**New Indian Restaurant in Sydney region – Analysis Report**

# Introduction

A famous Indian Restaurant chain is expanding business worldwide and looking for appropriate locations in Sydney Australia for opening a new Restaurant. Chief Financial Officer of the company is seeking help from a Data Science company for recommendations on locations where a new restaurant should be opened. The final output report should recommend suitable suburbs in Sydney area where new Restaurant should have chances of doing good business

# Data

## Suburbs & Region

Sydney is a large city. Details of al suburbs in Sydney area is required along with associated post codes and region. Post codes are numers assigned by Australia Post. Some examples of regions are like North Shore, Northern Beaches, Wester Suburbs etc. Sydney is capitol of NSW and NSW govt has a URL to get datasets.

[https://data.nsw.gov.au/data/dataset](https://data.nsw.gov.au/data/dataset/spatial-services-nsw-suburb)

These data sets are freely available to NSW Govt employees only. For private/ Commercial use approvals are required. In fact, Australis Post does have required data, but datasets are available for commercial use on chargeable basis. We have found another source to get required data easily. The data was downloaded from <https://www.prospectshop.com.au/Files/SydneyMetro_Postcodes.xls>

There are 202 suburbs in Sydney Area divided into 10 regions. Most prosperous regions are North Shore and Inner Sydney

## Crime Data

High crime rates can make potential customers uncomfortable. This will an important criterion in deciding new location for a new restaurant, Data can be sourced by

Since crime data is very complicated and has many attributes like crime type, number of incidents etc. It was decided to classify crimes in a suburb to 3 categories. High, Medium and Low. Stakeholders will be interested in opening a new restaurant in low/ medium crime rate suburbs. Western Suburbs and some of inner Suburbs have got relatively high crime rate. Otherwise, Sydney is considered to be safe city

https://data.nsw.gov.au/data/dataset/8a39cc2d-3c5a-4274-97de-a7f8ce5b24ee

## Income

Family income will also be an important consideration in deciding location of new restaurant. Data on average family income suburb wise will be required for assessment. Data can be sourced from

<http://abs.gov.au/household-income>

Family incomes is grouped into 3 categories. High, Medium and Low. Targeted customers are with family and with total family earnings in the range of Median earning of Sydney. People in very high earning bracket tends to go expensive restaurant in City

## Parking

People tends to visit more to Restaurants with ample Parking. In deciding recommended location. Aim is to get data on number of public parking spaces in each Sydney suburb

**Source :** [**https://data.gov.au/dataset?tags=parking**](https://data.gov.au/dataset?tags=parking)

It was difficult to get authentic parking data from NSW government sit. For the time being parking data is not considered for the assessment. It will include in the next phase

# Methodology

Research was done on existing Indian / Asian restaurant in Sydney’s suburbs. Main objective was to find number of restaurants so that this data can be used in the Analysis

Based on experience in other western countries and discussions with Stakeholders, following points were considered

1. One existing India restaurant and at least one other Asian Restaurant (Thai, Malaysian, Chinese, Japan
2. No Indian Restaurant and at least 5 Asian Restaurants
3. Areas with Medium family income should be preferred. High income families does not have liking to India Cuisines
4. Areas having low/ Medium crime rates should be considered

Get all required data from NSW Govt / Australian Bureau of Statistics sites

Where ever appropriate data cannot be found interpolate on base available data. Make assumptions and document these assumptions in final report

Latitude and Longitude of each suburb was acquired through geolocator.

Four Square API’s were used to get Indian/ Asian Restaurant details after getting co-ordinates of suburbs. The Foursquare API provides location-based experiences with diverse information about venues, users, photos, and check-ins.

Create a panda data frame with following features for each suburb in Sydney Area

1. Number of Indian Restaurants
2. Number of Asian Restaurants
3. No crime incidents
4. Average Family Income

Clustering modelling approach was used to group suburbs first and then find the most appropriate one. K-Means Algorithm was used to form clusters. Kmeans is vastly used for clustering in many data science applications, especially useful if you need to quickly discover insights from unlabelled data

Analyse each cluster and identify clusters with following characteristics

1 No Indian Restaurants

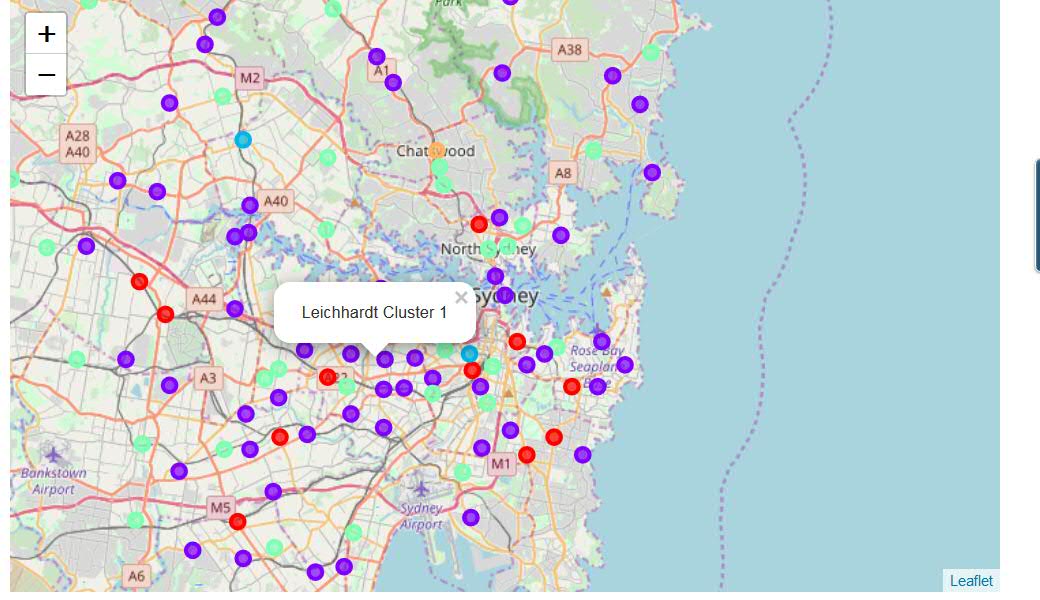
2 Many Asian Restaurants

4 suburbs with Medium Income

5 Less crime rate

# Results

Five clusters were made using k-Means. Below is the plot of these five clusters



Cluster 1: These are the suburbs in cluster 1. In this cluster, mostly there are existing Indian Restaurants and some Asian restaurants.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Ashfield | 2 | 11 | | Auburn | 1 | 10 | | Beverly Hills | 0 | 10 | | Bondi Junction | 3 | 7 | | Campsie | 1 | 9 | | Chippendale | 2 | 8 | | Crows Nest | 4 | 11 | | Kings Cross | 6 | 9 | | Kingsford | 0 | 11 | | Lidcombe | 1 | 7 | | Pyrmont | 1 | 7 | | Randwick | 5 | 6 | |
| Cluster 2:  |  |  |  | | --- | --- | --- | | Annandale | 1 | 0 | | Balmain | 2 | 1 | | Beecroft | 1 | 1 | | Belfield | 1 | 0 | | Belmore | 0 | 2 | | Belrose | 0 | 1 | | Blakehurst | 0 | 0 | | Bondi | 0 | 2 | | Botany | 0 | 0 | | Brookvale | 0 | 2 | | Cammeray | 0 | 1 | | Campbelltown | 2 | 1 | | Camperdown | 0 | 1 | | Canterbury | 0 | 2 | | Carlingford | 0 | 1 | | Carlton | 0 | 0 | | Chullora | 0 | 0 | | Concord | 0 | 0 | | Coogee | 1 | 1 | | Croydon | 0 | 1 | | Croydon Park | 0 | 1 | | Drummoyne | 0 | 1 | | Dulwich Hill | 0 | 2 | | Dural | 0 | 1 | | Edgecliff | 0 | 1 | | Erskine Park | 0 | 1 | | Fairfield | 0 | 1 | | Five Dock | 1 | 0 | | Forestville | 0 | 1 | | Gordon | 1 | 1 | | Granville | 1 | 1 | | Haberfield | 0 | 2 | | Harbord | 1 | 0 | | Heathcote | 0 | 1 | | Homebush | 1 | 0 | | Homebush Bay | 1 | 0 | | Hoxton Park | 0 | 1 | | Kensington | 0 | 2 | | Killara | 0 | 0 | | Kingsgrove | 0 | 2 | | Kogarah | 0 | 1 | | Leichhardt | 0 | 0 | | Liverpool | 4 | 0 | | Manly | 0 | 2 | | Marrickville | 1 | 1 | | Meadowbank | 0 | 2 | | Menai | 0 | 1 | | Milsons Point | 1 | 0 | | Minchinbury | 0 | 1 | | Minto | 0 | 0 | | Miranda | 0 | 2 | | Mona Vale | 0 | 1 | | Mosman | 0 | 0 | | Narellan | 0 | 2 | | Paddington | 1 | 0 | | Parramatta | 0 | 1 | | Peakhurst | 0 | 0 | | Pennant Hills | 2 | 1 | | Penrith | 2 | 1 | | Penshurst | 1 | 1 | | Petersham | 1 | 1 | | Punchbowl | 0 | 0 | | Redfern | 1 | 1 | | Regents Park | 0 | 0 | | Rhodes | 1 | 1 | | Richmond | 1 | 0 | | Rooty Hill | 0 | 1 | | Rose Bay | 0 | 0 | | Rosebery | 0 | 0 | | Rozelle | 0 | 0 | | Rydalmere | 0 | 2 | | Sans Souci | 0 | 1 | | Seven Hills | 2 | 0 | | Stanmore | 1 | 0 | | Sydney City | 0 | 0 | | Toongabbie | 0 | 1 | | Waverley | 0 | 0 | | Windsor | 0 | 2 |  Cluster 3 |
| Eastwood 0 25  Ultimo 0 24   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | Cluster 4 |  |  |  |  | |  |  |  |  |  |  | |
| |  |  |  | | --- | --- | --- | | Artarmon | 1 | 2 | | Avalon | 0 | 3 | | Balgowlah | 1 | 3 | | Bankstown | 1 | 3 | | Baulkham Hills | 1 | 2 | | Burwood | 0 | 4 | | Castle Hill | 2 | 3 | | Chester Hill | 0 | 3 | | Cremorne | 0 | 5 | | Cronulla | 1 | 3 | | Dee Why | 2 | 2 | | Double Bay | 1 | 2 | | Enfield | 0 | 3 | | Epping | 2 | 2 | | Gladesville | 0 | 3 | | Glebe | 0 | 3 | | Gymea | 1 | 3 | | Hurstville | 0 | 4 | | Jannali | 1 | 2 | | Kellyville | 1 | 3 | | Lakemba | 2 | 2 | | Mascot | 1 | 3 | | Merrylands | 0 | 3 | | Narrabeen | 0 | 3 | | Neutral Bay | 4 | 2 | | Newport | 1 | 3 | | Newtown | 0 | 4 | | North Ryde | 0 | 3 | | North Sydney | 0 | 5 | | Padstow | 1 | 2 | | Riverstone | 0 | 3 | | Rockdale | 4 | 3 | | Summer Hill | 1 | 2 | | Surry Hills | 3 | 4 | | Sylvania | 1 | 2 | | Turramurra | 2 | 2 | | Waterloo | 1 | 4 | | Wentworthville | 4 | 2 | | Willoughby | 3 | 4 | |
|  |

Cluster5:

Chatswood 5 16

**Suburbs in Cluster 2 and then in cluster 3 are the favourite ones for setting a new Indian Restaurant**

# Discussion & Recommendations

**Based on cluster modelling and results obtain, these suburbs are recommended for opening new Indian Restaurant:**

**Eastwood: This Sydney suburb is in “Gladesville-Ryde-Eastwood” region. This suburb has many Asian Restaurant but missing Indian Restaurant. It is medium Income area**

**Ultimo: This Sydney suburb is in “Central & Inner Metropolitan” area. It too has many Asian Restaurants but missing Indian. It is with people with high income earnings and low crime rate. It is expected that customer going to Asian restaurants will get attracted to Indian**

# Conclusion

There will be a second phase of the analysis, in which missing data on Parking and more accurate data on Income can be obtained and included

Above analysis has not considered Affordability to set up business. A separate report will be prepared by Accountants on setting up and operations cost. This report must be analysed in conjunction with findings here before decision is taken on location