

# Applied Data Science Capstone

## Data :

We will be dealing with location data for Bangalore. The following are the datasets I have used.

### Data 1: Contains regions in Bangalore and the neighborhoods within them

- First we need the various neighborhoods in the different regions in Bangalore.
- I have manually copied the region and neighborhoods into a list and turned it into a dataframe, one for each region.
- Finally I have concatenated the individual region dataframes to arrive at the final dataframe for locations, which looks like this:

	Region	Neighborhood
0	Central	Cantonment
1	Central	Domlur
2	Central	Indiranagar
3	Central	Jeevanbheemanagar
4	Central	Malleswaram
...	...	...
5	Western	Nandini Layout
6	Western	Nayandahalli
7	Western	Rajajinagar
8	Western	Rajarajeshwari Nagar
9	Western	Vijayanagar

65 rows × 2 columns

Source/Tool: [https://en.wikipedia.org/wiki/List\\_of\\_neighbourhoods\\_in\\_Bangalore](https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore)

## Data 2: Contains the co-ordinates of the neighborhoods above

- To get the latitude and longitude values for each neighborhood, I have made use of the python package geocoders in geopy.
- After using geocoders and getting the latitudes and longitudes, our dataframe looks like this:

	Region	Neighborhood	Latitude	Longitude
0	Central	Cantonment	12.979120	77.591300
1	Central	Domlur	12.962467	77.638196
2	Central	Indiranagar	12.973291	77.640467
3	Central	Jeevanbheemanagar	12.964200	77.658100
4	Central	Malleswaram	13.002735	77.570325
...	...	...	...	...
5	Western	Nandini Layout	13.010406	77.537803
6	Western	Nayandahalli	12.941325	77.521212
7	Western	Rajajinagar	12.988234	77.554883
8	Western	Rajarajeshwari Nagar	12.927441	77.515522
9	Western	Vijayanagar	12.971889	77.545789

65 rows × 4 columns

Source/Tool: Python package Nominatim in geopy.geocoders.

### Data 3: Contains the different venues in each of the neighborhoods

- To retrieve the venues in all the neighborhoods, I have used the Foursquare API.
- The final dataframe containing the regions, neighborhoods, their coordinates, the venues and the venue coordinates and category is as shown:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Cantonment	12.979120	77.591300	Vidhana Soudha	12.979027	77.591881	Capitol Building
1	Cantonment	12.979120	77.591300	Cubbon Park	12.977042	77.595277	Park
2	Cantonment	12.979120	77.591300	Lobby @ ITC Gardenia	12.976568	77.589033	Garden
3	Cantonment	12.979120	77.591300	Wine Board	12.975636	77.590058	Vineyard
4	Domlur	12.962467	77.638196	Lavonne	12.963909	77.638579	Café
...	...	...	...	...	...	...	...
658	Vijayanagar	12.971889	77.545789	Natti style, vijayanagar	12.972908	77.544356	Indian Restaurant
659	Vijayanagar	12.971889	77.545789	Adichunchanagiri Kalyana Mantap	12.973807	77.545510	Arcade
660	Vijayanagar	12.971889	77.545789	Ruchi Bhel Centre	12.969259	77.544368	Indian Restaurant
661	Vijayanagar	12.971889	77.545789	reliance fresh	12.974825	77.548303	Department Store
662	Vijayanagar	12.971889	77.545789	just bake	12.973699	77.541899	Dessert Shop

Source/Tool: Foursquare API

