

## DATA DESCRIPTION

**To solve this problem we need the following data**

- List of post offices and their pin codes in Bangalore, IN
- Location coordinates of these postal areas
- Venue data, we will use this data to perform the clustering technique.

## SOURCES OF DATA / DATA EXTRACTION

This page “<https://finkode.com/ka/bangalore.html>” contains a list of post offices and their pin code in the city of Bangalore, IN. We will be using web scraping techniques to extract the data from the html code using python requests and BeautifulSoup package. Then we will get the coordinates of these locations using the python geocoder package, which will give us the latitude and longitude values of each post office location.

After that, we will use the Foursquare API to access the venue data for these post office locations. Foursquare will provide many categories of venue data, we are particularly interested in the restaurant venues to proceed further in our analysis.