Data Collection and Preprocessing Phase

Date	26 July 2025
Team ID/ Skill Wallet ID	**SWUID20250195143**
Project Title	Anemia Sense: Leveraging Machine Learning For Precise Anemia Recognitions
Maximum Marks	2 Marks

Data Quality Report:

The Data Quality Report summarizes the assessment of the Kaggle dataset, noting that no missing values were found and identifying a moderate class imbalance between anemic and non-anemic records, which was resolved using random undersampling.

Data Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	No missing values	None	N/A
Kaggle Dataset	Imbalance in number of anemic and non-anemic records.	Moderate	Random Undersampling.