Recommending an Indian restaurant at Parramatta or Sutherland LGA

Aneesh Raj

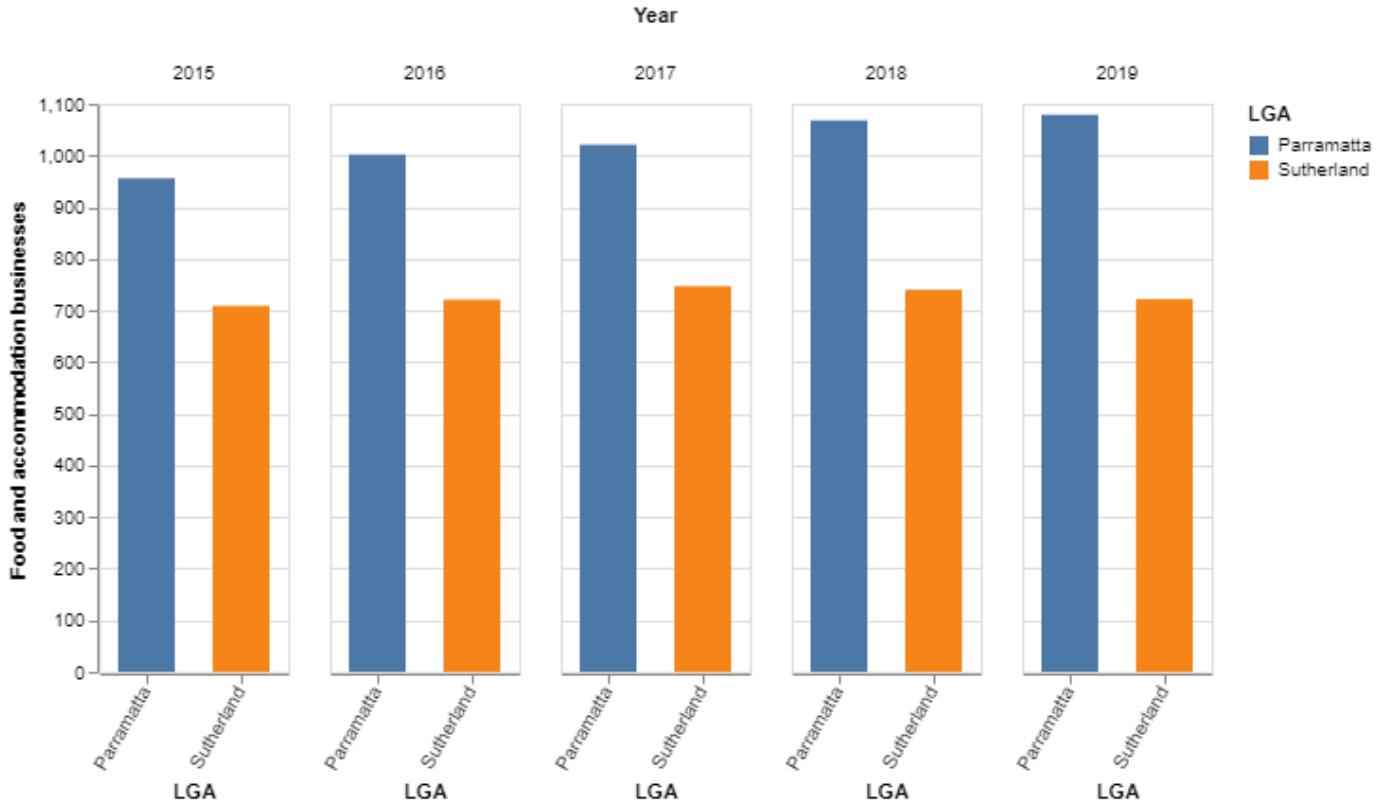
Business Problem

- The objective of this project is to analyse the two local government areas (LGA) in Sydney (Parramatta and Sutherland Shire) and recommend the three best suburbs in one of these LGAs to start an Indian restaurant.
- The target audience for this project is entrepreneurs who is thinking about opening an Indian restaurant in any of the two LGAs in Sydney.

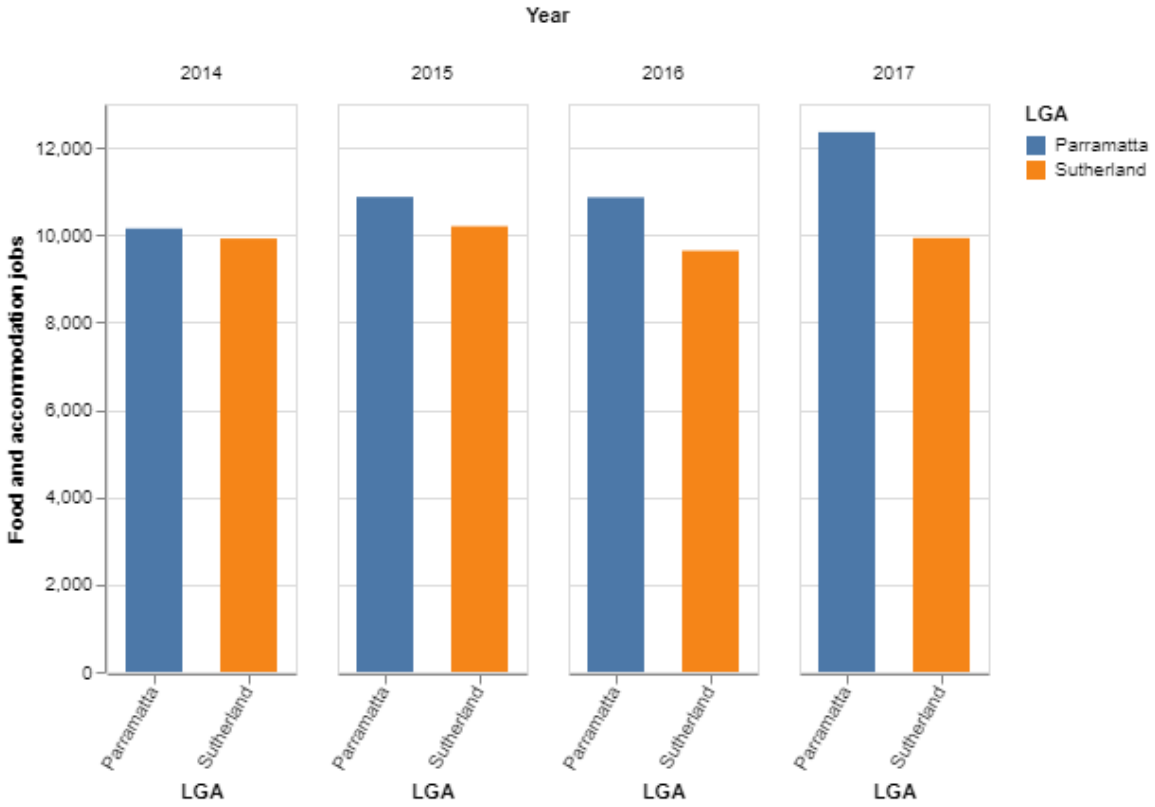
Data Acquisition and Cleaning

- Australian Bureau of Statistics pages are used to gather data related to Parramatta and Sutherland LGA.
- Exploratory data analysis is performed to choose the LGA that is best suited to start an Indian restaurant business.
- Suburb list of the chosen LGA is then collected by scraping the wiki page.
- Foursquare API is used to gather venue details in the suburb list.
- Geopy is used to get the coordinates of the suburbs.

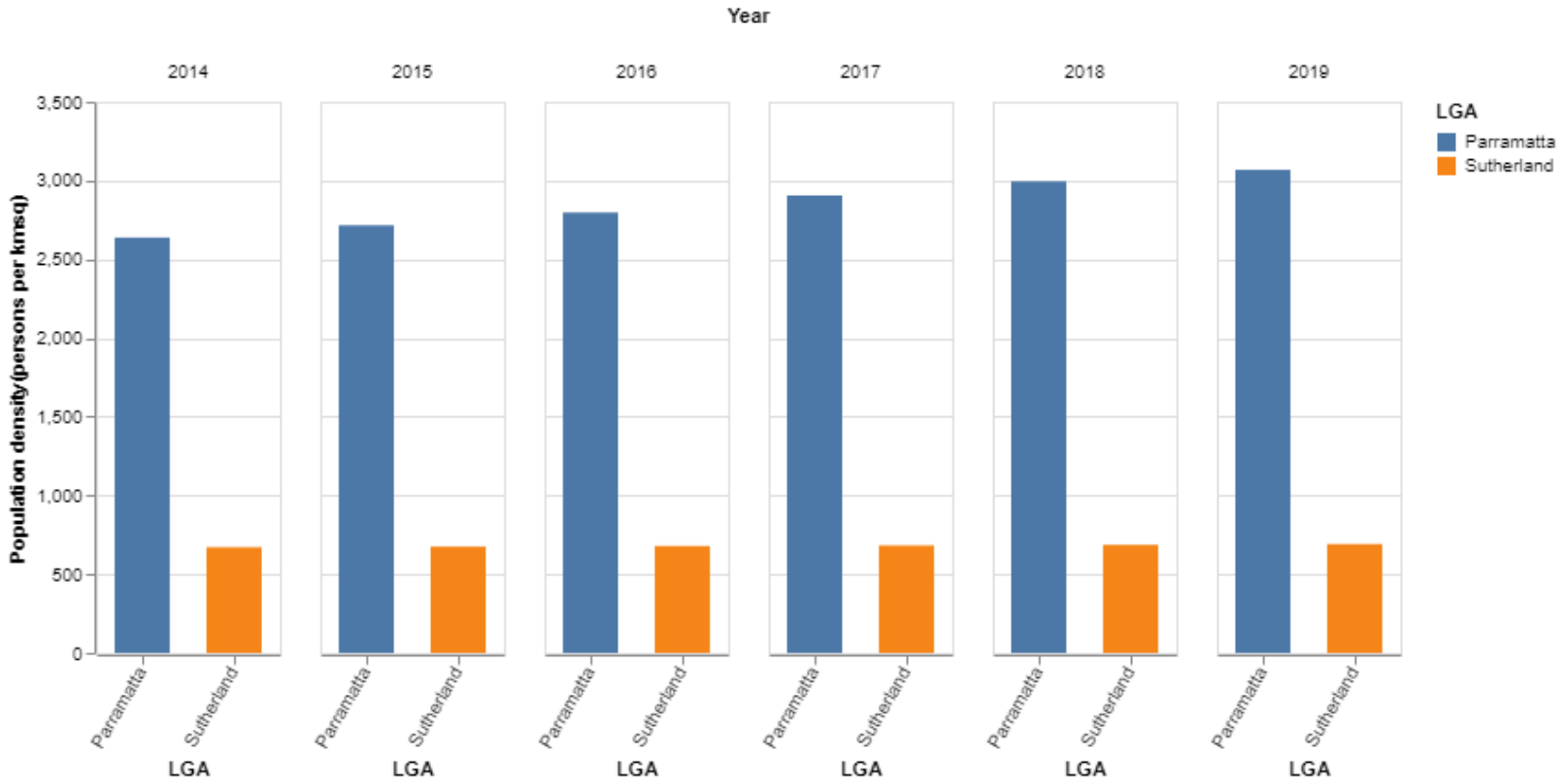
Exploratory data analysis



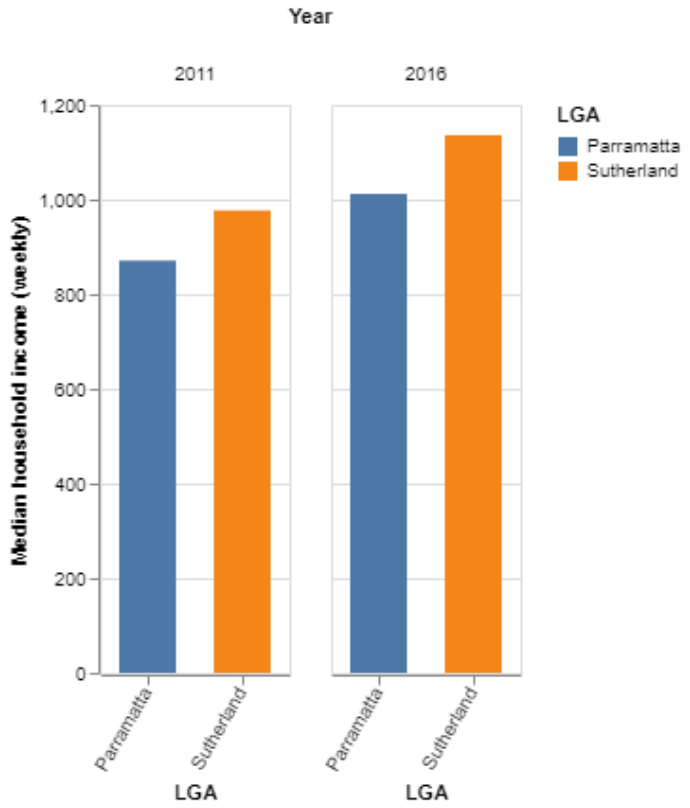
Exploratory data analysis



Exploratory data analysis



Exploratory data analysis

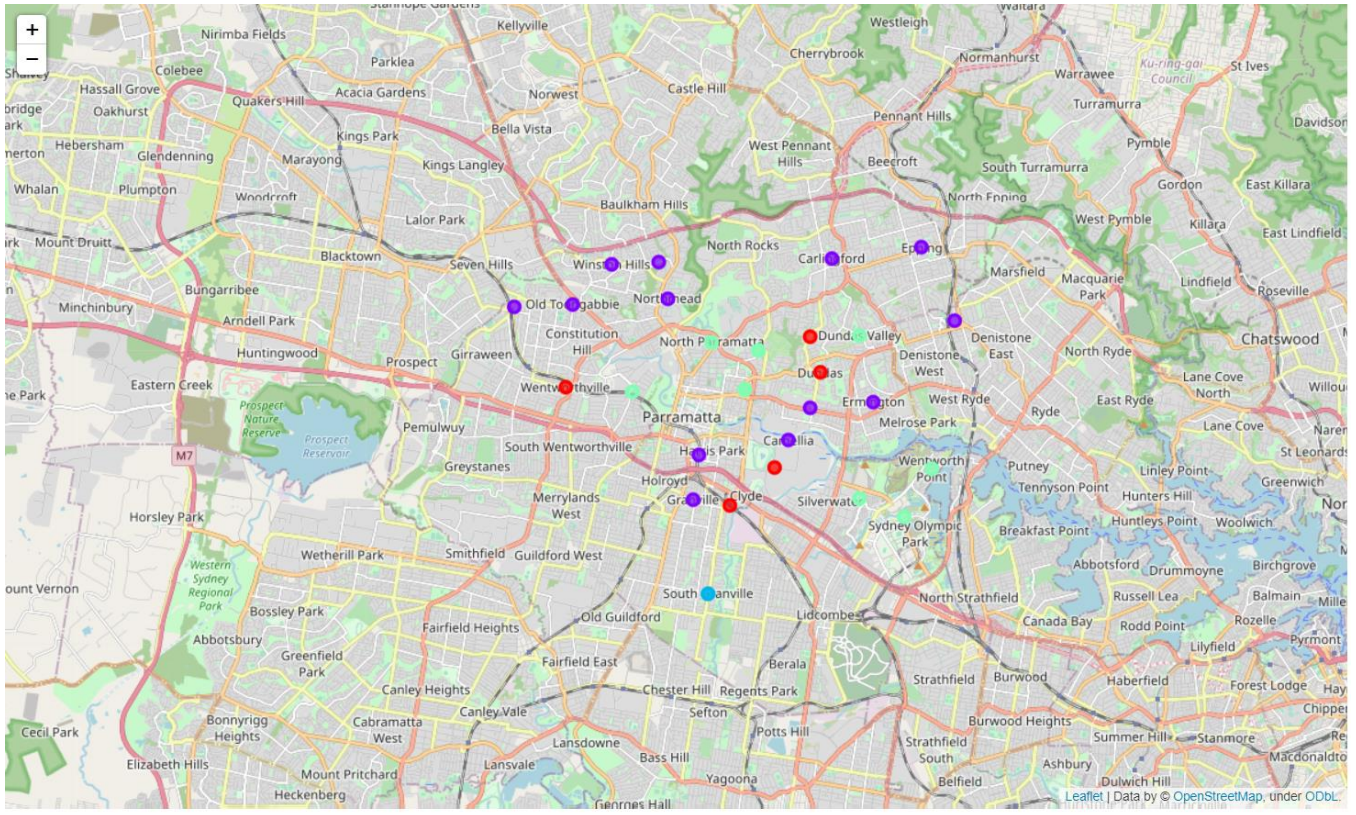


- From these EDAs performed, it is pretty evident that Parramatta LGA is best suited to start the Indian restaurant business.

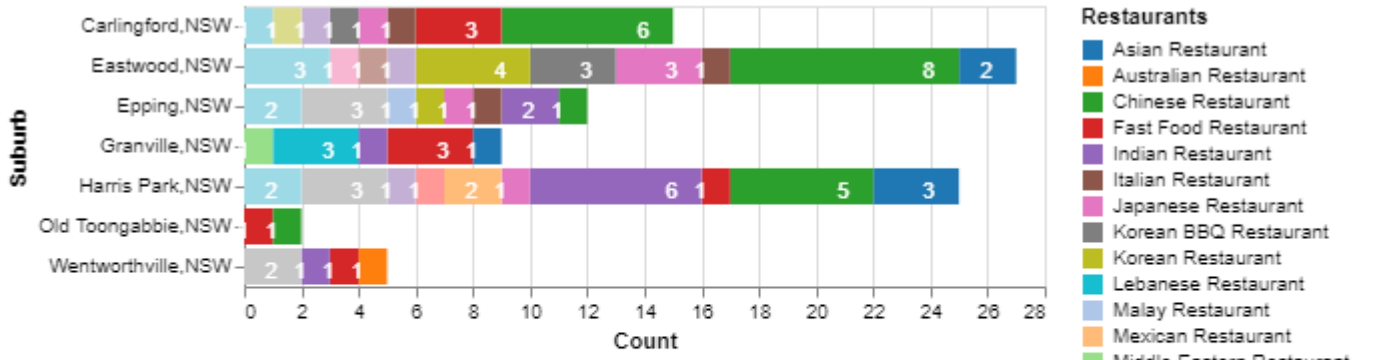
K means clustering

- K means clustering is performed to identify the suburbs in Parramatta LGA with restaurant as the common venues.
- Then those suburbs are filtered to identify the suburbs with restaurant as the venue in 1st, 2nd or 3rd most common venue.
- After that only the suburbs with restaurants are chosen.

Results and discussion



- The colours red, purple, and green represents cluster 0, 1, and 3 respectively.
- The clusters in green have the common venues as café.
- The cluster in red has common venues as train stations.
- Finally, Cluster in green has what is required, the most common venues as restaurants.
- A filter is applied to choose only the suburbs that has restaurants in at least two of the first three common venues columns. Then only the restaurant venues are included and all other venues are filtered out.



- The top 3 suburbs to start an Indian restaurant is Eastwood, Carlingford and Epping.
- Harris Park is ignored because it has the highest number of Indian restaurant and therefore it will be very highly competitive.
- Eastwood has 27 restaurants but there are no Indian restaurants.
- Carlingford and 15 restaurants and no Indian restaurant.
- Epping has 12 restaurants and only 2 Indian restaurants. So, it is recommended to open an Indian restaurant in any of these three suburbs.

Conclusion

- This project has made use of Australian Bureau of Statistics website to get data related to the food and accommodation industry in Parramatta and Sutherland LGA.
- Exploratory data analysis is performed on that data to choose the LGA that is best suited to start an Indian restaurant.
- Then Foursquare API is used to get the restaurant locations situated in Parramatta LGA as it seems to be the best to start an Indian restaurant.
- K-means clustering algorithm has been used to cluster the suburbs in Parramatta LGA to find out the suburbs that has restaurants as the 1st, 2nd and 3rd most common venue.
- Then best 3 suburbs are recommended to start an Indian restaurant based on the customer foot traffic (Based on highest concentration of restaurants) as well as least amount of direct competition (By eliminating suburbs with high number of Indian restaurants).