

IBM Data Science Professional Certificate

Capstone Project: The Battle of the Neighbourhoods

Segmenting and Clustering selected Neighbourhoods of Hyderabad

Johny Ijaq

July 18, 2020

1. Introduction

1.1. Background

Hyderabad city is the capital city of the Indian state of Telangana. The city has an estimated population of around 8 million, making it 4th largest city in India, while the population of the metropolitan area was estimated above 9 million. Hyderabad city is known for its rich history, food and its multi-lingual culture, both geographically and culturally. As a resident of this city, I decided to explore Hyderabad for my project.

Hyderabad, since the time of the Nizams, has been the seat of art, culture, literature, music, science, and a variety of other academic activities. In modern times as well, the city has managed to hold on to its rich academic tradition, with some of the most premier educational institutions in the country making their home on the city's landscape

The city has three central universities, including the top-ranked University of Hyderabad, The Maulana Abdul Kalam Azad National Urdu University and English and Foreign Languages University, along with the famous state university such as Osmania University. The city and surrounding areas are home to top engineering institutes such as Indian Institute of Technology - Hyderabad, International Institute of Information Technology, Hyderabad, Birla Institute of Technology & Sciences. Hyderabad also hosts big names in business schools such as Indian School of Business (ISB) and ICFAI Business School. Nalsar University of Law, Tata Institute of Social Science Research and Tata Institute of Fundamental Research are the other prestigious institutes in the city.

Hyderabad is also a major centre for research in pharma, biotech and defence-related science and technology. Apart from many defence research labs such as Defence Research and Development Organisation (DRDO) and space institutes, the city has prestigious research

laboratories such as the Centre for Cellular and Molecular Biology (CCMB), Indian Institute of Chemical Technology, Centre for DNA and Fingerprinting Diagnostics (CDFD), National Institute of Nutrition (NIN) etc.

Incorporating a large number of universities, management colleges, research centres, and technical institutes, the capital of Telangana, is undoubtedly shaping itself rapidly to be the country's higher education hub. This has led to the development of many business centres nearby these educational and research institutes. Exploring the venues around these centres could provide valuable insights about the business opportunities present in these areas.

1.2 Problem

I selected few neighbourhoods from Hyderabad that are hosting famous educational and research institutes. From this project I would like to explore all the great amenities and other types of venues that exist in the neighbourhood, such as restaurants, cafeterias, fast food joints, bakeries, stationaries, pharmacies, parks, schools and so on.

Main objective of this project is to find out the most common businesses in these areas, to explore the best business idea which one can start in these neighbourhoods and to select the area where the type of business one want to install is less intense. Finally, a combination of location data and machine learning was used to group the neighbourhoods into clusters.

1.3. Interest

The results from this project can be useful to anyone who is planning to start a business in the studied neighbourhoods. Anyone who is looking to find the best neighbourhood to open a shop of interest also finds this study useful. One can choose to compare different neighbourhoods in terms of a service, search for potential explanation of why a neighbourhood is popular, the cause of complaints in another neighbourhood, or anything else related to neighbourhoods.

2. Data sources

1. Latitude and longitude data for each neighbourhood was obtained from <https://www.gps-latitude-longitude.com/address-to-longitude-latitude-gps-coordinates>.
2. I used the Foursquare API to explore the nearby venues from the selected universities and research institutes in order to find the common venues or businesses. I searched for the venues up to 1 KM from the GPS coordinates (latitude and longitude) of the six neighbourhoods. Foursquare API helped me to search for a specific type of venues, explore a particular venue and to get trending venues around each neighbourhood. I retrieved the following information from the foursquare API for each neighbourhood.
 - A. Venue Name
 - B. Venue Category
 - C. Venue Latitude
 - D. Venue longitude

