Data Collection and Preprocessing Phase

Date	19 June 2024
Team ID	-
Project Title	Golden Harvesting: 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	This project aims to develop a machine learning model to predict the quality of apples based on various factors. By analyzing historical data on apple quality and environmental conditions during growth, the model will help farmers and distributors make informed decisions about harvesting and storage.
Data Collection Plan	Search for dataset with various factors such as size, color, firmness, and sugar content.
Raw Data Sources Identified	The raw data is obtained from Kaggle which contains information about various attributes of a set of fruits, providing insights into their characteristics

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	This dataset contains information about various attributes of a set of fruits, providing insights into their characteristics. The dataset includes details such as fruit ID, size, weight, sweetness, crunchiness, juiciness, ripeness, acidity, and quality.	https://www.kaggle.com/datasets/nelgiriyewithana/apple-quality/data	CSV	174 kB	Public