IBM Applied Data Science Capstone

The Battle of Neighborhoods:

Facilities Favoured by Singapore's HDB Residents

Heranshan Subramaniam

Business Problem

Singapore is a small city state with a land area of 728.6 km2, housing a total population of 5.76 million. This makes Singapore one of the most densely populated countries in the world with a population density of 8,041 ppl/km2. 81% of Singapore's population resides in public housing, known as Housing Development Board (HDB) flats. These flats are situated in 26 towns within Singapore.

This project intends to study the neighbourhood facilities favoured by the population residing in HDB flats. This will be done by studying the change in town population between 2018 and 2019 versus the facilities available in these towns.

Data

Estimated Singapore Resident Population in HDB Flats

Data will be retrieved from the Singapore Governments open dataset on Estimated Singapore Resident Population in HDB Flats from data.gov.sg website.

Raw data retrieved from the governments dataset includes town population spanning 11 years from 2008 to 2019. As only data from 2018 and 2019 is required, data is cleaned up to remove data from other years. A new column for population change is also added, taking the difference between 2019 and 2018 population.

Singapore Town Location Data

Data coordinated of Towns are retrieved using GeoPy.

Singapore Town Top Venue Recommendations

FourSquare API is used to retrieve the most common venue categories in each Town. This feature is used to group the neighbourhoods into clusters. The following data are retrieved on the first query:

- Venue ID
- Venue Name
- Coordinates
- Category Name