Data Analysis of Venues and Housing Reference Prices in Tirana

By: Ardit Curri

1. Introduction

1.1 Background

As a born and raised citizen of Tirana, I have decided to perform my analysis on the venues and housing prices of this city.

Tirana is the heart and capital of Albania, where over half a million people live. The city's territory has been inhabited since the Iron Age, but its history as a modern city began in 1614 when Sulejman Pasha Bargjini, a rich feudal from the village of Mullet, built four main buildings that became the centre of the new town: a mosque, a Turkish bath, a bakery, and some inns. Since 1920, this city has been the capital of Albania and is the most important economic, financial, political and trade centre in Albania.

It is becoming more challenging for investors to find the right place to invest for their businesses in Tirana, which is growing and getting crowded each year. Investors prefer to buy properties in neighbourhoods with a lower price that would fit to the type of business they plan to establish.

The aim of this project is to help investors understand the housing reference prices and categories of venues operating for each neighbourhood. To achieve this, a map of Tirana and its neighbourhoods will be created, reflecting information on the real estate index (Housing Reference Prices) and common venues operating in these areas.

1.2 Data sources

Tirana consists of 13 rural administrative units and city of Tirana, which includes eleven urban administrative units (referred to as neighbourhoods). For the purpose of this project, we will consider only urban administrative units, as the data availability on Foursquare is very limited.

To solve our challenge we will use the below data:

1. Administrative units in Municipality of Tirana [1] – Name and geo location for each administrative unit
2. Housing Reference Prices [2] – Average price per square meter for each administrative sub-unit
3. Administrative Units Geodata [3] – Geo location (polygon) data for each administrative unit
4. Venue data from Foursquare API [4] – Common venues for each administrative unit and information on their name, location and category.

2. Methodology

2.1 Analytical approach

The data on administrative units, their geo location and Housing Reference Prices, is available in the Municipality of Tirana webpage, and will be analysed and visualised using Folium.

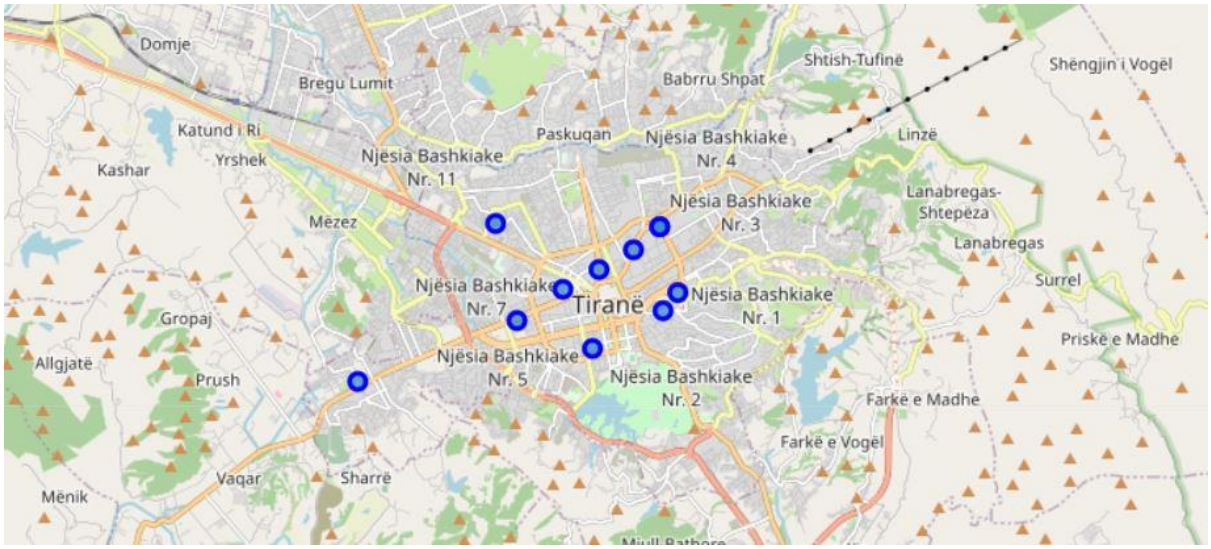
To understand the nature of the venues running in each neighbourhood, information will be retrieved from Foursquare API. The venues will be selected on a radius of 750 meters from the centre of the administrative unit, which is part of the information provided above. The neighbourhoods will be clustered into groups based on the category of venues operating in their respective areas.

2.2 Data exploration

A json file will be downloaded from Opendata Tirana to retrieve the information on each administrative unit and their geolocation (latitude and longitude).

	index	Neighborhood	Latitude	Longitude
0	2	Nj. Admin 2	41.325902	19.830347
1	3	Nj. Admin 1	41.328529	19.833451
2	4	Nj. Admin 9	41.332104	19.816886
3	5	Nj. Admin 8	41.335277	19.824183
4	6	Nj. Admin 3	41.338969	19.829606

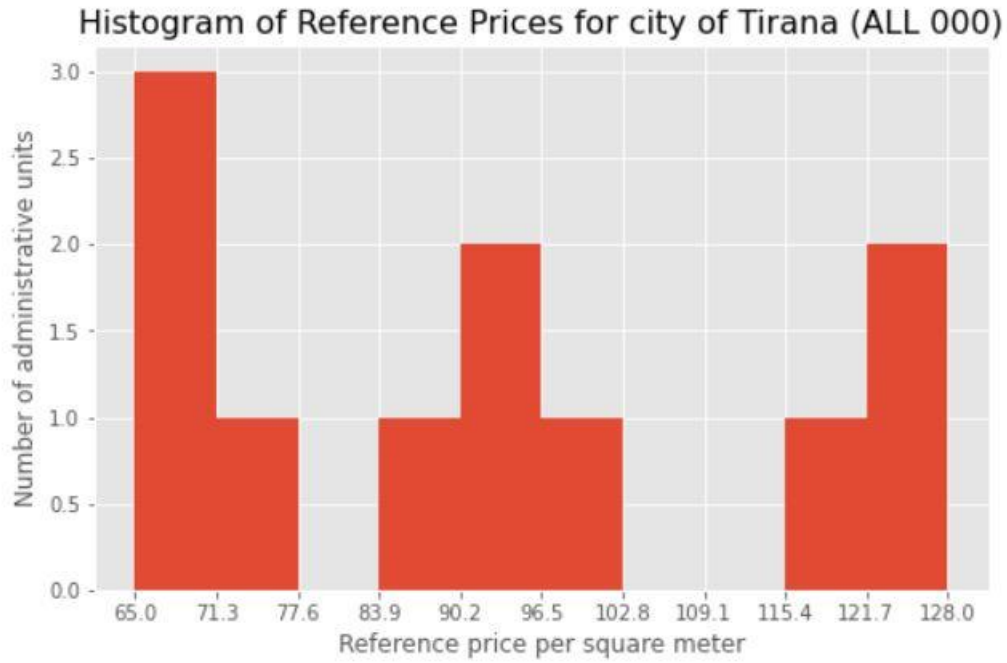
A map on Tirana and its administrative units is created in Folium:



Housing Reference Prices are published in the website of Municipality of Tirana and comprise the average price per square meter for each administrative sub-unit. The average price is calculated per neighbourhood.

	Neighborhood	RP_squaremeter
0	Nj. Admin 1	87850.000000
1	Nj. Admin 10	128833.333333
2	Nj. Admin 11	77333.333333
3	Nj. Admin 2	121300.000000
4	Nj. Admin 3	68500.000000

To understand the distribution of average HRP across neighbourhoods we visualise the data with Histogram.



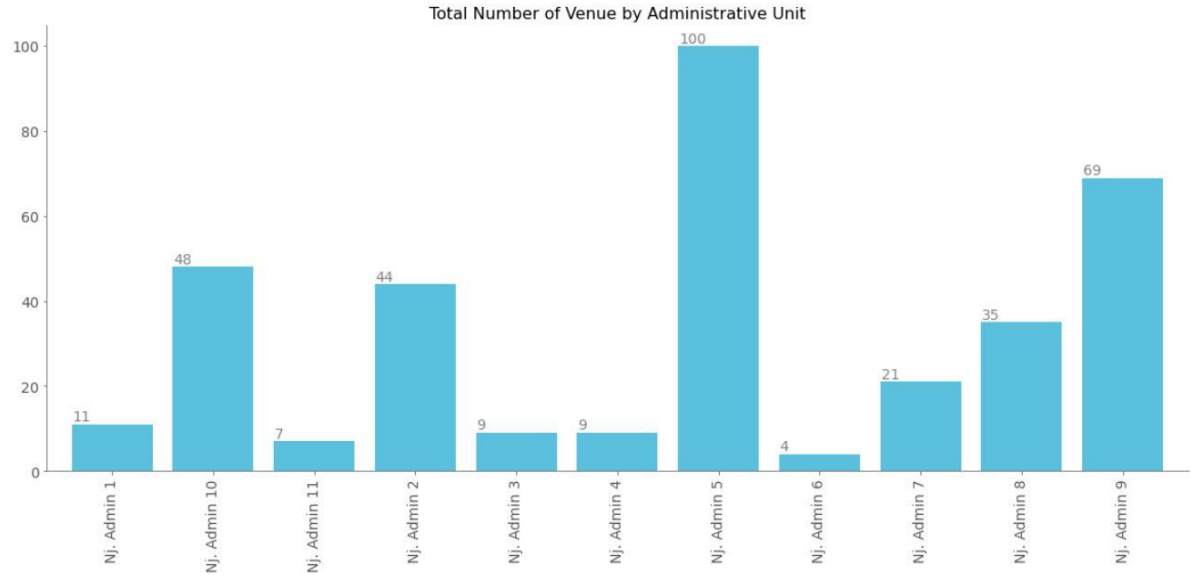
From the data, we can observe that there are three main categories of average HRP ranges:

1. Low HRP Level – price range between ALL 65’000-77’600
2. Medium HRP Level - price range between ALL 83’900-102’800
3. High HRP Level - price range between ALL 115’400-128’000

Foursquare API will be utilized to retrieve information on venues operating in each neighbourhood. The venues will be selected on a radius of 750 meters from the administrative unit centre, limited to 100 venues per neighbourhood. 357 venues are retrieved from Foursquare API providing information on their name, location (latitude and longitude) and category.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Nj. Admin 2	41.325902	19.830347	Fishop	41.322384	19.827218	Seafood Restaurant
1	Nj. Admin 2	41.325902	19.830347	Delano Lounge Restaurant	41.327002	19.824573	Restaurant
2	Nj. Admin 2	41.325902	19.830347	Mulliri Vjeter	41.323451	19.825499	Café
3	Nj. Admin 2	41.325902	19.830347	Edi Fishop	41.322189	19.826643	Fish & Chips Shop
4	Nj. Admin 2	41.325902	19.830347	Gourmet	41.323262	19.823910	Restaurant

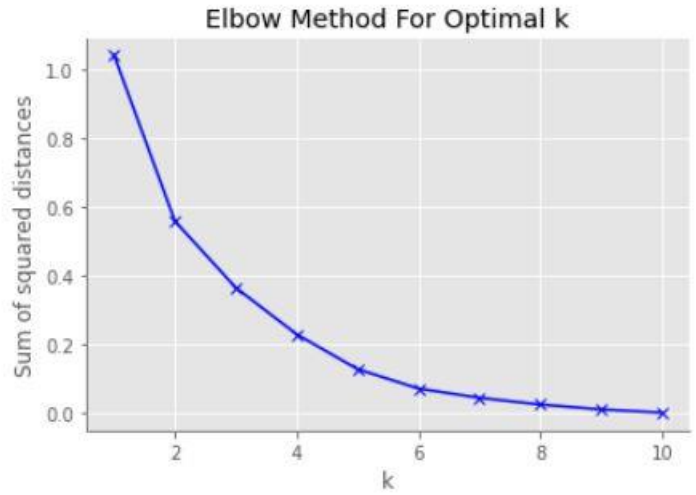
The information is very limited for venues in Tirana. The number of venues is visualized below for each administrative unit.



For each neighbourhood, the top 10 most common venues are identified.

	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	Nj. Admin 1	Hotel	Hostel	Seafood Restaurant	BBQ Joint	Coffee Shop	Dessert Shop	Mediterranean Restaurant	American Restaurant	Shopping Mall	Snack Place
1	Nj. Admin 10	Café	Bar	Hotel	Coffee Shop	Italian Restaurant	Cocktail Bar	Bakery	Pizza Place	Lounge	Market
2	Nj. Admin 11	Bar	Whisky Bar	Hotel	Diner	Electronics Store	Pub	Fast Food Restaurant	Eastern European Restaurant	Farmers Market	French Restaurant
3	Nj. Admin 2	Hotel	Café	Cocktail Bar	Seafood Restaurant	Pub	Restaurant	Fast Food Restaurant	Eastern European Restaurant	Hotel Bar	Hookah Bar
4	Nj. Admin 3	Hotel	Nightclub	Bar	Pop-Up Shop	Hotel Pool	Border Crossing	Bakery	Greek Restaurant	Gift Shop	Creperie

Unsupervised learning K-means algorithm will be used to cluster the neighbourhoods into groups that share the same characteristics on the venues that operate in their area. The Elbow Method will be used to find the optimal number of clusters.



A cluster value is assigned to each administrative unit.

	index	Neighborhood	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 Most Common Venue	10th Most Common Venue
0	2	Nj. Admin 2	41.325902	19.830347	0	Hotel	Café	Cocktail Bar	Seafood Restaurant	Pub	Restaurant	Fast Food Restaurant	Eastern European Restaurant	Hotel Bar	Hookah Bar
1	3	Nj. Admin 1	41.328529	19.833451	5	Hotel	Hostel	Seafood Restaurant	BBQ Joint	Coffee Shop	Dessert Shop	Mediterranean Restaurant	American Restaurant	Shopping Mall	Snack Place
2	4	Nj. Admin 9	41.332104	19.816886	0	Hotel	Café	Bar	Coffee Shop	Pizza Place	Hostel	History Museum	Seafood Restaurant	Italian Restaurant	Cocktail Bar
3	5	Nj. Admin 8	41.335277	19.824183	0	Hotel	Bar	Hostel	Café	Cocktail Bar	Coffee Shop	Seafood Restaurant	Pizza Place	Nightclub	Bed & Breakfast
4	6	Nj. Admin 3	41.338969	19.829606	2	Hotel	Nightclub	Bar	Pop-Up Shop	Hotel Pool	Border Crossing	Bakery	Greek Restaurant	Gift Shop	Creperie
5	7	Nj. Admin 4	41.339041	19.829566	2	Hotel	Nightclub	Bar	Pop-Up Shop	Hotel Pool	Border Crossing	Bakery	Greek Restaurant	Gift Shop	Creperie
6	8	Nj. Admin 5	41.319825	19.815625	0	Cocktail Bar	Café	Italian Restaurant	Bar	Coffee Shop	Hotel	Lounge	Mediterranean Restaurant	Restaurant	Pizza Place
7	9	Nj. Admin 10	41.329220	19.809519	3	Café	Bar	Hotel	Coffee Shop	Italian Restaurant	Cocktail Bar	Bakery	Pizza Place	Lounge	Market
8	11	Nj. Admin 11	41.339570	19.795195	4	Bar	Whisky Bar	Hotel	Diner	Electronics Store	Pub	Fast Food Restaurant	Eastern European Restaurant	Farmers Market	French Restaurant
9	14	Nj. Admin 6	41.314590	19.766441	1	Bar	Bistro	Wine Bar	Fish & Chips Shop	Diner	Eastern European Restaurant	Electronics Store	Farmers Market	Fast Food Restaurant	Gift Shop
10	19	Nj. Admin 7	41.324134	19.799690	3	Café	Hotel	Bar	Coffee Shop	Restaurant	Bus Station	BBQ Joint	Snack Place	Cocktail Bar	Pizza Place

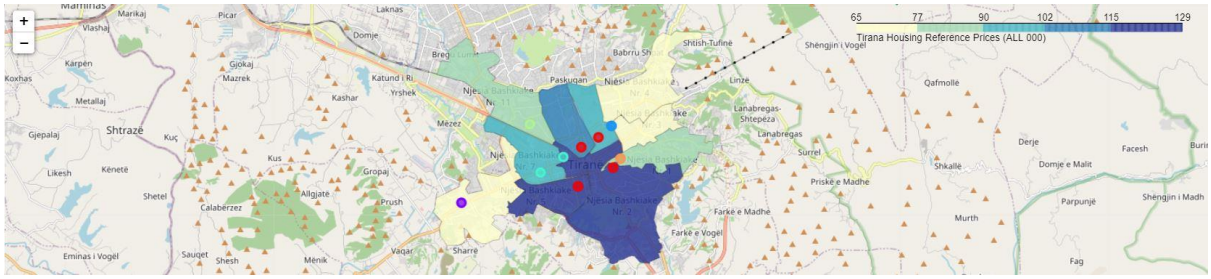
Data is examined and clusters are named on categories of venues.

- Cluster 0: Hotel; Café; Restaurants; Bar
- Cluster 1: Bar and Bistro
- Cluster 2: Multiple venues
- Cluster 3: Café; Bar; Hotel
- Cluster 4: Bar and Electronic store
- Cluster 5: Accommodations and Restaurants

3. Results

3.1 Create results

In the final visualization, a choropleth map of Tirana will show the average Housing Reference Price and each administrative unit will be clustered on the type of venues operating in their area.

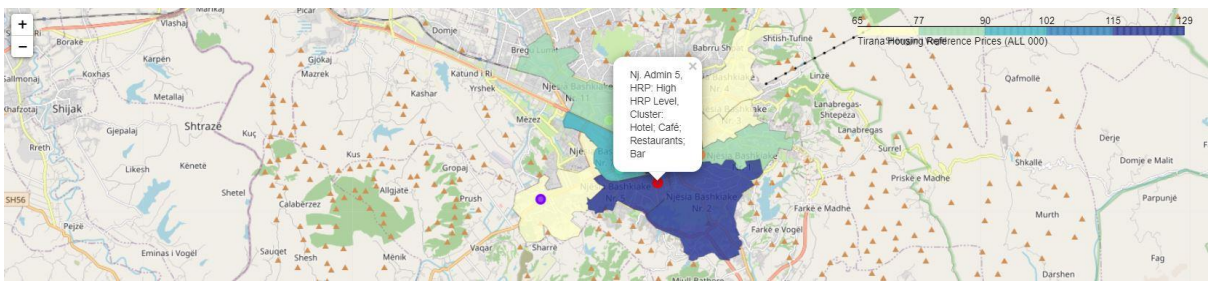


For each administrative unit the following are shown:

Administrative unit name,

Housing Reference Price range group

Cluster name



3.2 Discussion

Tirana is a medium sized city and the population is dense in the urban area. There are alternative approaches to cluster and classify areas to invest in real estate. Average Housing Reference Price and venues are the features selected for this analysis.

K-means algorithm is used to cluster the city administrative units and based on Elbow method we decided to cluster into 6 buckets. The information on venues in Tirana is very low and the analysis could be improved in the future, if there is more data available.

The study is completed by visualizing the urban area of Tirana with HRP range and Cluster for each Administrative unit and the map is coloured based on the HRP. In future studies, this information could be available to investors online and ready to use via webpages and/ or apps.

As a result from the analysis, we can observe the following:

1. Housing Reference Prices are highest in the city centre and the area near Artificial Lake of Tirana;
2. Housing Reference Prices are lowest in the most distant areas from city centre (Administrative Unit 3, 4 and 6);
3. Areas in the city centre are mostly used for Hotels, Cafe, Restaurants and Bars;
4. Administrative Unit 1 is mainly used for Accommodation (Hotel& Hostel) and Restaurants.

4. Conclusions

More and more nowadays, inventors are embracing data driven business models. The analysis above provides insights regarding the range of average Housing Reference Prices between different administrative units in Tirana and their respective common venues.

Therefore, using such insights will help investors take the right decisions when assessing business opportunities in city of Tirana.

A potential next step for the research is to research the correlation between different venues and the average Housing Reference Price in the nearby area.

Creating similar analysis in the future would be beneficial for investors, but also for local administration and businesses.

5. References

Administrative Units in Municipality of Tirana - https://opendata.tirana.al/sites/default/files/Selia_Nj_Admin_2019.geojson

Housing Reference Price in Tirana - https://www.tirana.al/artikull/informacion-mbi-tregun-e-pasurive-te-paluajtshme-ne-qytetin-e-tiranes

Administrative Units Geodata - https://opendata.tirana.al/sites/default/files/Density_NjA.txt

Foursquare API - https://developer.foursquare.com/