**SDG problem definition document, SQL scripts and Integration documentation**

**SDG: AFFORDABLE AND CLEAN ENERGY (SDG 7)**

**Problem Definition: Energy Wastage in Households**. Many households waste energy due to inefficient usage patterns, especially during peak hours. The goal is to use data to identify high-consumption times and activities, offer actionable insights, and suggest ways for households to reduce their energy consumption.

**DATABASE DESIGN**

The database was designed using MySQL database and exported to my computer from where it was imported to my google sheet

**ERD (Entity-Relationship Diagram):**

Households: Represents individual households with fields like household\_id, houdehold\_name, household\_size, and income\_level.

Devices: Listed energy-consuming devices with fields like device\_id, device\_type (e.g., air conditioning, lighting), wattage, and efficiency\_rating.

EnergyConsumption: Recorded energy usage per device with fields like consumption\_id, household\_id, device\_id, date, time\_period (e.g., peak, off-peak), and energy\_used (in kWh).

Tips: Contains energy-saving tips with fields like tip\_id, device\_type, and suggestion.

**Database Schema:**

**CREATE TABLE Households (**

**household\_id INT PRIMARY KEY,**

**household\_name VARCHAR(100),**

**household\_size INT,**

**income\_level VARCHAR(50)**

**);**

**CREATE TABLE Devices (**

**device\_id INT PRIMARY KEY,**

**device\_type VARCHAR(100),**

**wattage INT,**

**efficiency\_rating VARCHAR(20)**

**);**

**CREATE TABLE EnergyConsumption (**

**consumption\_id INT PRIMARY KEY,**

**household\_id INT,**

**device\_id INT,**

**date DATE,**

**time\_period VARCHAR(20),**

**energy\_used DECIMAL(10, 2),**

**FOREIGN KEY (household\_id) REFERENCES Households(household\_id),**

**FOREIGN KEY (device\_id) REFERENCES Devices(device\_id)**

**);**

**CREATE TABLE Tips (**

**tip\_id INT PRIMARY KEY,**

**device\_type VARCHAR(100),**

**suggestion TEXT**

**);**

**DATA GENERATION:**

The data was generated from 15 different households, living in different location in the same city.

**Households table**

Sql code:

**INSERT INTO Households (household\_id, household\_name, household\_size, income\_level) VALUES**

**(1, 'Cypee', 3, 'High');**

**EnergyConsumption Table**

This table records the energy consumption per device per household, across different dates and time periods.

SQL code:

**INSERT INTO EnergyConsumption (consumption\_id, household\_id, device\_id, date, time\_period, energy\_used) VALUES**

**(1, 1, 1, '2024-10-20', 'peak', 4.5);**

**Devices table**

SQL code:

**INSERT INTO Devices (device\_id, device\_type, wattage, efficiency\_rating) VALUES**

**(1, 'Air Conditioner', 2000, 'Low');**

**Tips Table**

This table provides energy-saving suggestions based on the device type.

Sql code:

**INSERT INTO Tips (tip\_id, device\_type, suggestion) VALUES**

**(1, 'Air Conditioner', 'Use ceiling fans to circulate air before turning on AC.');**

**SQL Programming**

Data Retrieval:

Retrieve energy consumption for each household by time period.

sql code:

**SELECT household\_id, time\_period, SUM(energy\_used) AS total\_energy\_used**

**FROM EnergyConsumption**

**GROUP BY household\_id, time\_period;**

Find the devices consuming the most energy.

sql code:

**SELECT device\_type, SUM(energy\_used) AS energy\_by\_device**

**FROM EnergyConsumption**

**JOIN Devices ON EnergyConsumption.device\_id = Devices.device\_id**

**GROUP BY device\_type**

**ORDER BY energy\_by\_device DESC;**

**Data Analysis:**

Identify households with high peak energy consumption and suggest energy-saving tips based on device types.

sql code:

**SELECT h.household\_id, h.location, d.device\_type, SUM(ec.energy\_used) AS peak\_usage, t.suggestion**

**FROM Households h**

**JOIN EnergyConsumption ec ON h.household\_id = ec.household\_id**

**JOIN Devices d ON ec.device\_id = d.device\_id**

**LEFT JOIN Tips t ON d.device\_type = t.device\_type**

**WHERE ec.time\_period = 'peak'**

**GROUP BY h.household\_id, d.device\_type**

**HAVING SUM(ec.energy\_used) > 50; -- Threshold for high consumption**

**Analyzing the data**

To identify households with high peak energy consumption and suggest energy-saving tips, we’ll use the following approach:

Calculate Peak Energy Consumption: Aggregate energy consumption data during peak hours for each household.

Identify High Consumption Households: Filter households with peak energy usage above a defined threshold (for example, greater than 5.0 kWh).

Suggest Energy-Saving Tips: Join the result with Devices and Tips tables to provide relevant tips based on device types that contribute to high consumption.

Here’s the SQL code to achieve this analysis:

Step 1: Calculate total peak energy consumption for each household

**WITH PeakConsumption AS (**

**SELECT**

**h.household\_id,**

**h.location,**

**SUM(ec.energy\_used) AS total\_peak\_consumption**

**FROM**

**Households h**

**JOIN**

**EnergyConsumption ec ON h.household\_id = ec.household\_id**

**WHERE**

**ec.time\_period = 'peak'**

**GROUP BY**

**h.household\_id, h.location**

**)**

**-- Step 2: Identify high consumption households (threshold set at >5.0 kWh for example)**

**, HighConsumptionHouseholds AS (**

**SELECT**

**pc.household\_id,**

**pc.location,**

**pc.total\_peak\_consumption**

**FROM**

**PeakConsumption pc**

**WHERE**

**pc.total\_peak\_consumption > 5.0**

**)**

**-- Step 3: List devices used by high-consumption households and suggest tips**

**SELECT**

**hch.household\_id,**

**hch.location,**

**hch.total\_peak\_consumption,**

**d.device\_type,**

**d.wattage,**

**t.suggestion AS energy\_saving\_tip**

**FROM**

**HighConsumptionHouseholds hch**

**JOIN**

**EnergyConsumption ec ON hch.household\_id = ec.household\_id**

**JOIN**

**Devices d ON ec.device\_id = d.device\_id**

**LEFT JOIN**

**Tips t ON d.device\_type = t.device\_type**

**WHERE**

**ec.time\_period = 'peak'**

**ORDER BY**

**hch.household\_id, d.device\_type;**

**Data Analysis Using google sheets**

Imported Data: Export query results to CSV or directly import into google sheet.

Analysis:

Used Tables and charts to view total energy consumption by device type, time period, and household.

Created table for peak vs. off-peak usage and average energy usage.

Dashboard:

Main Chart: Bar chart showing total energy usage by all households at peak and off peak periods.

Trend Analysis: Line chart showing energy usage by time (day).

Savings Potential: Highlight potential savings if high-consumption devices are used during off-peak hours.

**Integration and Testing**

Integration Documentation:

Document data import steps from SQL to Excel, including any formatting or pivot table settings.

Exported the xcreated data on mysql database to my computer from where it was imported to my google sheet.

Created a trend analysis for the energy consumption data using a line chart in Google Sheets, using these steps:

Step 1: Organized my data into a table format with the following columns:

Household (Household name)

Device ID (Device number)

Date (Date of consumption)

Time Period (Peak or Off-Peak)

Energy Used (kWh used)

Step 2: Created a Pivot Table for Trend Data

Selected my data range and then selected Pivot Table from Insert tab.

In the Pivot table editor:

Under Rows, added the Date field. Under Values, added Energy Used and select SUM as the summary function. Also, added Household to Rows to separate trend lines per household.

Step 3: Inserted a Line Chart by selecting the pivot table (or the summarized data) to be use for the trend line.

In the Chart editor, selected Line chart.

Set Date as the X-axis and Energy Used as the Y-axis.