Description of Data:

The data contains of the records of the restaurant (including café) in Mumbai. It would contain all the columns as Name, address, location, longitude, latitude, category ID and so on. We will be taking out the columns as our requirement.

Methodology:

We will first create a link (url) with our required data like limit, radius, version, categoryID etc. to get the records of the following requirement. This record will be in JSON format but for our use we need to change the following JSON data to Data Frame. Once we get data Frame, we will just manipulate it and just take out our required attributes.

These records would contain the data of name, address, longitude, latitude of the restaurant all over Mumbai. We will map this on Mumbai map to understand and check that our data is correct.

We will now use ‘Elbow Curve’ to understand how much clusters to do and according to this we will implement the following K-means on the following records which will give us cluster labels and using this label we can map on Mumbai map different cluster with different colour and also we can check what does the particular cluster contains.