

Data Selection

The first exercise of our project is to select our dataset. Our analysis exercise and final conclusion will largely depend upon the quality of data being used. For this purpose, I will use the dataset provided in previous labs, The New York neighborhood dataset. This dataset has a total of five boroughs and 306 neighborhoods that we will segment and process for the purpose of finding the best borough and neighborhood to open the new Mediterranean Restaurant. In addition, to the dataset, we will use **Foursquare**, one of most popular location based social networks (LBSN) so we can explore venues and especially Mediterranean restaurants in the NY neighborhoods dataset.