Data Collection & Preprocessing Phase

Date	09 February 2026
Team ID	LTVIP2026TMIDS66183
Project Title	Civil Engineering Insight Studio
Maximum Marks	6 Marks

Preprocessing

Data preprocessing is performed to ensure that the collected raw data is accurate, consistent, and suitable for analysis. Initially, the raw data from various sources is cleaned by removing duplicate records, correcting errors, and handling missing values. Data formats, units, and naming conventions are standardized to maintain consistency across datasets. Outliers and invalid entries are identified and treated appropriately to improve data reliability

Section	Description
Data Overview	The data used in this project consists of civil engineering project inputs such as project details, construction progress, cost, schedule, resource usage, and inspection records. The data is collected through user inputs and simulated project datasets.
Data Cleaning	Raw project data is cleaned by removing duplicate records, correcting inconsistencies, and handling missing or invalid values.
Input Validation	Ensures that mandatory project fields are completed, progress values are within valid ranges, cost data is realistic, and resource usage entries are non-negative.
Data Standardization	Standard units, formats, and naming conventions are applied across all datasets to maintain consistency.
Analytical Processing	Validated and standardized data is processed using analytical techniques to derive project performance metrics and trends.
Error Handling	The system handles invalid inputs and processing errors gracefully by displaying appropriate messages and preventing system failure.

Data Preprocessing Templates	
Loading Data	Project-related inputs such as project details, construction progress, cost, schedule, and resource usage are collected through the application interface using text and numeric input fields.
Input Validation	Validation logic ensures that all mandatory fields are filled and that numeric values (cost, progress percentage, resource quantities) fall within acceptable ranges before further processing.
Data Structuring	The validated inputs are organized into a structured data format suitable for analysis and storage.
Analytic Handling	The structured data is processed using analytical methods to compute project performance metrics, trends, and indicators
Output Handling	The generated insights, reports, and visual dashboards are displayed through the application interface for user review and decision-making.