



INNOVATION. AUTOMATION. ANALYTICS

Energy Consumption Analysis

SQL Project



About me

Background:

Gurubalan P

- I have completed my B.Sc.Animation, and I have 4 years of professional experience as Texturing artist at xentrix studios.



[linkedin.com/in/gurubalan-p](https://www.linkedin.com/in/gurubalan-p)



github.com/Guru1901

Why you want to learn Data Analyst :

- I am an aspiring Data Analyst and I enjoy uncovering insights from data to solve real-world challenges. I'm passionate about how data can guide decision-making, improve efficiency, and create value. Developing these skills will also help me move closer to my long-term goal of working in AI and advanced analytics.





Objective of the Project

Structure & Build

- Build a structured database to store global energy consumption, production, GDP, and emission data.

Query & Analyze

- Use SQL queries to extract, combine, and analyze country-wise and year-wise energy metrics.

Compare & Evaluate

- Identify trends in emissions, GDP growth, and energy usage across different countries and years.

Measure Key Ratios

- Calculate per-capita energy use and energy-to-GDP efficiency for deeper insights.

Generate Insights

- Highlight major emitters, efficient economies, and long-term global energy patterns.



Entity Relationship Diagram

- An ER Diagram is a diagram that shows how data is structured and how different data tables are related to each other.

Country (Main / Reference Table):

- This is the central table in the ER diagram.

Gdp_3 (GDP Table):

- One country has many GDP records (one-to-many).

Population (Population Table):

- One country - many population records.

Emission_3 (Emissions Table):

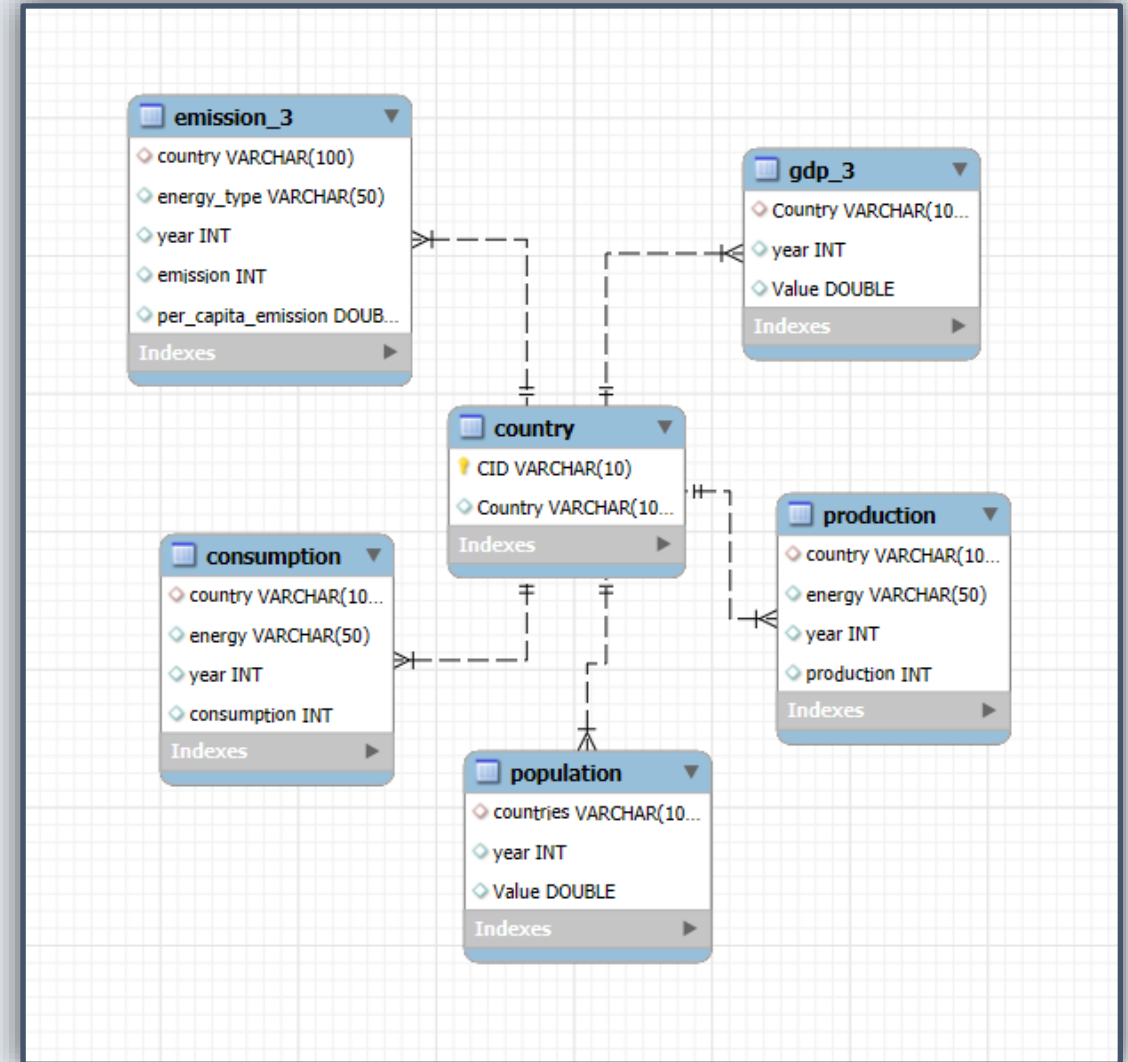
- A country can have: multiple energy types multiple years -- So this is a one-to-many relationship

production (Energy Production Table):

- One country - produces many types of energy over many years

consumption (Energy Consumption Table):

- A country consumes many types of energy - one-to-many





Key analysis questions



General & Comparative Analysis

1. What is the total emission per country for the most recent year available?
2. Compare energy production and consumption by country and year?
3. Which energy types contribute most to emissions across all countries?

Trend Analysis Over Time

1. How have global emissions changed year over year?
2. What is the trend in GDP for each country over the given years?
3. Has energy consumption increased or decreased over the years for major economies?

Ratio & Per Capita Analysis

1. What is the energy consumption per capita for each country over the last decade?
2. Which countries have the highest energy consumption relative to GDP?

Global Comparisons

1. What are the top 10 countries by population and how do their emissions compare?
2. Which countries have improved (reduced) their per capita emissions the most over the last decade?



General & Comparative Analysis

1. What is the total emission per country for the most recent year available?

```
select country,year,  
       sum(emission) as Total_Emission from EMISSION_3  
  
where year = (select max(year) from EMISSION_3)  
  
group by country,year  
  
order by Total_Emission desc;
```



country	year	Total_Emission
China	2023	24392
United States	2023	9590
India	2023	5642
Russia	2023	3688
Japan	2023	1920
Indonesia	2023	1659
Iran	2023	1646
Saudi Arabia	2023	1313
South Korea	2023	1288
Germany	2023	1200
Canada	2023	1171
South Africa	2023	894
Mexico	2023	882
Brazil	2023	876
Turkiye	2023	798
Australia	2023	789
United Kingdom	2023	682
Thailand	2023	674
Italy	2023	615
Vietnam	2023	611
France	2023	607
Taiwan	2023	578
United Arab ...	2023	544
Kazakhstan	2023	540
Poland	2023	528
Malaysia	2023	520
Spain	2023	509
Singapore	2023	478
Egypt	2023	474
Pakistan	2023	426
Argentina	2023	396
Iraq	2023	382
Netherlands	2023	376

Insights:

- A few major countries contribute most of the global emissions, while many others emit significantly less.
- China, the United States, and India are the top three emitters, clearly dominating global emission levels.
- Many countries - including Netherlands, Argentina, and Iraq - have much lower emissions, reflecting smaller populations or lower industrial activity.

2. Compare energy production and consumption by country and year?

```
select
  P.country,P.year,P.TOTAL_PRODUCTION,C.TOTAL_CONSUMPTION,(P.TOTAL_PRODUCTION - C.TOTAL_CONSUMPTION) as difference
  From (select country,year, sum(Production) as TOTAL_PRODUCTION from PRODUCTION group by country,year) as P join
  (select country,year, sum(Consumption) as TOTAL_CONSUMPTION from CONSUMPTION group by country,year) as C
  on P.country = C. Country AND P.Year = C.Year order by P.Country,P.year;
```



country	year	TOTAL_PRODUCTION	TOTAL_CONSUMPTION	difference
Argentina	2021	2	3	-1
Argentina	2022	3	4	-1
Argentina	2020	2	3	-1
Argentina	2023	4	4	0
Armenia	2020	0	0	0
Armenia	2023	0	0	0
Armenia	2022	0	0	0
Armenia	2021	0	0	0
Aruba	2020	0	0	0
Aruba	2023	0	0	0
Aruba	2022	0	0	0
Aruba	2021	0	0	0
Australia	2023	17	6	11
Australia	2020	18	6	12
Australia	2022	18	6	12
Australia	2021	17	6	11
Austria	2020	0	1	-1
Austria	2022	0	1	-1
Austria	2023	0	1	-1
Austria	2021	0	1	-1
Azerbaijan	2020	2	0	2
Azerbaijan	2022	2	1	1
Azerbaijan	2023	2	0	2
Azerbaijan	2021	3	0	3
Bahrain	2020	1	1	0
Bahrain	2022	1	1	0
Bahrain	2023	1	1	0
Bahrain	2021	1	1	0
Bangladesh	2020	1	1	0
Bangladesh	2022	1	2	-1
Bangladesh	2023	1	2	-1
Bangladesh	2021	1	1	0
Barbados	2022	0	0	0
Barbados	2020	0	0	0

Insights:

- Australia, Canada, and Brazil consistently produce more energy than they consume, indicating strong energy surplus.
- Austria and Azerbaijan consume more energy than they produce, showing dependency on energy imports.
- Many small or developing countries have very low production and consumption levels, reflecting limited industrial activity.

3. Which energy types contribute most to emissions across all countries?

```
select energy_type,  
       sum(emission) as Total_Emissions from EMISSION_3  
  
group by energy_type  
  
order by Total_Emissions desc limit 1 ;
```



Result Grid			Filter Rows:	Export
	energy_type	Total_Emissions		
▶	CO2 emissions (MMtonnes CO2)	142723		

Insights:

- CO₂-based energy sources are the highest contributors to global emissions.
- This indicates that countries still rely heavily on carbon-intensive fuels such as coal, oil, and natural gas.



Trend Analysis Over Time

1.How have global emissions changed year over year?

```
Select x.*,  
       Lag(Total_EMS_Year) over (order by year asc) as Pre,  
  
       round((((Total_EMS_Year - Lag(Total_EMS_Year) over (order by year asc))/  
  
       Lag(Total_EMS_Year) over (order by year asc))*100,2) as YOY_EMS_Change  
  
       from (select year, sum(emission) as Total_EMS_Year from EMISSION_3 group by year order by year ) x ;
```

Result Grid					Filter Rows:	Export
	year	Total_EMS_Year	Pre	YOY_EMS_Change		
▶	2020	67852	NULL	NULL		
	2021	70976	67852	4.60		
	2022	72445	70976	2.07		
	2023	74161	72445	2.37		

Insights:

- Global emissions have increased each year from 2020 to 2023, showing a steady upward trend.
- The largest rise occurred in 2021, with a 4.6% jump, likely driven by post-pandemic economic recovery.
- Emissions in 2022 and 2023 continued to grow at a slower but consistent rate (around 2%), indicating ongoing dependence on high-emission energy sources.

2.What is the trend in GDP for each country over the given years?

```
with yearly as(  
  select country,year, sum(value) as Total_GDP from GDP_3 group by country,year  
)  
  
  SELECT country,year,total_gdp,LAG(total_gdp) OVER (PARTITION BY country ORDER BY year) AS prev_year_gdp,  
  
  round((((Total_gdp - LAG(total_gdp) OVER (PARTITION BY country ORDER BY year))/Nullif(LAG(total_gdp)  
  
  OVER (PARTITION BY country ORDER BY year),0))*100,2) as YOY_Trend  
  
  from yearly;
```



country	year	total_gdp	prev_year_gdp	YOY_Trend
Afghanistan	2020	83.21645	NULL	NULL
Afghanistan	2021	65.95827	83.21645	-20.74
Afghanistan	2022	61.84236	65.95827	-6.24
Afghanistan	2023	58.90645	61.84236	-4.75
Afghanistan	2024	57.83112	58.90645	-1.83
Albania	2020	36.78752	NULL	NULL
Albania	2021	40.04762	36.78752	8.86
Albania	2022	41.99279	40.04762	4.86
Albania	2023	43.63172	41.99279	3.9
Albania	2024	45.2091	43.63172	3.62
Algeria	2020	531.9749	NULL	NULL
Algeria	2021	552.5308	531.9749	3.86
Algeria	2022	572.6695	552.5308	3.64
Algeria	2023	596.0418	572.6695	4.08
Algeria	2024	615.7547	596.0418	3.31
Angola	2020	215.9016	NULL	NULL
Angola	2021	218.2898	215.9016	1.11
Angola	2022	225.5227	218.2898	3.31
Angola	2023	227.5672	225.5227	0.91
Angola	2024	237.8399	227.5672	4.51
Antigua an...	2020	1.772876	NULL	NULL
Antigua an...	2021	1.887162	1.772876	6.45
Antigua an...	2022	2.057991	1.887162	9.05
Antigua an...	2023	2.174793	2.057991	5.68
Antigua an...	2024	2.260353	2.174793	3.93
Argentina	2020	866.9691	NULL	NULL
Argentina	2021	957.4966	866.9691	10.44
Argentina	2022	1007.055	957.4966	5.27

Insights:

- Most countries show steady GDP growth year over year.
- A few countries experience negative or unstable growth (e.g., Afghanistan), reflecting economic challenges.
- Countries like Albania, Algeria, Angola, and Argentina show steady GDP increases, highlighting consistent economic improvement.

3.Has energy consumption increased or decreased over the years for major economies?

```
WITH yearly AS
  (SELECT country,year,SUM(consumption) AS Total_Consumption FROM CONSUMPTION GROUP BY country, year)

SELECT country,year,Total_Consumption,

  LAG(Total_Consumption) OVER (PARTITION BY country ORDER BY year) AS prev_year_consumption,

  (Total_Consumption - LAG(Total_Consumption) OVER (PARTITION BY country ORDER BY year)) AS change_from_prev_year,

CASE
  WHEN LAG(Total_Consumption) OVER (PARTITION BY country ORDER BY year) IS NULL THEN 'first_year'

  WHEN Total_Consumption > LAG(Total_Consumption) OVER (PARTITION BY country ORDER BY year) THEN 'increased'

  WHEN Total_Consumption < LAG(Total_Consumption) OVER (PARTITION BY country ORDER BY year) THEN 'decreased'

  ELSE 'no_change'

END AS trend

FROM yearly

ORDER BY country, year;
```

country	year	Total_Consumption	prev_year_consumption	change_from_prev_year	trend
Afghanistan	2020	0	NULL	NULL	first_year
Afghanistan	2021	0	0	0	no_change
Afghanistan	2022	0	0	0	no_change
Afghanistan	2023	0	0	0	no_change
Albania	2020	0	NULL	NULL	first_year
Albania	2021	0	0	0	no_change
Albania	2022	0	0	0	no_change
Albania	2023	0	0	0	no_change
Algeria	2020	3	NULL	NULL	first_year
Algeria	2021	3	3	0	no_change
Algeria	2022	3	3	0	no_change
Algeria	2023	3	3	0	no_change
American S...	2020	0	NULL	NULL	first_year
American S...	2021	0	0	0	no_change
American S...	2022	0	0	0	no_change
American S...	2023	0	0	0	no_change
Angola	2020	0	NULL	NULL	first_year
Angola	2021	0	0	0	no_change
Angola	2022	0	0	0	no_change
Angola	2023	0	0	0	no_change
Antarctica	2020	0	NULL	NULL	first_year
Antarctica	2021	0	0	0	no_change
Antarctica	2022	0	0	0	no_change
Antarctica	2023	0	0	0	no_change
Antigua an...	2020	0	NULL	NULL	first_year
Antigua an...	2021	0	0	0	no_change
Antigua an...	2022	0	0	0	no_change
Antigua an...	2023	0	0	0	no_change
Argentina	2020	3	NULL	NULL	first_year
Argentina	2021	3	3	0	no_change
Argentina	2022	4	3	1	increased
Argentina	2023	4	4	0	no_change
Armenia	2020	0	NULL	NULL	first_year
Armenia	2021	0	0	0	no_change
Armenia	2022	0	0	0	no_change
Armenia	2023	0	0	0	no_change
Aruba	2020	0	NULL	NULL	first_year
Aruba	2021	0	0	0	no_change

Insights:

- Most countries show no major change in energy consumption across the years, indicating stable usage patterns.
- Only a few countries show clear increases or decreases, meaning consumption shifts are limited.

Ratio & Per Capita Analysis

1. What is the energy consumption per capita for each country over the last decade?

```
WITH last10 AS
    (SELECT DISTINCT year FROM CONSUMPTION ORDER BY CAST(year AS UNSIGNED) DESC LIMIT 10)

SELECT c.country, c.year,
    SUM(c.consumption) AS total_consumption, round(SUM(p.value), 2) AS total_population,
    ROUND(SUM(c.consumption) / NULLIF(SUM(p.value), 0), 6) AS consumption_per_capita
FROM CONSUMPTION c JOIN POPULATION p ON c.country = p.countries
AND c.year = p.year

WHERE c.year IN (SELECT year FROM last10)

GROUP BY c.country, c.year

ORDER BY c.country, CAST(c.year AS UNSIGNED);
```




country	year	total_consumption	total_population	consumption_per_capita
Afghanistan	2020	0	234413.88	0
Afghanistan	2021	0	200002.05	0
Afghanistan	2022	0	202894.2	0
Afghanistan	2023	0	207273.8	0
Albania	2020	0	17231.72	0
Albania	2021	0	17097.81	0
Albania	2022	0	16965.65	0
Albania	2023	0	16869.93	0
Algeria	2020	3	264252.54	0.000011
Algeria	2021	3	268566.6	0.000011
Algeria	2022	3	272864.28	0.000011
Algeria	2023	3	276985.32	0.000011
Angola	2020	0	200706.78	0
Angola	2021	0	207194.58	0
Angola	2022	0	213810.18	0
Angola	2023	0	220499.46	0
Antigua an...	2020	0	551.08	0
Antigua an...	2021	0	554.09	0
Antigua an...	2022	0	557.04	0
Antigua an...	2023	0	559.9	0
Argentina	2020	3	271151.76	0.000011
Argentina	2021	3	271873.68	0.000011
Argentina	2022	4	272447.46	0.000015
Argentina	2023	4	273230.4	0.000015
Armenia	2020	0	17345.35	0
Armenia	2021	0	17222.09	0
Armenia	2022	0	17285.24	0
Armenia	2023	0	17660.36	0
Aruba	2020	0	644.4	0
Aruba	2021	0	645.35	0
Aruba	2022	0	646.69	0
Aruba	2023	0	647.63	0
Australia	2020	6	153841.38	0.000039
Australia	2021	6	154318.5	0.000039
Australia	2022	6	156558	0.000038
Australia	2023	6	160297.98	0.000037
Austria	2020	1	53524.88	0.000019
Austria	2021	1	53769.48	0.000019
Austria	2022	1	54345.49	0.000018

Insights:

- Most countries have very low energy consumption per person, mainly due to large populations.
- Smaller-population countries show higher per-capita consumption when their overall usage is high.
- There is a wide global gap, with developed countries consuming more energy per person.

2. Which countries have the highest energy consumption relative to GDP?

```
SELECT c.country,  
  
SUM(c.consumption) AS total_consumption, round(SUM(g.value), 2) AS total_gdp,  
  
ROUND(SUM(c.consumption) / NULLIF(SUM(g.value), 0), 6) AS consumption_to_gdp_ratio  
  
FROM CONSUMPTION c JOIN GDP_3 g ON c.country = g.country AND c.year = g.year  
  
GROUP BY c.country ORDER BY consumption_to_gdp_ratio DESC LIMIT 5;
```



Result Grid					Filter Rows:	Export:	Wrap Cell Content:
	country		total_consumption	total_gdp	consumption_to_gdp_ratio		
▶	Trinidad and Tobago	4	784.32		0.0051		
	North Korea	4	889.49		0.004497		
	Turkmenistan	8	2364.17		0.003384		
	Bahrain	4	1854.49		0.002157		
	Kuwait	8	4686.51		0.001707		

Insights:

- Trinidad and Tobago has the highest energy use relative to GDP, indicating heavy energy-intensive industries.
- North Korea and Turkmenistan also show high ratios, suggesting lower economic output compared to their energy consumption.
- Bahrain and Kuwait rank high due to energy-driven economies, especially oil-related sectors.

Global Comparisons

1. What are the top 10 countries by population and how do their emissions compare?

```
SELECT P.countries,P.Total_Population,E.Total_EMS  
  
FROM (SELECT countries, SUM(value) AS Total_Population FROM POPULATION GROUP BY countries) AS P JOIN  
  
(SELECT country, SUM(emission) AS Total_EMS FROM EMISSION_3 GROUP BY country) AS E  
  
ON P.countries = E.country ORDER BY P.Total_Population DESC, E.Total_EMS DESC LIMIT 10;
```

	countries	Total_Population	Total_EMS
▶	India	7131251	20223
	China	7119629	92338
	United States	1673338.7	38453
	Indonesia	1395081.6	5313
	Pakistan	1216954	1640
	Nigeria	1116238.8	874
	Brazil	1051656.8	3405
	Bangladesh	848371.2000000001	894
	Russia	728048.29999999999	14481
	Mexico	643661	3416

Insights:

- India and China have huge populations and high emissions, showing strong energy demand.
- The United States shows very high emissions relative to its population, indicating high per-capita usage.
- Nigeria, Bangladesh, and Pakistan have large populations but low emissions, reflecting lower industrial or per-person energy use.

2. Which countries have improved (reduced) their per capita emissions the most over the last decade?

```
select x.*,
       lag(Total_Per_Capita) over (partition by country order by year) as PRE_Capita,
       round((((Total_Per_Capita - lag(Total_Per_Capita) over (partition by country order by year)) / lag(Total_Per_Capita) over (partition by country order by year)) * 100, 2) as `YOY_Changes%`,
       case
         when round((((Total_Per_Capita - lag(Total_Per_Capita) over (partition by country order by year)) / lag(Total_Per_Capita) over (partition by country order by year)) * 100, 2) > 0 then "improved"
         when round((((Total_Per_Capita - lag(Total_Per_Capita) over (partition by country order by year)) / lag(Total_Per_Capita) over (partition by country order by year)) * 100, 2) = 0 then "NO Changes"
         when round((((Total_Per_Capita - lag(Total_Per_Capita) over (partition by country order by year)) / lag(Total_Per_Capita) over (partition by country order by year)) * 100, 2) is null then "NO Value"
         else "reduced"
       end as Improvements_Status
from (select country, year, sum(per_capita_emission) as Total_Per_Capita from EMISSION_3 group by year, country order by year) x;
```

country	year	Total_Per_Capita	PRE_Capita	YOY_Changes%	Improvements_Status
Afghanistan	2020	0.028504384	NULL	NULL	NO Value
Afghanistan	2021	0.0213782880000000002	0.028504384	-25	reduced
Afghanistan	2022	0.0213782880000000002	0.0213782880000000002	0	NO Changes
Afghanistan	2023	0.0213782880000000002	0.0213782880000000002	0	NO Changes
Albania	2020	0.028504384	NULL	NULL	NO Value
Albania	2021	0.028504384	0.028504384	0	NO Changes
Albania	2022	0.028504384	0.028504384	0	NO Changes
Albania	2023	0.028504384	0.028504384	0	NO Changes
Algeria	2020	0.028504384	NULL	NULL	NO Value
Algeria	2021	0.028504384	0.028504384	0	NO Changes
Algeria	2022	0.028504384	0.028504384	0	NO Changes
Algeria	2023	0.028504384	0.028504384	0	NO Changes
American S...	2020	0.028504384	NULL	NULL	NO Value
American S...	2021	0.028504384	0.028504384	0	NO Changes
American S...	2022	0.028504384	0.028504384	0	NO Changes
American S...	2023	0.028504384	0.028504384	0	NO Changes
Angola	2020	0.028504384	NULL	NULL	NO Value
Angola	2021	0.028504384	0.028504384	0	NO Changes
Angola	2022	0.028504384	0.028504384	0	NO Changes
Angola	2023	0.028504384	0.028504384	0	NO Changes
Antarctica	2020	0.028504384	NULL	NULL	NO Value
Antarctica	2021	0.028504384	0.028504384	0	NO Changes
Antarctica	2022	0.028504384	0.028504384	0	NO Changes
Antarctica	2023	0.028504384	0.028504384	0	NO Changes
Antigua an...	2020	0.028504384	NULL	NULL	NO Value
Antigua an...	2021	0.028504384	0.028504384	0	NO Changes
Antigua an...	2022	0.028504384	0.028504384	0	NO Changes
Antigua an...	2023	0.028504384	0.028504384	0	NO Changes
Argentina	2020	0.028504384	NULL	NULL	NO Value
Argentina	2021	0.028504384	0.028504384	0	NO Changes
Argentina	2022	0.028504384	0.028504384	0	NO Changes
Argentina	2023	0.028504384	0.028504384	0	NO Changes
Armenia	2020	0.028504384	NULL	NULL	NO Value
Armenia	2021	0.028504384	0.028504384	0	NO Changes
Armenia	2022	0.028504384	0.028504384	0	NO Changes
Armenia	2023	0.028504384	0.028504384	0	NO Changes
Aruba	2020	0.028504384	NULL	NULL	NO Value
Aruba	2021	0.028504384	0.028504384	0	NO Changes
Aruba	2022	0.028504384	0.028504384	0	NO Changes

Insights:

- Most countries show “No Changes” in per-capita emissions, indicating stable emission levels over the years.
- A few countries, like Afghanistan, show a “reduced” trend in certain years, meaning their per-person emissions decreased.

Final Business Insights



1. Emission Concentration

- Emissions are highly concentrated in a few major economies like China, the U.S., and India.

2. Energy Imbalance

- Energy production and consumption are imbalanced, with some countries being surplus producers while others depend on imports.

3. Rising Global Emissions

- Global emissions continue to rise, showing limited progress toward clean energy transition.

4. Per-Capita Variations

- Per-capita energy use varies widely, highlighting inequality in energy access and efficiency.

5. GDP-Emission Link

- Economic growth strongly correlates with emissions, especially in industrialized nations.



Recommendations

1. Increase Renewable Adoption

- Accelerate renewable energy adoption to reduce dependence on fossil fuels.

2. Improve Energy Efficiency

- Improve energy efficiency in high-consumption and high-emission countries.

3. Support Developing Nations

- Support developing nations in scaling clean energy infrastructure.

4. Strengthen Emission Policies

- Implement stronger emission-reduction policies such as carbon pricing and efficiency standards.

5. Enhance Data Tracking

- Enhance data accuracy and tracking for better decision-making and monitoring.



Conclusion

- This project gives a clear, data-driven understanding of global energy patterns by analyzing consumption, production, GDP, population, and emissions across countries. The findings show that emissions are heavily concentrated in a few major economies, while many nations face energy imbalances and varying efficiency levels. Global emissions continue to rise, and per-capita energy use differs widely, reflecting unequal access to energy and development. Overall, the analysis highlights the need for cleaner energy sources, improved efficiency, and stronger global cooperation to support sustainable growth in the future.



Experience

Working With Multiple Datasets

- Gained hands-on experience integrating GDP, emissions, population, production, and consumption data for deeper analysis.

Advanced SQL Techniques

- Applied joins, window functions, CTEs, aggregations, and calculations to extract insights from large datasets.

Trend and Comparative Analysis

- Developed skills in identifying patterns, ratios, and year-over-year changes across countries.

Missing & Incomplete Data

- Several countries had gaps in historical data, affecting trend accuracy.

Complex Query Building

- Combining multiple large tables demanded careful query planning to avoid errors and maintain performance.

Challenges



Q&A

QUESTIONS AND ANSWERS





**THANK
YOU**

SAVE ENERGY

