Data Collection and Preprocessing Phase

Date	12 June 2024
Team ID	740112
Project Title	Life Style Change Due To Covid Prediction
Maximum Marks	2 Marks

Data Quality Report:

The data quality assessment process identified and addressed key issues related to completeness, consistency, accuracy, and relevance of the dataset. By ensuring data integrity and relevance, the dataset is now suitable for robust analysis and modeling to predict lifestyle changes due to COVID-19 effectively.

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 for numerical columns and mode imputation for categorical columns.
Kaggle Dataset	Categorical data in the dataset	Moderate	Encode categorical variables using appropriate encoding techniques such as one-hot encoding or label encoding.