

Data Collection

Data Sources:

WE have used few Neighborhood available present at below wikipedia link. This gives us the Neighborhood and Indian people living in them.

https://en.wikipedia.org/wiki/Ethnic_groups_in_London

Few of the Neighborhood from this have been taken for study. Also we have used foursquare api (<https://api.foursquare.com>) to fetch the details of each neighborhood which we will study.

Data Cleaning and feature engineering:

We have extractes the data from wikipedia which have only population and neighborhood.

Later we fetched the latitude and longitude for them.

Also all the venues have been fetched from four square api. Later these were cleaned, grouped. New feature have been derived from this data.