

## Data Collection and Preprocessing Phase

Date	4 <sup>th</sup> July 2024
Team ID	739754
Project Title	Food demand forecasting for food delivery company
Maximum Marks	2 Marks

### Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

### Data Collection Plan Template

Section	Description
Project Overview	Food Demand Forecasting for a Food Delivery Company using IBM Cloud involves leveraging advanced data analytics and machine learning algorithms to predict the future demand for various food items. By analyzing historical data such as order volumes, time of day, day of the week, weather conditions, and special events, the system can forecast the demand accurately. This helps the food delivery company optimize its operations by efficiently allocating resources, reducing wastage, and ensuring timely delivery to customers
Data Collection Plan	<p>Our base data consists of four csv files containing information about test data, train data and other required information</p> <ul style="list-style-type: none"> <li>•train.csv: Contains information like id, week, center id, meal id, checkout price, base price, emailer for promotion, homepage featured, number of orders. This file is used for training.</li> </ul>

	<ul style="list-style-type: none"> <li>• test.csv: Contains information like id, week, center id, meal id, checkout price, base price, emailer for promotion, homepage featured. This file is used for testing.</li> <li>• fulfilment_center_info.csv: Contains information of each fulfilment center.</li> <li>• meal_info.csv: Contains information of each meal being served.</li> </ul>
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle & UCI, the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information.

### Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	This dataset consists of various parameters like (id, week, center_id, meal_id, check_out price, num_of orders etc)	<a href="https://www.kaggle.com/datasets/kannanaikkal/food-demand-forecasting">https://www.kaggle.com/datasets/kannanaikkal/food-demand-forecasting</a>	CSV	337 kB	Public