

Battle of the Neighborhoods Capstone Project:

Data

2. DATA

2.1 Data Sources and Relevance

2.1.1 General Information

The first data source provides basic information about the neighborhoods¹ of Singapore scraped from Wikipedia ([here](#)), including:

- Neighborhood name
- Total area (km²)
- Residential area (km²)
- Number of dwelling units
- Projected ultimate number of dwelling units
- Population

2.1.2 Location Data

The second data source provides each neighborhood's name along with its latitude and longitude coordinates pulled from geographic data ([here](#)), including:

- Neighborhood name
- Latitude
- Longitude

2.1.3 Home Rental Pricing

The third data source provides the median monthly rent (in Singapore dollars, as of Q1 2020) for several different sizes of flat^{2,3} pulled from the Housing Development Board (HDB) website ([here](#)), including:

- Neighborhood name
- Median monthly rent for a 3-room flat
- Median monthly rent for a 4-room flat
- Median monthly rent for a 5-room flat

2.1.4 Venues Nearby

The fourth dataset is produced using the **Foursquare API** ([here](#)) to find all venues within a 500 meter radius of each neighborhood and examine what categories of venue are more common in each neighborhood; additional details about this dataset can be found in **Section 2.2 Data Cleaning**.

¹ It is worth noting that we will use Housing Development Board (HDB) Towns to constitute our neighborhoods as opposed to Districts or Planning Areas.

² Here we focus on Housing Development Board (HDB) flats as opposed to pricier private condos.

³ Note that there are some missing values in this dataset; no value is provided when there are less than 20 rental transactions in the quarter for that particular town and flat type.