Data Dictionary — Energy Poverty Prediction Project

Primary Source: World Bank — World Development Indicators (WDI)

Unit of Analysis: Country–Year (1990–2023)

Core Variables

attribute	dtype	type	subtype	num. nature	source	desc
country	STRING	Qual.	Nominal	–	World Bank WDI	Country or region name
country_code	STRING	Qual.	Nominal	–	World Bank WDI	ISO-3 country code
year	INT	Quant.	Interval	Discrete	World Bank WDI	Observation year
electricity_access	FLOAT	Quant.	Ratio	Continuous	World Bank WDI	Percentage of population with access to electricity
gdp_per_capita	FLOAT	Quant.	Ratio	Continuous	World Bank WDI	GDP per capita (current USD)
urban_population_pct	FLOAT	Quant.	Ratio	Continuous	World Bank WDI	Urban population as percentage of total
rural_population_pct	FLOAT	Quant.	Ratio	Continuous	World Bank WDI	Rural population as percentage of total
population_density	FLOAT	Quant.	Ratio	Continuous	World Bank WDI	Population per square kilometer
total_population	INT	Quant.	Ratio	Discrete	World Bank WDI	Total population count
renewable_energy_pct	FLOAT	Quant.	Ratio	Continuous	World Bank WDI (EG.FEC.RNEW.ZS)	Renewable energy consumption as percentage of total final energy consumption
government_effectiveness	FLOAT	Quant.	Interval	Continuous	World Bank WDI (GE.EST)	Government effectiveness index (estimate, typically ranges -2.5 to 2.5)

Derived Variables

attribute	dtype	type	subtype	num. nature	source	desc
risk_category	STRING	Qual.	Ordinal	–	Derived	Energy poverty classification based on electricity access thresholds (Severe, Moderate, Minimal)
prediction_year	INT	Quant.	Interval	Discrete	Derived	Future forecast year used for prediction (2024–2027)

attribute	dtype	type	subtype	num. nature	source	desc
predicted_electricity_access	FLOAT	Quant.	Ratio	Continuous	Model Output	Predicted percentage of population with access to electricity

Risk Category Definition

Category	Electricity Access
Severe	< 50%
Moderate	50% – 89%
Minimal	90% - 99%