INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Energy-Economic-Environmental Profiling using Global EIA Data

Analyzing energy production, consumption, emissions, GDP & sustainability indicators using SQL

Done By: Anuradha K

About me

- **Background ? (B-tech or M-tech)**

B.Arch Graduate

- **Why you want to learn Data Science**

I enjoy learning new things, and data science is a field that's always growing and changing. I want to add new skills like data analysis, programming, and machine learning to my skill set.
In architecture, we often make choices based on design principles. I'm excited to learn how to make decisions based on data, which can provide more precise and impactful results.

- **Any work experience**

No work experience.

- **Share your linkedin and github profile urls**

Anuradha Kilaparthi | LinkedIn
anuradhak0801 (Anuradha K) (github.com)

INNOMATICS RESEARCH LABS

INTRODUCTION

- Every country uses and produces energy — for electricity, transport, industries, and homes.
- But using energy also leads to **pollution**, especially when it comes from coal, oil, or gas. This pollution, called **carbon emissions (CO_2)**, is a big reason for **climate change**.
- At the same time, countries are trying to grow their **economy** (measured using GDP) and improve the quality of life for their people.
- This project looks at **how energy, environment, and economy are connected** using real-world data from different countries.

WHY IS THIS STUDY IMPORTANT?

- To compare countries based on sustainability
- To identify green growth trends
- To guide policy decisions on energy and climate

INNOMATICS RESEARCH LABS

DATA STRUCTURE

TABLE	WHAT IT CONTAINS
COUNTRY	Country Names, IDs
CONSUMPTION	Year-Wise Energy Usage
PRODUCTION	Year-Wise Energy Generation
EMISSION	CO_2 Emissions Data (Total & Per Capita)
GDP	GDP Values
POPULATION	Country-Wise Population Data

ABOUT THE DATA

country	cid
Afghanistan	Af0
Albania	Al1
Algeria	Al2
American Samoa	Am3
Angola	An4
Antarctica	An5
Antigua and Barbuda	An6
Argentina	Ar7
Armenia	Ar8
Aruba	Ar9
Australia	Au10
Austria	Au11
Azerbaijan	Az12
Bahrain	Ba13
Bangladesh	Ba14
Barbados	Ba15
Belarus	Be16
Belgium	Be17
Belize	Be18

Country IDs

country	energy_type	year	emission	per_capita_emission
American Samoa	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Antarctica	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Antigua and Barbuda	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Aruba	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Belize	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Bermuda	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Bhutan	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Botswana	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
British Virgin Islands	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Burkina Faso	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Burundi	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Cabo Verde	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Cambodia	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Cayman Islands	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096
Central African Rep...	Consumed natural gas (MMtonnes CO2)	2020	0	0.007126096

Per Capita Emissions

country	year	population_value
Afghanistan	2024	42647.49
Albania	2024	2791.765
Algeria	2024	46814.31
Angola	2024	37885.85
Antigua and Barbuda	2024	93.772
Argentina	2024	45696.16
Armenia	2024	2973.84
Aruba	2024	108.066
Australia	2024	27236.21
Austria	2024	9170.854
Azerbaijan	2024	10336.58
Bahrain	2024	1631.91
Bangladesh	2024	173562.4
Barbados	2024	282.5008
Belarus	2024	9056.696
Belgium	2024	11766.26
Belize	2024	417.072

Population of country

UNITS:
Emission: Million Metric Tons (eg. MMtonnes CO2), Per capita emission: Metric tons per person
Population Value: In Thousands

INNOMATICS RESEARCH LABS

ABOUT THE DATA

country	energy	year	production
Afghanistan	Nuclear (quad Btu)	2020	0
Albania	Nuclear (quad Btu)	2020	0
Algeria	Nuclear (quad Btu)	2020	0
American Samoa	Nuclear (quad Btu)	2020	0
Angola	Nuclear (quad Btu)	2020	0
Antarctica	Nuclear (quad Btu)	2020	0
Antigua and Barbuda	Nuclear (quad Btu)	2020	0
Aruba	Nuclear (quad Btu)	2020	0
Australia	Nuclear (quad Btu)	2020	0
Austria	Nuclear (quad Btu)	2020	0
Azerbaijan	Nuclear (quad Btu)	2020	0
Bahrain	Nuclear (quad Btu)	2020	0
Bangladesh	Nuclear (quad Btu)	2020	0
Barbados	Nuclear (quad Btu)	2020	0
Belize	Nuclear (quad Btu)	2020	0

Energy Production

country	year	gdp_value
Afghanistan	2020	83.21645
Albania	2020	36.78752
Algeria	2020	531.9749
Angola	2020	215.9016
Antigua and Barbuda	2020	1.772876
Argentina	2020	866.9691
Armenia	2020	33.00228
Aruba	2020	2.584789
Australia	2020	1219.307
Austria	2020	453.2275
Azerbaijan	2020	140.3826
Bahrain	2020	71.83838
Bangladesh	2020	922.9937
Barbados	2020	3.962201
Belarus	2020	176.2621
Belgium	2020	553.6443
Belize	2020	2.958894

GDP Value

country	energy	year	consumption
Afghanistan	Nuclear (quad Btu)	2020	0
Albania	Nuclear (quad Btu)	2020	0
Algeria	Nuclear (quad Btu)	2020	0
American Samoa	Nuclear (quad Btu)	2020	0
Angola	Nuclear (quad Btu)	2020	0
Antarctica	Nuclear (quad Btu)	2020	0
Antigua and Barbuda	Nuclear (quad Btu)	2020	0
Aruba	Nuclear (quad Btu)	2020	0
Australia	Nuclear (quad Btu)	2020	0
Austria	Nuclear (quad Btu)	2020	0
Azerbaijan	Nuclear (quad Btu)	2020	0
Bahrain	Nuclear (quad Btu)	2020	0
Bangladesh	Nuclear (quad Btu)	2020	0
Barbados	Nuclear (quad Btu)	2020	0
Belize	Nuclear (quad Btu)	2020	0

Energy Consumption

UNITS:
Production: Quadrillion British Thermal Units (i.e quad Btu)
Consumption: Quadrillion British Thermal Units (i.e quad Btu)
GDP Value: In Billions

INNOMATICS RESEARCH LABS

ANALYSIS - Energy & Emissions Profiling by Country

1A. How much energy is consumed and produced by each country?

This helps identify energy-rich countries (high production) and energy-dependent ones (high consumption). It is the first step in profiling national energy strategies.

```sql
-- Energy consumed and produced
SELECT c.country,
       SUM(cons.consumption) AS total_consumption,
       SUM(prod.production) AS total_production
FROM country c
LEFT JOIN consumption cons ON c.country = cons.country
LEFT JOIN production prod ON c.country = prod.country AND cons.year = prod.year
GROUP BY c.country
ORDER BY total_consumption DESC
LIMIT 10;
```

country	total_consumption	total_production
China	3996	3276
United States	2646	2712
Russia	840	1542
India	834	504
Japan	456	84
South Korea	324	84
Canada	324	570
Iran	318	414
Brazil	294	324
Germany	288	108

- China and the United States are the top energy consumers and producers.
- Japan and South Korea consume so much more than they produce → they are net energy importers.
- Russia has a significant surplus → major energy exporter.

ANALYSIS - Energy & Emissions Profiling by Country

1B. The relationship between energy consumption/production and CO_2 emissions.

This helps determine if emissions are proportional to energy use — useful for identifying clean vs. dirty energy economies.

```sql
-- Relationship b/w consumption/production and CO₂ emissions
SELECT e.country,
       SUM(cons.consumption) AS total_consumption,
       SUM(prod.production) AS total_production,
       SUM(e.emission) AS total_emission
FROM emission_3 e
LEFT JOIN consumption cons ON e.country = cons.country AND e.year = cons.year
LEFT JOIN production prod ON e.country = prod.country AND e.year = prod.year
GROUP BY e.country
ORDER BY total_emission DESC
LIMIT 10;
```

country	total_consumption	total_production	total_emission
China	15984	13104	3324168
United States	10584	10848	1384308
India	3336	2016	728028
Russia	3360	6168	521316
Japan	1824	336	292932
Iran	1272	1656	214776
Indonesia	768	1680	191268
Germany	1152	432	185724
South Korea	1296	336	182340
Saudi Arabia	1056	2640	180000

- China has the highest total emissions (over 3 billion metric tons CO_2).
- India produces less than it consumes but still has the 3rd highest emissions.
- Japan and Germany also show high emissions but lower consumption compared to China and the U.S.

ANALYSIS - Energy & Emissions Profiling by Country

1C. Comparison of countries in terms of their energy efficiency (e.g., emissions per unit of GDP or per unit of energy consumed).

Emissions per GDP → environmental cost of economic growth

```sql
-- Emissions per unit GDP
SELECT
    e.country,
    SUM(e.emission) / SUM(gdp.gdp_value) AS emission_per_gdp
FROM emission_3 e
JOIN gdp_3 gdp ON e.country = gdp.country AND e.year = gdp.year
GROUP BY e.country
ORDER BY emission_per_gdp DESC
LIMIT 10;
```

country	emission_per_gdp
Netherlands Antilles	2.1190909325935543
North Korea	0.7925853939935248
U.S. Virgin Islands	0.5389944277408574
Trinidad and Tobago	0.5316684283123918
Turkmenistan	0.515191683029713
Libya	0.3346498509172942
Somalia	0.3227252468928819
Bahrain	0.28956808800488626
South Africa	0.28904918548982156
Iran	0.28454308180666066

- Countries like Netherlands Antilles, North Korea, and U.S. Virgin Islands have very high emissions per GDP, meaning their economies are carbon intensive.
- South Africa and Iran are also high, despite being major producers.

INNOMATICS RESEARCH LABS

ANALYSIS - Energy & Emissions Profiling by Country

1C. Comparison of countries in terms of their energy efficiency (e.g., emissions per unit of GDP or per unit of energy consumed).

Emissions per unit of energy → carbon intensity of energy

```sql
-- Emissions per unit of energy consumed
SELECT
    e.country,
    SUM(e.emission) AS total_emission,
    SUM(c.consumption) AS total_consumption,
    ROUND(SUM(e.emission) / SUM(c.consumption), 3) AS emission_per_energy
FROM emission_3 e
JOIN consumption c
  ON e.country = c.country AND e.year = c.year
GROUP BY e.country
HAVING total_emission IS NOT NULL AND total_consumption IS NOT NULL
ORDER BY emission_per_energy DESC;
```

country	total_emission	total_consumption	emission_per_energy
Peru	2544	4	636.000
Azerbaijan	1818	4	454.500
Hong Kong	2934	8	366.750
Czechia	4260	12	355.000
South Africa	21432	80	267.900
Chile	4026	16	251.625
Colombia	3990	16	249.375
Indonesia	31878	128	249.047
Oman	3648	16	228.000
Bangladesh	5364	24	223.500
India	121338	556	218.234
Vietnam	13776	64	215.250
Mexico	20496	96	213.500

- Peru, Azerbaijan, Hong Kong, and Czechia have the highest CO_2 emissions per unit of energy consumed.
- This suggests they maybe heavily reliant on fossil fuels, or that energy is used in inefficient ways.

INNOMATICS RESEARCH LABS

ANALYSIS - Sustainability & Development Indicators

2A. GDP and population data with emissions and energy
Combining key indicators shows the balance between development and environmental impac

```sql
-- GDP and population data with emissions and energy
WITH gdp_data  AS (SELECT country, year, gdp_value FROM gdp_3),
     pop_data  AS (SELECT country, year, population_value FROM population),
     em_data   AS (SELECT country, year, SUM(emission) AS total_emission FROM emission_3 GROUP BY country, year),
     cons_data AS (SELECT country, year, SUM(consumption) AS total_consumption FROM consumption GROUP BY country, year),
     prod_data AS (SELECT country, year, SUM(production) AS total_production FROM production GROUP BY country, year)
SELECT
    gdp.country,
    gdp.year,
    gdp.gdp_value,
    pop.population_value,
    em.total_emission,
    cons.total_consumption,
    prod.total_production
FROM gdp_data gdp
JOIN pop_data pop ON gdp.country = pop.country AND gdp.year = pop.year
JOIN em_data em ON gdp.country = em.country AND gdp.year = em.year
LEFT JOIN cons_data cons ON gdp.country = cons.country AND gdp.year = cons.year
LEFT JOIN prod_data prod ON gdp.country = prod.country AND gdp.year = prod.year;
```

country	year	gdp_value	population_value	total_emission	total_consumption	total_production
China	2023	27313.92	1422585	24392	177	141
China	2022	25910.6	1425180	22968	168	139
China	2021	25123.27	1426437	22884	165	137
China	2020	23136.77	1426106	22094	156	129
United States	2022	21443.41	333765.9	9881	112	115
United States	2021	20917.86	332047.6	9812	112	111
United States	2023	22062.56	336445	9590	111	119
United States	2020	19723.57	331469.7	9170	106	107
India	2023	10937.52	1438070	5642	38	24
India	2022	10047.18	1425423	5158	36	22
India	2021	9387.808	1414204	4766	33	19
India	2020	8572.151	1402618	4657	32	19
Russia	2022	3954.023	145579.9	3771	36	64

- China’s GDP and population are growing yearly, but emissions remain very high, even slightly increasing.
- United States has consistent emissions despite high GDP.
- India shows lower emissions than China and the U.S., but a steep increase from 2020 to 2023.

ANALYSIS - Sustainability & Development Indicators

2B. Per capita emissions

This shows the average emissions caused by a citizen. Important for fair comparisons between large and small countries.

```sql
-- Per capita emissions
SELECT
    e.country,
    SUM(e.emission) AS total_emission,
    SUM(p.population_value) AS total_population,
    ROUND(SUM(e.emission) / SUM(p.population_value), 3) AS per_capita_emission
FROM emission_3 e
JOIN population p ON e.country = p.country AND e.year = p.year
GROUP BY e.country
HAVING total_emission IS NOT NULL AND total_population IS NOT NULL
ORDER BY per_capita_emission ASC;
```

country	total_emission	total_population	per_capita_emission
Afghanistan	72	522377.94999999995	0
Angola	141	561474.0000000002	0
Bangladesh	894	2699235.1999999993	0
Benin	48	217416.47999999998	0
Bhutan	5	12450.988000000001	0

country	total_emission	total_population	per_capita_emission
Netherlands Antilles	60	3051.464000000001	0.02
Qatar	933	45959.576000000015	0.02
Singapore	1800	90985.94399999999	0.02
Bahrain	358	24510.800000000003	0.015
United Arab Emirates	2099	160486.96799999996	0.013

- Countries like Afghanistan, Bangladesh, and Bhutan have the lowest per capita emissions, all below 1 ton CO_2/person, despite large populations and growing economies.
- Netherland Antilles and Qatar have very high per capita emissions.
- Vietnam, Philippines, and Turkey fall in the middle range.

INNOMATICS RESEARCH LABS

ANALYSIS - Sustainability & Development Indicators

2C. Emissions per GDP unit
Measures how much emission is generated per dollar of economic output — a good sustainability efficiency measure.

```sql
-- Emissions per GDP unit
SELECT
    e.country,
    SUM(e.emission) AS total_emission,
    SUM(g.gdp_value) AS gdp_value,
    ROUND(SUM(e.emission) / SUM(g.gdp_value), 3) AS emissions_per_gdp
FROM emission_3 e
JOIN gdp_3 g
  ON e.country = g.country AND e.year = g.year
GROUP BY e.country
HAVING total_emission IS NOT NULL AND gdp_value IS NOT NULL
ORDER BY emissions_per_gdp DESC;
```

country	total_emission	gdp_value	emissions_per_gdp
Netherlands Antilles	60	28.314027999999997	2.119
North Korea	470	592.99604	0.793
U.S. Virgin Islands	16	29.684908	0.539
Trinidad and Tobago	278	522.88228	0.532
Turkmenistan	812	1576.1123999999998	0.515
Libya	352	1051.8456800000001	0.335
Somalia	8	24.788888000000007	0.323
Bahrain	358	1236.3240799999999	0.29
South Africa	3572	12357.758399999999	0.289
Iran	5966	20966.948	0.285

- Netherlands Antilles, North Korea, and U.S. Virgin Islands have extremely high emissions per unit of GDP, indicating that their economies are heavily carbon-intensive.
- India and Iran are in the mid-range but still significantly lower.
- Switzerland, Sweden, Nepal, UK and Germany are some of the countries with the lowest emissions per GDP. These countries are economically efficient and environmentally cleaner, emitting less CO_2 per dollar of economic output.

ANALYSIS - Trends Over Time

3A. Trends in energy consumption, production, and emissions

Helpful for seeing how the world is progressing with energy needs and carbon footprints.

```sql
-- Trends in energy consumption, production, and emissions
SELECT
    cons.year,
    SUM(cons.consumption) AS global_consumption,
    SUM(prod.production) AS global_production,
    SUM(em.emission) AS global_emissions
FROM consumption cons
JOIN production prod ON cons.country = prod.country AND cons.year = prod.year
JOIN emission_3 em ON cons.country = em.country AND cons.year = em.year
GROUP BY cons.year
ORDER BY cons.year;
```

year	global_consumption	global_production	global_emissions
2020	13464	14112	2442672
2021	14256	14664	2554916
2022	14640	15048	2607822
2023	14688	14928	2669620

- Global energy consumption and production are rising year by year.
- Global GDP and energy use are both growing → development continues, but the efficiency is improving.

ANALYSIS - Trends Over Time

3B. Economic development trends (GDP growth vs. energy demand)

Measures whether energy use is growing with the economy — or if economies are becoming more energy-efficient.

```
-- Economic development trends (GDP growth vs. energy demand)
SELECT
    g.year,
    SUM(g.gdp_value) AS global_gdp,
    SUM(cons.consumption) AS global_energy_consumption
FROM gdp_3 g
JOIN consumption cons ON g.country = cons.country AND g.year = cons.year
GROUP BY g.year
ORDER BY g.year;
```

	year	global_gdp	global_energy_consumption
▶	2020	749171.7484362009	561
	2021	798238.0392941969	594
	2022	826257.7812101999	610
	2023	854001.8751344001	612

- As GDP increases, energy use also rises.
- Energy intensity per GDP unit is gradually improving → countries are producing more economic value for the same or slightly more energy.

INNOMATICS RESEARCH LABS

ANALYSIS - Dependency on Imports or Exports

4A. Whether a country is an energy importer or exporter by comparing its production vs. consumption. Comparing energy production vs consumption to determine dependency.

```sql
-- Whether a country is an energy importer or exporter by comparing its production vs. consumption.
SELECT
    c.country,
    SUM(prod.production) AS total_production,
    SUM(cons.consumption) AS total_consumption,
    CASE
        WHEN SUM(prod.production) > SUM(cons.consumption) THEN 'Exporter'
        WHEN SUM(prod.production) < SUM(cons.consumption) THEN 'Importer'
        ELSE 'Balanced'
    END AS energy_status
FROM country c
JOIN production prod ON c.country = prod.country
JOIN consumption cons ON c.country = cons.country AND prod.year = cons.year
GROUP BY c.country
ORDER BY energy_status;
```

country	total_production	total_consumption	energy_status
China	3276	3996	Importer
United States	2712	2646	Exporter
Russia	1542	840	Exporter
India	504	834	Importer
Japan	84	456	Importer
Canada	570	324	Exporter
South Korea	84	324	Importer
Iran	414	318	Exporter
Brazil	324	294	Exporter
Germany	108	288	Importer
France	192	282	Importer
Saudi Arabia	660	264	Exporter
Indonesia	420	192	Exporter

- Exporters: Russia, Iran, Canada, Saudi Arabia (production > consumption)
- Importers: China, India, Japan, Germany (consumption > production)
- Balanced: U.S. is closer to balance

INNOMATICS RESEARCH LABS

Your Experience/Challenges Working On The MySQL Project

Challenges:

- Data Inconsistency Across Tables. Required a lot of schema cleaning and standardization before analysis.
- Unit Mismatches and Threshold Calibration were tricky. GDP might be in billions, emissions in tons, population in millions. Queries needed careful tuning of thresholds to get valid results. Required multiple trial-and-error runs to find meaningful filtering conditions.
- Had difficulty understanding complex SQL joins and aggregations. Many queries involved 3–4 joins and different aggregations (`SUM`, `AVG`, `CASE`, `HAVING`). Used CTEs to break down queries step by step.
- Metrics like "emissions per GDP" or "per capita emissions" had to be calculated manually.
- Needed good understanding of both domain knowledge and SQL functions.
- Needed to understand what “PPP-adjusted GDP”, “energy efficiency”, or “carbon intensity” actually mean. Otherwise, hard to make meaningful interpretations from query results. Invested time in understanding energy economics and environmental indicators.

Your Experience/Challenges Working On The MySQL Project

My Experience:

- I learned how energy, emissions, GDP, and population data are connected in real life.
- I improved my SQL skills by writing queries using joins, filters, and calculations.
- Learned about the domain.
- This project helped me understand how to turn raw data into useful insights for sustainability.
- Overall, it was a good mix of technical learning and real-world understanding.

INNOMATICS
RESEARCH LABS

THANK YOU

INNOMATICS
RESEARCH LABS