

Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	740662
Project Title	Golden Harvest: A predictive model for apple quality assurance
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	A predictive model for apple quality assurance integrates various technologies and data analytics to optimize the harvesting process, ensuring that only high-quality apples are selected. The model leverages real-time data collection, advanced sensors, machine learning algorithms, and user interfaces to provide actionable insights to farmers

Data Collection Plan	Search for datasets related to predicting Apple quality Prioritize datasets with diverse demographic information
Raw Data Sources Identified	The raw data sources for this project include datasets obtained fromKaggle, 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	The dataset comprises Aid, Size, Weight, Sweetness, Crunchiness, Juiciness, Ripeness, Acidity, Quality	https://www.kaggle.com/datasets/nelgiriyeewithana/apple-quality	CSV	379kb	Public