

Data Warehouse Outline

STAR SCHEMA:

- 2 Fact Tables
 - Renewable Output – Renewable Electricity Dataset
 - Power Generation – Power Generation by Fuel Source Dataset
- 4 Dimension Tables
 - Date – Derived (Synthetic)
 - Geography – Renewable Electricity Dataset
 - Fuel Source – Power Generation by Fuel Source Dataset
 - Air Temperature – Average Mean Surface Air Temperature Dataset
- *Should we just modify the tables to have 1990 - 2023? For consistency reasons*

Fact Tables

RENEWABLE_OUTPUT (Primary)			
Attributes	Data Type	Key	Description
dateKey	INT	Foreign	FK to DIM_DATE
geoKey	INT	Foreign	FK to DIM_GEO
renewOutPct	DECIMAL(17, 15)	-	Renewable Electricity output as %

FUEL_GENERATION (Secondary)			
Attributes	Data Type	Key	Description
dateKey	INT	Foreign	FK to DIM_DATE
fuelKey	INT	Foreign	FK to DIM_FUEL_SOURCE
powGenGWH	INT	-	Power generation in Megawatt-hours (GWh) by fuel source

Dimension Tables

DIM_DATE			
Attributes	Data Type	Key	Description
dateKey	INT	Primary	Surrogate Key (Auto-INC)
calendarYear	INT	-	Year

DIM_GEO			
Attributes	Data Type	Key	Description
geoKey	INT	Primary	Surrogate Key (Auto-INC)
countryCode	VARCHAR(3)	-	Country Code (PH, USA, etc.)
countryName	VARCHAR(100)	-	Full Country Name
regionGrp	VARCHAR(50)	-	Denormalized high-level grouping (Arab World, Sub-Saharan Africa, etc.)

DIM_FUEL_SOURCE			
Attributes	Data Type	Key	Description
fuelKey	INT	Primary	Surrogate Key (Auto-INC)
fuelName	VARCHAR(50)	-	Energy Source Name
isRenewable	VARCHAR(1)	-	Yes (Y) or No (N)
energyCateg	VARCHAR(20)	-	Broad Categorization (Fossil, Renewable, Nuclear, etc.)

DIM_TEMP			
Attributes	Data Type	Key	Description
tempKey	INT	Primary	Surrogate Key (Auto-INC)
year	INT	-	Year
avgMeanTempDegC	DECIMAL(4, 2)	-	Average Mean Surface Air Temperature in Celsius

SCHEMA VISUALIZATION:

https://lucid.app/lucidchart/4f7c3a4c-bcc1-4fb8-8358-87ba2503e385/edit?viewport_loc=261%2C-66%2C3299%2C1364%2C0_0&invitationId=inv_c13fde24-a352-42a2-980b-5d17186ef064

In progress...