

# **Using Data Science to identify the optimal location for a new Italian restaurant in Sydney**

## **Data**

To solve this problem, we will need the following data:

- List of neighborhoods in Melbourne. This defines the scope of the project, as confined to the city of Melbourne, Australia
- Latitude and Longitude co-ordinates of these neighborhoods. We will need these to plot the map and also to get the venue data
- Venue data, related to restaurants, particularly Italian restaurants. We will use this data to perform clustering and other analysis on the neighborhoods

## **Sources and extraction of Data**

- We will use the data from the Australian government datasets located at [Data.com.au](http://Data.com.au) which is a geojson file containing the administrative boundaries of Sydney. We can extract the district names from there as well as any map boundaries if required
- Getting latitude and longitude data of these neighborhoods via geocoder package
- Using foursquare API to get details of venues in these neighborhood