Data description

For feasibility analysis of cafes in Manhattan

Available data from internet

We will have to use the json data of the city of New York to retrieve the location data for Manhattan. Fortunately this is available <u>here</u> and we can readily use it for our purpose. Secondly we will need information of various cafes and popular places of interests in the city and this can be retrieved using FourSquare developer API. The data will be retrieved using the personal login in Developer API and sorted and formatted. Lastly the population density of various Boroughs in the city of New York is available in the <u>internet</u> as well. We will have to scarp the data from different web pages and combine it with the data available from FourSquare for our requirement. In the final form the data will be a list of neighbourhoods sorted with respect to the number of cafes, popular places and its population density.

This formatted data will be used for the analysis of our problem using k-mean clustering.