

## Data Collection and Preprocessing Phase

Date	17 Dec 2025
Team ID	XXXXXX
Project Title	Global Energy Trends: A Comprehensive Analysis of Key Regions and Generation Modes using Power BI
Maximum Marks	3 Marks

### Data Quality Report

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle – Renewable Power Generation Dataset	Missing values for several years or energy types.	Moderate	Use <b>data imputation techniques</b> such as mean substitution or forward-fill for time-series; if too many values missing, remove those rows.
Kaggle – Renewable Power Generation Dataset	Inconsistent units (e.g., some values in GWh, others in MWh)	High	Standardize all units using <b>data transformation rules</b> (e.g., convert MWh → GWh by dividing by 1000).

Kaggle – Renewable Power Generation Dataset	Duplicate rows (same country- year-energy type repeated)	Low	Apply <b>deduplication</b> using Power Query / Python: drop_duplicates() or grouping rules.
---	---	-----	--