

Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	XXXXXX
Project Title	Chronic Kidney disease
Maximum Marks	2 Marks

Data Quality Report Template

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Column names where a bit too difficult to understand	Low	Replaced it with a set of new column names
Kaggle Dataset	Missing values in every column	Moderate	Using mean, median to fill in those values
Kaggle Dataset	Categorical Values where in strings	High	Using label encoding and numerical function to convert them into numerical values