

Data

Our goal is to compare in the structure of both cities so that we can draw conclusions about what characteristics make a city healthy to live in.

To do that we will have to segment the cities into their existing administrative districts and find out what venues are most common in each, clustering them to determine which districts are most similar within each city and then comparing the districts across the two cities.

Vienna is divided into 23 districts, a list of which I was able to scrape from [Austria's Open Data](#) site. This data frame contains the districts of all the cities in Austria but for our study we would only require the part that pertains to Vienna.

Singapore is divided into 55 planning areas as found on [Singapore's Planning Areas Wiki](#). The table contains lots of other information that we will not need for our segmentation.

We will only need the list of districts in each city in order to be able to geocode their co-ordinates and cluster them based on their most common venues

Foursquare API is used to retrieve the most common venues in each district