

Data Collection and Preprocessing Phase

Date	22 July 2025
Skillwallet ID	SWUID20250176043
Project Title	Global Energy Trends: A Comprehensive Analysis of Key Regions and Generation Modes using Power BI
Maximum Marks	10 Marks

Data Exploration and Preprocessing Template

Section	Description
Data Overview	The project utilizes six datasets covering global energy generation, energy usage, energy per capita, generation modes, and renewable vs non-renewable classification. These datasets were received in CSV format and imported into Power BI for analysis.
Data Cleaning	Removed unnecessary columns and blank rows, standardized naming conventions, corrected inconsistent text formatting (e.g., region names and energy types), and handled missing/null values by replacing them with zero or removing non-critical incomplete rows.
Data Transformation	Transformed region and mode columns into a consistent format (title case), removed duplicates, and extracted relevant date/year information where required. Some datasets were filtered to focus on specific years or geographic regions.
Data Type Conversion	Converted numeric fields (e.g., energy in GWh, per capita use) to decimal/whole numbers, and text columns like Region, Year, Mode to appropriate data types. Ensured "Year" was a whole number and energy metrics were in numeric format.
Column Splitting and Merging	No major splitting was required. Where applicable, columns were renamed or merged using Power Query to simplify column headers for easier visual development.
Data Modeling	Relationships were established using common keys such as Region and Year. Fact tables (e.g., generation data) were connected with dimension data (e.g., modes, regions) to allow flexible filtering and

	interactive visuals
Save Processed Data	All cleaned and transformed data tables were loaded into Power BI. Final Power BI file was saved as Global_Energy_Analysis_RishabhBhatt.pbix