

Data Collection and Preprocessing Phase

Date	6 Dec 2025
Team ID	RT
Project Title	Restaurant Recommendation System
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification

Ensures curated Bangalore restaurant data for TF-IDF cuisine matching and Flask deployment.

Data Collection Plan Template

Section	Description
Project Overview	ML recommender for personalized Bangalore restaurant suggestions via cuisine/review similarity; targets PS-1/PS-2 users (85%+ success)
Data Collection Plan	Single primary source: Kaggle public datasets focused on Bangalore Zomato listings (restaurants, cuisines, locations, ratings, costs).
Raw Data Sources Identified	Kaggle Zomato Bangalore: 51k+ restaurant records with 17 features for content-based filtering; supplemented by basic location validation if needed.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Zomato Bangalore Restaurants	Comprehensive Bangalore listings: name, location, cuisines (North Indian top), rate (mean 4.0), cost (~800 INR), reviews; ideal for TF-IDF soup features.	https://www.kaggle.com/datasets/rajeshhrampure/zomato-dataset/	CSV	574 MB	Public