

1. Data

Following data is needed to develop the projects:

- List of PAs in central Singapore: In this project, we specifically investigate central Singapore.
- The geographic coordinates of each PA: Geographic information of each PA can help to get the venue data.
- Venue data: Information help us to cluster the PA and thus determine the location of opening a bubble tea shop.

Data sources:

We get the list of PAs in central Singapore from Wikipedia (https://en.wikipedia.org/wiki/List_of_places_in_Singapore). We use Geocoder to gain the geographic information of the PA. Then Foursquare API is needed to discover the venue information of respective places.

We use the K-means method to cluster the PAs using machine learning method, and view the map use Folium.