

Data Collection and Preprocessing Phase

| | |
|---------------|----------------------------------|
| Date | 13 June 2025 |
| Name | Shreyas A Patil |
| Project Title | Restaurant Recommendation System |
| Maximum Marks | 2 Marks |

Data Collection Plan

| Section | Description |
|-----------------------------|---|
| Project Overview | Develop a restaurant recommendation system to assist users in finding dining options based on their preferences, location, and other relevant factors. By analyzing user preferences, restaurant ratings, and location data, this project aims to provide personalized recommendations that enhance the dining experience for users. |
| Data Collection Plan | The dataset used for this project was sourced from Kaggle and contains detailed information on over 9,000 restaurants in Bangalore, including attributes like name, location, cuisine, ratings, and pricing. This publicly available dataset was collected to support analysis and predictive modeling related to restaurant ratings and customer preferences. |
| Raw Data Sources Identified | The raw data for this project was obtained from the Kaggle dataset titled “Zomato Bangalore Restaurants” by Himanshu Poddar. The dataset is publicly available at https://www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants and includes key restaurant-related attributes such as restaurant names, locations, cuisines, average costs, online delivery availability, and user ratings. |

Raw Data Sources

| Source Name | Description | Location/URL | Format | Size | Access Permissions |
|-------------------------------|--|---|--------|--------|--------------------|
| SmartInternz Provided Dataset | Restaurant-level data including name, location, cuisines, rating and cost. | Data-Set zomato-bangalore-restaurants | CSV | ~ 93MB | Public |