* **SDG Selection:** Choose an SDG (e.g., SDG 3: Good Health, SDG 7: Affordable and Clean Energy).
* **Problem Definition:** Define a specific problem within your chosen SDG that can be addressed using data.

**SDG 7: Affordable and Clean Energy**

**Problem:**  
**Unreliable and Inadequate Electricity in Rural Areas**

**Specific Issue:**  
Many rural areas experience frequent power outages and limited electricity access, impacting daily life, education, and economic opportunities.

**Data-Driven Solution Approach:**

1. **Data Collection:**
   * **Electricity Usage**: Consumption patterns, outage frequency, and infrastructure details.
   * **Socioeconomic Impact**: Income, education, and healthcare data.
   * **Energy Source Availability**: Renewable energy adoption and potential.
2. **Data Analysis:**
   * **Identify Patterns**: Analyze outages, consumption, and infrastructure gaps.
   * **Geographic Mapping**: Map affected areas and overlay with socioeconomic data.
   * **Predictive Modeling**: Forecast outage patterns and evaluate interventions.
3. **Solution Design:**
   * **Infrastructure Upgrades**: Improve grid reliability and extend coverage.
   * **Renewable Energy**: Implement solar or wind solutions in underserved areas.
   * **Energy Efficiency**: Promote energy-efficient appliances and practices.
   * **Real-time Monitoring**: Develop systems for outage detection and alerts.
4. **Implementation and Evaluation:**
   * **Pilot Projects**: Test solutions in selected regions.
   * **Feedback Loop**: Gather input and refine interventions.
   * **Scaling Up**: Expand successful solutions to other areas.

**Outcome Goals:**

* **Improved Reliability**: More stable electricity supply.
* **Increased Access**: Better electricity access for households and businesses.
* **Enhanced Quality of Life**: Greater opportunities and improved living standards in rural areas.

**erDiagram**

Region {

int RegionID PK

string RegionName

int Population

decimal AreaSize

}

Household {

int HouseholdID PK

int RegionID FK

int HouseholdSize

}

School {

int SchoolID PK

int RegionID FK

string SchoolName

int StudentCount

}

Business {

int BusinessID PK

int RegionID FK

string BusinessName

int EmployeeCount

}

PowerOutage {

int OutageID PK

int RegionID FK

date OutageDate

int Duration

string Cause

}

Infrastructure {

int InfrastructureID PK

int RegionID FK

string Type

decimal Capacity

string Status

}

RenewableEnergy {

int RenewableEnergyID PK

int RegionID FK

string Type

decimal Capacity

date InstallationDate

}

EnergyConsumption {

int ConsumptionID PK

int EntityID FK

enum EntityType

date ConsumptionDate

decimal Amount

}

Region ||--o{ Household: has

Region ||--o{ School: has

Region ||--o{ Business: has

Region ||--o{ PowerOutage: has

Region ||--o{ Infrastructure: has

Region ||--o{ RenewableEnergy: has

Household ||--o{ EnergyConsumption: consumes

School ||--o{ EnergyConsumption: consumes

Business ||--o{ EnergyConsumption: consumes

**Part 3: SQL Programming**

* **Data Retrieval:** Write SQL queries to retrieve relevant data based on your problem definition.
* **Data Analysis:** Write SQL queries to analyze data and generate insights related to your SDG problem.

**Data Retrieval**

1. **Retrieve All Regions with Their Population and Area Size:**

SELECT RegionID, RegionName, Population, AreaSize

FROM Region;

1. **Get All Households in a Specific Region (e.g., RegionID = 1):**

SELECT HouseholdID, HouseholdSize

FROM Household

WHERE RegionID = 1;

1. **Find All Schools and Their Student Count in a Specific Region (e.g., RegionID = 2):**

SELECT SchoolID, SchoolName, StudentCount

FROM School

WHERE RegionID = 2;

1. **List All Power Outages with Duration Greater Than a Certain Number of Hours (e.g., 2 hours):**

SELECT OutageID, OutageDate, Duration, Cause

FROM PowerOutage

WHERE Duration > 2;

1. **Show All Renewable Energy Installations with Their Capacity and Installation Date:**

SELECT RenewableEnergyID, Type, Capacity, InstallationDate

FROM RenewableEnergy;

1. **Retrieve Energy Consumption Data for a Specific Entity Type and Date (e.g., Household for 2024-07-01):**

SELECT EntityID, Amount

FROM EnergyConsumption

WHERE EntityType = 'Household' AND ConsumptionDate = '2024-07-01';

#### Data Analysis Queries

1. **Calculate Total Energy Consumption by Region**

SELECT R.RegionName, SUM(E.Amount) AS TotalConsumption

FROM Region R

JOIN Household H ON R.RegionID = H.RegionID

JOIN EnergyConsumption E ON H.HouseholdID = E.EntityID AND E.EntityType = 'Household'

GROUP BY R.RegionName;

1. **Find Average Household Size in Each Region**

SELECT R.RegionName, AVG(H.HouseholdSize) AS AvgHouseholdSize

FROM Region R

JOIN Household H ON R.RegionID = H.RegionID

GROUP BY R.RegionName;

1. **Identify the Region with the Highest Number of Power Outages**

SELECT R.RegionName, COUNT(P.OutageID) AS NumberOfOutages

FROM Region R

JOIN PowerOutage P ON R.RegionID = P.RegionID

GROUP BY R.RegionName

ORDER BY NumberOfOutages DESC

LIMIT 1;

1. **Analyze Renewable Energy Capacity by Type**

SELECT Type, SUM(Capacity) AS TotalCapacity

FROM RenewableEnergy

GROUP BY Type;

1. **Determine the Region with the Highest Total Energy Consumption Across All Entities**

SELECT R.RegionName, SUM(E.Amount) AS TotalConsumption

FROM Region R

LEFT JOIN Household H ON R.RegionID = H.RegionID

LEFT JOIN EnergyConsumption E ON H.HouseholdID = E.EntityID AND E.EntityType = 'Household'

GROUP BY R.RegionName

UNION ALL

SELECT R.RegionName, SUM(E.Amount) AS TotalConsumption

FROM Region R

LEFT JOIN School S ON R.RegionID = S.RegionID

LEFT JOIN EnergyConsumption E ON S.SchoolID = E.EntityID AND E.EntityType = 'School'

GROUP BY R.RegionName

UNION ALL

SELECT R.RegionName, SUM(E.Amount) AS TotalConsumption

FROM Region R

LEFT JOIN Business B ON R.RegionID = B.RegionID

LEFT JOIN EnergyConsumption E ON B.BusinessID = E.EntityID AND E.EntityType = 'Business'

GROUP BY R.RegionName

GROUP BY R.RegionName

ORDER BY TotalConsumption DESC

LIMIT 1;

1. **Analyze Average Duration of Power Outages by Region**

SELECT R.RegionName, AVG(P.Duration) AS AvgOutageDuration

FROM Region R

JOIN PowerOutage P ON R.RegionID = P.RegionID

GROUP BY R.RegionName;