

INSTITUTO PRESBITERIANO MACKENZIE

TECNOLOGIA EM CIÊNCIA DE DADOS

Ciência de Dados

**Um Projeto Colaborativo de Análise Exploratória e
Visualização de Dados**

Projeto Aplicado I

Professor

Lucas Cerqueira Figueiredo

Nome: Felipe Israel dos Santos

R.A.: 10729570

Nome: Gabriel Henrique Titanegro Zanelatto

R.A.: 10730604

Nome: Juliana de Oliveira Sato

R.A.: 10727587

Nome: Yuri dos Santos Radziwill

R.A.: 10730741

Link para o repositório:

<<https://github.com/YuriRadzi/ProjetoAplicado1>>

Fonte da base: Consumo Mundial de Energia

<https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption/data>

Sumário

- 1. Cronograma**
- 2. Pipeline do Projeto**
- 3. Introdução**
 - 3.1. Objetivo Geral
 - 3.2. Objetivos Específicos
- 4. Metadados**
 - 4.1. Atributos Principais
 - 4.2. Desafios Identificados
- 5. Lista de Figuras**
- 6. Lista de Tabelas**
- 7. Análise exploratória**
- 8. Gráficos e Análises**
 - 8.1. Análise Sobre os Dados
 - 8.2. Gráficos Elaborados
- 9. Storytelling**
 - 9.1. Abertura
 - 9.2. Introdução ao texto
 - 9.3. Contextualização
 - 9.4. Conclusão e encerramento
- 10. Referências Bibliográficas**

CRONOGRAMA

Etapa 1 – Planejamento e Estruturação

10/02/2025 a 10/03/2025:

- Definição de grupos e responsabilidades (10/02): Divisão de tarefa entre os membros
 - **Todo o grupo**
- Início do Projeto: (10/02 a 19/02): Identificando os principais objetivos e metas.
 - Responsável: **Juliana e Yuri**
- (24/02): Encontro síncrono.
- Definição do tema, objetivo e cronograma. (27/02): Estabelecer metas para cada fase do projeto.
 - Responsável: **Gabriel e Felipe**
- (09/03): Entrega do Aplicando o Conhecimento 1.

Milestone 1: Estrutura do projeto definida.

Etapa 2 – Definição do Projeto e Primeiras Análises

11/03/2025 a 03/04/2025:

- Análise exploratória dos dados (11/03 a 15/03): Realizar análise exploratória inicial para identificar padrões nos dados.
 - Responsável: **Felipe**
- Limpeza e tratamento das informações (16/03 a 20/03): Reunir e limpar os dados fornecidos pelo Kaggle para a análise.
 - Responsável: **Gabriel**
- Escolha das bibliotecas no Python (Pandas, Matplotlib, Seaborn, Scikit-Learn).
(20/03 a 25/03): Primeiras análises e testes preditivos.

- Responsável: **Juliana e Yuri**
- Entrega da segunda etapa do Projeto (02/04)

Milestone 2: Conjunto de análises iniciais definidas.

Etapa 3 – Storytelling e Documentação

10/04/2025 a 28/04/2025:

- Desenvolvimento da Estratégia de Apresentação: (10/04 a 15/04): Consolidação das informações obtidas.
 - Responsável: **Felipe e Yuri**
- 18/04/2025 a 01/05/2025: Redação da documentação e revisão crítica.
 - Responsável: **Gabriel e Juliana**
- Preparação da Apresentação Final (25/04 a 27/04): Preparar a apresentação final
 - Responsável: **Todo o grupo**

Milestone 3: Documento estruturado e complementado com fontes externas.

Etapa 4 – Refinamento e Apresentação

29/04/2025 a 26/05/2025:

- Ajustes finais: (30/04 a 10/05): Refinamento final e ajustes.
 - Responsável: **Juliana e Felipe**
- (11/05 a 24/05): Criação da apresentação, gravação do vídeo e entrega final.
 - Responsável: **Gabriel e Yuri**

Milestone 4: Projeto finalizado e apresentado.

Pipeline de Dados

Este documento descreve as principais etapas envolvidas na construção e execução do pipeline do projeto, desde a coleta de dados até a análise exploratória final, destacando as principais atividades em cada etapa.

O processo analítico proposto neste projeto segue as seguintes etapas:

1. Coleta dos Dados:

- A base foi obtida através da plataforma Kaggle, com dados disponibilizados pela organização Our World in Data.

2. Leitura e Carregamento dos Dados

- Utilização da biblioteca Pandas para leitura dos arquivos .csv.

3. Limpeza e Tratamento dos Dados

- Verificação e tratamento de dados ausentes (NAs)
- Conversão e padronização de tipos de dados
- Normalização para comparações entre países

4. Análise Exploratória dos Dados

- Geração de estatísticas descritivas (média, mediana, desvio padrão)
- Detecção e análise de outliers
- Identificação de padrões e variações anuais

5. Visualização dos Dados

- Uso das bibliotecas **Matplotlib** e **Seaborn** para construir gráficos de linha, boxplots, histogramas e heatmaps de correlação.

6. Primeiras Inferências e Considerações

- Interpretação dos dados visualizados para elaboração de hipóteses e identificação de pontos de interesse para análises futuras.

INTRODUÇÃO

Estudo sobre o Consumo de Energia Elétrica Mundial: Uma Análise Histórica de 1965 a 2022

O consumo de energia elétrica é um dos pilares da sociedade moderna, mas sua dependência de fontes não renováveis, como o petróleo, levanta desafios ambientais e econômicos críticos. Nos últimos anos, o consumo de energia elétrica aumentou significativamente, levando governos e acadêmicos a buscarem alternativas sustentáveis, como energia nuclear e eólica.

Através dos dados obtidos no dataset World Energy Consumption, é possível identificar os países com maiores e menores consumos, analisar padrões históricos e entender os impactos de diferentes fontes energéticas.

Organização: A plataforma *Our World in Data* é uma organização dedicada à divulgação de dados globais sobre diversos aspectos sociais, ambientais e econômicos. No contexto energético, fornece um dataset abrangente sobre o consumo de energia ao longo dos anos, permitindo análises comparativas entre diferentes regiões e fontes energéticas.

Descrição: O dataset World Energy Consumption contém informações sobre o consumo energético global entre 1965 e 2022, abrangendo diferentes países, regiões e especificações das fontes de energia.

OBJETIVOS DO PROJETO

Objetivo Geral

O presente estudo tem como propósito elaborar novos relatórios analíticos, a serem publicados pela própria *Our World in Data*, com foco especial na transição para fontes de energia limpa e no impacto ambiental da geração de eletricidade ao redor do mundo. A intenção é transformar os dados já disponíveis na plataforma em informações organizadas, acessíveis e relevantes, que contribuam diretamente para o avanço da agenda energética global.

Entre os principais objetivos desses relatórios, estão:

- Identificar os países e continentes que mais evoluíram na substituição de fontes fósseis por energias renováveis
- Analisar o consumo energético per capita em diferentes regiões
- Apontar os maiores emissores de poluentes provenientes da geração de energia, permitindo uma visão clara da contribuição de cada economia para as emissões globais

Ao reunir e interpretar essas informações, o estudo busca evidenciar as desigualdades regionais, os avanços já conquistados e os principais gargalos que ainda limitam uma transição energética sustentável e equitativa. Espera-se que os relatórios sirvam como ferramenta de apoio para o planejamento de ações concretas, baseadas em dados, que promovam investimentos mais eficazes em tecnologias limpas e redução de emissões no setor energético.

Objetivos Específicos

Para alcançar o objetivo geral proposto, o projeto visa:

- Analisar o consumo mundial de energia elétrica entre 1965 e 2022 com base nos dados do dataset *World Energy Consumption*;

- Investigar padrões históricos de consumo de energia por país e por região, identificando tendências e mudanças significativas;
- Avaliar a participação de fontes renováveis e não renováveis no consumo energético global ao longo do tempo;
- Medir e interpretar as emissões de CO₂ associadas ao consumo de energia elétrica;
- Identificar outliers e anomalias relevantes nos dados que possam indicar situações atípicas;
- Produzir visualizações informativas que facilitem a compreensão dos dados analisados;
- Fornecer recomendações baseadas nos dados para possíveis melhorias em políticas energéticas sustentáveis.

METADADOS

- **Nome do Dataset:** World Energy Consumption
- **Período de Coleta:** 1900-2022
- **Fonte dos Dados:** Our World in Data
- **Formato dos Arquivos:** CSV / XLSX / JSON
- **Frequência de Atualização:** Anual
- **Unidade de Medida:** Terawatt-hora (TWh) para consumo de energia, quilowatt-hora per capita (kWh) para consumo individual e megatoneladas de CO₂ (MtCO₂) para emissões
- **Total de Registros:** ~5.000 linhas (estimado)
- **Total de Colunas:** 10 principais

Atributos principais do dataset

Country: Nome do país ou região analisada (Dado Categórico).

Year: Ano do registro (Dado Numérico - Inteiro).

Population: População estimada do país no ano correspondente (Dado Numérico - Inteiro).

Electricity consumption: Consumo total de eletricidade em Terawatt-hora (TWh) (Dado Numérico - Float).

Electricity per capita: Consumo médio de eletricidade por habitante em quilowatt-hora (kWh) (Dado Numérico - Float).

Share of renewables (%): Percentual de fontes renováveis no consumo total (Dado Numérico - Float).

CO₂ emissions from electricity (MtCO₂): Emissão de CO₂ associada ao consumo de eletricidade em megatoneladas (Dado Numérico - Float).

Primary energy consumption (EJ): Consumo primário de energia em exajoules (Dado Numérico - Float).

Fossil fuels consumption (%): Percentual do consumo energético baseado em combustíveis fósseis (Dado Numérico - Float).

GDP per capita (USD): Produto Interno Bruto per capita, ajustado pela inflação (Dado Numérico - Float).

Desafios Identificados:

- Dados ausentes para alguns países em períodos específicos.
- Dados efetivos de consumo apenas disponíveis a partir do ano 1965
- Na coluna Country também possui registros de continentes, regiões e mundial
- Necessidade de padronização das unidades de medida.
- Possível necessidade de normalização dos dados para comparação entre países.

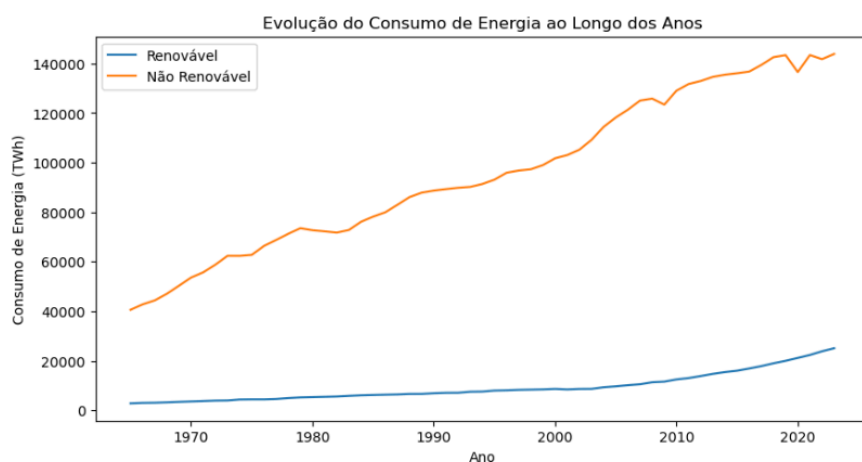
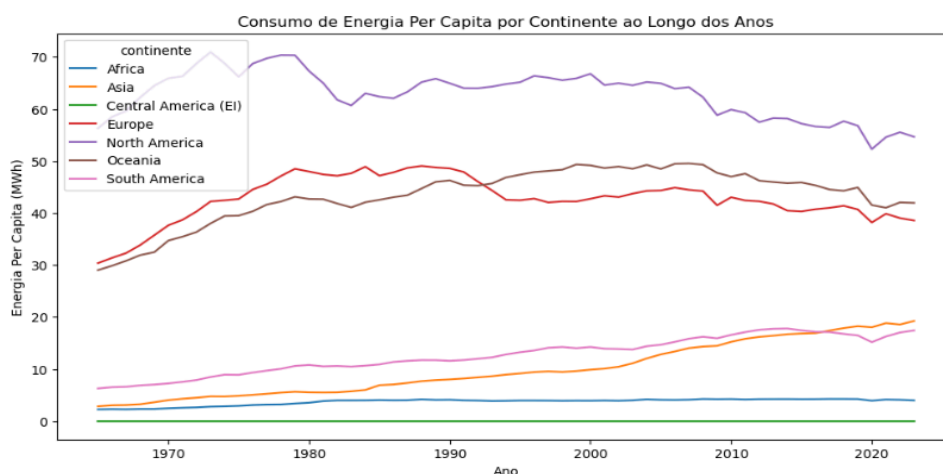
LISTA DE FIGURAS

- Figura 1: Consumo de Energia Per Capita por Continente ao Longo dos Anos
- Figura 2: Evolução do consumo de Energia ao longo dos anos
- Figura 3: Consumo de Energia por Continente ao Longo dos Anos
- Figura 4: Países com maior crescimento no consumo de energia per capita
- Figura 5: Correlação entre PIB e Percentual de Energia Renovável (Média dos últimos 10 anos)
- Figura 6: Top 10 países com maior % de energia renovável
- Figura 7: Evolução das emissões de CO2 por continente (2000 a 2022)
- Figura 8: Top 10 países emissores de gases de efeito estufa
- Figura 9: Evolução do Consumo de Diferentes Fontes de Energia Renovável

Crescimento Acelerado do Consumo Energético

Ao longo das últimas décadas, identificamos um aumento exponencial no consumo global de energia elétrica, especialmente a partir da segunda metade do século XX. Esse crescimento acompanha os avanços da industrialização, o crescimento populacional e a expansão de tecnologias de informação, principalmente em países desenvolvidos.

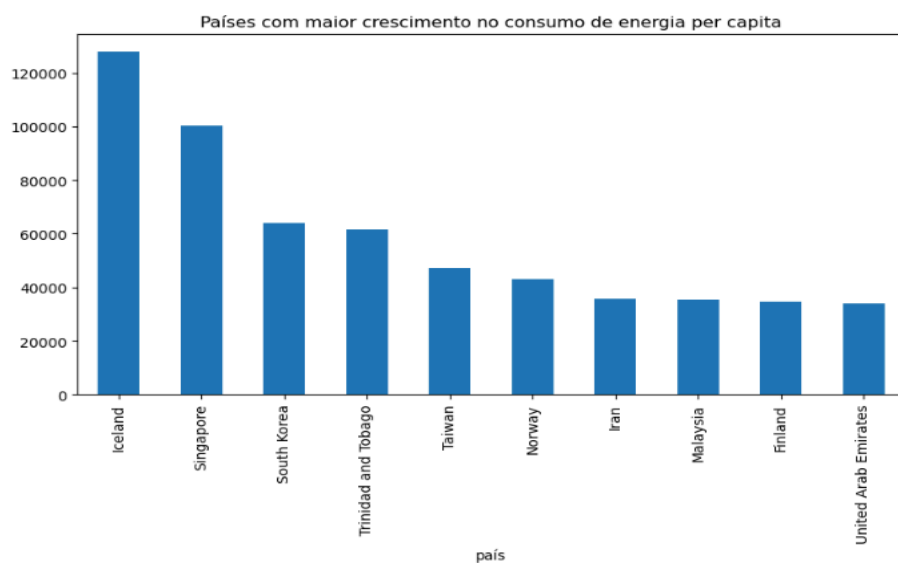
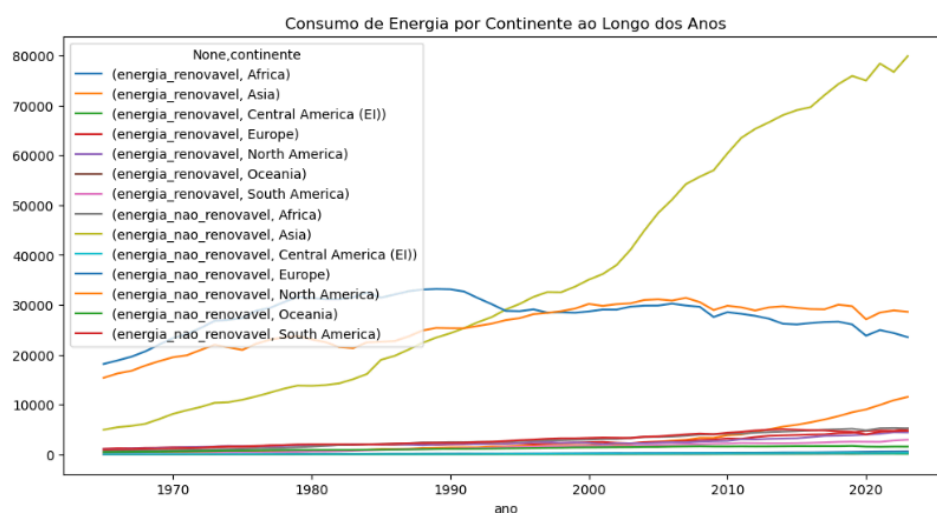
- **Insight:** países da América do Norte, Europa e partes da Ásia foram os maiores responsáveis por esse aumento, com destaque para os Estados Unidos e a China.
- **Métrica observada:** médias e totais de consumo anual, mostrando uma variação significativa entre 1965 e 2022.



Desigualdade no Consumo por Região

A análise revelou grandes disparidades no consumo energético entre regiões. Países desenvolvidos apresentam um consumo per capita muito maior do que países em desenvolvimento, o que levanta questionamentos sobre acesso à energia, políticas públicas e transições justas.

- **Insight:** a desigualdade energética continua sendo um desafio global que precisa ser enfrentado com políticas sustentáveis.
- **Métrica observada:** desvio padrão alto entre os países, indicando essa disparidade.

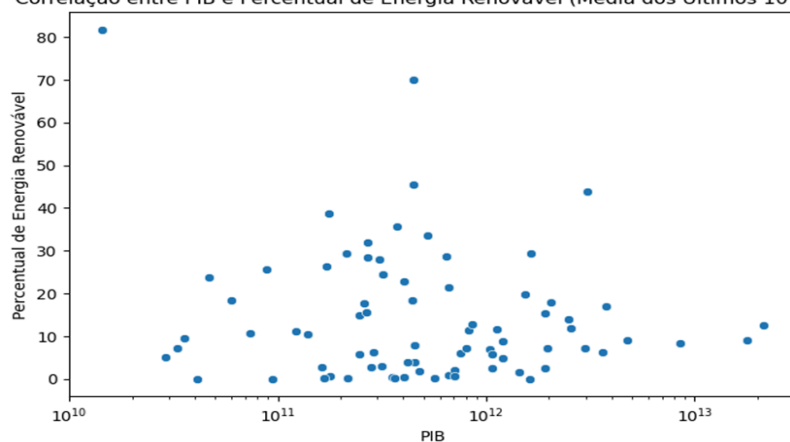


Tendência de Aumento nas Fontes Renováveis

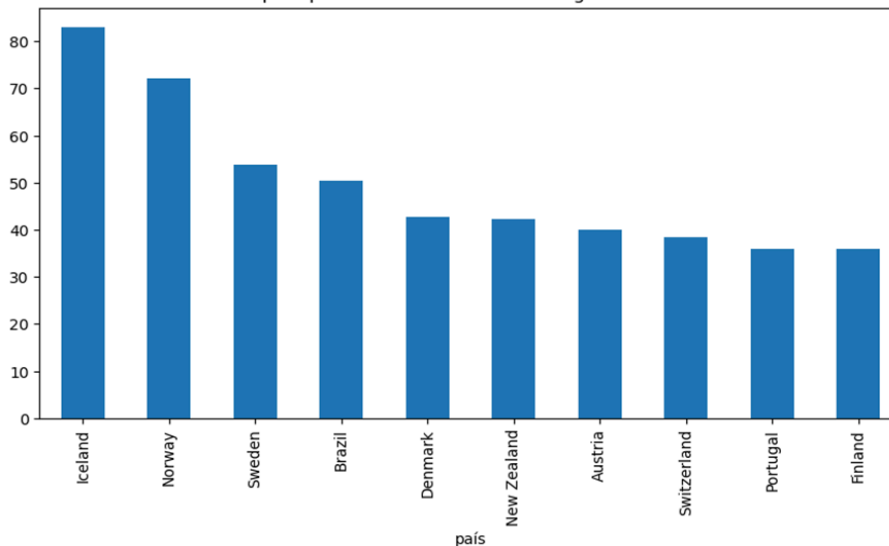
Embora o consumo total tenha aumentado, observou-se um crescimento expressivo no uso de fontes renováveis, especialmente após os anos 2000. Isso sugere um movimento em direção à transição energética sustentável.

- **Insight:** a Europa e partes da Ásia lideram o uso de energia solar e eólica.
- **Métrica observada:** distribuição percentual por tipo de fonte energética, demonstrando a mudança gradual na matriz energética global.

Correlação entre PIB e Percentual de Energia Renovável (Média dos Últimos 10 Anos)



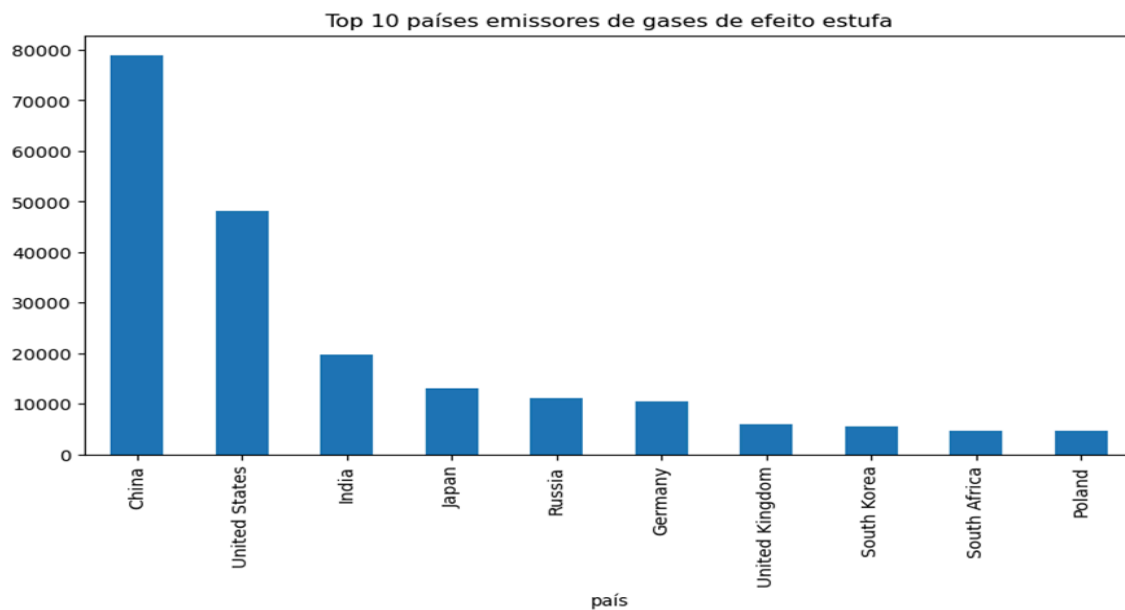
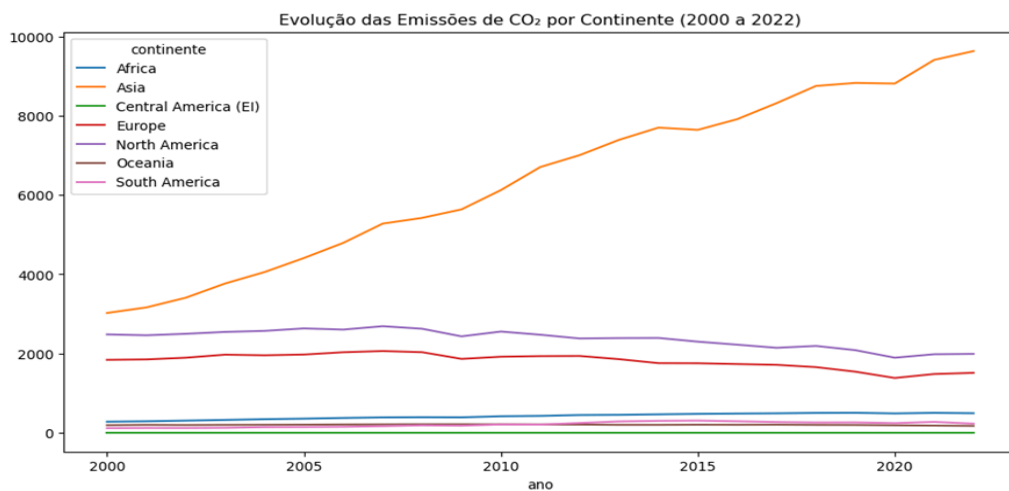
Top 10 países com maior % de energia renovável



Presença de Outliers e Padrões Sazonais

Alguns anos apresentam valores fora do padrão, possivelmente causados por eventos globais como guerras, crises econômicas ou pandemias. Essas anomalias foram identificadas durante a verificação dos outliers em séries temporais.

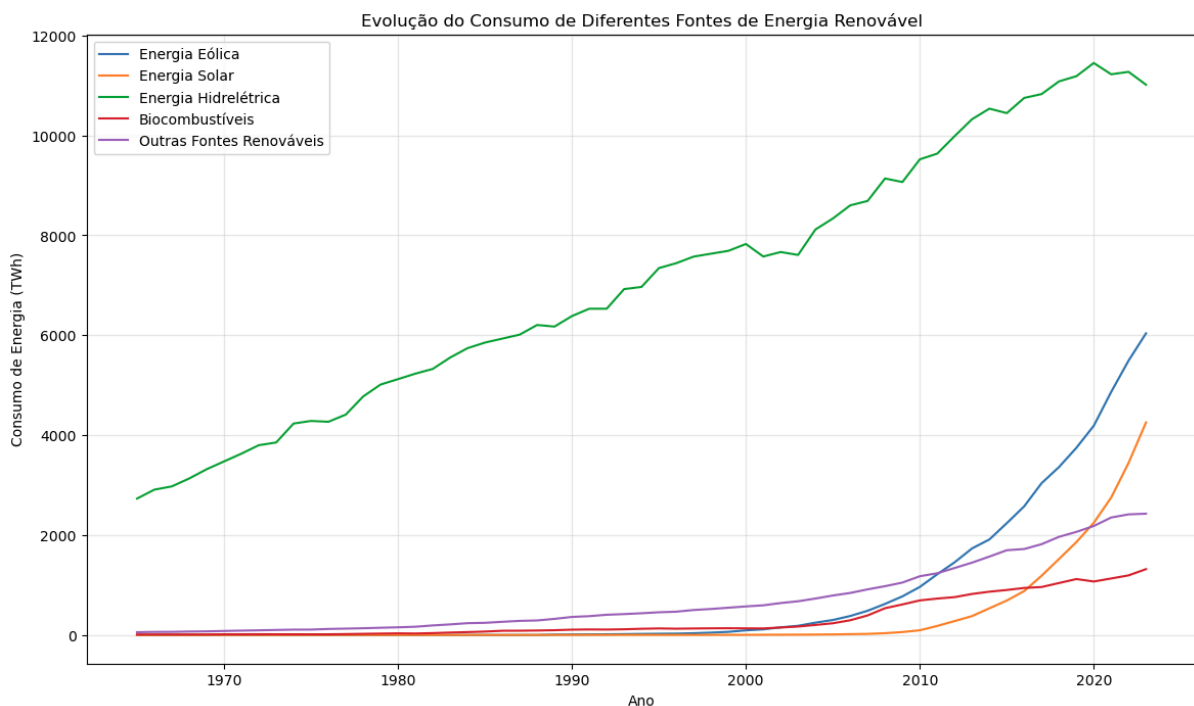
- **Insight:** Períodos como a pandemia de COVID-19 influenciaram diretamente o consumo em determinados países.
- **Métrica observada:** Gráficos de barras e gráficos de linhas com grandes variações de valores.



Evolução de energias renováveis

Ao analisarmos exclusivamente as fontes de energia renovável, é possível observar um padrão claro de crescimento ao longo das últimas décadas, impulsionado principalmente pelo avanço tecnológico e pela redução dos custos de implementação.

- **Insight:** A geração de energia hidrelétrica apresenta períodos de estagnação ou queda, especialmente nos anos mais recentes. Em contraste, as fontes solar e eólica vêm demonstrando um crescimento contínuo e acelerado, consolidando-se como as principais apostas do setor energético renovável no cenário global.
- **Métrica observada:** Através da análise dos gráficos de linha é possível verificar que determinadas fontes energéticas ganharam grande espaço nos últimos 15 anos



ANÁLISE EXPLORATÓRIA

Como podemos observar nos gráficos anteriores, o consumo mundial de energia tem aumentado de forma ininterrupta, algo que podemos observar devido ao crescimento populacional, à industrialização de países em desenvolvimento e ao aumento da demanda por tecnologias e serviços que dependem de eletricidade. Esse crescimento é especialmente impulsionado por setores como transporte, produção industrial e infraestrutura digital.

Outro ponto que podemos observar é referente à Figura 3, que nos mostra a correlação entre o PIB e o percentual de energia renovável utilizado pelos países. A dispersão dos pontos indica que não há uma correlação clara entre essas variáveis, sugerindo que um PIB mais alto não necessariamente implica em um maior uso de fontes renováveis de energia. Isso pode estar relacionado a fatores como políticas energéticas, disponibilidade de recursos naturais e investimentos em infraestrutura sustentável.

Do mesmo modo, na figura 4 podemos observar uma emissão maior de CO₂ no continente asiático, que possui um consumo extremamente elevado de combustíveis fósseis, além de abrigar os maiores polos industriais do mundo e possuir uma grande concentração populacional. Onde para agravar ainda mais a situação, o crescimento excessivo da população faz com que a necessidade por energia seja maior do que a das capacidades do continente de gerar energia limpa.

GRÁFICOS E ANÁLISES

Da mesma maneira que obtemos informações individuais ao analisarmos as linhas de um dataset, também podemos extrair outros tipos de informações ao examinarmos suas colunas. No dataset selecionado, é possível comparar as fontes de energia mais utilizadas mundialmente, bem como analisar aspectos como o uso total, a variação no consumo, o uso per capita e o custo associado.

Além disso, a análise pode ser realizada tanto em nível continental quanto nacional. Para isso, é comum remover linhas de dados que contenham diversas colunas com valores ausentes, restringindo a análise aos anos que apresentam maior disponibilidade de informações, a fim de garantir maior precisão.

Dessa forma, conseguimos obter uma visualização mais completa dos dados de determinado país. Contudo, é fundamental atentar-se à separação correta dos tipos de dados a serem analisados, pois, do contrário, corre-se o risco de misturar informações de natureza distinta, como dados de contagem, consumo ou valores monetários.

Ao segmentar colunas específicas, conseguimos apresentar informações com maior precisão e clareza. Assim, é possível analisar, por exemplo, o crescimento ou a redução do consumo per capita de cada tipo de combustível ao longo dos anos, bem como o aumento ou a diminuição dos custos associados. Além disso, essa análise permite observar como o crescimento do uso de uma determinada fonte de energia pode impactar na redução do consumo ou no preço das demais. As possibilidades de interpretação e análise são diversas.

REFERÊNCIAS

POUDAL, Pralabh. *World Energy Consumption*. Disponível em: <https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption/data>.

Acesso em: 27/02/2025.

AQUINO, Wellington. *Lista de tabelas: como criar corretamente no Word*. Disponível em:

[https://regrasabnt.com.br/listas-de-tabela/#O que e uma lista de tabelas ABNT](https://regrasabnt.com.br/listas-de-tabela/#O%20que%20e%20uma%20lista%20de%20tabelas%20ABNT).

Publicado em: 1 dez. 2024. Acesso em: 28/03/2025.

PIXABAY. *Pesquisa por "gráfico"*. Disponível em: <https://pixabay.com/pt/images/search/gráfico/>. Acesso em: 28/03/2025.

PIXABAY. *Imagem: Borrão, Gráfico, Computador, Dados*. Disponível em: <https://pixabay.com/pt/photos/borrão-gráfico-computador-dados-1853262/>. Acesso em: 28/03/2025.

PIXABAY. *Imagem: Computador, Resumo, Gráfico, O Negócio*. Disponível em: <https://pixabay.com/pt/photos/computador-resumo-gráfico-o-negócio-767776/>.

Acesso em: 28/03/2025.

RB ENERGIA. *Produção mundial de energia: veja como funciona o mercado*. Disponível em:

<https://www.rbenergia.com.br/blog/producao-mundial-de-energia-eletrica/#:~:text=Panorama%20da%20produção%20mundial%20de,que%20naturalmente%20ocorrem%20na%20transmissão>. Publicado em: 27 jul. 2022. Acesso em: 28/03/2025.

INTERNATIONAL ENERGY AGENCY. *Latin America Energy Outlook 2023 – Executive Summary*. Paris: IEA, 2023. Disponível em:

<https://www.iea.org/reports/latin-america-energy-outlook-2023/executive-summary?language=pt>. Acesso em: 29 abr. 2025.

BRASIL. Empresa de Pesquisa Energética. *Plano Nacional de Energia 2030 – Geração Hidrelétrica*. Brasília: EPE, 2007. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-165/topico-173/PNE%202030%20-%20Gera%C3%A7%C3%A3o%20Hidrel%C3%A9trica.pdf>. Acesso em: 29 abr. 2025.

GOMES, João Paulo Pombeiro; VIEIRA, Marcelo Milano Falcão. O campo da energia elétrica no Brasil de 1880 a 2002. *Revista de Administração Pública*, Rio de Janeiro, v. 43, n. 2, p. 347-372, abr. 2009. Disponível em: <https://www.scielo.br/j/rap/a/NWxd9HmK8wJBGKMPq6GcLqz/?format=html&lang=pt>. Acesso em: 29 abr. 2025.

ARIAS, Leonel Marques; SELLITTO, Miguel Afonso. Uso da energia ao longo da história: evolução e perspectivas futuras. *Revista Liberato*, Novo Hamburgo, v. 12, n. 17, p. 7–16, 2011. Disponível em: <https://revista.liberato.com.br/index.php/revista/article/view/164>. Acesso em: 29 abr. 2025