

Data Collection and Preprocessing Phase

Date	2 july 2024
Team ID	SWTID1720086522
Project Title	Forecasting Economic Prosperity: Leveraging Machine Learning For GDP Per Capita Prediction
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	Missing Data:Some features may have missing values (NaNs)	Moderate	Impute missing data using techniques like mean, median, or regression imputation.
Kaggle Dataset	Outliers:Extreme values in certain features can affect model performance.	low	Apply robust statistical methods (e.g., winsorization) or remove outliers if they significantly impact results.