Recommending Places To Stay in Doha, Qatar
Applied Capstone Project

Noora Al-Mana

1. INTRODUCTION

1.1 Background:
Qatar is quickly positioning itself as a leading tourist destination in the Middle East. With its growing tourism economy, as well as hosting the FIFA World Cup in 2022, it is expected to host millions of people from around the world. Just a few years ago, the world has not heard of this tiny, oil rich country. In recent years, it has vastly developed to be a diverse and culture-rich city in the region, thus attracting many expatriates and tourists to the country. As such, many new parks, attractions, museums, restaurants, and hotels have opened to cater to the growing demand.

1.2 The Problem:
Foreigners visiting the country have no understanding of the city or key places of interest to see or stay in. Hence, the challenge to resolve is to:
   (1) help tourists visualize the zones and cities in the country,
   (2) gain an understanding of what each area has to offer,
   (3) group them based on their similarities of venues, and finally,
   (4) find hotels to stay in based on their preferences and chosen area.

2. DATA ACQUISITION AND CLEANING

2.1 Data Sources:
In order to complete this project, I would need data from multiple sources. The first is a list of zones in Qatar. This can be found in the Qatar Open Data Portal. The dataset that has the information we need is called, "Population, Area and Population Density Per Square Kilometers By Zone in April 2015" and can be downloaded as a csv file. Because we are only interested in the zones data, it does not matter that the dataset is a little outdated. (https://www.data.gov.qa/explore/dataset/population-area-and-population-density-per-square-kilometers-by-zone-in-april-20/information/).

I will then use Geocoder from the Geopy library in Python to find the latitude and longitude values of each zone to plot.

Then, I will use the Foursquare API to find the most common venues in each zone to understand what is readily accessible in each, and cluster them based on their similarities.

Noora Al-Mana

Lastly, I will use BeautifulSoup4 to scrape Booking.com to obtain the data of hotels in our zone of choice, and compare the hotels available based on their location, star class, rating, and price.

2.2 Data Cleaning:
The dataset obtained from Qatar Open Data Portal has multiple names for the zones in Qatar as well as redundant rows for the same value. Below is a snapshot of the original dataset:

	No. of Zone	Zone	Population	Area in KM	Population Density Km2
1	4	Mushaireb	9,791	0.326	30,003.094
2	13	Mushaireb	18,278	0.674	27,114.656
3	14	Fereej Abdel Aziz	15,706	0.532	29,513.417
4	16	Old Al Ghanim	16,334	0.412	39,661.660
5	17	Al Rufaa / Old Al Hitmi	6,026	0.396	15,230.862
6	20	Wadi Al Sail	240	1.314	182.645
7	21	Rumaila	1,590	0.790	2,012.045
8	23	Fereej Bin Mahmoud	17,086	1.162	14,707.121
9	26	Najma	28,228	1.136	24,855.415
10	32	Madinat Khalifa North / Dahl Al Hamam	12,364	2.419	5,110.710

I therefore cleaned the dataset by dropping redundant rows and renaming the zones correctly.

|	No. of Zone	Zone Number	Zone Name	Zone	Population	Area in KM	Population Density Km2
0	4	Zone 4, Qatar	Mushaireb, Qatar	Mushaireb	9791	0.326333	30003.093700
1	5	Zone 5, Qatar	Al Najada, Qatar	Al Najada	2086	0.219777	9491.432859
2	6	Zone 6, Qatar	Old Al Ghanim, Qatar	Old Al Ghanim	1124	0.284367	3952.637541
3	7	Zone 7, Qatar	Al Souq, Qatar	Al Souq	297	0.280341	1059.425683
4	13	Zone 13, Qatar	Mushaireb, Qatar	Mushaireb	18278	0.674100	27114.656400

Lastly, we get the latitude and longitude values of each zone from Geocoder, and drop all the unnecessary columns, leaving us with the below data frame:

	Zone	latitude	longitude
0	Abu Samra	25.285061	51.552715
1	Al Adaid	25.285061	51.552715
2	Al Dafna	25.319584	51.536284
3	Al Doha Al Jadeeda	25.277164	51.532286
4	Al Gharrafa	25.327770	51.454294

Noora Al-Mana

3. METHODOLOGY

3.1 Exploratory Data Analysis:
First, I visualized the zones in Doha to understand their geography before I can cluster them.
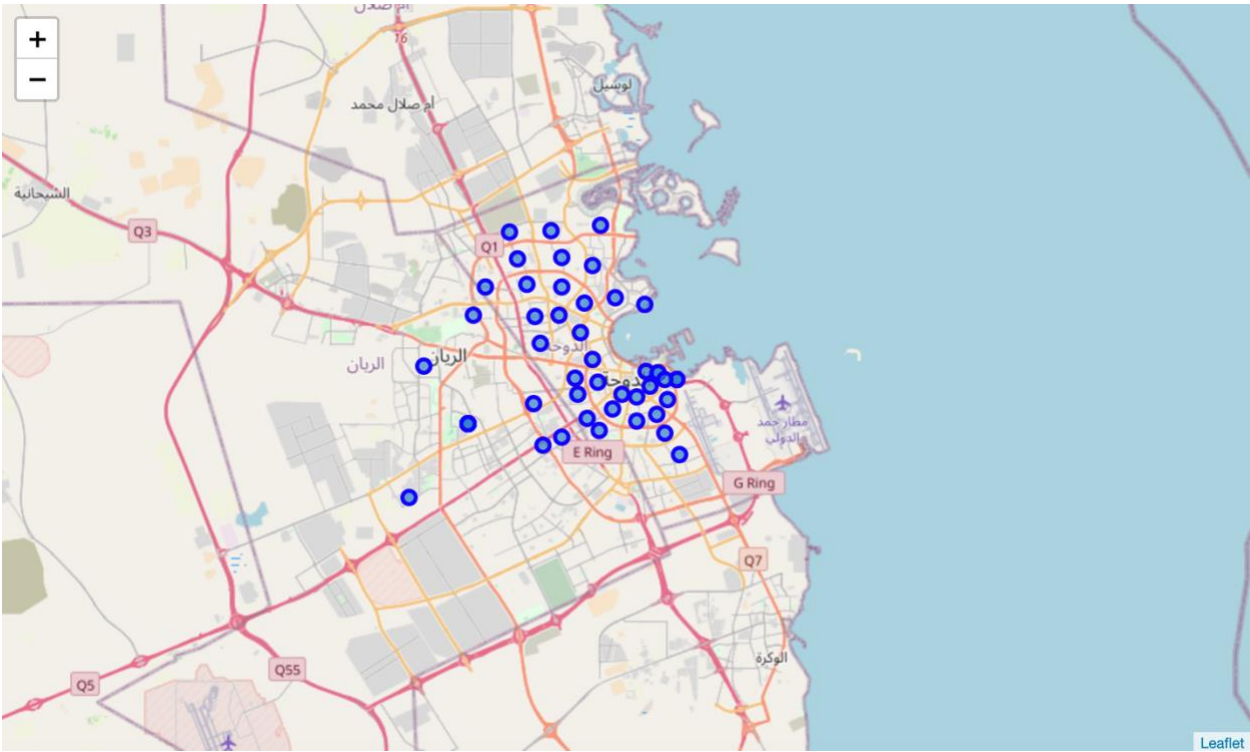
3.2 Finding Venues In Each Cluster:
I used the Foursquare API to find the different types of venues in each zone. I then used one hot encoding to create dummy variables of each category of venue returned. Lastly, I grouped them by zone and got the mean of each category, leaving me with the below data frame:

	Zone	Accessories Store	Afghan Restaurant	African Restaurant	Airport	Airport Lounge	Airport Service	American Restaurant	Aquarium	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	A Entertainr
0	Abu Samra	0.01	0.0	0.00	0.0	0.0	0.0	0.040000	0.0	0.00	0.00	0.0	0.0	
1	Al Adaid	0.01	0.0	0.00	0.0	0.0	0.0	0.040000	0.0	0.00	0.00	0.0	0.0	
2	Al Dafna	0.00	0.0	0.01	0.0	0.0	0.0	0.000000	0.0	0.01	0.00	0.0	0.0	
3	Al Doha Al Jadeeda	0.00	0.0	0.01	0.0	0.0	0.0	0.000000	0.0	0.00	0.01	0.0	0.0	
4	Al Gharrafa	0.00	0.0	0.00	0.0	0.0	0.0	0.024691	0.0	0.00	0.00	0.0	0.0	

Finally, I sorted the venues in descending order to get the top 10 most common venue types in each zone.

	Zone	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	Abu Samra	Coffee Shop	Café	Fast Food Restaurant	American Restaurant	Ice Cream Shop	Clothing Store	Hotel	Burger Joint	Movie Theater	Thai Restaurant
1	Al Adaid	Coffee Shop	Café	Fast Food Restaurant	American Restaurant	Ice Cream Shop	Clothing Store	Hotel	Burger Joint	Movie Theater	Thai Restaurant
2	Al Dafna	Hotel	Coffee Shop	Café	Italian Restaurant	Lounge	Restaurant	Bar	Steakhouse	Asian Restaurant	French Restaurant
3	Al Doha Al Jadeeda	Hotel	Café	Middle Eastern Restaurant	Restaurant	Bakery	Italian Restaurant	Mediterranean Restaurant	Coffee Shop	Turkish Restaurant	Seafood Restaurant
4	Al Gharrafa	Café	Restaurant	Burger Joint	Clothing Store	Coffee Shop	Fast Food Restaurant	Shopping Mall	Bakery	Sandwich Place	Grocery Store

3.3 K-Means Clustering:

To cluster the zones based on their similarities, I first had to find the optimal number of K clusters. I used the Elbow method to do so.
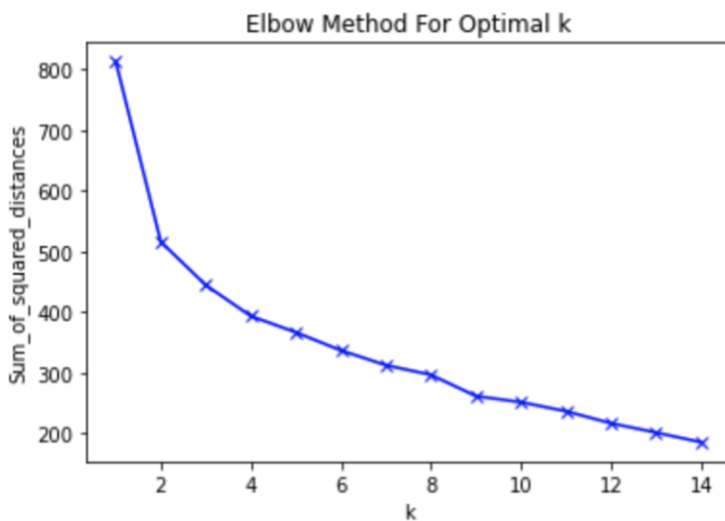
Although it is ambiguous, I determined the optimal K to be 4 clusters.

I clustered the zones in Qatar into 4 clusters, and mapped them using Folium to help me visualize their locations:

	Zone	latitude	longitude	Cluster Labels	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 M... Comm Ven...
0	Abu Samra	25.264376	51.445310	0.0	Coffee Shop	Café	Fast Food Restaurant	American Restaurant	Ice Cream Shop	Clothing Store	Hotel	Burger Joint	Mc Thea...
1	Al Adaid	25.264376	51.445310	0.0	Coffee Shop	Café	Fast Food Restaurant	American Restaurant	Ice Cream Shop	Clothing Store	Hotel	Burger Joint	Mc Thea...
2	Al Dafna	25.319584	51.536284	3.0	Hotel	Coffee Shop	Café	Italian Restaurant	Lounge	Restaurant	Bar	Steakhouse	As Restaur...
3	Al Doha Al Jadeeda	25.277164	51.532286	3.0	Hotel	Café	Middle Eastern Restaurant	Restaurant	Bakery	Italian Restaurant	Mediterranean Restaurant	Coffee Shop	Turk Restaur...
4	Al Gharrafa	25.327770	51.454294	2.0	Café	Restaurant	Burger Joint	Clothing Store	Coffee Shop	Fast Food Restaurant	Shopping Mall	Bakery	Sandw Pl...

Analyzing the venues and locations of each zone, I have determined that my preferred zone to stay in during my vacation is Al Dafna, which is locally and more commonly referred to as West Bay.

### 3.4 Scraping Hotels From Booking.com Using BeautifulSoup4:

Finally, I scraped the information about hotels in Qatar from Booking.com using BeautifulSoup4 and saved them in a Pandas data frame. I saved the hotel name, star class, rating, type of room available, price for my stay, as well as a link to their website.

	Hotel	Stars	Rating	Price in CAD $	Room Type	Link
0	The St. Regis Doha	5-star hotel	8.5	1426	Superior Double, Guest room, 2 Double	https://www.booking.com/hotel/qa/the-st-regis-...
1	The Ritz-Carlton, Doha	5-star hotel	8.7	1505	Deluxe King Room	https://www.booking.com/hotel/qa/the-ritz-carl...
2	Four Seasons Hotel Doha	5-star hotel	8.9	1762	Deluxe King Room with City View	https://www.booking.com/hotel/qa/four-seasons-...
3	Grand Hyatt Doha Hotel & Villas	5-star hotel	8.3	1314	Grand 2 Twins	https://www.booking.com/hotel/qa/grand-hyatt-d...
4	Magnum Hotel & Suites West Bay	5 stars	7.7	1253	Three-Bedroom Suite	https://www.booking.com/hotel/qa/governor-west...
5	Hilton Doha The Pearl Residences	5-star hotel	8.3	983	King Studio	https://www.booking.com/hotel/qa/hilton-doha-t...
6	Hilton Doha	5-star hotel	8.3	1569	Twin Room	https://www.booking.com/hotel/qa/hilton-doha.e...
7	Sheraton Grand Doha Resort & Convention Hotel	5 stars	8.9	1433	Deluxe Twin Room	https://www.booking.com/hotel/qa/sheraton-doha...
8	Intercontinental Doha - The City	5-star hotel	8.5	1044	Deluxe Room	https://www.booking.com/hotel/qa/intercontinen...
9	InterContinental Doha Hotel	5-star hotel	8.8	1355	Superior King Room with City View	https://www.booking.com/hotel/qa/intercontinen...
10	Kempinski Residences & Suites, Doha	5 stars	8.6	1101	Superior One-Bedroom Suite	https://www.booking.com/hotel/qa/kempinski-res...
11	Centara West Bay Residences & Suites Doha	5-star hotel	8.4	826	Standard Studio - Queen Bed	https://www.booking.com/hotel/qa/centara-west-...

4. RESULTS

4.1 K-Means Clustering:
I found that Qatar is segmented into 70 zones. In these 70 zones, there are 200 unique categories of venues in Qatar.
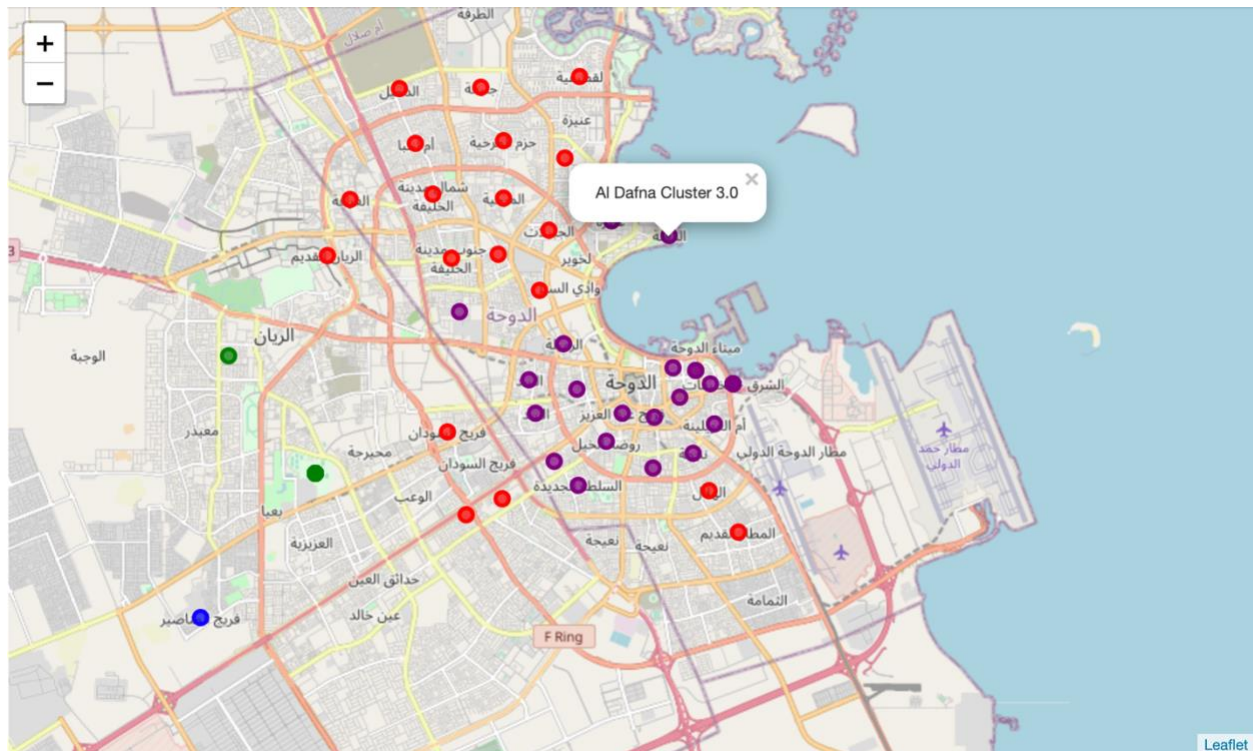
Clustering them into 4 groups, I found that Clusters 0 (Green) and 1 (Blue) are very suburban and are far away from the city.

Cluster 2 (Red) is the biggest and consists of many cafes, restaurants, and shopping malls, while being very residential with not many hotels.

Cluster 3 (Purple) is in the city, has many hotels, and a diverse set of restaurants and eateries.

Now that I have a bit more of an understanding of the layout of the city, I would like to stay in a hotel in Cluster 3 to experience the diversity in the city. Moreover, I also would like to be near the beach to take advantage of the hot weather in the area. Hence, Al Dafna (also commonly known as West Bay by the locals) is my ideal area to stay in during my visit.

4.2 Scraping Hotels:
Using BeautifulSoup4 to scrape hotels from Booking.com, I have found 25 hotels in West Bay I could stay in during my trip from December 1st, 2020 – December 6th, 2020.

Noora Al-Mana

I narrowed it down by filtering hotels with a rating higher than 8.6 and was left with my top 10 hotels.

	Hotel	Stars	Rating	Price in CAD $	Room Type	Link
0	The Ritz-Carlton, Doha	5-star hotel	8.7	1505.0	Deluxe King Room	https://www.booking.com/hotel/qa/the-ritz-carl...
1	Four Seasons Hotel Doha	5-star hotel	8.9	1762.0	Deluxe King Room with City View	https://www.booking.com/hotel/qa/four-seasons-...
2	Sheraton Grand Doha Resort & Convention Hotel	5 stars	8.9	1433.0	Deluxe Twin Room	https://www.booking.com/hotel/qa/sheraton-doha...
3	InterContinental Doha Hotel	5-star hotel	8.8	1355.0	Superior King Room with City View	https://www.booking.com/hotel/qa/intercontinen...
4	Rabban Suites West Bay Doha	5-star hotel	8.7	806.0	Deluxe Room	https://www.booking.com/hotel/qa/fraser-suites...
5	W Doha	5-star hotel	8.9	1139.0	Spectacular, Queen Room, Guest Room, 2 Double ...	https://www.booking.com/hotel/qa/w-doha-qatar....
6	JW Marriott Marquis City Center Doha	5-star hotel	8.8	1557.0	Deluxe King Room, Guest room, 1 King	https://www.booking.com/hotel/qa/jw-marriott-m...
7	City Centre Rotana Doha	5-star hotel	8.8	918.0	King Classic Room	https://www.booking.com/hotel/qa/city-centre-r...
8	Marriott Marquis City Center Doha Hotel	5-star hotel	8.8	1234.0	Deluxe Room, Guest room, 1 King	https://www.booking.com/hotel/qa/marriott-marq...
9	Dusit Doha Hotel	5-star hotel	9.0	1091.0	Superior King Room	https://www.booking.com/hotel/qa/dusit-doha.en...

Hotel Prices in West Bay for 5 Nights in December 2020

Noora Al-Mana
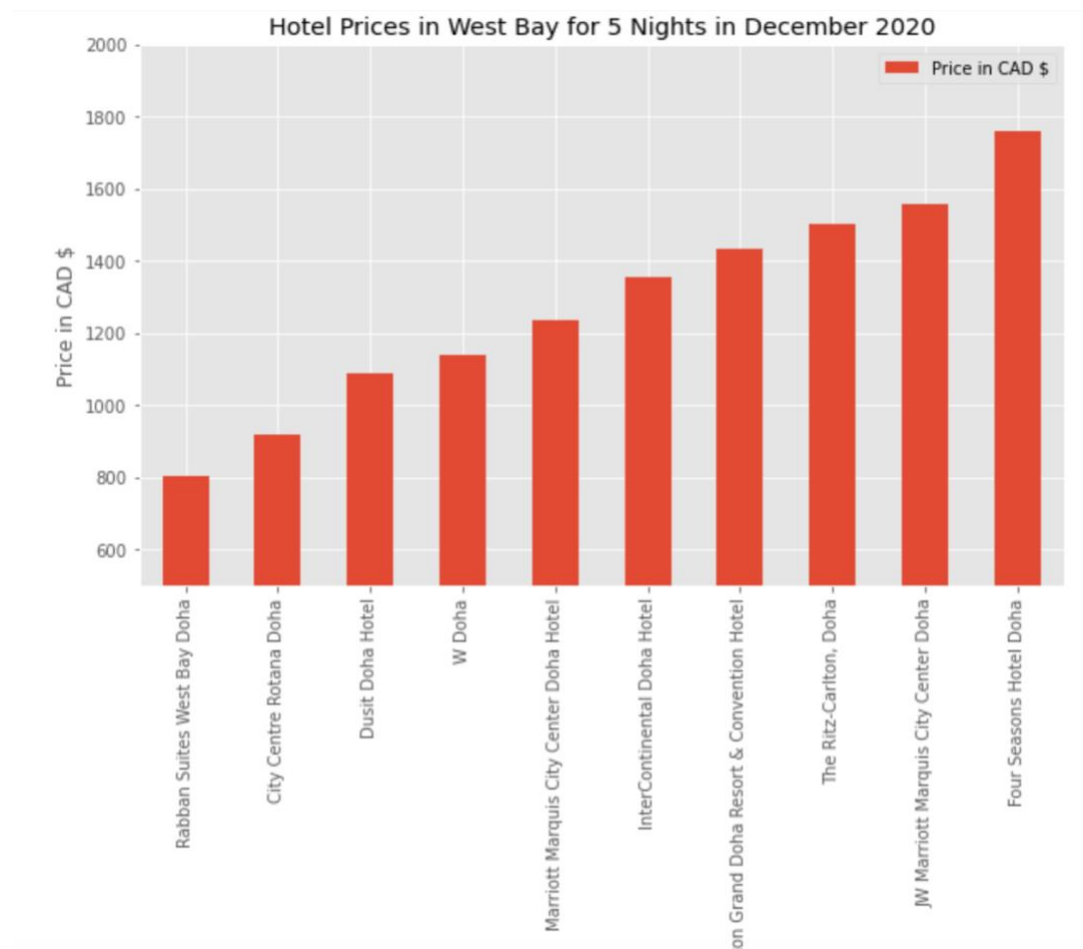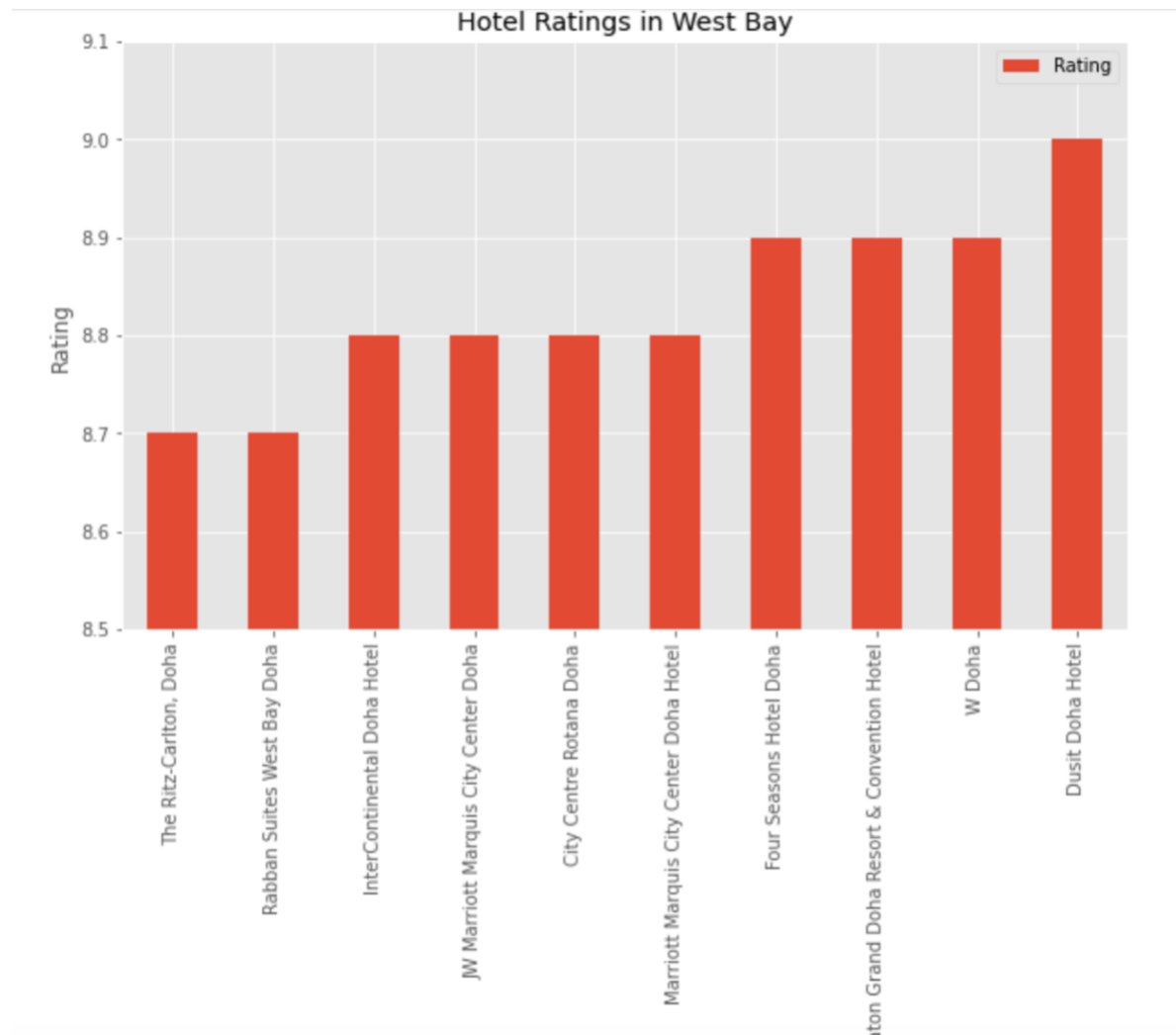
The average price for a 5-night stay in a 5-star hotel in West Bay in December is CAD $1280.

The highest rated 5-star hotel is Dusit Doha Hotel (9.0) and the most expensive one is the Four Seasons ($1762). For my trip, I will also consider W Doha as it has the second highest rating while also being averagely priced, which could be a great deal for luxury and affordability for our trip. Therefore, my top 3 choices are Dusit Doha, The Four Seasons, and W Doha.

I plotted these 3 hotels on a map to help me visualize their locations in the city:

5. DISCUSSION

The goal of this project is to help me decide where to stay during my vacation to Qatar from December 1st, 2020 to December 6th, 2020. The first observation I had was that choosing an optimal K was rather ambiguous from the Elbow Method, which leads me to believe that it is rather difficult to differentiate between the clusters.

I would like to treat myself to the ultimate luxury vacation, as such, West Bay is the best place to stay during my trip. It boasts many hotels, a diverse array of restaurants and cafes, shopping, as well as accessible clean beaches. It also has the highest concentration of 5-star hotels, which have an average price of $1280 for a 5-night stay.

Based on the prices and ratings, I have narrowed it down to my top 3 choices in hotels: Dusit Doha Hotel, W Doha, and the Four Seasons Doha. The W Hotel is far from West Bay though so it is not ideal for me. The remaining two are both great options however, The Four Seasons is the best choice in terms of luxury, as well as being directly on the beach which is important to me during my vacation.

Noora Al-Mana

6. CONCLUSION

This project aimed to help me understand the layout and geography of Doha, Qatar. It helped me visualize the zones and cluster them based on their similarities. Finally, it helped me choose my top 3 hotels to stay in while visiting in December of 2020.

Although The Four Seasons Hotel is my final choice, many other factors can influence my trip, including fluctuating prices as we grow closer to the dates of travel, as well as type of rooms available. Moreover, my dates of choice are during peak season of tourism in Qatar as many people choose to travel during the winter months to take advantage of the moderate climate.