The Battle of Neighborhoods - Singapore

1. Data acquisition and cleaning

1.1. Data sources

Locations in Singapore may be classified based on postal codes which are administered by Singapore post. The postal district, postal sector which is associated with general locations may be accessed from https://en.wikipedia.org/wiki/Postal_codes_in_Singapore. Additional information regarding the amenities near those general locations may be access through FourSquare API.

1.2. Data cleaning

Data scraped from the Wikipedia URL above was saved into a pandas data frame. Redundant data such as null data and duplicates were removed.

1.3. Feature selection

We will be looking at common venues in each district.