

Battle of Neighborhoods | Exploring Islamabad

Anees Akhtar

April 05, 2021

1. Introduction:

Islamabad is the capital city of Pakistan. The city's master plan, designed by Greek architect *Constantinos Apostolou Doxiadis*, divides the city into zones, including administrative, diplomatic enclave, residential areas, educational sectors, industrial sectors, commercial areas, and rural and green areas.

Islamabad is divided into five major zones: Zone-I, Zone-II, Zone-III, Zone-IV, and Zone V. Out of these, Zone IV is the largest in area. Zone I consists mainly of all the developed residential sectors while Zone II consists of the under-developed residential sectors. Each residential sector is identified by a letter of the alphabet and a number, and covers an area of approximately $2 \text{ km} \times 2 \text{ km}$. The sectors are lettered from A to I, and each sector is divided into four numbered sub-sectors.

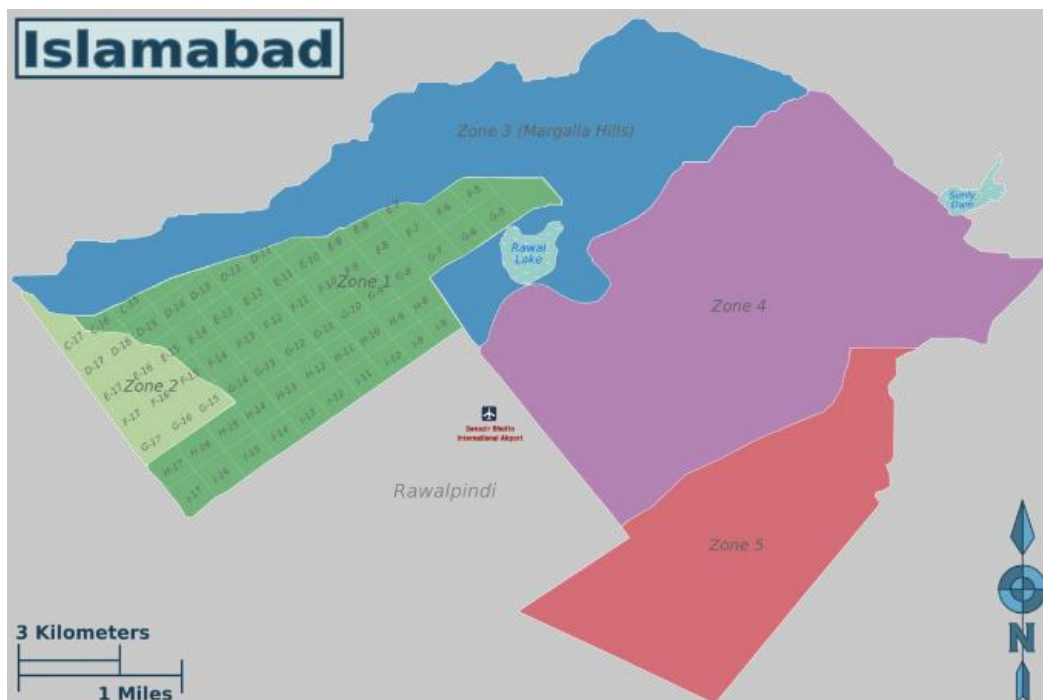


Figure 1: Masterplan of Islamabad City

2. Business Problem:

In this project, we will explore, segment, and cluster the Zone-1 sectors in the city of Islamabad based on most common venue information. We will also try to answer following stakeholder questions so they can make investment decision:

1. What are the most common venues in Zone-I?
2. Which sectors in Zone-I have highest venues?
3. Which sectors in Zone-I are yet to be developed based on most common venues?

3. Data:

Following data sources will be needed to extract/generate the required information:

- General information on Islamabad will be gained from [Wikipedia](#)
- Zones and sectors information will be gained from [here](#)
- Most common venues, their categories and location in every sector will be obtained using **Foursquare API**
- Coordinate of sectors will be obtained using **Geopy** library

4. Methodology:

In this project, we will explore, segment, and cluster the Zone-1 sectors in the city of Islamabad based on most common venue information. We will also try to answer following stakeholder questions so they can make investment decision:

1. What are the most common venues in Zone-I?
2. Which sectors in Zone-I have highest venues?
3. Which sectors in Zone-I are yet to be developed based on most common venues?

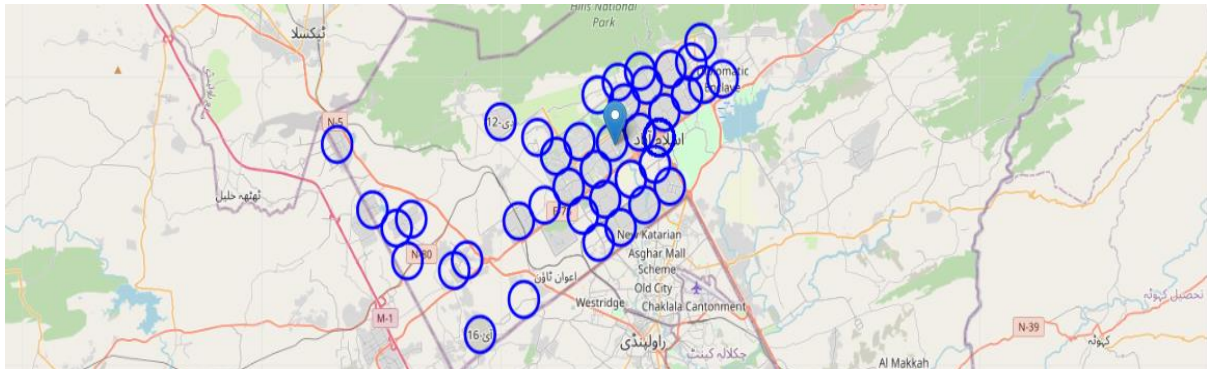
In first step, we have collected the required data: location and type (category) of most common venues within 2km from center of each sector

Second step in our analysis will be calculation and exploration of 'venue density' across different sectors of Islamabad - we will use bar charts to identify a few promising areas to answer questions.

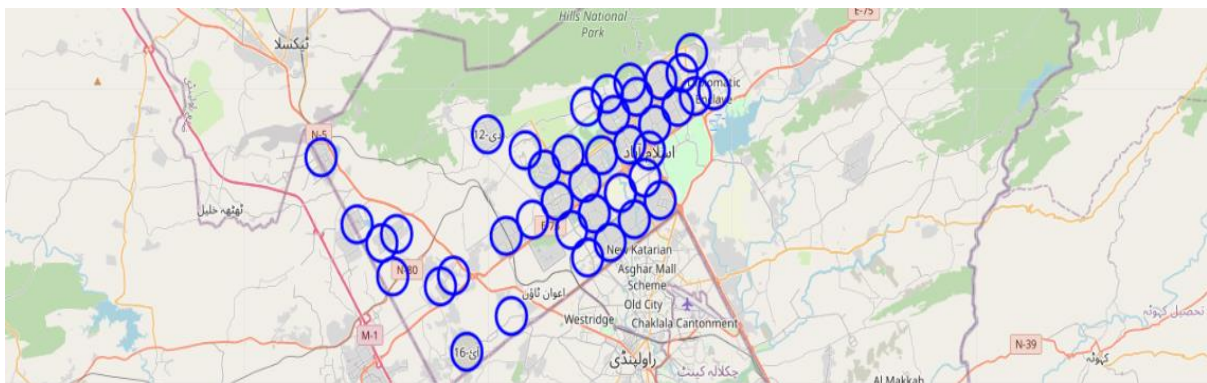
In third and final step we will focus on detecting outliers in Zone-I i.e. clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with least number of most common venues in Zone-I.

5. Exploratory Data Analysis:

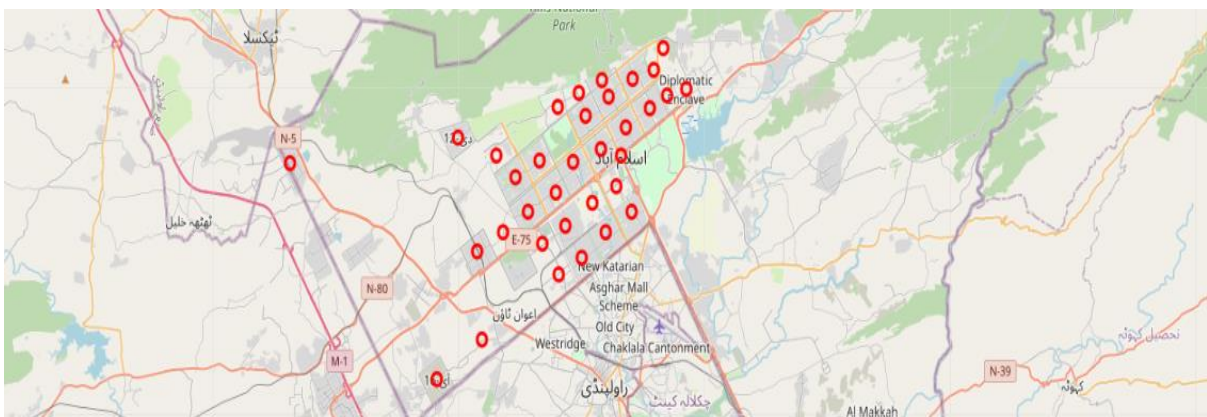
In first step, we used Wikipedia to identify the Zones & Sectors and then used Geopy API to get coordinates for these sectors.



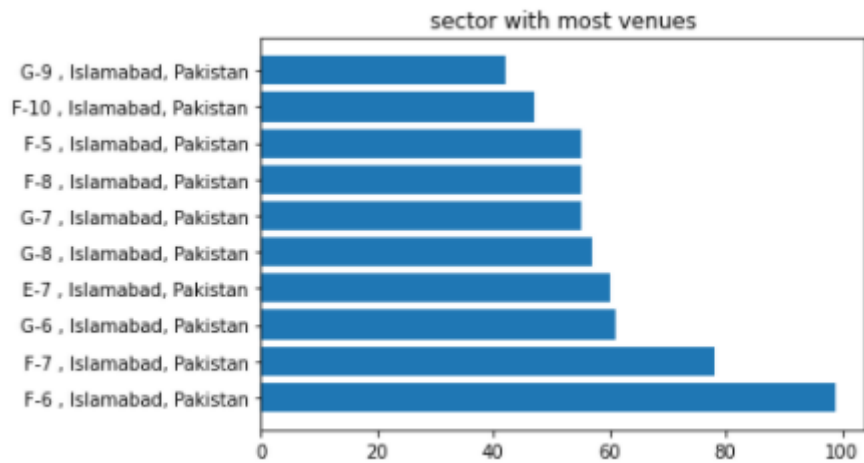
Then we removed duplicate coordinates and sectors.



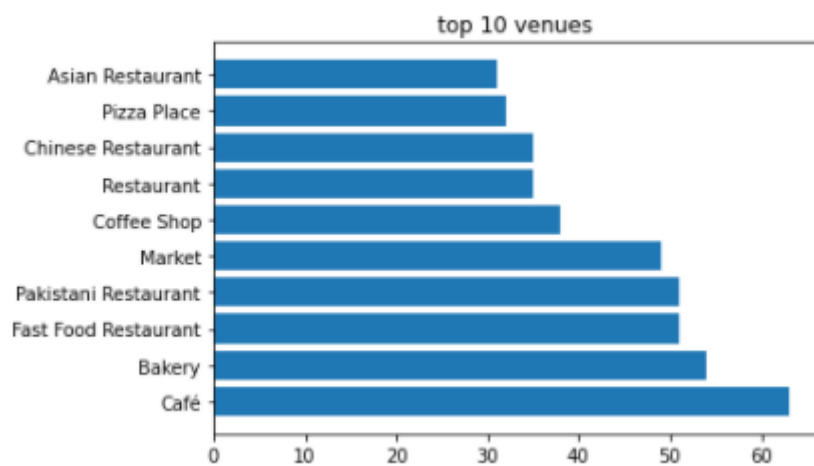
After that, we dropped sectors of Zone-II.



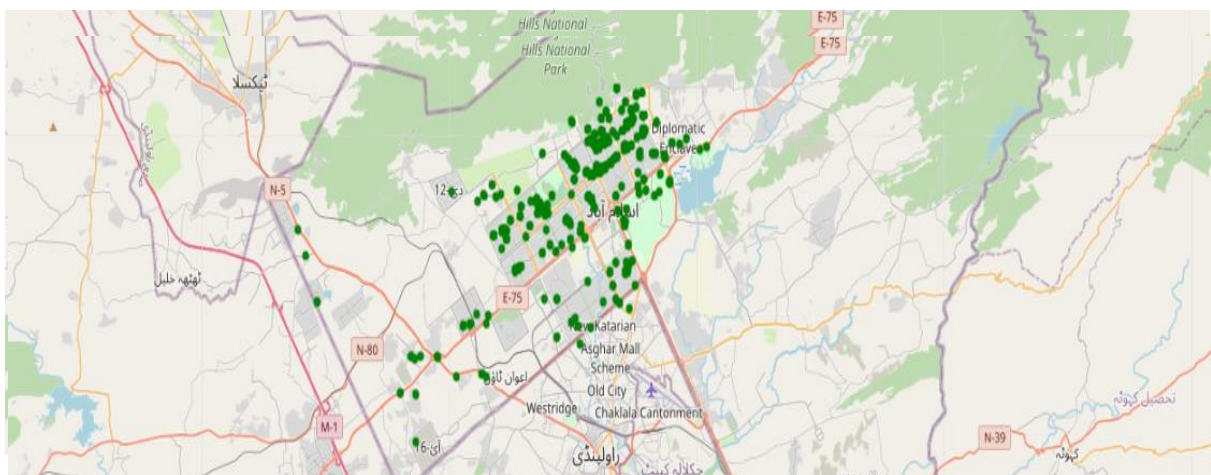
Following are top 10 sectors having most common venues in Zone-I.



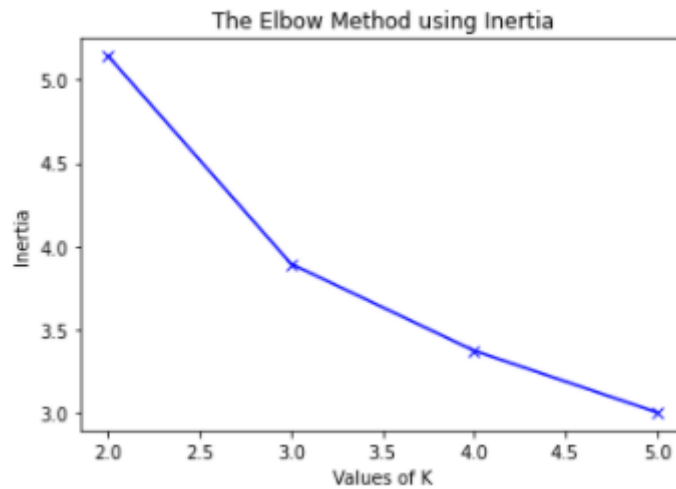
Following are top 10 most common venues in sectors of Zone-I.



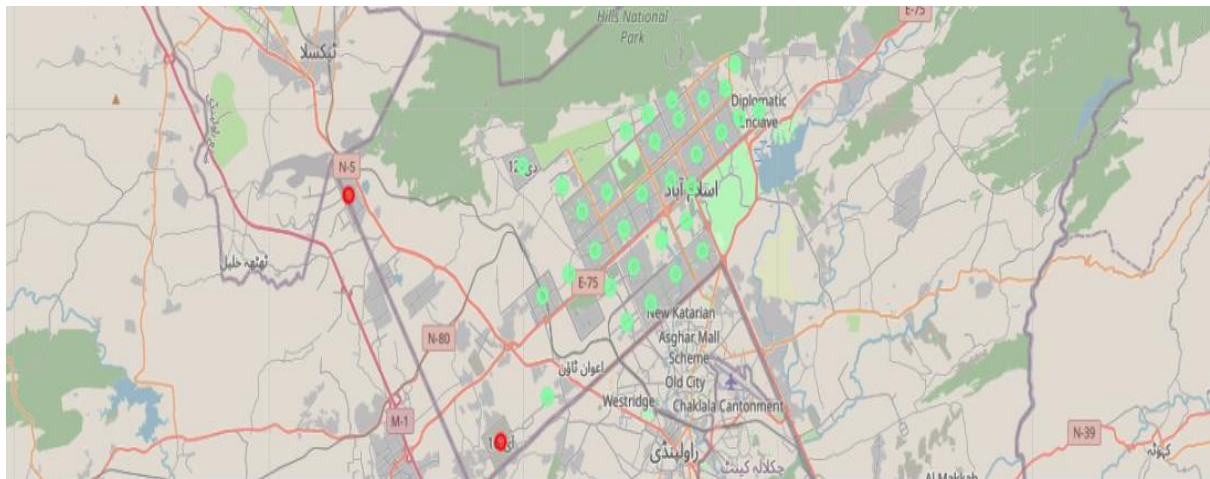
All common venues Zone-1.



We used elbow method to find best value of k for K-means clustering.



After clustering, we got two types of sectors in Zone-I w.r.t most common venues.



6. Results and Discussion:

What are the most common venues in Zone-I?

Therefore, during EDA, we have observed that top venues of Zone-1 are following:

1. Cafe
2. Bakery
3. Fast Food Restaurant
4. Pakistani Restaurant
5. Markets
6. Coffee Shop

7. Restaurant
8. Chinese Restaurant
9. Pizza Place
10. Asian Restaurant

It can be noted that 9 out of 10 most common venues are Food places. Particularly Chinese & Fast Food/Pizza restaurants are topping the list alongside traditional Pakistani restaurants & teahouses.

Which sectors in Zone-I have highest venues?

We have observed that following sectors have most venues in Zone-1:

1. F-6 , Islamabad
2. F-7 , Islamabad
3. G-6 , Islamabad
4. E-7 , Islamabad
5. G-8 , Islamabad
6. G-7 , Islamabad
7. F-8 , Islamabad
8. F-5, Islamabad
9. F-10, Islamabad
10. G-9 , Islamabad

We can notice that 5 out of 10 sectors belong to 'F' belt so we can say that these sectors are most developed in Islamabad.

Which sectors in Zone-I are yet to be developed based on most common venues?

After clustering most common venues in Zone-I, results show that all sectors of Zone-I have similar venues particular food places except sector I-16 & B-17. I-16 & B-17 are still under development and does not have much of common venues yet. This, of course, does not imply that these sectors are actually optimal locations for a new restaurant! We tried to provide a starting point for more detailed analysis, which could eventually result in location, which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

7. Conclusion:

Purpose of this project was to explore, segment, and cluster the Zone-1 sectors in the city of Islamabad based on most common venue information. We tried to answer following stakeholder questions so they can make investment decision:

1. What are the most common venues in Zone-I?
2. Which sectors in Zone-I have highest venues?
3. Which sectors in Zone-I are yet to be developed based on most common venues?

Location & venue data is collected using Geopy & Foursquare API respectively. Exploratory data analysis is performed to get answer of first and second question. Clustering of common venues was then performed in order to answer third question and create major zones of interest (containing least common venues).

Final decision on optimal business location will be made by stakeholders based on specific characteristics of locations in every recommended sector, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.