Data Collection and Preprocessing Phase

Date	07 February 2026
Team ID	LTVIP2026TMIDS65517
Project Title	Advancing Nutrition Science through GeminiAI – NutriGen
Maximum Marks	2 Marks

Data Quality Report

The **NutriGen Data Quality Report** summarizes potential data quality issues related to user-provided health and nutrition inputs. Since the project does not rely on an external dataset, data quality management focuses on ensuring accurate, valid, medically reasonable, and meaningful user inputs before processing by the Gemini AI model.

Data Source	Data Quality Issue	Severity	Resolution Plan
User Input (Age)	Age field empty or unrealistic (e.g., 0, 200)	High	Apply validation to ensure age is within a medically reasonable range (e.g., 5–100 years)
User Input (Weight/Height)	Unrealistic or negative values	Moderate	Validate numeric ranges and prevent submission of negative or extreme values.
User Input (Health Conditions)	Missing critical health information (e.g., diabetes not specified)	High	Mandatory selection of health conditions with "None" option if not applicable.

User Input (Dietary Preference)	Conflicting or incomplete dietary choices	Moderate	Use dropdown selections (Vegetarian, Vegan, Non-Veg, etc.) to standardize inputs.