

## Data Collection and Preprocessing Phase

Date	19 June 2025
Team ID	SWTID1750050475
Project Title	SmartLender - Applicant Credibility Prediction for Loan Approval
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 'Gender', 'Married', 'Dependents', 'Self_Employed', 'LoanAmount', 'Loan_Amount_Term', and 'Credit_History' columns.	Moderate	Use mean/median imputation.
Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data.