Text

Description automatically generated



|  |  |
| --- | --- |
| **Team Members** | **Matric Number** |
| Swastik Mishra | G2200524L |
| Prasanthi Ramaswamy | G2202081D |
| Divya Gupta | G2200597K |
| Xia Zhiyang (Michael) | G2201771F |
| Goh Pei Xuan | G2200605C |

**AN6001 GROUP PROJECT BY TEAM 8**

**Restaurantra: Restaurant Recommendation System**

**Dataset and Data Dictionary**

The data used for the visualizations and models were obtained from the Kaggle dataset section. This contains various datasets related to restaurants and the dataset pertaining to US was chosen for this project. The variables in the dataset are as follows:

1. Name :- Represents the name of the restaurant
2. Street Address:- Represents the street address of the restaurant
3. Location :- Represents the state, city and zip code
   1. 20 Cities
   2. 5 States (Washington, Texas, New York, New Jersey, and California)
4. Type :- Represents the type of cuisine served
5. Reviews :- Represents star ratings out of 5
6. No of Reviews :- Represents number of people who have rated
7. Comments :- Represents customer reviews
8. Contact Number :- Represents phone number
9. Trip\_advisor Url :- URL of the restaurant on Trip Advisor
10. Menu :- Represents URL of the menu on Trip Advisor
11. Price Range: Denoted by a ‘$’, representing costs. The more ‘$’ there are, the more expensive the restaurant is.
    1. $: Least Expensive
    2. $$ - $$$: Mid – Tier
    3. $$$$: Most Expensive