# Data and methodology

In this exercise, I will explore, segment, and cluster the areas in Singapore. For the Singapore area data, a Wikipedia page exists that has all the information I need to explore and cluster. The link is as following: <https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore>. I am going to scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas dataframe so that it is in a structured format. Of course, foursquare location data will be extracted to visualize Singapore areas and venues. Once the data is in a structured format, I am going to conduct the segmenting and clustering analysis to explore the areas in Singapore. k-means clustering algorithm is going to be used in this analysis. Furthermore, the Folium library is applied to visualize the areas.