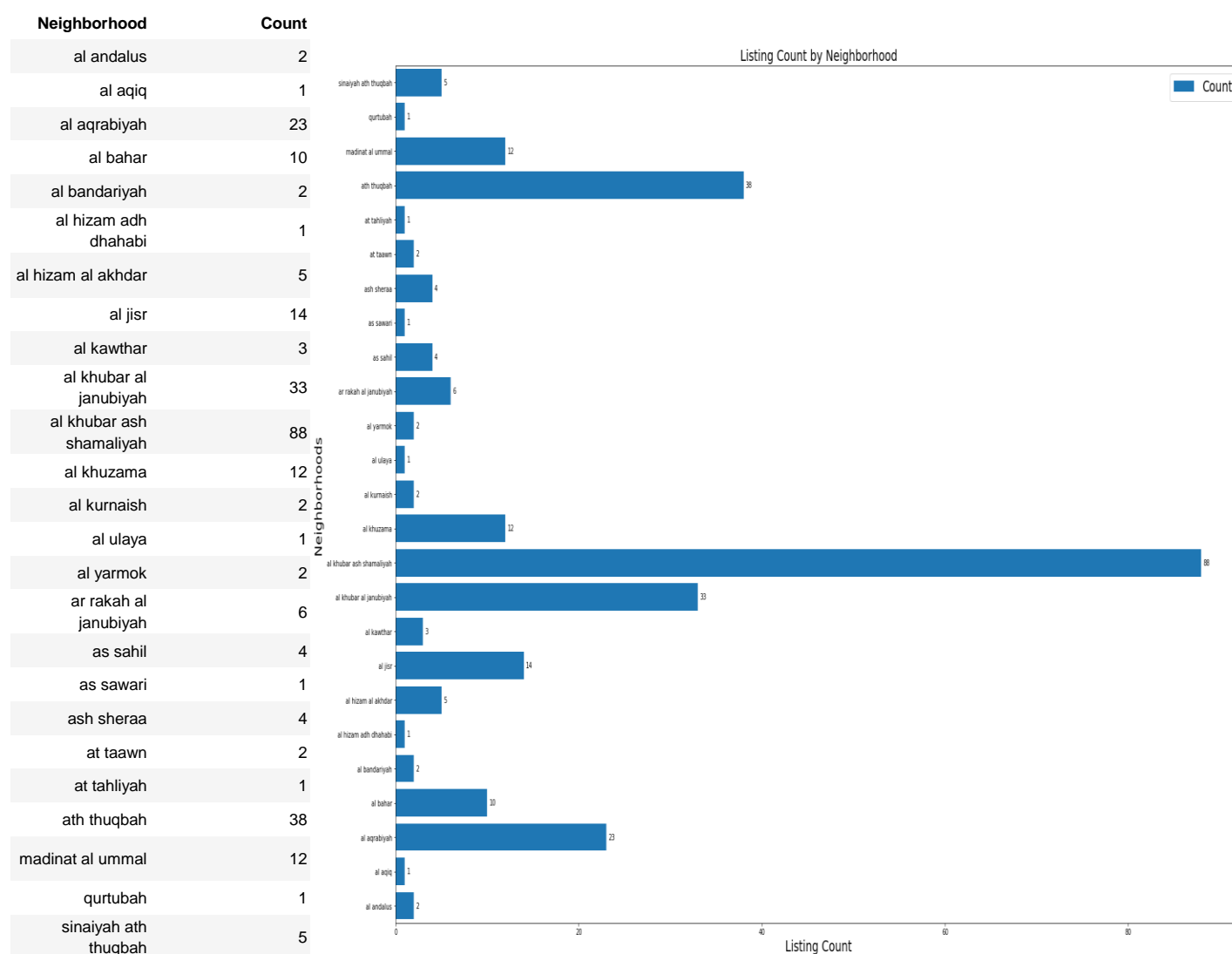
Now let us look at the supply data, places where its still available for rent on Aqar website. Let us look at the average price per meter for every neighborhood.



Investigating the highest which is Al Murjan, which is not that expensive, yielded that the place is in Adh Dhahran which is outside Al Khobar so we can drop it. The algorithm matched it to Al Murjan because Adh Dhahran and Al Murjan are close in Arabic. Investigating the lowest which is At Tahliyah yielded that the average price per

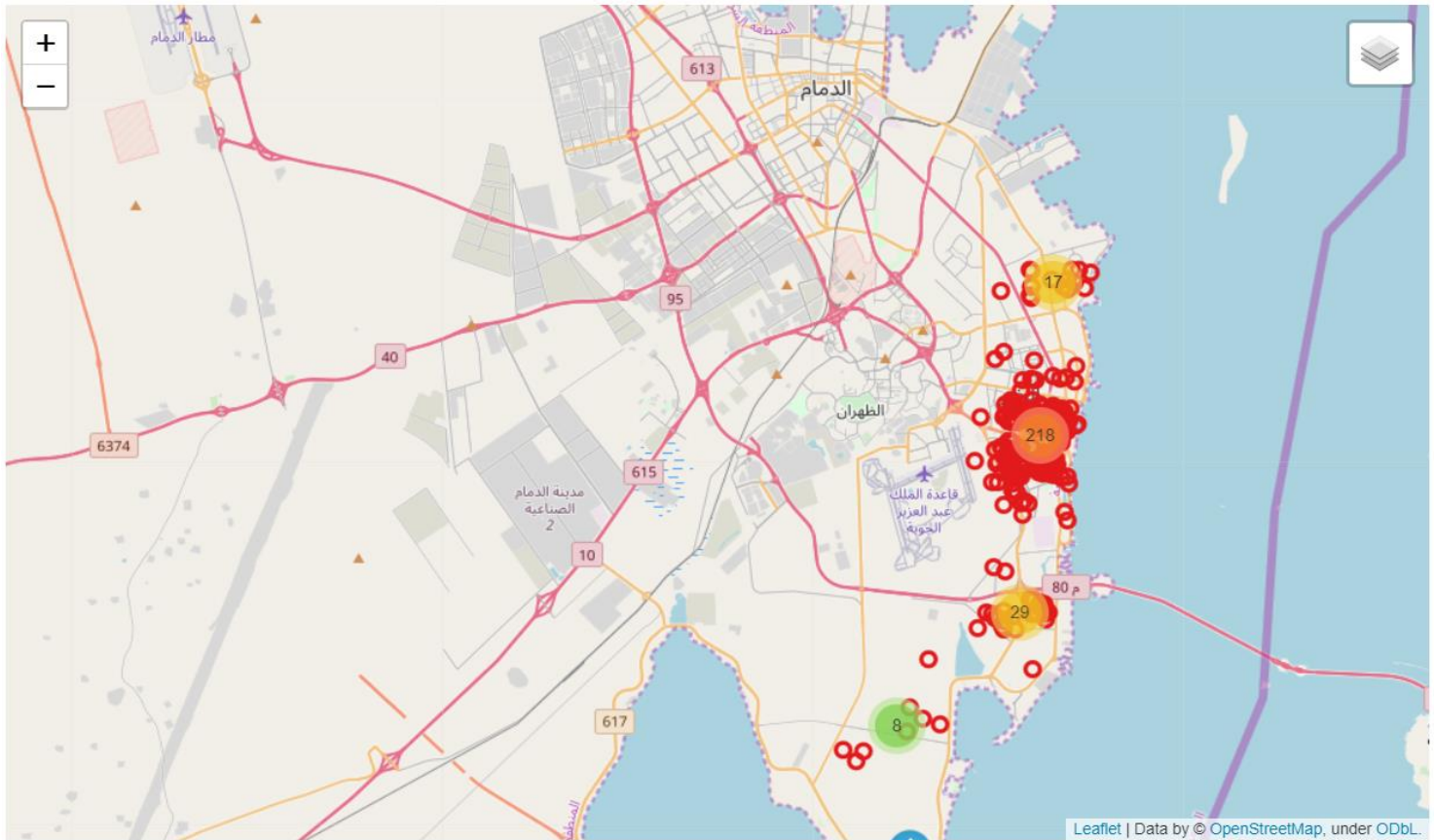
meter is one-thousand but the data was entered wrongly. Since the price is not a factor in our modeling, we shall keep it as is.

Let us look at the count of the listings per neighborhood



The highest number of listings is in Al Khubar Ash Shamaliyah with eighty-eight listing, which makes sense especially after the tighten up of the regulations concerning labor laws in attempt to make some retail trades to have more Saudis in it which resulted in a lot of empty stores because of the high cost of the Saudi labor.

Now, let us visualize it.



As you can see, most of the listings are concentrated in center of the city as expected.

Models

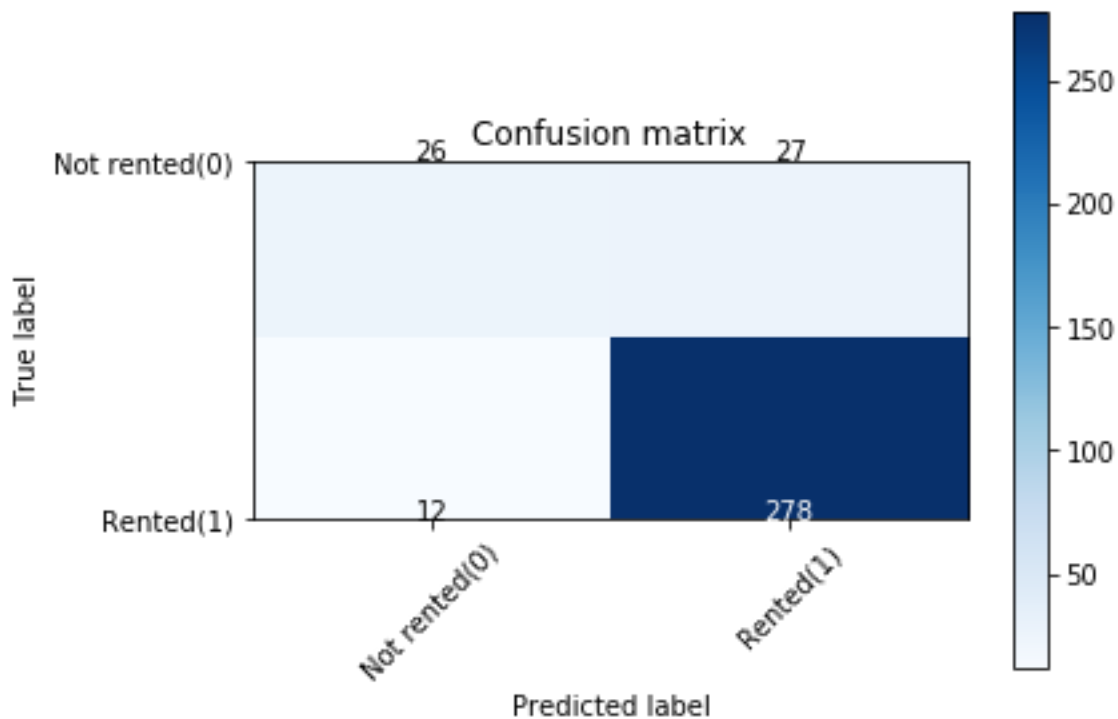
We are going to model the data using latitude and longitude data and rented or unrented (target). I expect that the best model would be K Nearest Neighbor would produce the best results as we only have the location data, we are going to model the data using several models to test whether any other model would produce any better results.

Results

Summary results is provided below:

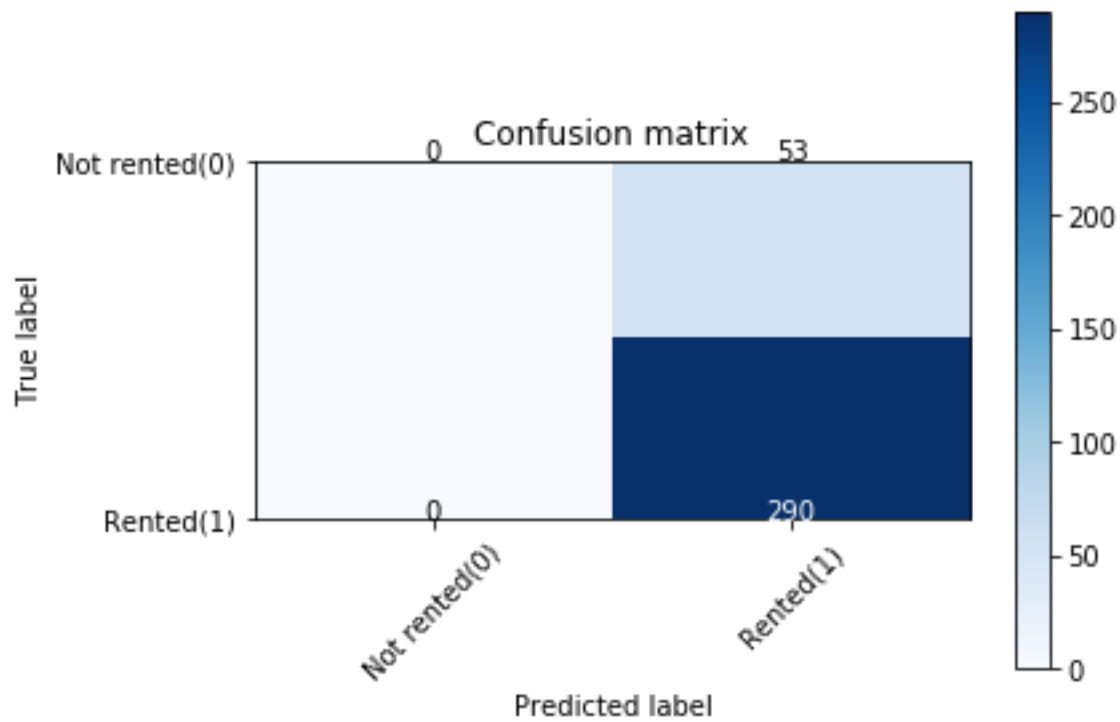
Model	KNN	Logistic Regression	SVM
Model Specifications	algorithm='auto', leaf_size=30, metric='euclidean', metric_params=None, n_jobs=None, n_neighbors=3, p=2, weights='distance'	C=0.0001, class_weight=None, dual=False, fit_intercept=True, intercept_scaling=1, l1_ratio=None, max_iter=100, multi_class='warn', n_jobs=None, penalty='l2', random_state=None, solver='liblinear', tol=0.0001, verbose=0, warm_start=False	C=10, cache_size=200, class_weight=None, coef0=0.0, decision_function_shape='ovr', degree=3, gamma=1, kernel='rbf', max_iter=-1, probability=False, random_state=None, shrinking=True, tol=0.001, verbose=False
Accuracy	0.8862973760932945	0.8454810495626822	0.8483965014577259
F1	0.934453781512605	0.9162717219589257	0.9174603174603174
Jaccard Similarity	0.8862973760932945	0.8454810495626822	0.8483965014577259

Now, let us explore the confusion metrics for each model. We will begin with the KNN model:



As you can see the model is excellent in predicting true values and moderately good in predicting negatives. Since the negatives are randomize points and are not actual points, the model is very good.

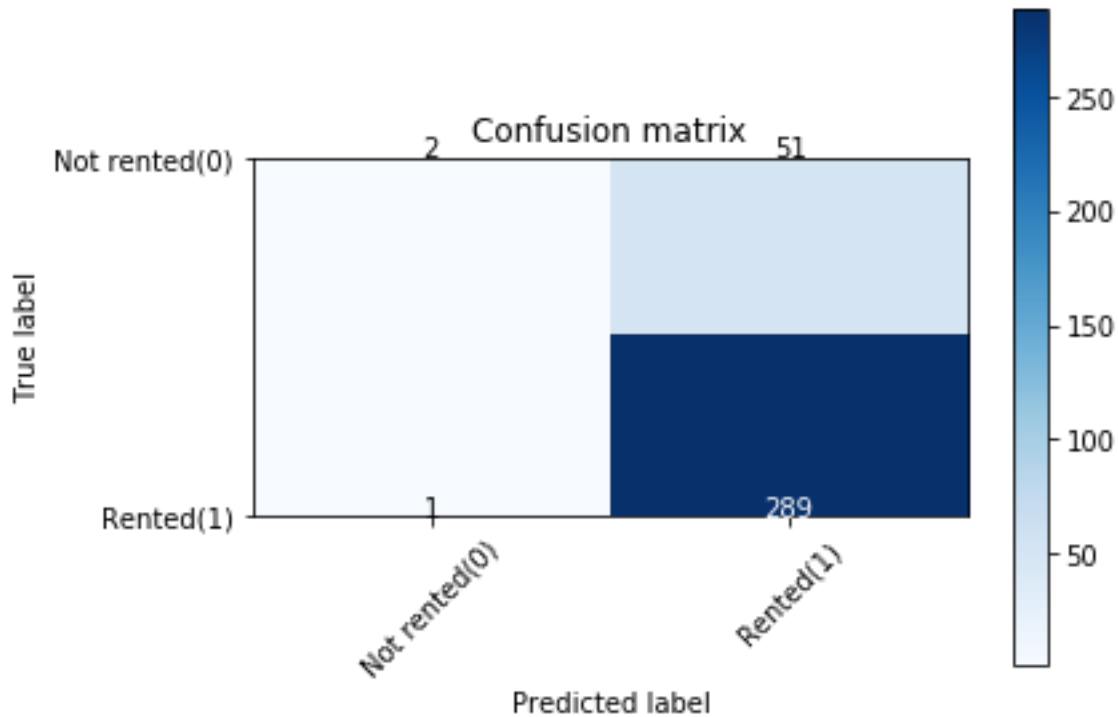
Let us explore the Logistic Regression confusion matrix:



The model seems to predict everything as rented (positive) which is not good.

Discussion

Let us explore the Support Vector Machine convolution matrix:



The model seems to predict mostly positive values which is not good.

Discussion

As expected, the best performing model is the K Nearest Neighbor with an accuracy of 88.6% and an acceptable confusion matrix. It was very good that we ended up with a model with such an accuracy given that the unrented points are randomized based on the neighborhood they are located in as mentioned in the data processing section of this report.

Conclusion

We started with three data sets in order to obtain the location data for the rented places (Foursquare data) and unrented data (Aqar data) which was fuzzy matched with the neighborhood data in order to obtain there location and then they were randomized based on a normal distribution with a mean point as the center of the neighborhood and a standard deviation equal to the third of the distance between the center and the edge of the neighborhood. We then modeled the data using KNN model which was the best model with acceptable results given that the only inputs are the location data and those are randomize for the unrented location.

Contact Information

**Name**

Bader Ababtain

Tel +966562333199

b.ababtain@gmail.com