1. **Data Acquisition and Cleaning**

**Data sources**

For the above said business problem data is collected from below link and Web scraping is done

<https://www.geonames.org/postal-codes/AU/VIC/victoria.html>

Based on Latitude and Longitude values obtained for Neighbourhoods, we use Foursquare API to get

1. Venue Name

2. Venue ID

3. Venue Location

4. Venue Category

5. Number of Likes

6. Top 10 Venues

**Data Cleaning**

Using Beautiful Soup and Requests I scraped the data from the above link and converted to json file. Once it is done, I took table from the website having Postal code of Melbourne, Latitude, Longitude, Neighbourhoods, State and Country.

Since we do not require Country for our data as we already have other features to determine venues, I dropped that column and converted to a data frame having meaningful information.

I created account in [www.foursquare.com](http://www.foursquare.com) and got client ID and client secret to connect API to get the venues and clustered those venues based on similarities.