Introduction

Ever since Crown Prince Mohammad bin Salman (MBS) has become second in line to the Saudi Arabian Crown, many have hoped the Kingdom would open up to modernity and liberalize their society and economy. While many initiatives have been launched under MBS like the Vision 2030 plan there are many reasons to also be skeptical about the reform narrative. While the Vision 2030 plan offers many ways for investors to make money in Saudi Arabia, recent controversies make it hard to know what is actually happening inside the kingdom.

By using Foursquare data and KNN algorithms, I hope to analyze how similar or dissimilar the major cities in Saudi Arabia are. My hypothesis is Mecca has to be the most internationalized city within the Kingdom. Because over two million pilgrims visit the city every year for the Hajj, the venues in the city have to serve many international clients. If many of the major cities in Saudi Arabia such as Riyadh and Jeddah are similar to Mecca we can conclude that they too are internationalizing.

Who is this Project For?

Because Saudi Arabia is so unlike Western societies and economies it can be hard for Western investors to make informed decisions about deploying capital. Specifically investing in Saudi Arabia is difficult because investors are essentially betting on whether MBS is genuine in his drive for reform. Without any supporting data it can seem that investing within the Kingdom is based solely on a promise from a young Crown Prince.

By analyzing the most popular venues in each major city and using KNN to group them we can find unbiased data that can elucidate what is actually happening in Saudi Arabia. If many of the cities are similar to Mecca which has to be the most international then this will provide unbiased data for investors to consider.

Data

I will create a list of the most populous cities in Saudi Arabia and then will query the Foursquare API to return the 10 most popular venues in each city. These two type of data will be combined into a data frame which will be put through the KNN algorithm.

Methodology

Using the Foursquare API I will query the most popular venues in the 11 most populated cities inside Saudi Arabia. I will then rank these venues based on their popularity and will use KNN to determine how similar or dissimilar each city is with one another. These data will be visualized using the Folium library for python as well as data frames that explain each final cluster of data.

If many of the other major cities within the kingdom have similar popular venues as Mecca we can conclude that the whole Kingdom is opening up similarly to Mecca.

Results

KNN Cluster 1

	City	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
1	Jeddah	Pizza Place	Waterfront	Clothing Store	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant
3	Medina	Department Store	Waterfront	Watch Shop	Convenience Store	Cosmetics Shop	Cupcake Shop	Dessert Shop	Diner	Doner Restaurant	Donut Shop
4	Dammam	Boutique	Waterfront	Coffee Shop	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop
8	Abha	Pizza Place	Waterfront	Clothing Store	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant
9	Al-Khobar	Coffee Shop	Watch Shop	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop
11	Najran	Shopping Mall	Clothing Store	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop

KNN Cluster 2

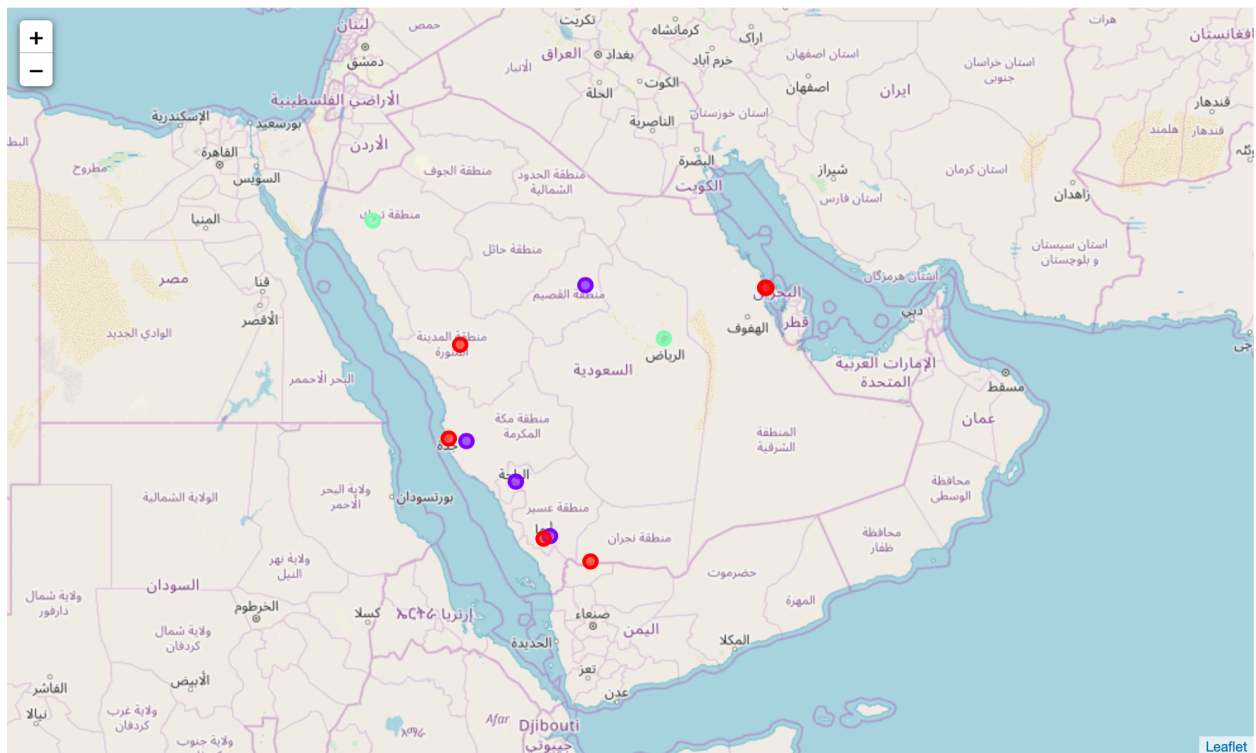|	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Mecca	Coffee Shop	Watch Shop	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop
6	Buraidah	Men's Store	Waterfront	Garden	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant
7	Khamis Mushait	Coffee Shop	Watch Shop	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop
10	Al Bahah	Gym / Fitness Center	Waterfront	Eastern European Restaurant	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant

KNN Cluster 3

	City	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	Riyadh	Coffee Shop	Watch Shop	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop
5	Tabuk	Electronics Store	Watch Shop	Convenience Store	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Doner Restaurant	Donut Shop

Using Folium library to visualize each cluster
        Red is Cluster 1
        Purple is Cluster 2
        Green is Cluster 3

Discussion

Based off the KNN analysis of the venues within each city, neither Jeddah nor Riyadh are in the same cluster as Mecca. Furthermore, looking at the Folium library output, it looks instead cities are more related based on their relationship to the coast. Coastal cities are more similar to other coastal cities, slightly inland cities are related to each other and further inland cities are related.

Given the hypothesis about Mecca's popular venues versus the other major cities it looks like this data does not support the reform narrative.

Conclusion

Based off the data from Saudi Arabia and the Foursquare API, it does not look like the other major cities are similar to Mecca. While the Cluster 2 has 3 other cities along with Mecca in it, these look more related due to their distances from the coast.

This data does not support the reform narrative on its own. A further analysis could be using historical venue data and then could be continued into the future. This could offer more insights on the changing venues within Saudi Arabia.