Data Collection and Preprocessing Phase

Date	08 August 2025
Skill Wallet ID	**SWUID20250188325**
Project Title	Predictive Pulse: Harnessing Machine Learning for Blood Pressure Analysis
Maximum Marks	2 Marks

Data Quality Report:

The Data Quality Report 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 Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Missing values in the 'C', 'Age', 'History', 'Patient', 'TakeMedication', 'Severity', 'BreathShortness', 'VisualChanges', 'NoseBleeding', 'Whendiagnoused', 'Systolic', 'Diastolic', 'ControlledDiet', 'Stages' columns.	Moderate	Use mean/median imputation.
Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data.