

2. Data Collection

The data required for this project is collected from various resources. The data and their source is described below.

2.1 Neighborhoods Data:

The neighborhood or location data is scrapped from website <https://www.zricks.com/Locality/Jaipur> and parsed using html parser using BeautifulSoup. The neighborhood column of the pandas dataframe is created by this location data.

2.2 Geographical Coordinates:

The geographical coordinates of the neighborhoods are obtained from GeoPy library in Python. The address or location is provided in the geocode function in the library and it returns the latitude and longitude of the provided address. We pass the locations that we scrapped from <https://www.zricks.com/Locality/Jaipur> website and we get latitude and longitude for every address. If any location in geopy library returns none value that we return 'Nan' for it. This data is relevant for plotting the map of Jaipur using the Folium library in python.

2.3 Venues Data:

The venue data has been extracted using the Foursquare API. This data contains venue recommendations for all neighborhoods in Jaipur and is used to study the popular venues of different neighborhoods.