Data

1.Below is the required data:

- List of neighborhoods in the city of Hyderabad. This will give us the scope of the areas to open shopping malls near to the residential areas and meet the demand.
- Latitude and Longitudes of the extracted neighborhood data. This will act as the input parameters to the foursquare api to explore a neighborhood.
- Shopping malls data, this will be the output from the foursquare api and help us to create clusters on the neighborhoods depending on the frequency of malls in a given radius of a particular neighborhood.

2.Sources and methods to get the data:

- The list of neighborhoods in Hyderabad is taken from a wikipedia page (https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Hyderabad), contains 31 neighborhoods in total, combining all the zones. By using web scraping technique we will get the neighborhood data from the wikipedia page.
- We will then get the Latitude and Longitude of each neighborhood using geocoder library and attach these coordinates to our neighborhood data
- Then with the help of Foursquare api calls we will send the coordinates of each neighborhoods and get the venues details of shopping malls and create the shopping malls data of each neighborhood.
- This project will be completed with the help of data science skills, data cleaning, exploring, analyzing, visualizing using folium maps in particular. Also takes help from foursquare api to get the shopping mall details which are near to each neighborhood. And finally using a machine learning technique 'K-means clustering' to cluster malls into different categories.

3.Snapshot of datasets

Below is the screenshot of the head of the dataset that consists of list of neighborhoods in the column 1 and its coordinates fetched with thet help of geocoder as Latitude and Longitude in columns 2 and 3.This data will be used to send to foursquare api to fetch the venue categories that includes shopping mall categories as well if found any in particular neighborhood

	Neighborhood	Latitude	Longitude
0	Sanathnagar	17.456965	78.443478
1	Ghatkesar	NaN	NaN
2	Mehdipatnam	17.394263	78.434251
3	Balanagar	17.476746	78.422108
4	Malkajgiri	17.448344	78.528973

However we have dropped the rows that contains NaN's as those are anyways outskirt locations and had not found coordinates using geocoder library.

With the help of Foursquare api we are able to fetch the nearby venues and filtered shopping mall categories out of those venues and the frequency/mean of those malls present in the entire neighborhoods is shown as below.

	Neighborhoods	Shopping Mall
0	Alwal	0.000000
1	Amberpet	0.047619
2	Ameerpet	0.010000
3	Balanagar	0.142857
4	Dilsukhnagar	0.076923