

Data

Data Requirement

Based on our business problem ie. – “Identifying a suitable location to build a Shopping Mall in City of Mangalore.” we need following data.

1. We need a list of neighborhoods in Mangalore to identify locations with Residential buildings, Hotels, Colleges, parks, transportation etc.
2. We need the latitude and longitude coordination of these neighborhoods for different analysis and visualization.
3. Analyze the obtained data to identify the best locations for coming up with new Shopping Mall.

Data Source

Following are the data sources needed for the analysis

1. **Google Map** to get coordinates of Mangalore City
2. **Wikipedia** to collect information about different neighborhoods (wards) in Mangalore.
3. Using **Foursquare** API's to find out the best location for the mall. Also analyze the different neighborhoods to find out the number of Residential building, Hotels, arks etc.
4. Will be using **Geopy Nominatim** to get the latitude and longitude of different locations/Wards.
5. Pandas will be used to clean up the data as we may have to add/remove/fix coordinates of different locations/wards. Data collection and cleaning will be an important activity in this project as there may not be enough information available from one single source.