Clustering Apartment Communities by Surrounding Venues

Data

The data needed for this project, along with how the data will be obtained are described as follows:

- **List of apartment communities**

A full list of the apartment communities in Houston can be obtained by web scraping using Requests and BeautifulSoup. Specifically, the website of Apartment List (https://www.apartmentlist.com/tx/houston) will be scraped. And the name and the address of each community will be retrieved.

- **Geospatial data (latitude and longitude) of apartment communities**

The geospatial data of each apartment community can be acquired by converting the address of each community into the corresponding coordinates using the geocoders provided by the geopy package. If an address in our apartment list cannot be recognized by geopy, that apartment community will be deleted from our dataset.

- **Information of surrounding venues**

Information of the surrounding venues can be conveniently obtained through the Foursquare API. The explore endpoint will be used. And venues within 1000 meters of each community will be retrieved.