How the above data will be used to solve the problem?

Many small suburbs in Melbourne city have a similar postcode. All such suburbs will be combined to condense the data. Then all the suburbs of Melbourne will be displayed on a map using the Folium library of Python. The purpose of displaying the suburbs on the map is to see the vicinity of the suburbs.

In the next step, for each suburb, the ten most famous venues will be fetched using Foursquare API. The newly created data will be clustered using the scikit-learn library of Python. The optimized number of clusters will be created by running the clustering function for different values of the number of clusters.

These clusters will be displayed on the map to visually observe, how these clusters are covering the Melbourne region. Each cluster will be individually inspected for a variety of venues. Also, each cluster will be inspected for having Thai restaurants.

The cluster having most Thai restaurants and other kinds of dining places in different suburbs will be finalized for further evaluation. The selected cluster will be further summarized to the suburbs having a Thai restaurant in five topmost famous venues. All these suburbs will be displayed on the map to see the distance of these suburbs from the busiest destinations in Melbourne city.

In the last step, these suburbs will be ranked based on the vicinity and the presence of fine dining places in the suburb. The top three suburbs will be suggested to James to open his new Thai restaurant.