

# **RECOMMENDATION OF OFFICE OR RENTAL SPACES FOR ENTREPRENEURS IN SINGAPORE**

## **2. DATA PREPARATION**

### **2.1 Data Sources**

The main dataset is downloaded from Urban Redevelopment Authority (URA) website. The dataset contains the median rental of office and retail space in quarter 3 of 2019. The dataset contains the Street name and the associated rent in Singapore dollars (\$) per square meter (psm) per month. These rental values are categorized separately for Retail and Office space. Not all the street locations can be used for both office and retail businesses. Some of the street locations can be used for both office and retail spaces.

### **2.2 Data Cleaning**

There are total 152 street locations in the dataset. Some of the values in Office and Retails spaces are NaN because those locations may not be used for office or retail spaces. So, these NaN values are replaced with value 0 and consider this value when segmenting the locations based on the criteria.

The Foursquare API is used to collect that data based on the locations. Foursquare is a technology company that built a massive dataset of accurate location data. Foursquare were very smart about building their dataset. They crowd-sourced their data and had people use their app to build their dataset and add venues and complete any missing information they had in their dataset.

Currently its location data is the most comprehensive out there, and quite accurate that it powers location data for many popular services. Their API and location data are currently being used by over 100,000 developers. Foursquare allows you to make requests through its web platform in order to get familiarized with the API, without complex steps such as user authentication.

This Foursquare API is used to generate the Venue Categories near each street location using the Coordinates and get the relevant venues such as Bus Station, Train Station and Food Court.