Coursera Capstone

Capstone Project - The Battle of Neighborhoods

Author: Varun Panditpautra

1. Data Section

Following sources of data are used while executing the Capstone Project: -

- **Data title: -**
  Open Data published by Government of UK under the section **HM Land Registry: Price Paid Data**

  Type of data: -
  Dataset in form of CSV file

  Duration: -
  August 2018 data

  Description of the dataset: -
  Price Paid Data includes information on all property sales in England and Wales that are sold for full market value and are lodged with us for registration.

  The dataset includes the transactions received at HM Land Registry in the period from the first to the last day of August 2018.

  This dataset was downloaded and later hosted on https://labs.cognitiveclass.ai/ for ease of use.

  Source: -
  http://landregistry.data.gov.uk/

- **Data title: -**
  Google Maps Geocoding API

  Type of data: -
  JSON

  Duration: -
  N/A

  Description of the data: -
  Location coordinates obtained by Gmaps API calls.

  Location Information obtained from Price Paid Dataset is used to obtain the location coordinates from Google Maps.

  A separate Python script has been developed to extract the unique street names, district names from the Price Paid Dataset and embed those in the GMaps API calls to obtain the required information.

  Source: -
  Google Cloud Platform/ Google Maps

- **Data title: -**
  Foursquare location data

  Type of data: -
  JSON

  Duration: -
  N/A

  Description of the data: -
  Location coordinates obtained by Foursquare API calls.

  To determine the proximity of various amenities as per the client's requirement, Foursquare location data is used.

  Source: -
  https://foursquare.com/