

Fundação Getúlio Vargas

Ciência de Dados

Análise de dados do Consumo de Energia Mundial

Beatriz Marques
Luísa Costa
Maria Clara de Souza
Raphaella Roma

Orientador:
Prof. Dr. Rafael Pinho

Rio de Janeiro
2024

Resumo

Este trabalho utiliza o dataset World Energy Consumption, mantido pelo Our World in Data, para explorar quatro hipóteses relacionadas à distribuição de energia e à correlação entre variáveis energéticas e aspectos sociais dos países no período de 1900 a 2022. Os dados foram tratados com Python, e gráficos foram gerados utilizando as bibliotecas Matplotlib e Seaborn. Os resultados, que variam conforme cada hipótese, serão apresentados ao longo deste estudo.

Palavras-chave: Dataset, Energia, Países, Python, Gráficos.

Abstract

This work uses the World Energy Consumption dataset, maintained by Our World in Data, to explore four hypotheses related to energy distribution and the correlation between energy variables and social aspects of countries from 1900 to 2022. The data was processed using Python, and graphs were generated with the Matplotlib and Seaborn libraries. The results, which vary depending on each hypothesis, will be presented throughout this study.

Keywords: Dataset, Energy, Countries, Python, Graphs.

1 Introdução

A crescente crise ambiental faz lembrar da necessidade de um monitoramento detalhado do consumo de energia mundial, e torna a análise de fontes renováveis um tema urgente da atualidade. Este presente trabalho visa investigar o comportamento da distribuição de energia ao longo dos anos, com foco nas fontes renováveis e sua relação com fatores socioeconômicos.

O objetivo geral do estudo é compreender os padrões de comportamento das diferentes fontes de energia, com um recorte socioeconômico. Para isso, quatro hipóteses principais foram formuladas, incluindo a análise da proporção de energia renovável por país e continente, a correspondência entre produção e demanda, e a dependência de combustíveis fósseis em relação ao PIB. Considerando a extensão temporal dos dados e a perda intrínseca dos registros integrais, valores nulos foram tratados e a análise foi concentrada nos anos mais recentes, período em que as energias renováveis começaram a ganhar maior relevância.

Sobre os Dados:

A base de dados utilizada para este estudo foi obtida no Kaggle, intitulada World Energy Consumption, disponibilizada por Poudel (2023). Esta base contém informações detalhadas sobre o consumo de energia em diferentes países ao longo do tempo, abrangendo diversas fontes de energia, como solar, eólica, hidrelétrica, carvão, gás, petróleo, nuclear etc. Os dados contém diversas informações sobre consumo e produção de energia, de diversos países ao longo do tempo. Todas as informações foram produzidas pelo Our World in Data.

Análises:

A análise foi feita utilizando a biblioteca Pandas e Seaborn para visualizações, os dados foram processados para remover outliers e valores faltantes que pudessem interferir nas análises.

2 Metodologia

Inicialmente foi feita uma limpeza nos dados. Cada hipótese precisa de dados diferentes, então foi feita uma função para cada.

Na hipótese 1 o primeiro tratamento foi selecionar apenas as colunas necessárias para a análise considerando todo o período 1965 a 2022. Em seguida, foram excluídas as linhas com valores nulos. Adicionalmente, para análise de demanda de energia elétrica, foi considerado o período a partir do ano 2000 pela indisponibilidade desta informação nos anos anteriores.

Foram escolhidos os gráficos de linha para mostrar a evolução histórica dos indicadores e os gráficos de dispersão para análise comparativa, permitindo avaliar as correlações entre os dados.

Já para a hipótese 2, criamos um dicionário com os países por continente e depois definimos uma função para mapear os dados e alocá-los em uma nova coluna no dataset que foi gerado para desenvolver essa hipótese. Além disso, selecionamos apenas as colunas que seriam usadas, criando uma para a soma das colunas sobre o consumo de todas as energias renováveis da tabela. Para finalizar a limpeza, filtramos os dados no intervalo de tempo de 2001 a 2021.

Na hipótese 3 o primeiro tratamento foi eliminar os não países do dataset, ou seja, continentes e blocos econômicos. A seguir, selecionamos apenas as colunas necessárias: país, ano, PIB, demanda e produção. Depois excluímos linhas que contêm valores nulos nas colunas de demanda ou produção. Por fim, filtramos os dados para o período de 2000 a 2021, que é onde a maioria dos países têm dados completos.

Para a hipótese 4, foram selecionados os dados correspondentes ao consumo de combustível fóssil, de energia proveniente de carvão mineral, de petróleo, e de gás, de diversos países ao longo dos anos, assim como o seu PIB. Foram realizados alguns processos de limpeza para remover ocorrências de dados faltantes ou insuficientes para uma análise mais precisa, removendo também dados que não correspondiam à análise como de continentes por exemplo, onde o PIB não consta.

Depois da limpeza de dados, cada hipótese recebeu seu próprio módulo para a criação dos gráficos.

3 Desenvolvimento

3.1 Aumento da população implica em maior demanda de energia e uma possível dependência da importação?

3.1.1 Análise da Relação entre o Consumo de Energia e a População:

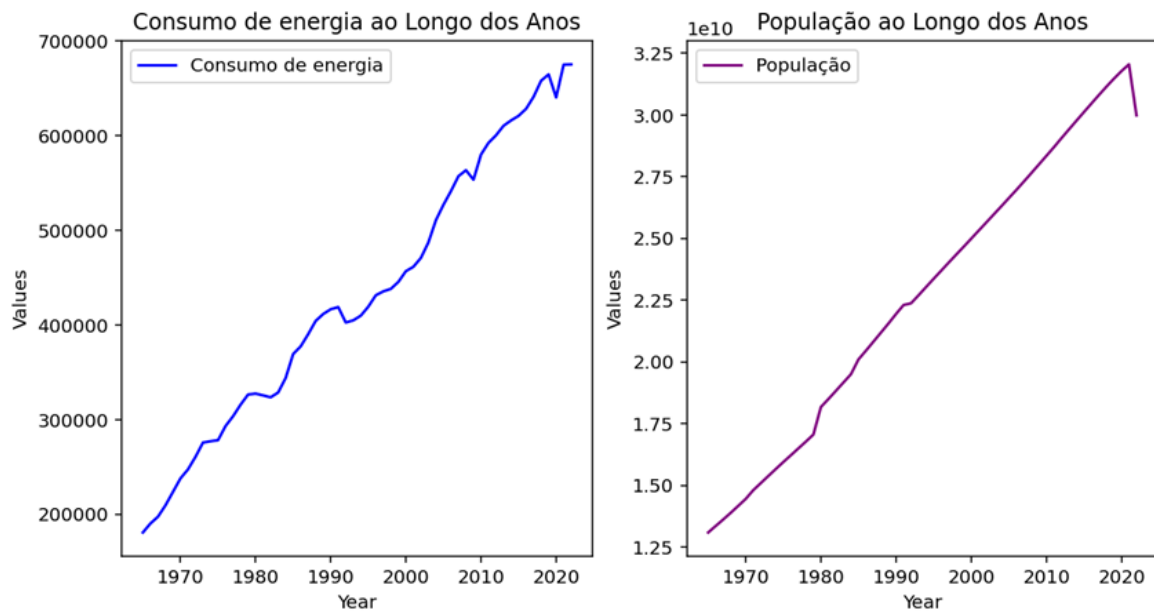


Figure 1: Evolução do consumo de energia e da população

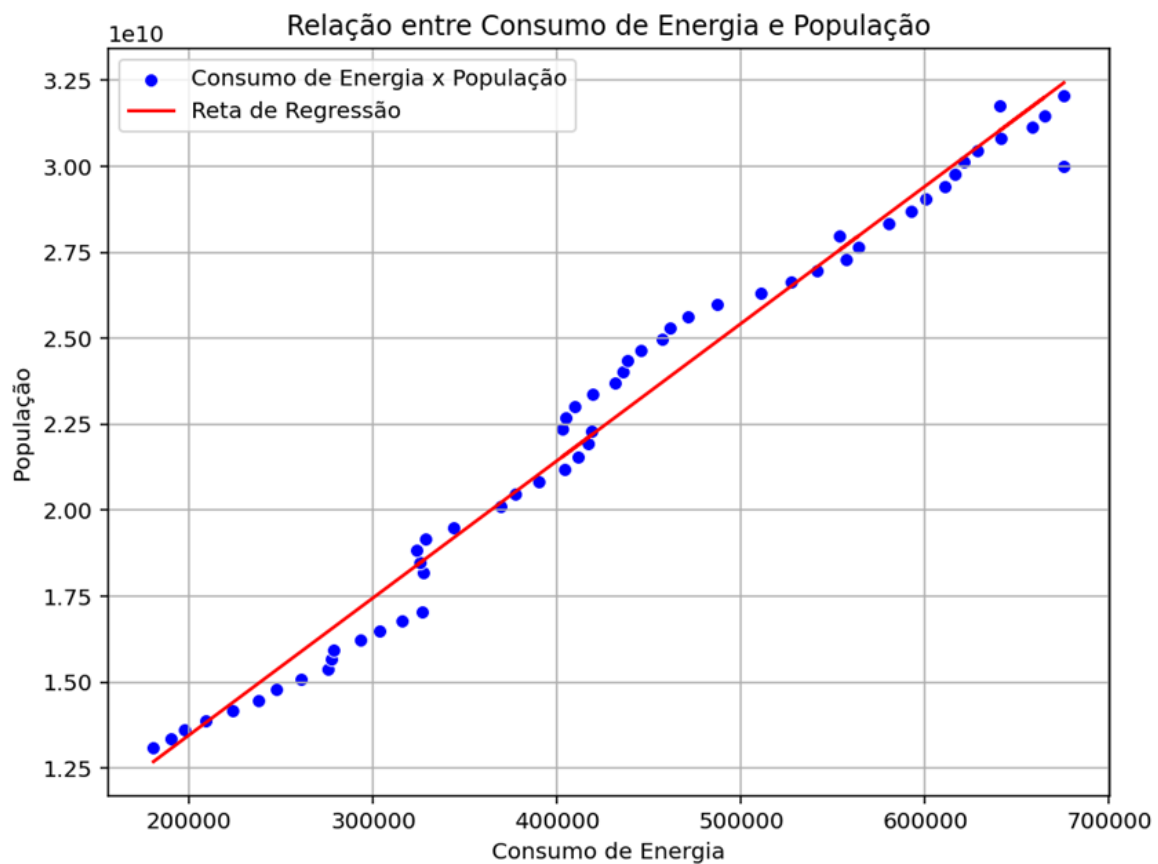


Figure 2: Relação entre consumo de energia e população

Os gráficos mostram a evolução da população e do consumo de energia do período de 1965 a 2022. Verifica-se que o crescimento do consumo de energia acompanha o crescimento da população, indicando ter uma correlação forte.

3.1.2 Análise da Relação entre o balanço de energia elétrica (diferença entre a geração e a demanda de energia elétrica) e as importações líquidas:

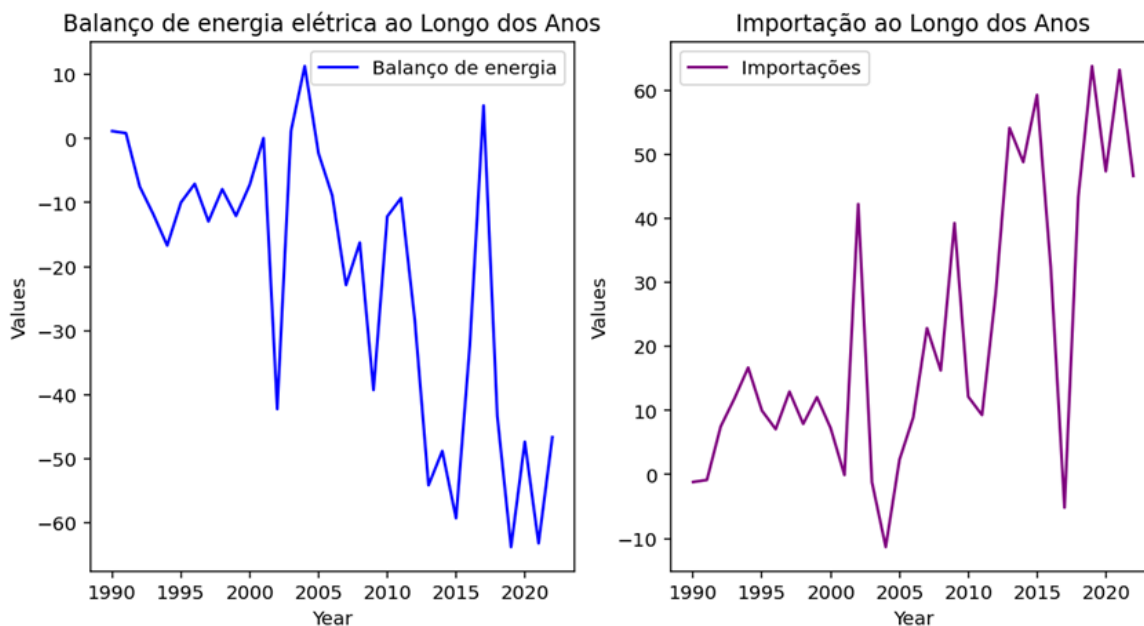


Figure 3: Evolução do balanço de energia elétrica e das importações líquidas

Pela análise dos gráficos, verifica-se que em vários momentos a demanda de eletricidade não é suprida pela geração dos países, sendo necessárias as importações para atendimento da população.

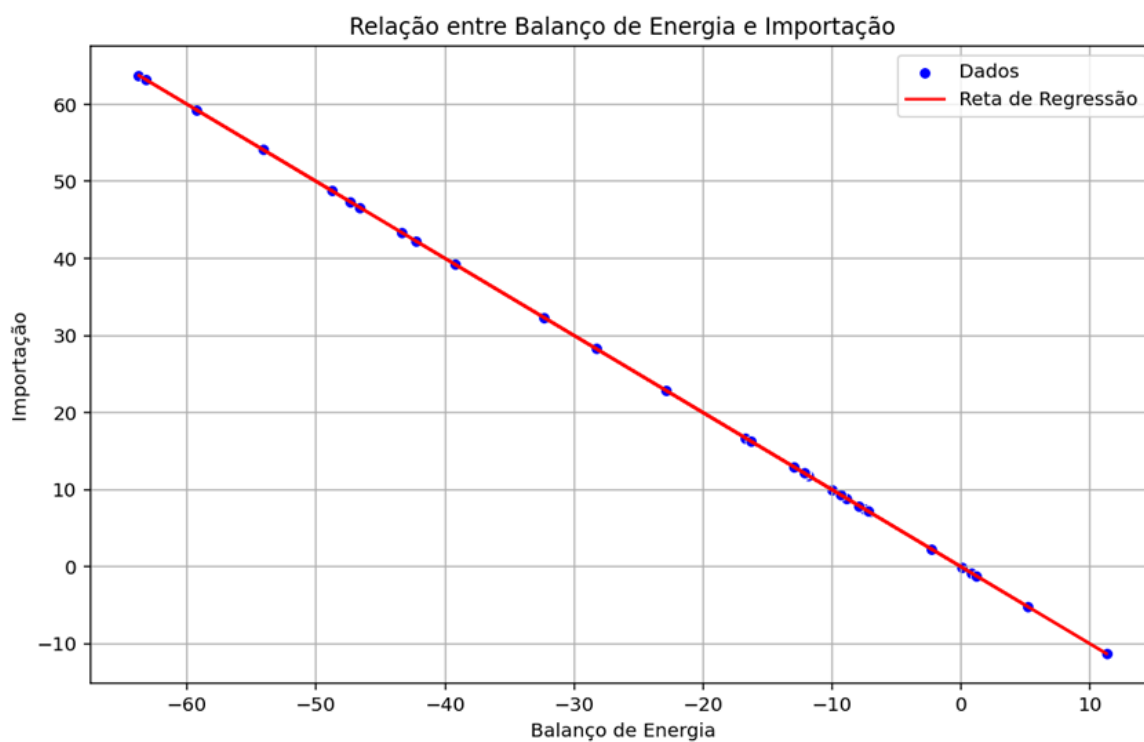


Figure 4: Evolução do balanço de energia e das importações líquidas

Confirmando a análise anterior, as importações são totalmente dependentes do balanço de energia elétrica, ocorrendo sempre em que a geração não é suficiente para atendimento da demanda.

3.1.3 Análise da relação entre a demanda de eletricidade e as importações líquidas:

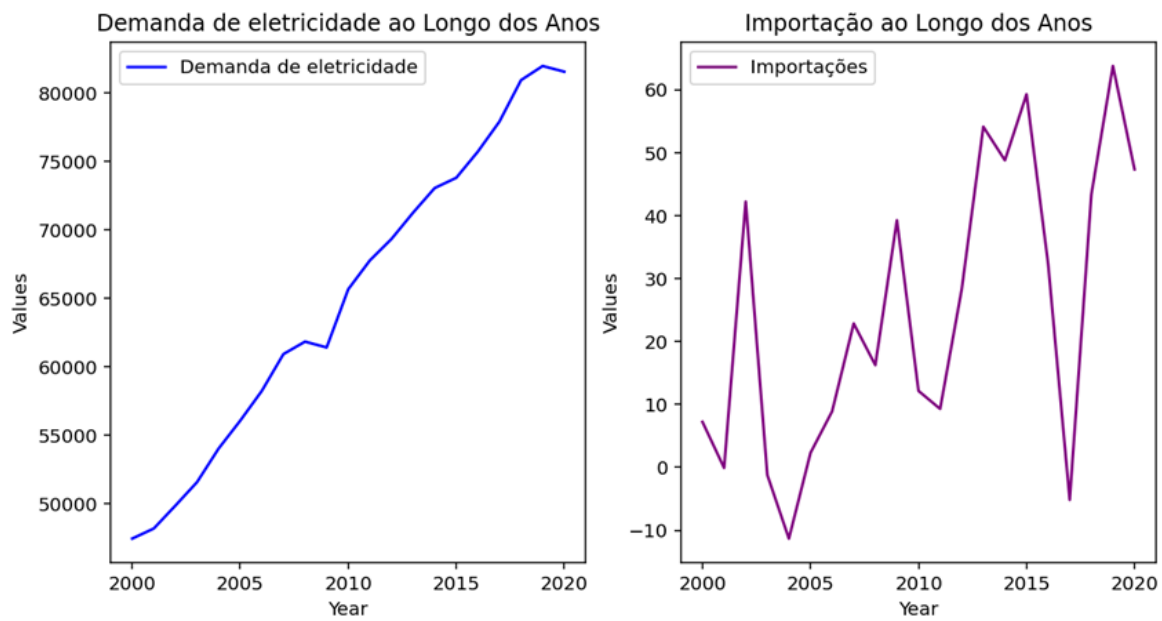


Figure 5: Evolução da demanda de eletricidade e das importações líquidas

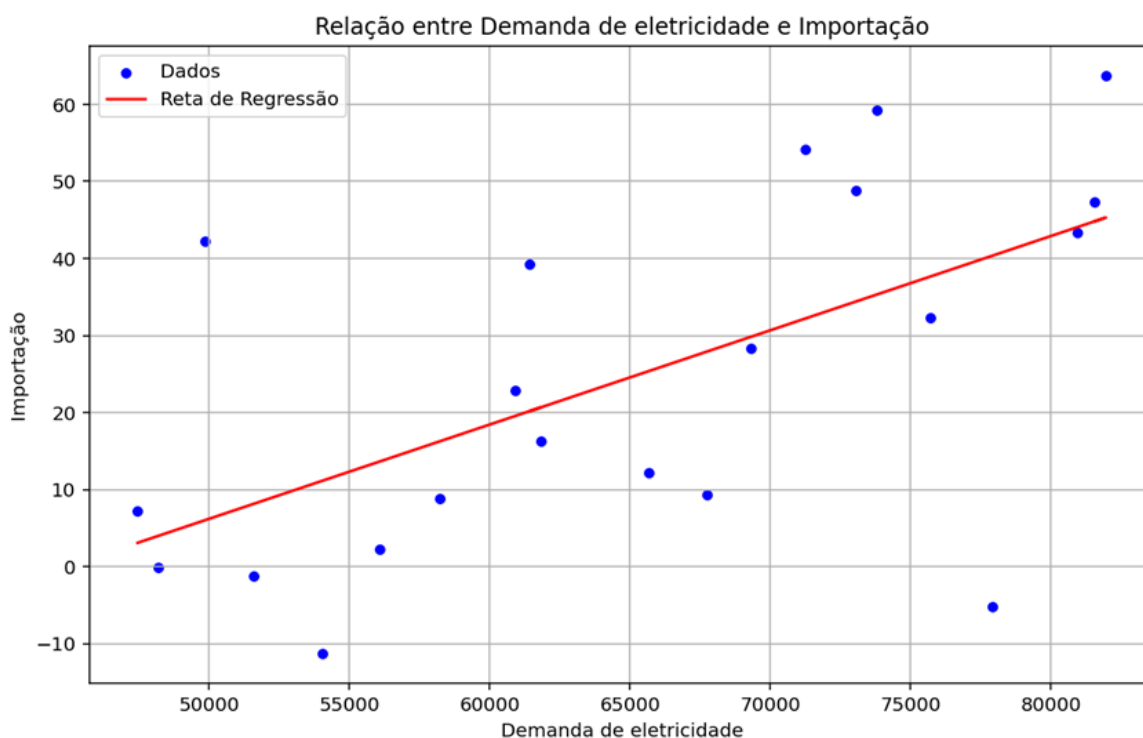


Figure 6: Relação da demanda de eletricidade e das importações líquidas

Pela análise dos gráficos, verifica-se que há uma correlação entre crescimento da demanda e da importação de eletricidade, mas essa correlação é fraca porque as importações dependem também da relação com a geração (balanço de energia elétrica), como visto na análise anterior, e não apenas com a demanda.

Observação: Nesta análise foi considerado o período posterior ao ano 2000, pois não havia informações disponíveis de demanda anteriores a esse ano.

3.2 Como a distribuição de fontes renováveis de energia varia por continente ao longo dos anos?

O objetivo dessa hipótese é investigar os padrões de consumo da energia renovável entre os continentes, focando na sua evolução ao longo do tempo.

Para obtenção de um quadro geral da situação mundial, somamos os valores válidos das colunas correspondentes às energias de fontes renováveis para visualizarmos a tendência no decorrer das últimas duas décadas que, como explicitado anteriormente, são as que possuem mais dados. Com a linha da média global plotada, é possível perceber as discrepâncias entre os números dos consumidores, principalmente entre a Ásia (positivamente) e Oceania, África e América Central (baixo consumo).

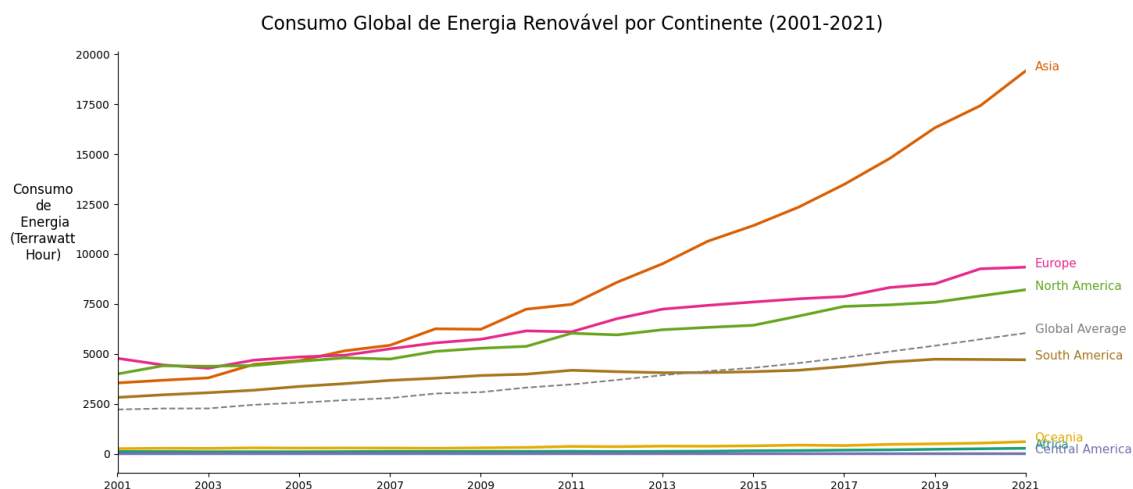


Figure 7: Tendências do Consumo de Energia Renovável por Continente

Um mapa de calor a partir das porcentagens permite uma análise mais individualizada e precisa das taxas de crescimento dos continentes. Segue o mesmo:

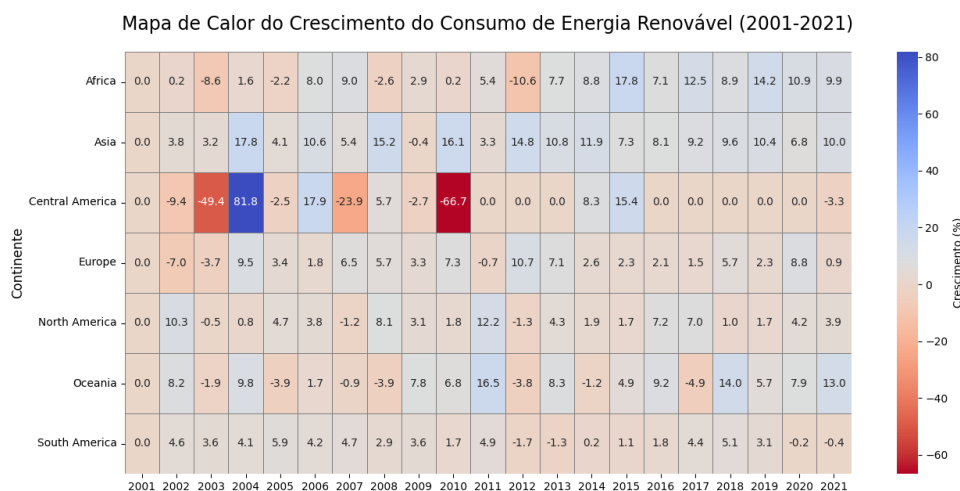


Figure 8: Mapa de Crescimento do Consumo de Energia Renovável

Há um destaque para a variação da América Central, enquanto os demais apresentam certa homogeneidade. Em uma avaliação mais profunda notamos que a Ásia mantém um crescimento acelerado e estável, seguida de outros continentes com taxas moderadas que são também considerados mais desenvolvidos, como Europa e Oceania.

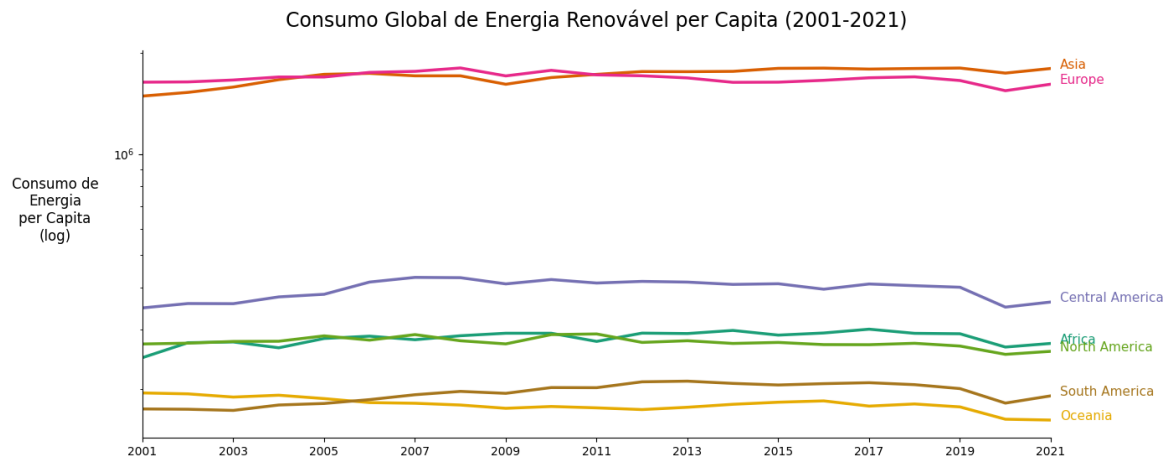


Figure 9: Consumo de Energia Renovável Per Capita

Esse gráfico foi gerado em escala logarítmica para uma melhor visualização devido a existência de discrepâncias consideráveis entre os continentes no consumo per capita. A escala logarítmica ajuda a evidenciar as diferenças, permitindo que as variações menores também sejam visíveis.

Ressalta-se também que os dados per capita foram retirados de uma coluna pré-calculada do dataset original, por estar mais completa.

A visualização revela uma liderança clara no consumo per capita da Europa e da Ásia por anos. Todos os continentes não apresentam grande volatilidade no decorrer do tempo. Os continentes com consumo per capita mais baixo sugerem uma disparidade no acesso a energias renováveis no continente.

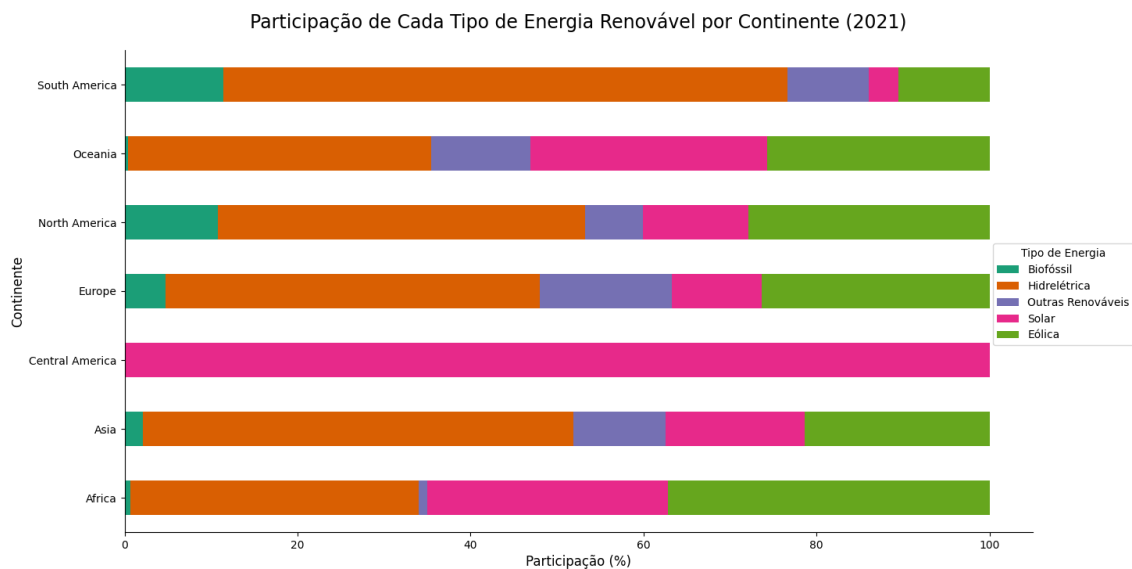


Figure 10: Fontes de Energia Renovável por Continente

As porcentagens de participação do tipo de energia, dentre as fontes renováveis, demonstra uma prevalência da hidrelétrica em todos os continentes (exceto a América Central, que possui dados apenas de energia solar). A América do Norte, a Europa e a Oceania apresentam certo equilíbrio e diversidade quanto as fontes de energia, o que pode ser reflexo de incentivos fiscais nessa área.

3.3 A produção de energia por fontes renováveis acompanha a demanda de energia globalmente?

Para responder a essa pergunta, analisamos o comportamento da produção de energia por fontes renováveis em relação à demanda global de energia ao longo dos anos. Inicialmente, foi construído um gráfico de barras empilhadas comparando a demanda global de energia com a produção de energia renovável em cada ano.

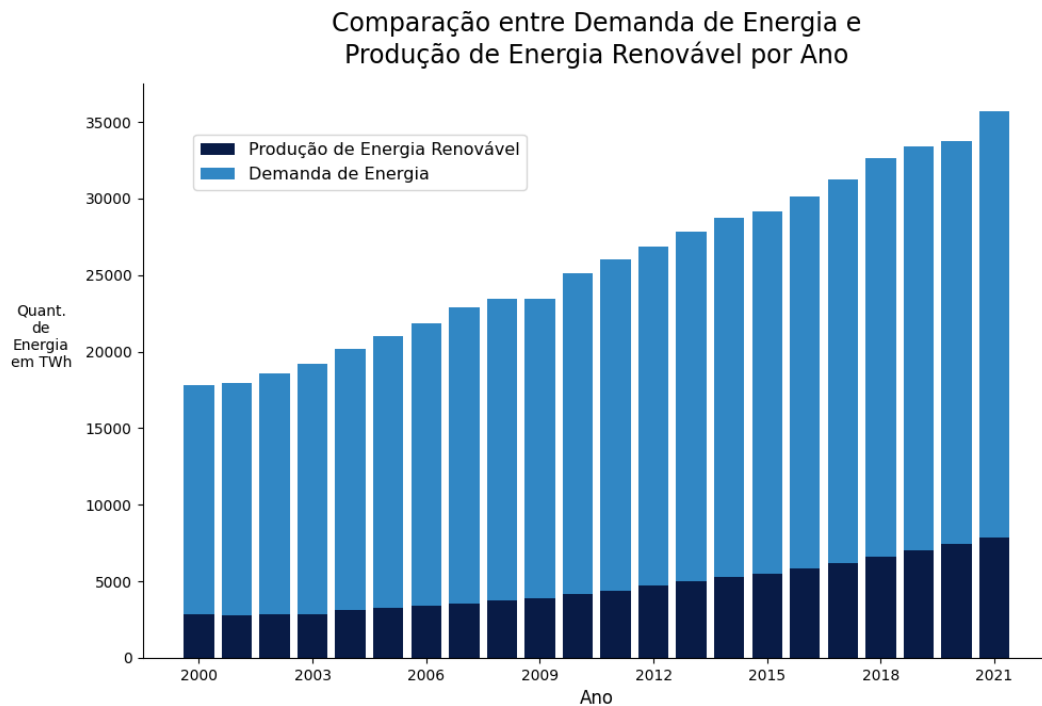


Figure 11: Gráfico de barras empilhadas que compara a demanda e a produção por ano

No gráfico, observamos que, de maneira geral, tanto a demanda quanto a produção de energia renovável cresceram ao longo do tempo. No entanto, para entender se há uma relação direta entre essas variáveis, foi calculado o coeficiente de correlação no módulo "hipótese 3". O resultado foi aproximadamente 0,97, indicando uma forte correlação positiva: quando a demanda aumenta, a produção de energia renovável também tende a aumentar.

Adiante, surge uma curiosidade: a produção de energia renovável cresce mais rápido que a demanda? Afinal, tal análise pode contribuir para importantes estudos futuros como, será que a produção de tal energia poderá um dia chegar bem próximo de atender toda a demanda, já que no gráfico acima ela não representa nem 50 por cento? O gráfico de linhas a seguir mostra a taxa de variação das duas variáveis:

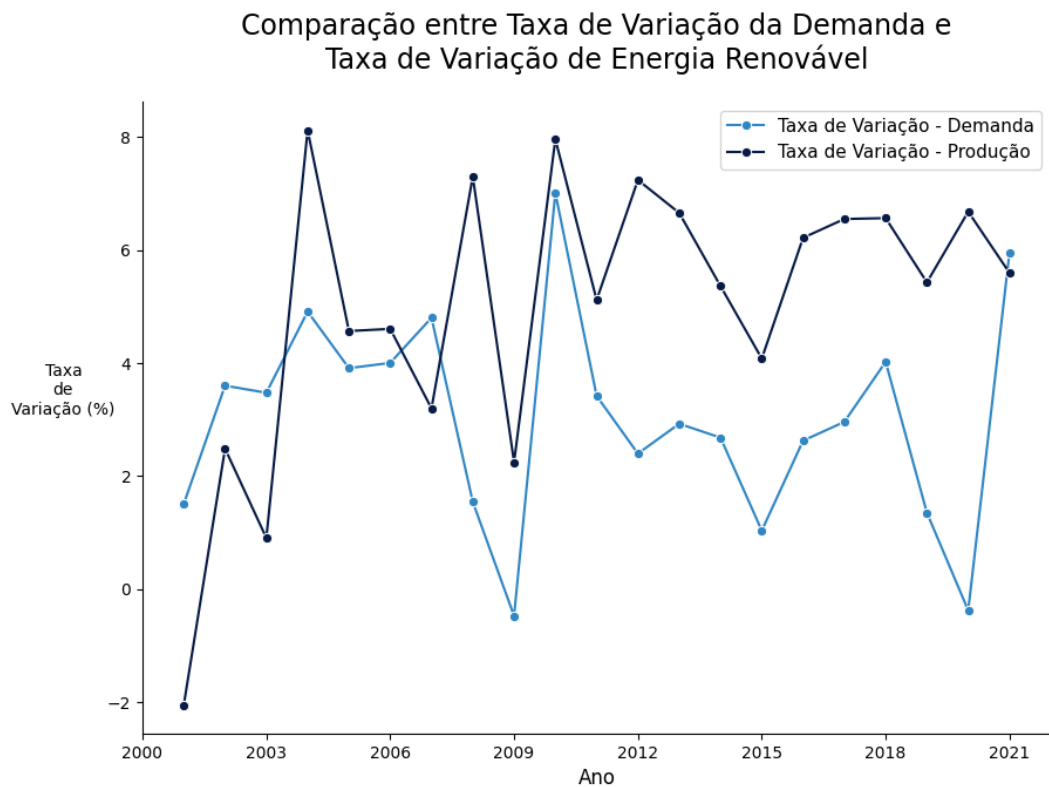


Figure 12: Line Plot que compara a taxa de variação da demanda com a taxa de variação da produção de energia renovável por ano

Os resultados mostram que não há um padrão consistente: em alguns anos, a taxa de variação da demanda foi maior; em outros, a produção renovável cresceu mais rapidamente. Isso sugere que a análise apenas das taxas anuais não é suficiente para tirar conclusões definitivas. Para obter uma visão mais precisa, foi criado um gráfico de linhas que apresenta a proporção entre a produção de energia renovável e a demanda global de energia ao longo dos anos.

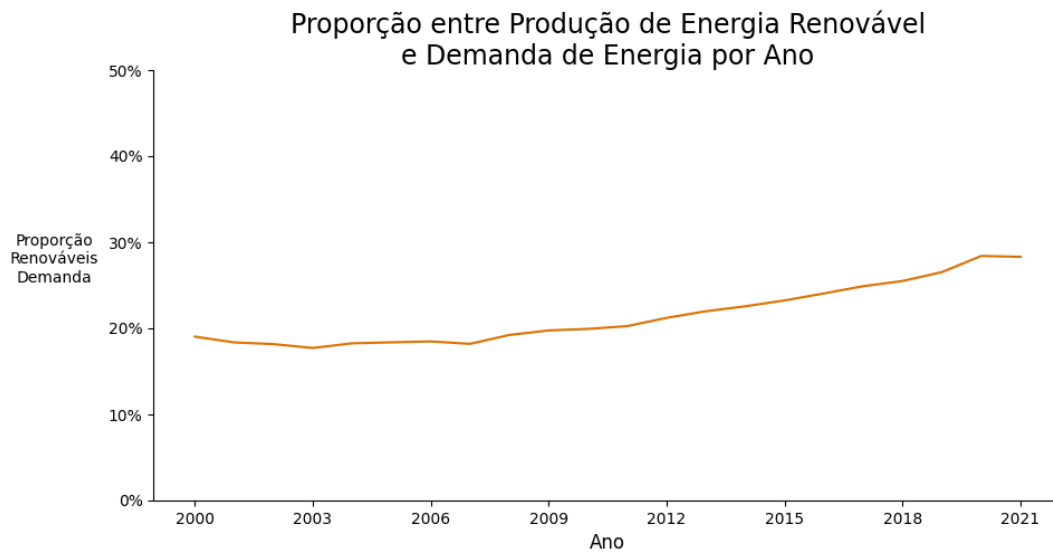


Figure 13: Line Plot que compara a proporção entre a produção de energia renovável e a demanda de energia ao longo dos anos

Os dados indicam que, de forma geral, essa proporção tem aumentado ao longo do tempo, sugerindo que a produção de energia renovável está crescendo em um ritmo mais rápido que a demanda de energia. Isso é um indicativo positivo, mas vale destacar que, mesmo em 2020, essa proporção não chega a 40 por cento, evidenciando que mais da metade da demanda ainda é atendida por fontes não renováveis.

Para aprofundar a análise, avaliamos individualmente a proporção entre a produção de energia renovável e a demanda de energia nos três países mais ricos do dataset (com maior média de PIB): Estados Unidos, China e Índia.

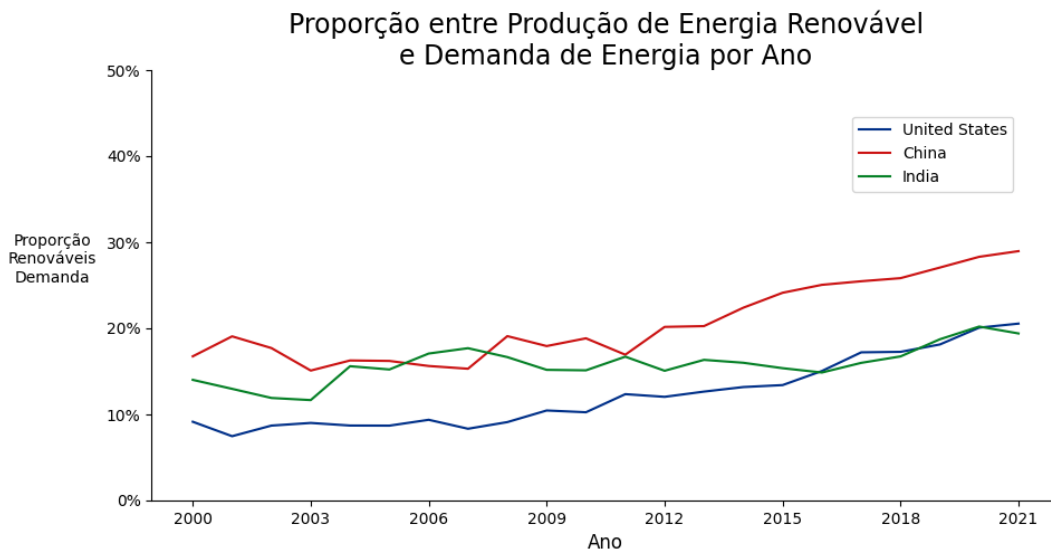


Figure 14: Line Plot que faz a comparação da proporção para os 3 países mais ricos

O país mais rico, EUA, possui a menor proporção em praticamente todos os anos analisados. Em 2000 a proporção era de 10 por cento e em 2022 foi para aproximadamente 20 por cento, tendo o crescimento mais uniforme entre os 3 países (quase sempre crescente).

O segundo mais rico, China, possui a maior proporção em quase todos os anos, sendo 18,5 por cento em 2000 e 29 por cento em 2022.

A Índia apresenta um comportamento mais irregular, com períodos de crescimento na proporção alternados com quedas.

Não há um padrão claro entre os três países, o que se explica por diversos fatores que influenciam a relação entre produção renovável e demanda, como: políticas governamentais, investimentos em infraestrutura energética, disponibilidade de recursos naturais para energias renováveis e crescimento econômico e populacional.

3.4 Países com maior PIB tem maior dependência de combustíveis fósseis em comparação com aqueles de menor PIB?

3.4.1 Análise da relação de crescimento e decrescimento do PIB em comparação com as variações no consumo de combustível fóssil dos países com Maior consumo:

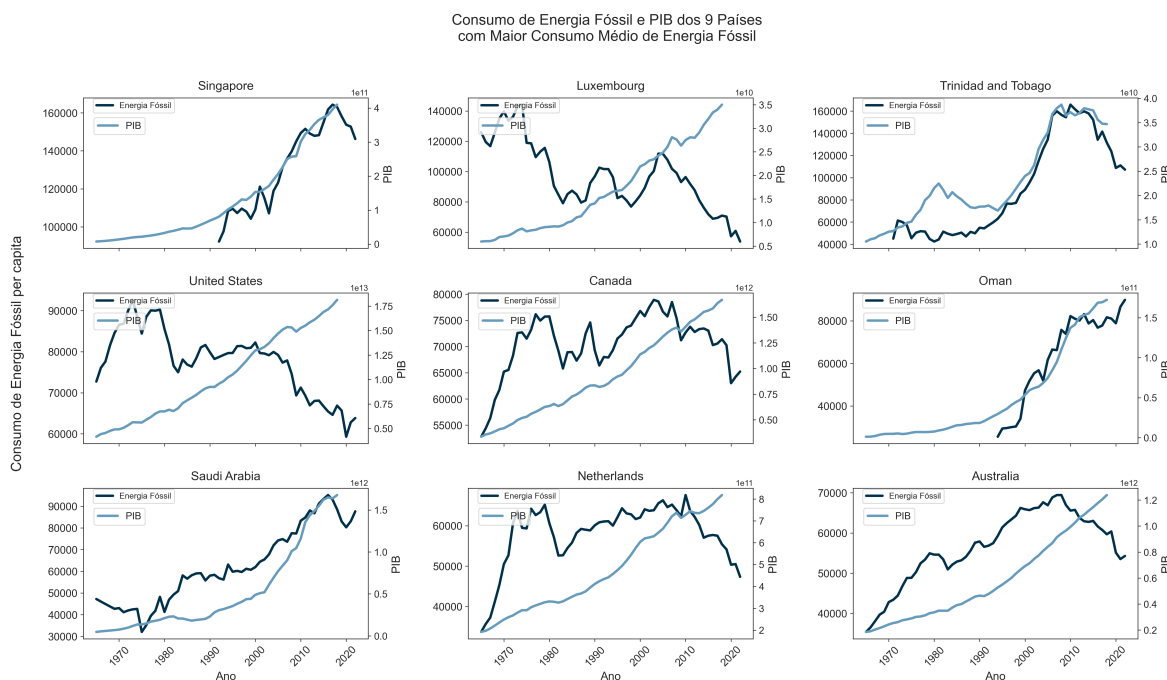


Figure 15: Comparação entre PIB e consumo de Combustível fóssil dos países com maior consumo médio ao longo do tempo.

Vemos uma certa relação entre as variáveis mais acentuada entre os países Cingapura, Trinidad e Tobago, Omã e Arábia Saudita, enquanto países como Estados Unidos e Luxemburgo apresentam comportamento muito diferente. A partir disso podemos realizar análises mais profundas buscando por dados específicos dos países listados para compreender melhor os dados observados. Foi analisado que entre os países que representaram uma forte correlação entre os dados temos alguns fatores relevantes em comum:

- A Arábia Saudita e Omã são tidas como grandes exportadoras de petróleo e gás e o setor energético tem grande influência no PIB desses países. No caso de Trinidad e Tobago temos que a economia depende principalmente de seus recursos naturais, já que o país fica nas maiores reservas de petróleo e gás natural do Caribe. O consumo interno está diretamente ligado à extração e processamento de petróleo, refletindo uma forte correlação com o crescimento econômico. No caso de Singapura, apesar de não ser produtor de petróleo é um importante centro de refinamento e comércio do petróleo. Por ter uma população pequena e alta densidade o transporte e a indústria intensiva contribuem para o PIB.

- Apesar de ser um dos maiores produtores de petróleo e gás natural, a economia dos EUA é altamente diversificada com serviços, tecnologia e manufatura desempenhando papéis mais significativos no PIB, logo, as mudanças nas variáveis não afetam diretamente uma na outra. Além disso o país tem investido em energia renovável, eficiência energética e redução de emissões, permitindo crescimento econômico com menor dependência de combustíveis fósseis, o que pode explicar uma correlação mais fraca ou até negativa entre PIB e consumo de combustíveis fósseis. No caso de Singapura, o país tem sua economia voltada para serviços financeiros e tecnologia, com uma população pequena e alta eficiência energética, investindo significativamente em fontes renováveis e estratégias de descarbonização. Isso resulta em um consumo de combustíveis fósseis relativamente baixo. Como o setor energético desempenha um papel secundário na economia, há pouca ou nenhuma correlação positiva entre consumo de combustíveis fósseis e crescimento econômico.

3.4.2 Análise da relação de crescimento e decrescimento do PIB em comparação com as variações no consumo de combustível fóssil dos países com Menor consumo:

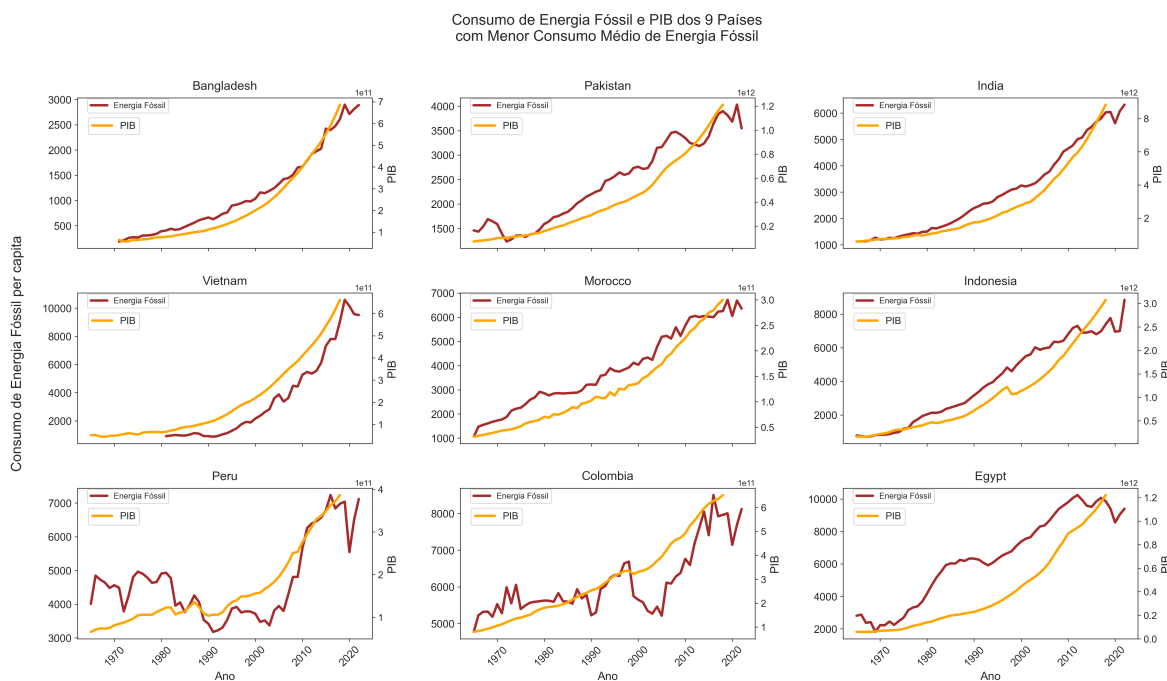


Figure 16: Comparação entre PIB e consumo de Combustível fóssil dos países com menor consumo médio ao longo do tempo.

Vemos que entre esses países parece haver uma maior correlação entre os dois aspectos analisados. Para ter uma visão melhor das informações observadas podemos buscar uma contextualização.

A relação geral entre PIB e consumo de energia fóssil nesses países está diretamente ligada ao grau de industrialização, transição energética e dependência de combustíveis fósseis.

Países como Índia, Indonésia, e Vietnã mostram uma correlação positiva mais forte entre PIB e consumo de energia fóssil, refletindo economias em crescimento e maior dependência de combustíveis fósseis.

Países como Peru e Colômbia têm uma relação um pouco menos linear, indicando economias mais diversificadas e o uso de fontes alternativas de energia.

O crescimento econômico de Bangladesh está ligado ao aumento da atividade industrial e ao desenvolvimento de infraestrutura, que dependem de combustíveis fósseis, assim como o Paquistão, que tem uma economia em desenvolvimento, com forte dependência de combustíveis fósseis para geração de energia elétrica e transporte. Isso reflete a correlação positiva entre PIB e consumo de energia.

3.4.3 Análise de Correlação

É possível ver que entre os países que consomem em média mais combustível fóssil não parece haver correspondência em termos de crescimento e decrescimento de ambas as variáveis, diferentemente do que acontece entre os países que consomem menos energia fóssil em que parece haver uma correlação positiva entre as variáveis. Podemos então analisar estatisticamente o grau de correlação que é demonstrado entre ambas as variáveis para cada país.

Crítérios de Correlação:

Para analisar a correlação positiva entre os dados, foi definido:

- $> 0,7$: correlação alta
- $0,3 < \text{correlação} < 0,7$: moderada
- $< 0,3$: correlação baixa

O gráfico a seguir mostra a contagem dos países de acordo com a sua respectiva classificação com base no cálculo da correlação entre as variáveis PIB e consumo de energia fóssil:

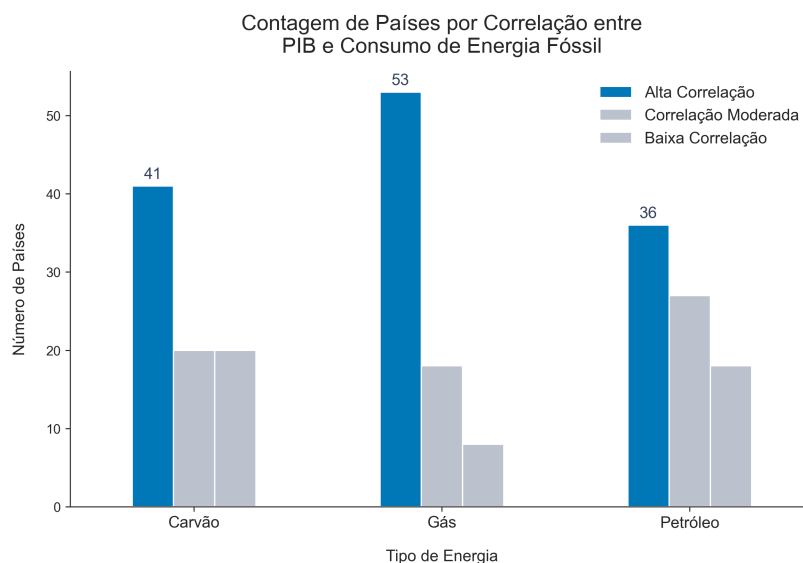


Figure 17: Contagem de países com base em sua classificação de correlação entre PIB e consumo de Combustível Fóssil.

O gráfico indica que a maior parte dos países apresentam uma correlação alta entre as variáveis. É importante ressaltar que dado o contexto, altos índices de correlação não indicam causalidade dado que ambas as variáveis são influenciadas por fatores externos. A comparação visa encontrar padrões de consumo de energia entre países com alto e baixo PIB. Podemos ver que pode existir uma certa tendência a consumir energia proveniente de fontes fósseis quando há o aumento do PIB (crescimento econômico). Apesar dessa relação variar de acordo com cada país, é possível inferir que um avanço econômico muitas vezes pode acarretar em um aumento da demanda de energia, elevando o consumo de combustíveis fósseis que ainda são a principal fonte de energia global. Isso vem do fato de que conforme um país cresce, as indústrias passam a produzir mais e a população a consumir mais, podendo aumentar a demanda de energia. Mas vemos que isso não é regra, ao menos quando se trata de energia fóssil. Apesar de muitos países apresentarem essa tendência, vemos que em muitos casos o efeito é oposto e a correlação é negativa, o que parece acontecer com Luxemburgo e Estados Unidos, por exemplo (Fig.15). Isso se deve ao fato de que cada país tem suas particularidades e um desenvolvimento diferente, tanto do ponto de vista econômico quanto de matriz energética.

O aumento do consumo pode ser causado por diversos fatores, principalmente conforme há o desenvolvimento dos países, como o crescimento de indústrias, utilização de veículos de transporte, infraestrutura, padrões de consumo, etc. Sendo assim essas variáveis podem sim estar (em alguns casos) relacionadas de certa forma, mas sabemos também que conforme os países passam a adotar outras estratégias de consumo e eficiência energética e com os atuais avanços tecnológicos, o consumo de energia proveniente de fontes fósseis passe a diminuir. Através disso podemos compreender os resultados obtidos e concluir que apesar dessa relação ser válida para diversos países, existem diversos fatores que podem afetar essa

relação, logo isso nem sempre será verdade.

4 Conclusão

Relação entre população, consumo de energia, balanço de energia e importação líquida

Verificou-se que a demanda de energia elétrica é crescente e fortemente correlacionada ao crescimento da população.

Os países têm como desafio suprir a demanda crescente de energia de sua população. Quando sua produção interna não consegue atender a essa demanda, há necessidade das importações de energia.

Distribuição de consumo por continente

Conclui-se que a Ásia e a Europa lideram o consumo de energia renovável de maneira bem distribuída (considerando a tendência per capita), enquanto os demais continentes apresentam um crescimento mais lento e instável, inclusive com certo grau de dependência aparente nas hidrelétricas como única fonte de energia limpa.

Demanda de energia e produção de energia renovável

A produção de energia renovável não só tem acompanhado o aumento da demanda global de energia, como indicado pela forte correlação (0,97), bem como está crescendo em um ritmo mais rápido. No entanto, mesmo em 2022, essa proporção não chega a 40 por cento. Esse aumento da proporção ao longo do tempo é promissor, caso continue no mesmo ritmo, mas ainda há um longo caminho para que fontes renováveis assumam um papel dominante na matriz energética global.

PIB e Consumo de Combustível Fóssil

Com base nos resultados, podemos inferir que o crescimento econômico (aumento do PIB) tem uma certa relação com o aumento do consumo de energia fóssil. No entanto, mais estudos são necessários para entender as particularidades dessa relação e o impacto de fontes de energia renovável e outros fatores externos.

5 Referências

O conjunto de dados utilizado nesta análise está disponível na plataforma Kaggle, acessível pelo seguinte link: Kaggle - World Energy Consumption.

Os dados e visualizações foram produzidos pela Our World in Data, com coleta, agregação e documentação realizadas por Hannah Ritchie, Pablo Rosado, Edouard Mathieu e Max Roser. O conjunto de dados foi elaborado com base nas seguintes fontes principais:

5.1 Fontes de dados e etapas de processamento

Statistical Review of World Energy: Energy Institute (EI);

International Energy Data: U.S. Energy Information Administration (EIA);

Energy from Fossil Fuels: The Shift Dataportal;

Yearly Electricity Data: Ember;

European Electricity Review: Ember;

Combined Electricity: Our World in Data com base nos dados da Ember;

Energy Mix: Our World in Data com base nos dados da EI;

Fossil Fuel Production: Our World in Data com base nos dados da EI e The Shift Dataportal;

Primary Energy Consumption: Our World in Data com base nos dados da EI e EIA;

Electricity Mix: Our World in Data com base nos dados da EI e Ember;

Energy Dataset: Our World in Data com base em todas as fontes acima

5.2 Adaptação para variáveis regionais e per capita

Além das fontes mencionadas, variáveis regionais, per capita e ajustadas por PIB foram construídas utilizando:

Regions: Our World in Data;

Population: Our World in Data, baseado em diversas fontes;

Income Groups: World Bank;

GDP: University of Groningen GGDC's Maddison Project Database (Bolt e van Zanden, 2020)

5.3 Descrição das variáveis

A descrição detalhada das colunas, incluindo unidade de medida e fonte de cada variável, está documentada no repositório do GitHub da Our World in Data: [GitHub - OWID Energy Codebook](#).