

Coursera Capstone Project: Office Supplier Expansion - New Location Data Analysis

Data description

Data collection will utilise Foursquare data. The API will be used to gather potential customer data in both cities.

For example: Foursquare data could be filtered to find potential customers using category ID's such as Offices, Business Centres etc. The extracted data can then be imported as .json files into data frames ready for preparation.

Data extracted will include features such as city/ borough location, business type, business name along with their longitude and latitude co-ordinates. Borough postal codes will also be scraped from wikipedia which can then be merged with longitude and latitude co-ordinates using geopy library

Data will then be prepared and cleaned before analysis using data frames and mapping techniques using Folium. These processes will establish which city has a greater number of potential customers.

For the second part of the Data Science problem, potential customers within the selected city will then be segmented and clustered. The results overlaid on a map of the city. The data maps should establish the ideal borough within the city for the clients business expansion.