



Data Description:

Data Link: <https://api.foursquare.com/v2/venues/explore?> NEAR_LOCATION = 'Santa Clara, CA'

Will use Santa Clara dataset which we scrapped from foursquare. Dataset consisting of ZipCode, Business Name, Street, Business Category, Latitude and Longitude.

Foursquare API Data:

We will need data about different businesses with their categories in different neighborhoods of Santa Clara in a radius of 10km. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

Yelp Business Search API Data:

<https://api.yelp.com/v3/businesses/search>

After finding the list of businesses, we then connect to the Yelp API to gather information about the business rating about the venues. We draw a bounding box of 5kms around the businesses based on their latitude and longitude.

Bing API Data:

<https://dev.virtualearth.net/REST/v1/Traffic/Incidents/>

We then connect this boundingbox data to bing to get the traffic incident data.

The data retrieved from a combination of FourSquare, Yelp and Bing contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. ZipCode
2. Street
3. Business Category
4. Business Name
5. Business Rating
6. Business Latitude and Longitude
7. BoundingBox Latitude and Longitude around the Business
8. Traffic Incident Counts in the BoundingBox