



**UNIVERSIDADE CATÓLICA DE SANTOS**

**CIÊNCIA DA COMPUTAÇÃO**

**Pesquisa Curricularização de Graduação**  
**Análise de correlação entre dois Data Sets**  
**Análise da Transparência nos Gastos de Energia**

*Felipe Barbosa dos Santos - 4324223*  
*João Gabriel Henriques Cardoso - 6100450*  
*Lucas Carmona Neto - 8342055*  
*Lucas Cerqueira Galvão - 4477944*  
*Pedro Henrique Bonifacio Martins - 2502334*

Santos  
2024

## Sumário

<b>1. Resumo.....</b>	<b>2</b>
<b>2. Introdução.....</b>	<b>3</b>
<b>3. Problematização.....</b>	<b>4</b>
3.1. Problema da Pesquisa.....	4
3.2. Justificativa da pesquisa.....	4
<b>4. Objetivos.....</b>	<b>5</b>
4.1. Objetivo Geral.....	5
4.2. Objetivos específicos e questões levantadas.....	5
<b>5. Metodologia.....</b>	<b>6</b>
5.1. Metodologia geral.....	6
5.2. Metodologias específicas.....	7
5.3. Descrição das Colunas do Dataset.....	11
5.4. Códigos utilizados.....	18
<b>6. Cronograma Final.....</b>	<b>21</b>
6.1. Justificativa do Cronograma.....	21
<b>7. Referencial Teórico.....</b>	<b>23</b>
7.1. Energia Renovável.....	23
7.2. Energia Não Renovável.....	24
7.3. Teorias e Métodos de Análise de Dados.....	25
7.4. Casos de Uso de Análise de Dados no Setor Energético.....	25
7.5. Impacto de Análise de Dados na Política e Gestão Energética.....	26
<b>8. Resultados e Discussões.....</b>	<b>27</b>
8.1. Coleta e Tratamento de Dados.....	27
8.2. Análise Exploratória de Dados.....	27
8.3. Desenvolvimento de Gráficos e Visualizações.....	27
8.4. Comparação entre Diferentes Aspectos.....	29
8.5. Predição de Dados.....	29
8.6. Uso de Conjuntos de Dados Diferentes.....	30
8.7. Proposição de Medidas para Aprimoramento.....	30
<b>9. Conclusão.....</b>	<b>33</b>
<b>10. Referências.....</b>	<b>34</b>

## **1. Resumo**

Este estudo investiga a transparência dos dados de gastos de energia elétrica através de uma análise exploratória de conjuntos de dados abertos disponíveis em portais de transparência. Reconhecendo os desafios enfrentados pela gestão energética contemporânea — tais como a falta de padronização, dificuldades de acesso e complexidades na interpretação dos dados —, esta pesquisa propõe superar tais barreiras com o uso de tecnologias e técnicas avançadas de análise de dados. Utilizando a biblioteca Python Pandas, o projeto envolve a coleta rigorosa de dados, seu tratamento e análise exploratória para revelar padrões, tendências e correlações significativas que possam informar e influenciar políticas públicas e práticas sustentáveis. O trabalho começa com uma coleta de dados focada, seguida por uma fase de tratamento para garantir a comparabilidade e integridade dos dados. As etapas de análise exploratória são acompanhadas pelo desenvolvimento de visualizações gráficas que facilitam a interpretação dos dados e destacam as variações nos gastos de energia ao longo do tempo e entre diferentes regiões. O estudo se propõe não apenas a analisar os dados, mas também a oferecer recomendações para melhorar a transparência e a acessibilidade dos dados.

## **2. Introdução**

A gestão de energia elétrica é um pilar central para a sustentabilidade ambiental e eficiência econômica global. Em um cenário de crescente demanda energética e urgência climática, a transparência nos dados de consumo e gastos de energia torna-se crucial para o desenvolvimento de políticas energéticas eficazes e sustentáveis. Este trabalho foca na análise exploratória de dados abertos disponíveis em portais de transparência, especificamente aqueles relacionados aos gastos de energia elétrica, com o objetivo de melhorar a acessibilidade e compreensão desses dados para promover uma gestão energética mais eficiente e participativa e desenvolver as habilidades relacionadas a esse tipo de análise.

A transparência nos dados enfrenta desafios significativos, que vão desde a falta de padronização e acessibilidade até a complexidade na interpretação das informações disponíveis. Essas barreiras impedem não apenas a análise eficaz, mas também limitam a capacidade dos cidadãos e dos formuladores de políticas de tomar decisões informadas. Portanto, este estudo utiliza métodos avançados de coleta, tratamento e análise de dados para superar esses obstáculos. Através do uso da biblioteca Python Pandas, o projeto implementa uma abordagem rigorosa que permite a exploração detalhada dos padrões de consumo e gastos de energia, facilitando a identificação de tendências e a comparação entre diferentes regiões e períodos.

Além disso, a análise dos dados de gastos de energia elétrica oferece insights valiosos que podem orientar políticas públicas e promover práticas sustentáveis. Assim, ao correlacionar os conjuntos de dados de gastos de energia com princípios de dados abertos e transparência, este estudo destaca como o desenvolvimento de habilidades de análise refinadas trazem insights poderosos para a tomada de decisão.

### **3. Problematização**

No contexto da gestão de energia, a transparência dos dados é fundamental para garantir a eficiência, a sustentabilidade e a participação cidadã. Entretanto, mesmo com os dados disponíveis nos portais de transparência, enfrentamos desafios significativos. Desde a falta de padronização e acessibilidade dos dados até a complexidade na interpretação das informações, surge a seguinte questão: como assegurar uma transparência efetiva dos dados sobre gastos de energia por meio da análise de conjuntos de dados disponíveis nos portais de transparência?

#### **3.1. Problema da Pesquisa**

Ao lidar com conjuntos de dados sobre gastos de energia disponibilizados nos portais de transparência, surge a necessidade de compreender os desafios específicos que afetam a transparência e a efetividade desses dados. Como garantir que os dados relacionados aos gastos de energia sejam transparentes e acessíveis, considerando as nuances técnicas e de interpretação, para promover uma gestão mais eficiente e participativa nesse setor?

#### **3.2. Justificativa da pesquisa**

A investigação sobre a transparência dos dados referentes aos gastos de energia e a importância dos dados abertos nesse contexto é crucial diante dos desafios enfrentados na gestão energética contemporânea. Compreender e superar as barreiras que dificultam a transparência e a acessibilidade dos dados nesse domínio é essencial para promover a eficiência energética, a sustentabilidade e a participação cidadã. Além disso, a análise desses dados pode revelar insights valiosos para orientar decisões políticas importantes, incentivar práticas sustentáveis e empoderar os consumidores na tomada de decisões informadas. Assim, a pesquisa se justifica como um meio para fomentar o debate e propor soluções que ampliem o acesso, a compreensão e a utilidade dos dados de gastos de energia, fortalecendo, portanto, a governança e a responsabilidade no setor energético.

## 4. Objetivos

### 4.1. Objetivo Geral

Realizar uma análise abrangente e exploratória que correlacione conjuntos de dados abertos provenientes de portais de transparência, especificamente focados nos gastos de energia. Através dessa análise, buscamos compreender como a transparência dos dados relacionados ao consumo e gastos de energia pode ser aprimorada e efetivada. Nosso intuito é identificar padrões, tendências e insights relevantes que possam contribuir para uma maior compreensão e utilização dos dados públicos disponibilizados nos portais de transparência de energia. Com isso, pretendemos fornecer insights que podem ser usados para decisões políticas importantes, incentivando a eficiência energética, a sustentabilidade e a participação cidadã na gestão dos recursos energéticos.

### 4.2. Objetivos específicos e questões levantadas

- **Coleta e Tratamento de Dados:** Realizar a coleta dos conjuntos de dados de gastos de energia disponíveis nos portais de transparência e efetuar o tratamento inicial utilizando a biblioteca Pandas do Python.
- **Análise Exploratória de Dados:** Explorar e analisar os dados coletados utilizando a biblioteca Pandas, identificando padrões, tendências e anomalias nos gastos de energia.
- **Desenvolvimento de Gráficos e Visualizações:** Desenvolver gráficos e visualizações informativas dos dados utilizando recursos oferecidos por ferramentas como o PowerBI.
- **Comparação entre Diferentes Aspectos:** Realizar comparações entre diferentes períodos de tempo, regiões geográficas ou setores de consumo de energia, empregando técnicas de análise comparativa disponíveis no Pandas.
- **Predição de Dados:** Aplicar diferentes métodos de previsão, como treinamento procedimental e regressão linear, para prever tendências e mudanças futuras nos dados de energia.
- **Uso de Conjuntos de Dados Diferentes:** Usar conjuntos de dados de energia renovável e de energia não renovável para análises e comparações.
- **Proposição de Medidas para Aprimoramento:** Propor medidas e recomendações para aprimorar a transparência e acessibilidade dos dados de gastos de energia nos portais de transparência, com base nas análises realizadas.

## **5. Metodologia**

### **5.1. Metodologia geral**

Iniciaremos o processo de coleta de dados acessando o site [www.kaggle.com](https://www.kaggle.com). Neste site, selecionamos datasets que compilam várias informações originadas de diversos sites de transparência. Essas informações são descritas e vinculadas diretamente na publicação dos dados no Kaggle, permitindo uma verificação detalhada de sua confiabilidade. Os arquivos foram baixados em formato CSV e, posteriormente, convertidos para Excel para facilitar as visualizações iniciais.

Escolhemos o Kaggle devido à sua capacidade de oferecer uma ampla gama de datasets consolidados, o que é essencial para uma análise abrangente e precisa dos gastos energéticos. Após a obtenção dos dados, empregamos a biblioteca Pandas do Python para realizar o tratamento inicial. Este tratamento incluirá a organização dos dados, seleção de colunas e linhas relevantes, remoção de entradas desnecessárias e padronização das unidades de medidas.

Com os dados devidamente preparados, procederemos à análise exploratória. Esta etapa envolverá o desenvolvimento de visualizações como histogramas e gráficos de dispersão, que nos ajudarão a visualizar melhor a distribuição dos dados e identificar padrões. Além disso, realizaremos cálculos estatísticos, incluindo média, mediana e desvio padrão, para obter um entendimento mais profundo das tendências nos dados.

As análises e visualizações geradas nos permitirão realizar comparações efetivas, prever tendências futuras e propor medidas de aprimoramento. Estas propostas se basearão nas insights obtidos e visarão melhorar a transparência dos dados e a eficiência energética através de recomendações políticas e práticas.

## 5.2. Metodologias específicas

### 5.2.1. Coleta e Tratamento de Dados

- **Fontes de Dados:** Iniciaremos a coleta de dados acessando o site [www.kaggle.com](http://www.kaggle.com), onde encontramos datasets compilados de várias informações originadas de diversos sites de transparência. Esses datasets incluem dados detalhados sobre gastos de energia por região e setor, disponibilizados em formatos que facilitam o acesso e a análise.
- **Justificativa da Escolha:** optamos por este site devido à sua capacidade de consolidar informações de múltiplas fontes de transparência, cuja confiabilidade foi verificada através dos links e descrições disponíveis na publicação dos dados no Kaggle. Isso nos permite ter uma base de dados abrangente e detalhada, essencial para uma análise precisa dos gastos energéticos.
- **Formato dos Dados:** os dados foram inicialmente baixados em formato CSV, o qual é preferido por sua simplicidade e facilidade de manipulação em análises subsequentes. Para visualizações iniciais e manipulações mais simples, os dados foram convertidos para o formato de planilha (.xlsx).
- **Preparação Inicial dos Dados:** após a obtenção dos dados, utilizaremos a biblioteca Pandas do Python para realizar a limpeza e preparação inicial. Este processo incluirá a seleção de colunas e linhas que são mais relevantes para nossos objetivos de estudo, a remoção de entradas que não contribuem para a análise e a padronização das unidades de medidas. Essas ações são fundamentais para garantir que os dados estejam em condições ótimas para análises mais profundas e complexas.
- **Segurança da Informação:** no dataset escolhido, não temos informações sensíveis ou sigilosas, garantindo a privacidade e a segurança dos dados utilizados na pesquisa. Para reforçar a segurança da informação, seguimos as diretrizes da norma ISO 27001, que estabelece os requisitos para um sistema de gestão da segurança da informação (SGSI). Isso assegura que todas as medidas de segurança necessárias são implementadas para proteger a integridade e a confidencialidade dos dados.

### 5.2.2. Análise Exploratória de Dados:

- **Técnicas Estatísticas:** A análise começará com uma descrição estatística básica para entender as características centrais dos dados, seguida de testes de correlação para investigar as relações entre diferentes variáveis e análise de variância (ANOVA) para comparar grupos de dados, especialmente variações de consumo e gastos por fontes energéticas.



- **Ferramentas de Visualização:** Para visualizar os resultados das análises, utilizaremos Matplotlib e Seaborn para criar gráficos detalhados, como gráficos de linha para tendências ao longo do tempo e histogramas e box plots para distribuições de gastos. Além disso, o PowerBI será empregado para desenvolver dashboards interativos, facilitando a visualização de padrões complexos e análises geográficas através de mapas de calor.
- **Estrutura da Análise:**
  - **Análise de Tendências:** Utilizar gráficos de linha para examinar como os gastos de energia evoluem ao longo do tempo, identificando tendências significativas.
  - **Distribuições de Gastos:** Explorar as distribuições de gastos de energia utilizando histogramas e box plots para entender a variabilidade e identificar outliers.
  - **Análise Geográfica:** Empregar mapas de calor para visualizar geograficamente os dados, permitindo uma compreensão mais profunda de como os gastos de energia variam por região.
- **Exploração de Dados com Pandas:** Usaremos a biblioteca Pandas para realizar essas análises, aproveitando sua capacidade de manejar grandes volumes de dados e executar complexas manipulações de dados, que são essenciais para detectar padrões, tendências e anomalias nos dados energéticos.

### 5.2.3. Desenvolvimento de Gráficos e Visualizações:

- **Implementação no PowerBI:** Utilizaremos o PowerBI para desenvolver dashboards interativos e avançados, que permitirão aos usuários filtrar os dados visualizados por período, região e tipo de energia. Isso facilitará análises específicas e permitirá que diferentes stakeholders explorem os dados conforme suas necessidades.
- **Justificativa da Escolha:** A escolha do PowerBI como ferramenta de visualização deve-se à sua capacidade de criar interfaces interativas e visualmente atraentes que podem ser facilmente manipuladas e compartilhadas. Além disso, a ferramenta suporta uma ampla gama de visualizações, tornando-a ideal para apresentar complexidades de dados de forma compreensível.
- **Gráficos e Visualizações:**
  - **Visualização de Tendências:** Desenvolveremos gráficos de linha no PowerBI para mostrar as tendências de consumo e gastos com energia ao longo do tempo, identificando padrões sazonais ou anomalias.

- **Comparação de Dados:** Utilizar gráficos de barras para comparar o consumo de energia entre diferentes regiões ou categorias, como fontes renováveis e não renováveis, destacando disparidades e pontos de interesse.

#### 5.2.4. Comparação entre Diferentes Aspectos:

- **Métodos Comparativos:** Para realizar comparações eficazes, empregaremos gráficos de barra e box plots, que são excelentes para visualizar diferenças entre grupos. Essas visualizações serão usadas para comparar o consumo de energia entre diferentes regiões, períodos e setores industriais.
- **Processo de Análise:**
  - **Seleção de Variáveis Relevantes:** Identificar e selecionar as variáveis mais pertinentes que influenciam os gastos e consumo de energia. Isso inclui fatores como tipo de energia, volume de consumo, custos associados e dados demográficos regionais.
  - **Testes Estatísticos:** Aplicar testes estatísticos adequados, como a análise de variância (ANOVA), para determinar se as diferenças observadas entre as variáveis selecionadas são estatisticamente significativas. Esses testes ajudarão a confirmar padrões e discrepâncias identificados nas comparações visuais.
- **Comparação Regional e Temporal:** Utilizar a biblioteca Pandas para analisar como os gastos de energia variam entre diferentes regiões e ao longo de diferentes períodos. Essa análise ajudará a identificar tendências regionais e temporais, assim como discrepâncias significativas que podem indicar áreas de preocupação ou oportunidades para intervenção.
- **Análise Setorial:** A análise também abrangerá diferentes setores industriais, comparando o consumo energético para avaliar a eficiência energética e identificar setores com maior potencial de melhoria. Isso não só informa sobre a eficiência atual, mas também sugere áreas onde políticas ou tecnologias de economia de energia podem ser aplicadas para reduzir custos e impacto ambiental.

#### 5.2.5. Predição de Dados:

- **Modelos Utilizados:** Para a modelagem preditiva, empregaremos uma combinação de técnicas estatísticas avançadas adaptadas às características dos dados de energia. Utilizaremos a regressão linear para identificar tendências lineares, árvores de decisão para modelar decisões baseadas em condições complexas, e modelos ARIMA para análise de séries temporais, permitindo prever tendências futuras com base em dados históricos.
- **Implementação:**

- **Códigos em Python:** A implementação dos modelos será realizada usando Python, com a utilização de bibliotecas especializadas como Scikit-learn para regressão e árvores de decisão, e Statsmodels para a implementação de modelos ARIMA. Essas ferramentas fornecem funcionalidades robustas para modelagem estatística e são ideais para lidar com a complexidade dos dados de energia.
- **Validação dos Modelos:**
  - **Divisão dos Dados:** Os dados serão divididos em conjuntos de treinamento e teste para garantir que os modelos sejam avaliados de maneira justa e eficaz. Esta abordagem permite validar a performance do modelo em dados não vistos durante o treinamento.
  - **Métricas de Avaliação:** Utilizaremos o RMSE (Root Mean Square Error) para avaliar a precisão dos modelos. Esta métrica é crucial para entender o quão próximas as previsões do modelo estão dos valores reais, ajudando a identificar a melhor abordagem e ajustar os parâmetros conforme necessário.
- **Avaliação de Modelos:** A precisão dos modelos preditivos será rigorosamente testada para garantir sua confiabilidade. Os resultados ajudarão a refinar e ajustar os modelos, melhorando assim sua precisão e utilidade para prever gastos de energia futuros.

#### 5.2.6. Uso de Conjuntos de Dados Diferentes:

- **Dados Renováveis vs. Não Renováveis:** Analisar conjuntos de dados separados de fontes de energia renováveis e não renováveis para entender o impacto ambiental e a sustentabilidade.

#### 5.2.7. Proposição de Medidas para Aprimoramento:

- **Base para Recomendações:** As recomendações serão fundamentadas nas análises e insights gerados nas etapas anteriores do estudo. Utilizaremos os dados coletados, as análises comparativas, e os modelos preditivos para identificar áreas críticas onde melhorias podem ser implementadas, tanto em termos de práticas de consumo quanto em políticas públicas.
- **Propostas Baseadas em Análise:**
  - **Melhoria de Transparência e Acessibilidade:** Propor medidas específicas para aprimorar a transparência e acessibilidade dos dados energéticos disponibilizados nos portais de transparência. Isso pode incluir recomendações para a padronização dos formatos de dados e a implementação de interfaces mais intuitivas para os usuários.

- Eficiência Energética: Sugerir iniciativas e práticas para incentivar o uso mais eficiente de energia, como programas de educação para consumidores, incentivos para tecnologias de baixo consumo e a implementação de padrões mais rigorosos para aparelhos e máquinas.
- **Recomendações Políticas:**
  - **Incentivo a Fontes Renováveis:** Formular políticas públicas que promovam a adoção e o desenvolvimento de fontes de energia renováveis, como solar e eólica, incluindo subsídios, incentivos fiscais e apoio a pesquisas e desenvolvimento.
  - **Regulações e Legislações:** Sugerir alterações ou novas legislações que visem melhorar a eficiência energética geral e reduzir o impacto ambiental, focando em setores industriais de alto consumo e em áreas urbanas.
- **Estrutura do Relatório:**
  - **Formato:** O relatório será estruturado conforme as normas da ABNT, incluindo uma introdução que contextualiza o problema e os objetivos, seguido pelo desenvolvimento dos insights, as recomendações práticas detalhadas e uma conclusão que ressalta os benefícios esperados das recomendações propostas.

### 5.3. Descrição das Colunas do Dataset

Utilizamos um dataset detalhado de consumo de energia mundial. Abaixo, descrevemos brevemente cada coluna presente no dataset, junto com seu tipo de dado e unidade de medida, quando aplicável.

- **country:** País ou região geográfica. (Tipo: String)
- **year:** Ano da observação. (Tipo: Int)
- **iso\_code:** Código ISO - Códigos de três letras conforme a ISO 3166-1 alpha-3. (Tipo: String)
- **population:** População - População por país, disponível de 10.000 AEC a 2100, baseado em dados e estimativas de diferentes fontes. (Unidade: pessoas; Tipo: Float)
- **gdp:** Produto Interno Bruto (PIB) - Dados ajustados pela inflação e diferenças no custo de vida entre países. (Unidade: dólares internacionais de 2011; Tipo: Float)
- **biofuel\_cons\_change\_pct:** Mudança anual percentual no consumo de biocombustíveis - Inclui biogás e biodiesel, ajustados para o conteúdo energético. (Tipo: Float)
- **biofuel\_cons\_change\_twh:** Mudança anual no consumo de biocombustíveis - Inclui biogás e biodiesel, ajustados para o conteúdo energético. (Unidade: terawatt-horas; Tipo: Float)

- **biofuel\_cons\_per\_capita**: Consumo de biocombustíveis per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **biofuel\_consumption**: Consumo de energia primária de biocombustíveis - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **biofuel\_elec\_per\_capita**: Geração de eletricidade a partir de bioenergia per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **biofuel\_electricity**: Geração de eletricidade a partir de bioenergia - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **biofuel\_share\_elec**: Participação da bioenergia na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **biofuel\_share\_energy**: Participação do consumo de energia primária que vem de biocombustíveis - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **carbon\_intensity\_elec**: Intensidade de carbono na geração de eletricidade - Gases de efeito estufa emitidos por unidade de eletricidade gerada, medidos em gramas de CO2 equivalente por kilowatt-hora. (Unidade: gramas de CO2 equivalente por kilowatt-hora; Tipo: Float)
- **coal\_cons\_change\_pct**: Mudança anual percentual no consumo de carvão. (Tipo: Float)
- **coal\_cons\_change\_twh**: Mudança anual no consumo de carvão. (Unidade: terawatt-horas; Tipo: Float)
- **coal\_cons\_per\_capita**: Consumo de carvão per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **coal\_consumption**: Consumo de energia primária de carvão - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **coal\_elec\_per\_capita**: Geração de eletricidade a partir de carvão per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **coal\_electricity**: Geração de eletricidade a partir de carvão - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **coal\_prod\_change\_pct**: Mudança anual na produção de carvão - Medido como uma porcentagem da produção do ano anterior. (Tipo: Float)
- **coal\_prod\_change\_twh**: Mudança anual na produção de carvão - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **coal\_prod\_per\_capita**: Produção de carvão per capita - Medido em kilowatt-horas por capita. (Unidade: kilowatt-horas; Tipo: Float)
- **coal\_production**: Produção de carvão - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **coal\_share\_elec**: Participação do carvão na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **coal\_share\_energy**: Participação do consumo de energia primária que vem do carvão - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **electricity\_demand**: Demanda de eletricidade - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)

- **electricity\_generation**: Geração total de eletricidade - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **electricity\_share\_energy**: Geração total de eletricidade como participação da energia primária - Medido como percentagem do consumo total direto de energia primária. (Tipo: Float)
- **energy\_cons\_change\_pct**: Mudança anual no consumo de energia primária. (Tipo: Float)
- **energy\_cons\_change\_twh**: Mudança anual no consumo de energia primária. (Unidade: terawatt-horas; Tipo: Float)
- **energy\_per\_capita**: Consumo de energia primária per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas por pessoa; Tipo: Float)
- **energy\_per\_gdp**: Consumo de energia primária por PIB - Medido em kilowatt-horas por dólar internacional. (Unidade: kilowatt-horas por \$; Tipo: Float)
- **fossil\_cons\_change\_pct**: Mudança anual percentual no consumo de combustíveis fósseis. (Tipo: Float)
- **fossil\_cons\_change\_twh**: Mudança anual no consumo de combustíveis fósseis. (Unidade: terawatt-horas; Tipo: Float)
- **fossil\_elec\_per\_capita**: Geração de eletricidade a partir de combustíveis fósseis per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **fossil\_electricity**: Geração de eletricidade a partir de combustíveis fósseis - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **fossil\_energy\_per\_capita**: Consumo de combustíveis fósseis per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **fossil\_fuel\_consumption**: Consumo de energia primária de combustíveis fósseis - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **fossil\_share\_elec**: Participação dos combustíveis fósseis na geração de eletricidade - Medido como percentagem do total de eletricidade. (Tipo: Float)
- **fossil\_share\_energy**: Participação do consumo de energia primária que vem de combustíveis fósseis - Medido como percentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **gas\_cons\_change\_pct**: Mudança anual percentual no consumo de gás natural. (Tipo: Float)
- **gas\_cons\_change\_twh**: Mudança anual no consumo de gás natural. (Unidade: terawatt-horas; Tipo: Float)
- **gas\_consumption**: Consumo de energia primária de gás natural - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **gas\_elec\_per\_capita**: Geração de eletricidade a partir de gás natural per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **gas\_electricity**: Geração de eletricidade a partir de gás natural - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **gas\_energy\_per\_capita**: Consumo de gás natural per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **gas\_prod\_change\_pct**: Mudança anual na produção de gás natural - Medido como uma percentagem da produção do ano anterior. (Tipo: Float)

- **gas\_prod\_change\_twh**: Mudança anual na produção de gás natural - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **gas\_prod\_per\_capita**: Produção de gás natural per capita - Medido em kilowatt-horas por capita. (Unidade: kilowatt-horas; Tipo: Float)
- **gas\_production**: Produção de gás natural - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **gas\_share\_elec**: Participação do gás natural na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **gas\_share\_energy**: Participação do consumo de energia primária que vem do gás natural - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **greenhouse\_gas\_emissions**: Emissões de gases de efeito estufa pela geração de eletricidade - Medido em megatoneladas de CO<sub>2</sub> equivalente. (Unidade: megatoneladas de CO<sub>2</sub> equivalente; Tipo: Float)
- **hydro\_cons\_change\_pct**: Mudança anual percentual no consumo de hidreletricidade. (Tipo: Float)
- **hydro\_cons\_change\_twh**: Mudança anual no consumo de hidreletricidade - A energia equivalente de entrada é baseada na geração bruta e não contabiliza o fornecimento de eletricidade transfronteiriço. (Unidade: terawatt-horas; Tipo: Float)
- **hydro\_consumption**: Consumo de energia primária de hidreletricidade - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)
- **hydro\_elec\_per\_capita**: Geração de eletricidade a partir de hidreletricidade per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **hydro\_electricity**: Geração de eletricidade a partir de hidreletricidade - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **hydro\_energy\_per\_capita**: Consumo de hidreletricidade per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **hydro\_share\_elec**: Participação da hidreletricidade na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **hydro\_share\_energy**: Participação do consumo de energia primária que vem da hidreletricidade - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **low\_carbon\_cons\_change\_pct**: Mudança anual percentual no consumo de energia de baixo carbono. (Tipo: Float)
- **low\_carbon\_cons\_change\_twh**: Mudança anual no consumo de energia de baixo carbono - A energia equivalente de entrada é baseada na geração bruta e não contabiliza o fornecimento de eletricidade transfronteiriço. (Unidade: terawatt-horas; Tipo: Float)
- **low\_carbon\_consumption**: Consumo de energia primária de fontes de baixo carbono - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)

- **low\_carbon\_elec\_per\_capita:** Geração de eletricidade a partir de fontes de baixo carbono per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **low\_carbon\_electricity:** Geração de eletricidade a partir de fontes de baixo carbono - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **low\_carbon\_energy\_per\_capita:** Consumo de energia de baixo carbono per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **low\_carbon\_share\_elec:** Participação das fontes de baixo carbono na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **low\_carbon\_share\_energy:** Participação do consumo de energia primária que vem de fontes de baixo carbono - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **net\_elec\_imports:** Importações líquidas de eletricidade - Importações menos exportações de eletricidade, medido em TWh. (Unidade: terawatt-horas; Tipo: Float)
- **net\_elec\_imports\_share\_demand:** Importações líquidas de eletricidade como participação da demanda - Importações menos exportações de eletricidade, medida como porcentagem da demanda total de eletricidade. (Tipo: Float)
- **nuclear\_cons\_change\_pct:** Mudança anual percentual no consumo de energia nuclear. (Tipo: Float)
- **nuclear\_cons\_change\_twh:** Mudança anual no consumo de energia nuclear - A energia equivalente de entrada é baseada na geração bruta e não contabiliza o fornecimento de eletricidade transfronteiriço. (Unidade: terawatt-horas; Tipo: Float)
- **nuclear\_consumption:** Consumo de energia primária de energia nuclear - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)
- **nuclear\_elec\_per\_capita:** Geração de eletricidade a partir de energia nuclear per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **nuclear\_electricity:** Geração de eletricidade a partir de energia nuclear - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **nuclear\_energy\_per\_capita:** Consumo de energia nuclear per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **nuclear\_share\_elec:** Participação da energia nuclear na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **nuclear\_share\_energy:** Participação do consumo de energia primária que vem da energia nuclear - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **oil\_cons\_change\_pct:** Mudança anual percentual no consumo de petróleo. (Tipo: Float)
- **oil\_cons\_change\_twh:** Mudança anual no consumo de petróleo. (Unidade: terawatt-horas; Tipo: Float)
- **oil\_consumption:** Consumo de energia primária de petróleo - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **oil\_elec\_per\_capita:** Geração de eletricidade a partir de petróleo per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)



- **oil\_electricity**: Geração de eletricidade a partir de petróleo - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **oil\_energy\_per\_capita**: Consumo de petróleo per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **oil\_prod\_change\_pct**: Mudança anual na produção de petróleo - Medido como uma porcentagem da produção do ano anterior. (Tipo: Float)
- **oil\_prod\_change\_twh**: Mudança anual na produção de petróleo - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **oil\_prod\_per\_capita**: Produção de petróleo per capita - Medido em kilowatt-horas por capita. (Unidade: kilowatt-horas; Tipo: Float)
- **oil\_production**: Produção de petróleo - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **oil\_share\_elec**: Participação do petróleo na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **oil\_share\_energy**: Participação do consumo de energia primária que vem do petróleo - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **other\_renewable\_consumption**: Consumo de energia primária de outras renováveis - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)
- **other\_renewable\_electricity**: Geração de eletricidade a partir de outras renováveis, incluindo bioenergia - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **other\_renewable\_exc\_biofuel\_electricity**: Geração de eletricidade a partir de outras renováveis, excluindo bioenergia - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **other\_renewables\_cons\_change\_pct**: Mudança anual percentual no consumo de outras fontes renováveis. (Tipo: Float)
- **other\_renewables\_cons\_change\_twh**: Mudança anual no consumo de outras fontes renováveis - A energia equivalente de entrada é baseada na geração bruta e não contabiliza o fornecimento de eletricidade transfronteiriço. (Unidade: terawatt-horas; Tipo: Float)
- **other\_renewables\_elec\_per\_capita**: Geração de eletricidade a partir de outras renováveis, incluindo bioenergia, per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **other\_renewables\_elec\_per\_capita\_exc\_biofuel**: Geração de eletricidade a partir de outras renováveis, excluindo bioenergia, per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **other\_renewables\_energy\_per\_capita**: Consumo de outras fontes renováveis per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **other\_renewables\_share\_elec**: Participação de outras renováveis na geração de eletricidade, incluindo bioenergia - Medido como porcentagem do total de eletricidade. (Tipo: Float)

- **other\_renewables\_share\_elec\_exc\_biofuel**: Participação de outras renováveis na geração de eletricidade, excluindo bioenergia - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **other\_renewables\_share\_energy**: Participação do consumo de energia primária que vem de outras renováveis - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **per\_capita\_electricity**: Geração total de eletricidade per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **primary\_energy\_consumption**: Consumo de energia primária - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **renewables\_cons\_change\_pct**: Mudança anual percentual no consumo de fontes renováveis. (Tipo: Float)
- **renewables\_cons\_change\_twh**: Mudança anual no consumo de fontes renováveis. (Unidade: terawatt-horas; Tipo: Float)
- **renewables\_consumption**: Consumo de energia primária de fontes renováveis - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)
- **renewables\_elec\_per\_capita**: Geração de eletricidade a partir de fontes renováveis per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **renewables\_electricity**: Geração de eletricidade a partir de fontes renováveis - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **renewables\_energy\_per\_capita**: Consumo de fontes renováveis per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **renewables\_share\_elec**: Participação das fontes renováveis na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **renewables\_share\_energy**: Participação do consumo de energia primária que vem de fontes renováveis - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **solar\_cons\_change\_pct**: Mudança anual percentual no consumo de energia solar. (Tipo: Float)
- **solar\_cons\_change\_twh**: Mudança anual no consumo de energia solar - A energia equivalente de entrada é baseada na geração bruta e não contabiliza o fornecimento de eletricidade transfronteiriço. (Unidade: terawatt-horas; Tipo: Float)
- **solar\_consumption**: Consumo de energia primária de energia solar - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)
- **solar\_elec\_per\_capita**: Geração de eletricidade a partir de energia solar per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **solar\_electricity**: Geração de eletricidade a partir de energia solar - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **solar\_energy\_per\_capita**: Consumo de energia solar per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)

- **solar\_share\_elec**: Participação da energia solar na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **solar\_share\_energy**: Participação do consumo de energia primária que vem da energia solar - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)
- **wind\_cons\_change\_pct**: Mudança anual percentual no consumo de energia eólica. (Tipo: Float)
- **wind\_cons\_change\_twh**: Mudança anual no consumo de energia eólica - A energia equivalente de entrada é baseada na geração bruta e não contabiliza o fornecimento de eletricidade transfronteiriço. (Unidade: terawatt-horas; Tipo: Float)
- **wind\_consumption**: Consumo de energia primária de energia eólica - Medido em terawatt-horas, usando o método de substituição. (Unidade: terawatt-horas; Tipo: Float)
- **wind\_elec\_per\_capita**: Geração de eletricidade a partir de energia eólica per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **wind\_electricity**: Geração de eletricidade a partir de energia eólica - Medido em terawatt-horas. (Unidade: terawatt-horas; Tipo: Float)
- **wind\_energy\_per\_capita**: Consumo de energia eólica per capita - Medido em kilowatt-horas por pessoa. (Unidade: kilowatt-horas; Tipo: Float)
- **wind\_share\_elec**: Participação da energia eólica na geração de eletricidade - Medido como porcentagem do total de eletricidade. (Tipo: Float)
- **wind\_share\_energy**: Participação do consumo de energia primária que vem da energia eólica - Medido como porcentagem do total de energia primária, usando o método de substituição. (Tipo: Float)

## 5.4. Códigos utilizados

Neste trabalho, utilizamos a linguagem de programação Python, conhecida por sua flexibilidade e vasto ecossistema de bibliotecas voltadas para análise de dados e machine learning. Empregamos o Pandas para análise e manipulação de dados, o NumPy para operações avançadas de cálculo científico, e o Matplotlib para a criação de visualizações gráficas. Adicionalmente, utilizamos o Scikit-learn, uma ferramenta essencial que facilita a implementação de técnicas de machine learning, como a Regressão Linear, permitindo construir modelos preditivos robustos de forma eficiente. Os códigos completos e scripts utilizados neste projeto estão disponíveis para consulta no nosso repositório GitHub. Acesse através do link: <https://github.com/lucascarmon4/pcg-analise-de-dados>.

### 5.4.1. Coleta e Tratamento de Dados

A coleta de dados foi realizada através da importação de um arquivo CSV contendo informações detalhadas sobre o consumo de energia mundial. O tratamento desses dados incluiu várias etapas cruciais para garantir a qualidade e precisão das análises subsequentes.

```
1 df = pd.read_csv('Consumo Energia Mundial.csv') # Carregando o arquivo CSV
2 countries_to_remove = [
3     'South America', 'North America' ... # Lista de regiões que não são países
4 ]
5 df = df[~df['country'].isin(countries_to_remove)] # Removendo regiões que não são países
```

#### 5.4.2. Desenvolvimento de Gráficos e Visualizações

```
1 # TOP 20 países com maior consumo de energia não renovável
2 df_non_renewables_grouped = df_non_renewables.groupby('country').sum()
3 df_non_renewables_grouped = df_non_renewables_grouped.sort_values(by='fossil_fuel_consumption', ascending=False).head(20)
4
5 plt.figure(figsize=(12, 8))
6
7 plt.barh(df_non_renewables_grouped.index, df_non_renewables_grouped['coal_consumption'], color='black', label='Carvão')
8 plt.barh(df_non_renewables_grouped.index, df_non_renewables_grouped['oil_consumption'], left=df_non_renewables_grouped['coal_consumption'], color='blue', label='Óleo')
9 plt.barh(df_non_renewables_grouped.index, df_non_renewables_grouped['gas_consumption'], left=df_non_renewables_grouped['coal_consumption'] + df_non_renewables_grouped['oil_consumption'], color='red', label='Gás')
10 plt.barh(df_non_renewables_grouped.index, df_non_renewables_grouped['nuclear_consumption'], left=df_non_renewables_grouped['coal_consumption'] + df_non_renewables_grouped['oil_consumption'] + df_non_renewables_grouped['gas_consumption'], color='green', label='Nuclear')
11 plt.legend()
12 plt.title('Consumo de energia não renovável por país')
13 plt.xlabel('Consumo de energia não renovável (TWh)')
14 plt.ylabel('País')
15 plt.show()
```

**Linhas 1-3:** Seleccionamos e agrupamos os dados para os 20 países com o maior consumo de energia não renovável.

**Linhas 5-14:** Criamos um gráfico de barras horizontais empilhadas que mostra a contribuição de cada tipo de energia não renovável no total consumido por cada país. Cada camada representa um tipo diferente de energia, facilitando a visualização da distribuição do consumo entre carvão, óleo, gás e nuclear.

**Linhas 7-10:** Configuramos as barras do gráfico para cada tipo de energia, usando cores distintas e legendas apropriadas para facilitar a leitura e interpretação do gráfico.

**Linha 11:** Adicionamos uma legenda para identificar cada tipo de energia no gráfico.

**Linhas 12-14:** Definimos títulos e rótulos para os eixos, e exibimos o gráfico.

### 5.4.3. Predição de Dados

Para a predição do consumo de energia, desenvolvemos um modelo de regressão linear que estima o consumo futuro baseado em dados históricos. Este modelo é aplicado de forma individual para cada tipo de energia em cada país, proporcionando previsões específicas que podem informar políticas energéticas e estratégias de sustentabilidade.

```
1 def predict_energy_consumption(df, country, energy_type):
2     country_data = df[(df['country'] == country) & df[energy_type].notna() & df['year'].notna()]
3     if country_data.empty:
4         return None, None
5
6     X = pd.DataFrame(country_data['year'], columns=['year'])
7     y = country_data[energy_type]
8
9     model = LinearRegression()
10    model.fit(X, y)
11
12    future_years = pd.DataFrame(range(X['year'].max() + 1, X['year'].max() + 51), columns=['year'])
13
14    predictions = model.predict(future_years)
15
16    return future_years['year'].values, predictions
```

**Linhas 1-4:** A função `predict_energy_consumption` recebe um DataFrame, o nome de um país e o tipo de energia como parâmetros. A função filtra os dados para esse país e tipo de energia, verificando também se há dados suficientes para uma análise.

**Linhas 6-7:** As variáveis `X` e `y` são definidas para representar os anos (como variável independente) e o consumo de energia (como variável dependente), respectivamente.

**Linhas 9-10:** Um modelo de regressão linear é instanciado e treinado com os dados históricos.

**Linhas 12-13:** São criados dados para os próximos 50 anos após o último ano disponível no conjunto de dados, e o modelo é usado para prever o consumo de energia nesses anos futuros.

**Linha 15:** A função retorna os anos futuros e suas previsões correspondentes.

## 6. Cronograma Final

Atividade	14/04 - 16/04	01/05 - 08/05	11/05 - 16/05	17/05- 19/05	20/05- 25/05	26/05- 27/05
Coleta e tratamento de dados	X					
Análise exploratória de dados		X				
Desenvolvimento de gráficos e visualização			X	X		
Comparação entre diferentes aspectos				X		
Predição de Dados				X	X	
Uso de conjuntos de dados diferentes					X	X
Proposição de medidas para aprimoramento					X	X

### 6.1. Justificativa do Cronograma

- **Período de coleta e tratamento de dados** é um período inicial crucial para iniciar a coleta de dados e realizar os procedimentos preliminares de tratamento. É essencial estabelecer uma base sólida de dados bem organizada e limpa para garantir que as análises subsequentes sejam precisas e eficientes.
- **A análise exploratória de dados** é uma fase fundamental do projeto, pois permite compreender melhor a estrutura e as características dos dados. Esta etapa requer tempo para explorar os dados de forma abrangente, identificar padrões, tendências e possíveis problemas nos dados que exigem correção.
- **O desenvolvimento de gráficos e visualizações** é essencial para comunicar os insights obtidos durante a análise de dados de forma clara e eficaz. Essa etapa exige tempo para criar visualizações detalhadas e interativas que ajudarão a apresentar os resultados de forma compreensível para os stakeholders.

- **A comparação entre diferentes aspectos dos dados**, como regiões, setores industriais e fontes de energia, é uma parte crucial da análise. Essa etapa requer tempo para realizar comparações detalhadas e identificar insights significativos que possam orientar as recomendações finais do projeto.
- **A predição de dados** é uma etapa avançada do projeto que requer tempo para selecionar e ajustar os modelos estatísticos adequados, além de validar sua precisão. Essa etapa é essencial para prever tendências futuras nos gastos de energia e fornecer insights valiosos para orientar as recomendações de políticas e práticas.
- **A análise de conjuntos de dados diferentes**, como dados de fontes renováveis e não renováveis, requer tempo adicional para explorar e comparar os dados de forma abrangente. Essa abordagem permite uma compreensão mais holística dos padrões de consumo de energia e suas implicações.
- **A proposição de medidas para aprimoramento** é uma parte crítica do projeto que requer tempo para revisar os insights obtidos durante a análise de dados e formular recomendações práticas e acionáveis. Essa etapa é essencial para traduzir os resultados da análise em iniciativas tangíveis que possam gerar impacto real no uso de energia e na sustentabilidade.

## **7. Referencial Teórico**

A energia é fundamental na sociedade moderna, servindo como base para o desenvolvimento econômico, social e tecnológico, sendo essencial para a vida contemporânea.

A energia é a única moeda universal: uma de suas muitas formas deve ser transformada para realizar qualquer coisa. As manifestações universais dessas transformações variam desde as enormes rotações de galáxias até reações term nucleares em estrelas. (SMIL, 2017, p.13)

Sua disponibilidade e sustentabilidade são críticas para garantir um futuro próspero e seguro para as gerações futuras, e diante da latente necessidade, o setor de energia vem crescendo mais e mais. Atualmente, dispomos de energia oriunda de fontes diversas, como petróleo, hidroeletricidade, gás natural, carvão, entre outras. Frente às diversas novas tecnologias que surgem a todo momento e ao estilo de vida das pessoas cada vez mais focado em modernização e automação de processos, fica clara a importância do setor de energia no desenvolvimento da sociedade moderna.

Destarte, todo o estudo da evolução é crucial para entendermos melhor o papel da energia na sociedade moderna, e a partir disso, o grupo tem a proposta de fazer a análise de dados de Datasets sobre energias renováveis e não renováveis, a fim de compreender alguns panoramas energéticos, fornecendo insights de como as fontes de energias são utilizadas, ajudando a compreender o equilíbrio entre diferentes fontes de energia, tendências importantes, sua participação de mercado, avanços tecnológicos e principalmente seus impactos ambientais, fazendo uma análise de dados detalhada de energia renovável e não renovável, permitindo uma avaliação mais precisa da sustentabilidade das fontes de energia.

Para isso, é necessário o entendimento mais preciso sobre energia renovável e não renovável, vantagens e desvantagens e principais fontes de energia.

### **7.1. Energia Renovável**

A energia renovável é aquela que é gerada a partir de fontes naturalmente reabastecidas e que são virtualmente inesgotáveis no contexto humano. Essas fontes de energia são consideradas sustentáveis porque não esgotam os recursos naturais finitos da Terra e geralmente têm impactos ambientais menores em comparação com as fontes de energia não renováveis, além de que elas são a parte mais importante da transição para um sistema de energia que abandona os combustíveis fósseis, combatendo assim o aquecimento global. Além disso, são energias limpas que protegem a saúde humana e o meio ambiente. Produzir cada vez mais energia renovável e abandonar as fontes convencionais é uma necessidade compartilhada por todos os países do mundo. De acordo com dados mais



recentes do relatório da Agência Internacional de Energia Renovável (IRENA), “em 2022, até mesmo 83% de toda a capacidade de eletricidade adicionada foi proveniente de fontes renováveis. Enquanto em 2021, as energias renováveis geraram 38% da eletricidade do mundo” (ENEL, 2024).

Ademais, a ENEL ainda afirma,

O destino das energias renováveis é se tornar a fonte mais vantajosa de eletricidade para o planeta e desenvolvimento econômico. [...] Porque a energia renovável, quando é produzida com uma visão integrada, capaz de atravessar toda a cadeia de valor, resulta ser totalmente sustentável (ENEL, 2024).

São exemplos de fontes de energia renováveis: solar (energia do sol), eólica (vento), hídrica (energia da água dos rios), geotérmica (energia do interior da Terra), biomassa (matéria orgânica), oceânica (energia proveniente das marés e ondas) e hidrogênio.

## **7.2. Energia Não Renovável**

Toda energia não renovável é derivada de fontes que não se reabastecem em uma escala de tempo humana significativa, o que significa que são esgotáveis em um determinado momento. Essas fontes de energia são geralmente extraídas do subsolo e sua disponibilidade depende de reservas finitas.

A energia não renovável é amplamente utilizada em todo o mundo devido à sua alta densidade energética e à infraestrutura existente para sua produção e distribuição. No entanto, seu uso tem implicações significativas para o meio ambiente, incluindo a emissão de gases de efeito estufa e outros poluentes atmosféricos, bem como questões relacionadas à segurança e à gestão de resíduos, especialmente no caso da energia nuclear. Além disso, como essas fontes são finitas, há preocupações sobre a segurança energética e a estabilidade dos preços à medida que as reservas se esgotam.

As fontes não renováveis são versáteis e apresentam elevada produtividade, sendo essas suas principais vantagens. Como desvantagem estão os impactos ambientais ocasionados pela queima de combustíveis fósseis e pelo manejo inadequado dos elementos radioativos e rejeitos das usinas.

O aumento das emissões de gases causadores do aquecimento global é uma preocupação central, [...] Esse fenômeno também gera outros efeitos negativos, como o aumento do nível do mar e perturbações nos padrões climáticos (MOREIRA, 2024).

São fontes não-renováveis: petróleo, carvão mineral, xisto, gás natural e nuclear. São exemplos de fontes de energia renováveis: solar (energia do sol), eólica (vento), hídrica

(energia da água dos rios), geotérmica (energia do interior da Terra), biomassa (matéria orgânica), oceânica (energia proveniente das marés e ondas) e hidrogênio. São fontes não-renováveis: petróleo, carvão mineral, xisto, gás natural e nuclear.

### **7.3. Teorias e Métodos de Análise de Dados**

A crescente preocupação com a sustentabilidade e a mitigação das mudanças climáticas têm impulsionado a transição global para fontes de energia renovável. Nesse contexto, a análise de dados desempenha um papel crucial na compreensão do panorama energético, tanto em relação às fontes renováveis quanto às não renováveis. Este projeto tem como objetivo explorar as teorias e métodos de análise de dados aplicados a conjuntos de dados sobre energia renovável e não renovável, buscando extrair insights valiosos para orientar políticas, investimentos e estratégias de transição energética. A partir disso, como citado no parágrafo de objetivos, o grupo vai utilizar do Power BI como uma ferramenta muito ampla para a análise de dados do setor de energia renovável e não renovável, oferecendo recursos avançados de visualização de dados, como gráficos interativos, mapas geoespaciais e análises mais aprofundadas e uma comunicação mais eficaz dos resultados.

Com sua capacidade de visualização intuitiva, análise em tempo real e integração de diferentes fontes de dados, o Power BI torna-se uma solução essencial para obter insights valiosos e tomar decisões estratégicas no setor. (AWARI, 2023).

Com isso, os dados de produção de energia de fontes renováveis, como solar e eólica, podem ser importados e analisados para identificar padrões sazonais, tendências de produção e eficiência operacional. Ao mesmo tempo, os dados relacionados a fontes não renováveis, como petróleo, gás natural e carvão, podem ser igualmente explorados para compreender a demanda, os preços e os impactos ambientais.

### **7.4. Casos de Uso de Análise de Dados no Setor Energético**

A análise de dados no setor energético oferece uma gama diversificada de casos de uso, desde a previsão de demanda até a gestão de riscos e a otimização da produção. Um dos principais usos é a previsão de demanda de energia, onde técnicas avançadas são empregadas para projetar necessidades futuras com base em fatores como histórico de consumo e condições climáticas. Além disso, essa análise de dados desempenha um papel crucial na manutenção preditiva de ativos, como turbinas eólicas e painéis solares, permitindo que as empresas identifiquem potenciais falhas antes que ocorram, reduzindo custos e minimizando tempo de inatividade.

Para fontes de energia renováveis, a análise de dados é fundamental na otimização da produção, considerando variáveis como padrões climáticos e eficiência dos equipamentos. Isso possibilita o máximo aproveitamento dessas fontes limpas de energia, além da gestão de energia, permitindo monitoramento em tempo real do consumo e identificação de

oportunidades para redução de custos e impacto ambiental. “Existe uma preocupação muito grande em como os meios de gerar energia impactam no meio ambiente e suas consequências. A busca por recursos renováveis que trazem menos danos é um ponto comum entre a maior parte dos países e gera debates intensos” (RBE, 2022).

A análise de mercado e precificação é outra área importante, fornecendo insights sobre tendências do mercado de energia e comportamento do consumidor, essenciais para estratégias de precificação e investimento. Ademais, a análise de dados também desempenha um papel vital na gestão de riscos e conformidade, detectando anomalias nos dados de produção, avaliando impactos externos, como desastres naturais, e garantindo conformidade com regulamentações do setor.

### **7.5. Impacto de Análise de Dados na Política e Gestão Energética**

Em primeiro lugar, a análise de dados permite uma compreensão mais profunda das necessidades de energia. Através da mineração de dados históricos e padrões de consumo, podemos prever com precisão a demanda futura, fornecendo uma base sólida para o planejamento estratégico da capacidade de geração e distribuição, além de comparações com países como o Brasil, que é um bom exemplo de energia renovável para o mundo todo.

O Brasil chama atenção de maneira positiva por ser um exemplo. A matriz energética do país é composta por 83,4% de fontes limpas, segundo o Ministério de Minas e Energia. Isso engloba hidrelétricas, eólica, biomassa/biogás e solar centralizada. (RBE, 2023)

Além disso, a gestão eficaz da energia é aprimorada pela otimização dos processos de produção. Ao analisar dados operacionais detalhados, podemos identificar áreas de ineficiência e oportunidades para aumentar a eficiência e reduzir os custos, como ocorreu na Alemanha, onde “o governo alemão investiu em subsídios a fim de estimular a instalação de painéis fotovoltaicos nas residências e empresas.” (RBE, 2023).

A ascensão das energias renováveis é grandemente facilitada pela análise de dados. Através da identificação de recursos eólicos e solares ideais, podemos maximizar o potencial dessas fontes limpas e sustentáveis, integrando-as de forma mais eficaz na matriz energética.

No campo da política energética, a análise de dados fornece uma base objetiva para a tomada de decisões. Os formuladores de políticas podem agora fundamentar suas escolhas em evidências sólidas, avaliando o impacto potencial de diferentes medidas e desenvolvendo estratégias mais eficazes para enfrentar desafios energéticos complexos.

# 8. Resultados e Discussões

## 8.1. Coleta e Tratamento de Dados

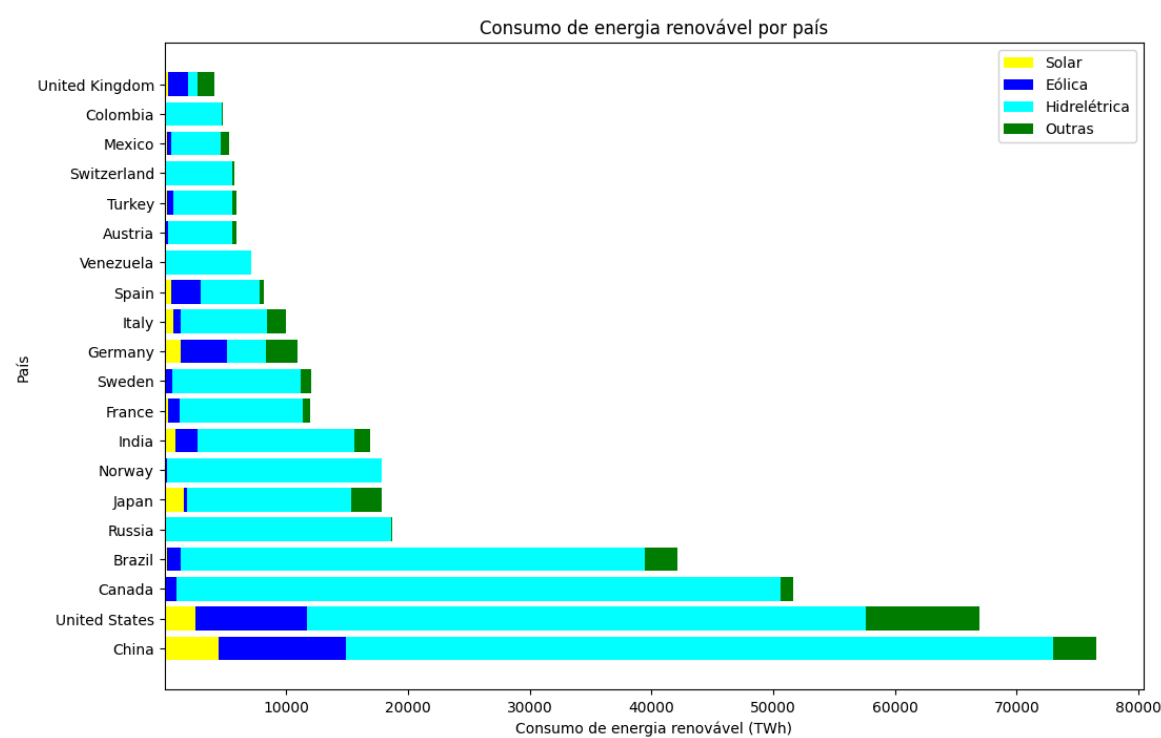
Para a análise dos gráficos relacionados ao consumo de energia renovável e não renovável em diferentes países, utilizamos dados que correlacionam o PIB com o consumo energético. Estes dados foram coletados de fontes confiáveis e tratados para garantir comparabilidade e integridade, facilitando assim uma análise acurada das tendências observadas.

## 8.2. Análise Exploratória de Dados

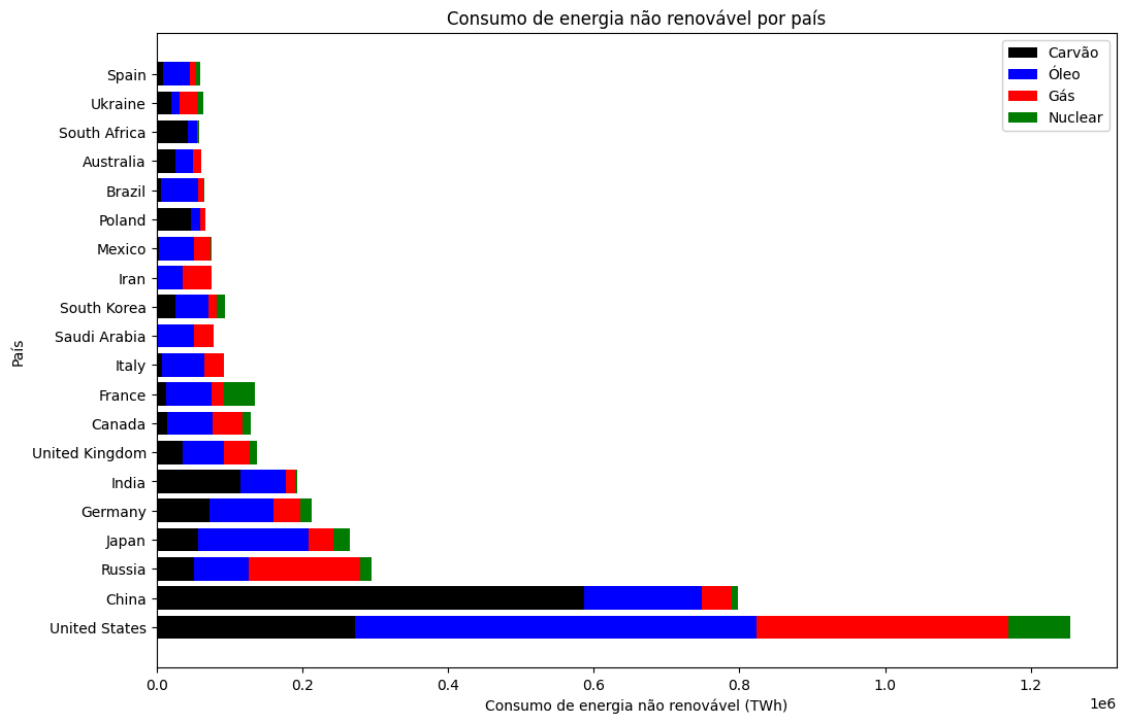
A análise exploratória revelou uma relação direta entre o aumento do PIB e o aumento do consumo de energia não renovável, sugerindo uma dependência contínua de fontes energéticas menos sustentáveis por economias mais desenvolvidas. Simultaneamente, observou-se um incremento no uso de energias renováveis, embora em menor escala.

## 8.3. Desenvolvimento de Gráficos e Visualizações

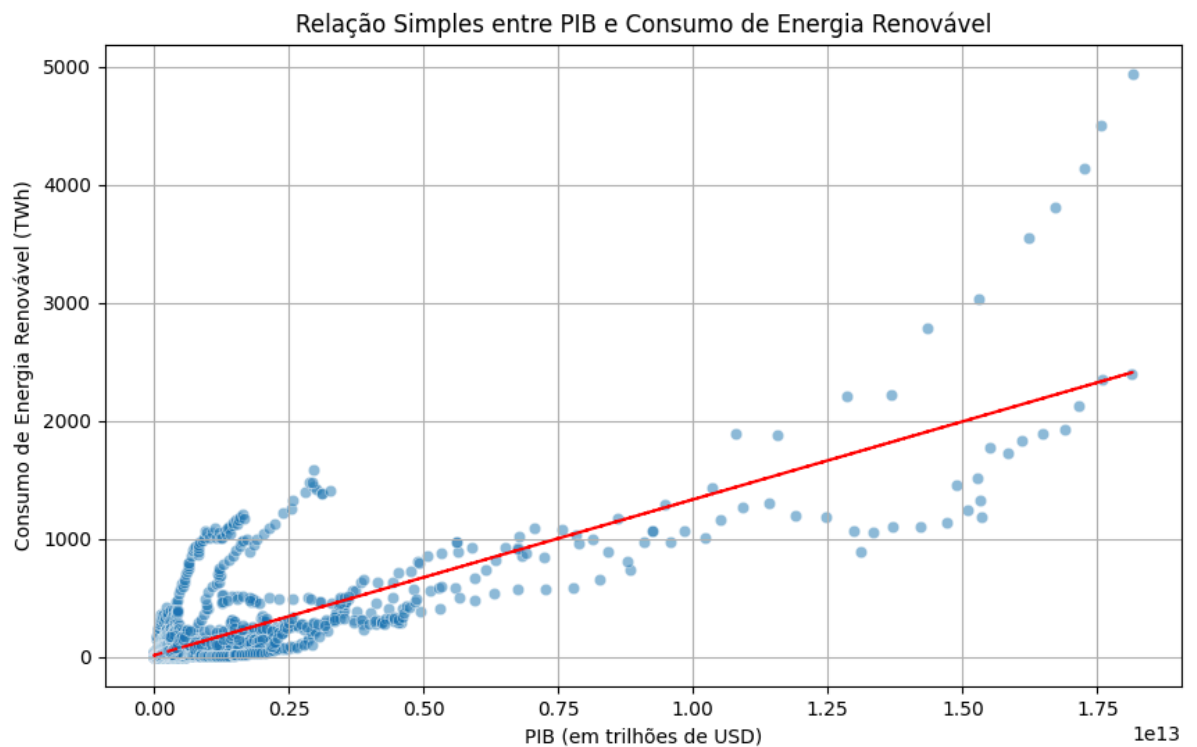
### 8.3.1. Diversidade no Consumo de Energia Renovável por País (Top 20)



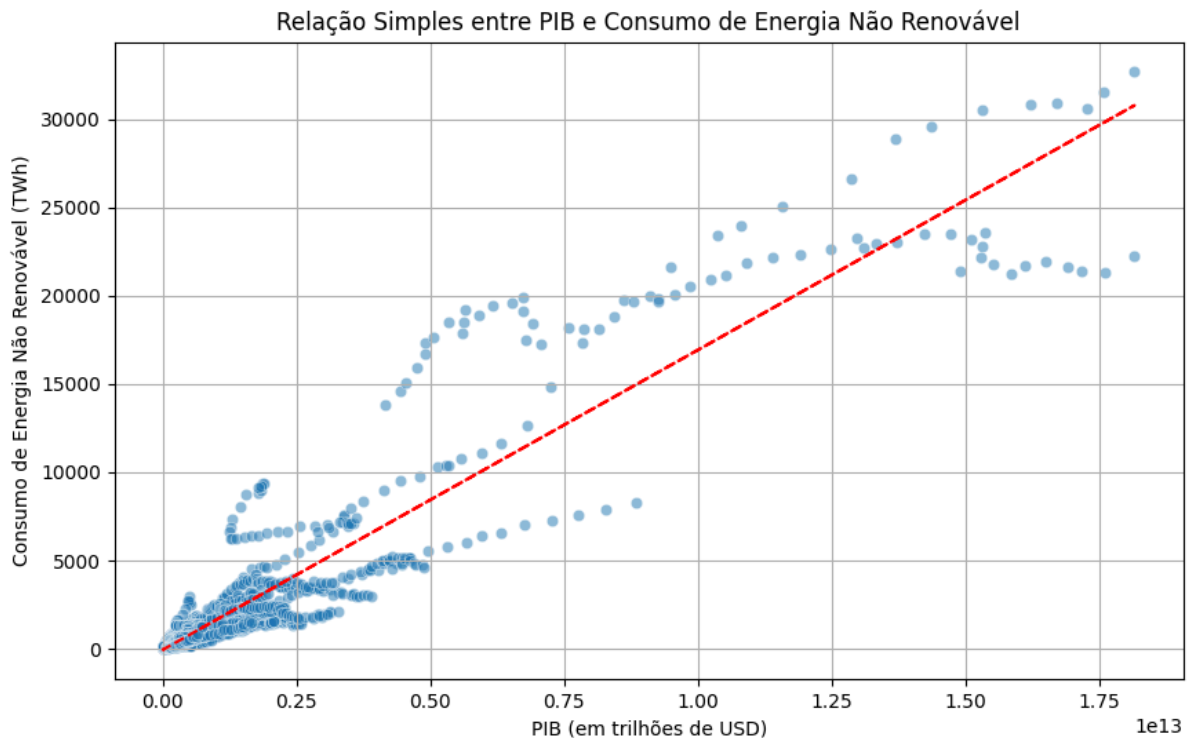
### 8.3.2. Diversidade no Consumo de Energia Não Renovável por País (Top 20)



### 8.3.3. Correlação entre PIB e Consumo de Energia Renovável



### 8.3.4. Correlação entre PIB e Consumo de Energia Não Renovável



### 8.3.5. Dashboard Interativo em Power BI

Para preparar os dados para o Power BI, criamos diversas etapas em Python com o objetivo de otimizar a forma como os dados estavam estruturados. A seguir, descrevo as etapas principais que realizamos para alcançar um modelo de dados eficiente e interativo no Power BI:

- **Limpeza e Transformação dos Dados:**

Inicialmente, utilizamos Python para limpar e transformar os dados brutos. Essa etapa incluiu a remoção de valores nulos, correção de inconsistências e normalização dos dados para garantir que estivessem prontos para análise, também escolhemos apenas as colunas mais relevantes para a nossa análise, essas colunas são:

**country**

**year**

**gdp**

**biofuel\_electricity**

**hydro\_electricity**

**other\_renewable\_electricity**

**solar\_electricity**

**wind\_electricity**

**coal\_electricity**

**gas\_electricity**

**nuclear\_electricity**

**oil\_electricity**

**electricity\_generation**

**renewables\_electricity**

**renewables\_electricity\_%**

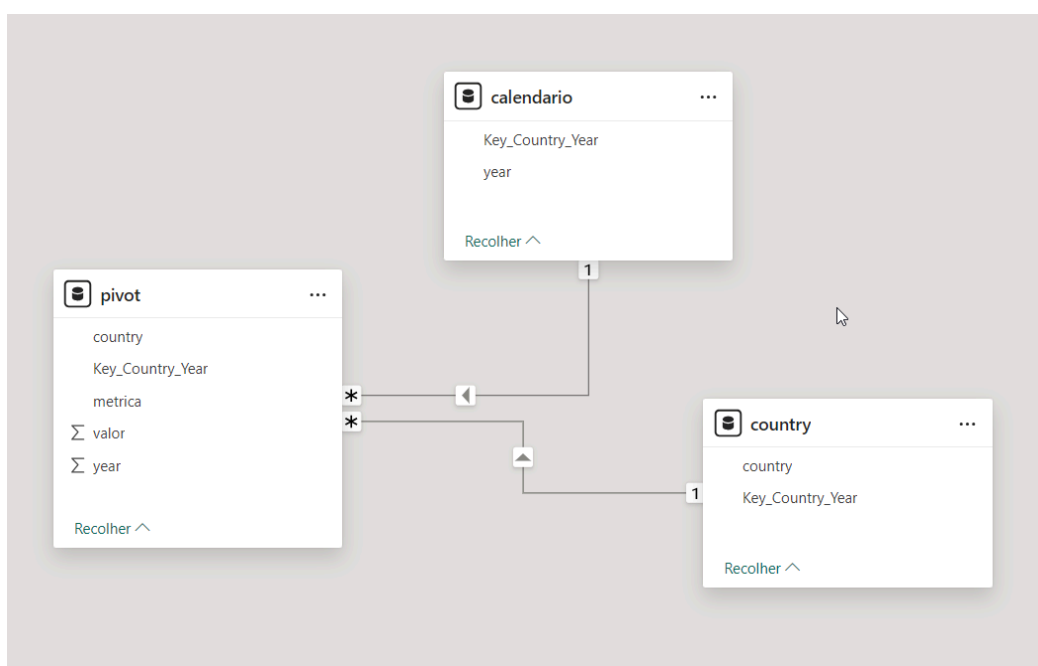
**norenewables\_electricity**

**non\_renewables\_electricity\_%**

Essas métricas foram tiradas do dataset inicial, cujas medidas foram explicadas no capítulo 5.3, outras partes foram calculadas e colocadas em conformidade com procedimentos triviais em Python e próprias manipulações pelo Excel e Power BI

- **Criação de Chaves de Relacionamento:**

Para facilitar os relacionamentos no Power BI, criamos uma chave composta chamada Key\_Country\_Year. Essa chave foi gerada concatenando os valores de país e ano, permitindo um relacionamento unívoco entre as tabelas.



- **Geração de Tabelas Auxiliares:**

Criamos duas tabelas auxiliares: uma para países (country) e outra para métricas (pivot). Essas tabelas foram projetadas para armazenar informações específicas de cada país e métrica, facilitando a criação de filtros e a segmentação dos dados nos dashboards.

- **Transposição e Concatenação:**

Para garantir que as tabelas auxiliares se encaixassem corretamente no modelo de dados, realizamos uma transposição e concatenação dos dados. Esse processo envolveu a reorganização das tabelas para que cada entrada tivesse a mesma estrutura, permitindo um relacionamento preciso e eficaz no Power BI.

- **Importação e Criação dos Relacionamentos no Power BI:**

Após a preparação dos dados em Python, importamos as tabelas resultantes para o Power BI. No Power BI, estabelecemos os relacionamentos entre as tabelas utilizando a chave Key\_Country\_Year. Isso permitiu criar uma estrutura de dados relacional que suporta a análise dinâmica e interativa.

- **Configuração dos Dashboards:**

Com os dados e relacionamentos configurados, construímos dashboards interativos que permitem aos usuários modificar os gráficos com apenas um clique nas opções de países e métricas. Essa interatividade é crucial para explorar os dados de forma eficaz e obter insights valiosos.

## **8.4. Comparação entre Diferentes Aspectos**

Na nossa análise, realizamos uma comparação entre o consumo de diferentes tipos de energia, destacando especialmente a contribuição de energias renováveis como solar, eólica e hidrelétrica em contraste com energias não renováveis como carvão, óleo, gás e nuclear. Utilizando gráficos que ilustram o consumo por país, essa comparação nos permitiu visualizar não apenas o volume total de consumo, mas também a proporção de diferentes fontes de energia, ressaltando as prioridades energéticas de cada nação.

## **8.5. Predição de Dados**

Através do modelo de regressão linear implementado, a predição de dados foi realizada para projetar o consumo futuro de energia em diversos países, baseando-se em dados históricos de consumo e correlação com o PIB. Estas previsões são visualizadas em



gráficos que mostram as tendências de consumo de 2023 a 2073, permitindo uma visualização clara das expectativas futuras de demanda energética. Essa abordagem preditiva é essencial para planejamento e tomadas de decisão informadas em políticas energéticas.

## **8.6. Uso de Conjuntos de Dados Diferentes**

Nossa análise aproveitou a diversidade de conjuntos de dados para obter uma visão abrangente sobre o consumo de energia. Além dos dados principais sobre consumo de energia renovável e não renovável, foram incorporados detalhes sobre a economia dos países (PIB) e demografia (população). Esse uso integrado de dados econômicos e energéticos enriqueceu a análise, permitindo correlacionar o desenvolvimento econômico com as práticas energéticas e destacar como as tendências econômicas influenciam o consumo energético.

## **8.7. Proposição de Medidas para Aprimoramento**

### **8.7.1. Padronização dos Formatos de Dados**

- **Desafio:** A falta de padronização nos formatos de dados dificulta a comparação e análise eficiente.
- **Medida:** Implementar um formato de dados padrão para todos os conjuntos de dados de energia disponibilizados em portais de transparência. Este formato deve ser aberto, como CSV ou JSON, e seguir uma estrutura uniforme que inclua metadados detalhados sobre a origem e a metodologia de coleta dos dados.

### **8.7.2. Adoção de Tecnologias de Big Data e Machine Learning**

- **Desafio:** A complexidade e o volume dos dados de energia podem dificultar a análise manual.
- **Medida:** Utilizar tecnologias de Big Data e Machine Learning para processar grandes volumes de dados e identificar padrões, tendências e anomalias. Ferramentas como Hadoop, Spark, e bibliotecas de Machine Learning como Scikit-learn e TensorFlow podem ser integradas aos processos de análise.

### **8.7.3. Melhoria na Acessibilidade dos Dados**

- **Desafio:** A dificuldade de acesso aos dados de energia impede uma análise mais ampla e participativa.
- **Medida:** Desenvolver plataformas web intuitivas que permitam a fácil navegação, download e visualização dos dados. Essas plataformas devem incluir filtros

interativos para facilitar a busca e segmentação dos dados por ano, região e tipo de energia.

#### **8.7.4. Transparência e Disponibilidade em Tempo Real**

- Desafio: A falta de atualizações frequentes e em tempo real limita a relevância dos dados disponíveis.
- Medida: Implementar sistemas de atualização em tempo real ou em intervalos regulares para garantir que os dados de consumo de energia estejam sempre atualizados. APIs públicas podem ser utilizadas para fornecer acesso em tempo real aos dados mais recentes.

#### **8.7.5. Educação e Capacitação dos Usuários**

- Desafio: A complexidade dos dados de energia pode ser uma barreira para o entendimento por parte do público geral.
- Medida: Oferecer programas de capacitação e recursos educacionais para usuários, incluindo tutoriais, webinars e cursos online sobre como interpretar e utilizar os dados de energia. Isso aumentará a capacidade dos cidadãos e dos formuladores de políticas de fazer análises informadas.

#### **8.7.6. Incentivo à Utilização de Fontes Renováveis**

- Desafio: A dependência de fontes de energia não renováveis persiste em muitos países.
- Medida: Formular políticas públicas que incentivem a adoção de fontes de energia renováveis, como solar e eólica, por meio de subsídios, incentivos fiscais e apoio à pesquisa e desenvolvimento. Implementar programas de financiamento para projetos de energia renovável, especialmente em regiões com alto potencial para essas fontes.

#### **8.7.7. Implementação de Padrões de Eficiência Energética**

- Desafio: A eficiência energética varia significativamente entre diferentes setores e regiões.
- Medida: Desenvolver e implementar padrões de eficiência energética para indústrias, edifícios e eletrodomésticos. Esses padrões devem ser obrigatórios e monitorados regularmente para garantir conformidade e promover a redução do consumo de energia.

#### **8.7.8. Fomento à Pesquisa e Desenvolvimento**

- Desafio: A inovação tecnológica é essencial para a transição para uma matriz energética mais sustentável.

- Medida: Investir em pesquisa e desenvolvimento de novas tecnologias de energia limpa e eficiente. Parcerias público-privadas podem ser estabelecidas para acelerar o desenvolvimento e a implementação de tecnologias inovadoras no setor energético.

#### **8.7.9. Auditoria e Fiscalização dos Dados**

- Desafio: A credibilidade dos dados pode ser questionada se não houver um processo robusto de verificação.
- Medida: Estabelecer mecanismos de auditoria e fiscalização dos dados de energia para garantir sua precisão e integridade. Organismos independentes podem ser responsáveis por auditar os dados regularmente e publicar relatórios de conformidade.

#### **8.7.10. Dispersão de Dados e Foco no Desenvolvimento Sustentável**

A análise dos dados de consumo de energia e PIB revelou uma dispersão significativa na relação entre o desenvolvimento econômico e o uso de energia renovável. Embora seja intuitivo pensar que países com PIB mais alto tenham maior capacidade e recursos para investir em fontes de energia renovável, nossos dados mostram que essa relação não é linear e, muitas vezes, não se confirma.

Os países desenvolvidos tendem a ter maior PIB, mas isso não implica necessariamente que eles utilizem mais energia renovável. Muitos países com economias avançadas continuam a depender fortemente de fontes de energia não renováveis devido a diversos fatores, incluindo infraestrutura existente, políticas energéticas e interesses econômicos. Por outro lado, alguns países em desenvolvimento têm investido de forma significativa em energia renovável, motivados por iniciativas internacionais de sustentabilidade e pela necessidade de diversificar suas fontes de energia.

Com base na dispersão dos dados, é claro que os países com PIB alto são os principais alvos para medidas de desenvolvimento sustentável por várias razões: Países com PIB elevado possuem os recursos financeiros necessários para planejar e implementar mudanças significativas em suas matrizes energéticas. Eles têm a capacidade de investir em pesquisa e desenvolvimento, infraestrutura e tecnologias de energia renovável. Esses países são os maiores produtores de energia em números absolutos. Portanto, qualquer mudança na sua matriz energética pode ter um impacto global substancial na redução das emissões de gases de efeito estufa e na mitigação das mudanças climáticas. Países com economias desenvolvidas têm a capacidade de liderar pelo exemplo, influenciando políticas energéticas globais e incentivando outras nações a adotarem práticas sustentáveis. Seu papel de liderança pode acelerar a transição energética global. Países desenvolvidos geralmente possuem uma infraestrutura robusta e avançada, que pode ser adaptada para integrar novas tecnologias de

energia renovável. Isso facilita a transição em comparação com países que precisam construir essa infraestrutura do zero.

Diante dessas observações, propomos as seguintes medidas específicas para países com PIB elevado: Implementar políticas que incentivem a modernização da infraestrutura energética para integrar mais fontes renováveis, como solar, eólica e hídrica. Oferecer subsídios e incentivos fiscais para empresas e consumidores que adotem tecnologias de energia limpa. Isso pode incluir desde grandes projetos industriais até sistemas solares residenciais. Aumentar o financiamento para pesquisa e desenvolvimento de novas tecnologias de energia sustentável, visando melhorar a eficiência e reduzir os custos de produção e armazenamento de energia renovável. Implementar regulamentações rigorosas de eficiência energética para indústrias, edifícios e transportes. Estabelecer padrões mínimos obrigatórios que incentivem a adoção de práticas e tecnologias mais eficientes. Promover campanhas de conscientização pública sobre os benefícios da energia renovável e a importância da eficiência energética. Programas educacionais podem ajudar a mudar comportamentos e atitudes em relação ao consumo de energia. Fomentar parcerias entre o setor público e privado para desenvolver projetos de energia renovável. Isso pode incluir a criação de consórcios para grandes projetos de infraestrutura ou incentivos para startups inovadoras no setor energético.

Focar em países com PIB alto para a implementação de medidas de desenvolvimento sustentável é estratégico tanto pela disponibilidade de recursos quanto pelo impacto significativo que essas nações podem ter na produção global de energia. Ao direcionar esforços e investimentos para essas economias, é possível acelerar a transição para uma matriz energética mais sustentável, promovendo benefícios ambientais, econômicos e sociais em escala global.

## 9. Conclusão

Este estudo ressaltou a importância crucial da transparência nos dados de gastos de energia elétrica, explorando como a melhoria da acessibilidade a essas informações pode beneficiar tanto formuladores de políticas quanto o público geral. As discussões teóricas conduzidas até o momento enfatizam a necessidade de adotar abordagens mais sistemáticas e abertas para a disponibilização dos dados, essenciais para uma gestão energética eficaz e sustentável.

A partir da problematização inicial do trabalho, a elaboração dos objetivos foi conduzida de maneira clara e direcionada, refletindo um entendimento profundo das necessidades do setor e das limitações atuais na transparência dos dados. Este processo meticuloso de definição de objetivos, alinhado com uma revisão literária abrangente, culminou na criação de um cronograma bem estruturado, que delineia as fases subsequentes do estudo e garante uma abordagem sistemática para alcançar os resultados esperados.

Ao longo deste trabalho, discutimos as melhores ferramentas de programação, previsão e visualização, identificando aquelas que mais se alinham aos nossos objetivos de pesquisa. Ferramentas como Python, para programação e tratamento de dados, e plataformas de visualização como Power BI, assim como métodos preditivos avançados como machine learning, se mostraram particularmente adequadas para o tratamento, a análise e a previsão de tendências nos dados energéticos. Estas ferramentas não apenas facilitam a manipulação dos dados, mas também promovem uma visualização clara e acessível das informações, essenciais para a compreensão e tomada de decisões informadas. A escolha das referências teóricas que fundamentaram o uso destas ferramentas e métodos, tais como os trabalhos de Smil (2017) e Moreira (2024) sobre as dinâmicas energéticas e suas implicações, pareceu particularmente assertiva, oferecendo um suporte robusto para as técnicas aplicadas.

A escolha dessas ferramentas e referências teóricas é justificada pela sua capacidade de fornecer análises detalhadas e insights preditivos, que são fundamentais para superar as barreiras à transparência dos dados. Este é um passo inicial crucial que define a direção para futuras investigações e intervenções práticas. A implementação dessas ferramentas será vital para avançar no estudo e testar a eficácia de diferentes abordagens na prática.

Por fim, espera-se que as bases estabelecidas e as ferramentas selecionadas neste trabalho inicial ajudem a moldar estratégias mais eficientes para a disseminação e uso dos dados energéticos, promovendo uma gestão mais sustentável dos recursos energéticos. As próximas etapas do estudo buscarão desenvolver essas estratégias e avaliar sua eficácia em contextos práticos, sempre com o objetivo de fortalecer a transparência e a responsabilidade no setor energético.

## 10. Referências

AWARI. Power BI Aneel: Análise de Dados para o Setor de Energia com Power BI, 2023.

Disponível em:

<https://awari.com.br/power-bi-aneel-analise-de-dados-para-o-setor-de-energia-com-power-bi/>.

Acesso em: 25 abr. 2024.

ENEL GREEN POWER. Energias Renováveis, 2024. Disponível em:

<https://www.enelgreenpower.com/pt/learning-hub/energias-renoveis>. Acesso em: 24 abr.

2024.

MOREIRA, Thalles. Fontes de energia não renováveis e seus impactos, eCycle, 2023.

Disponível em: <https://www.ecycle.com.br/fontes-de-energia-nao-renovaveis/>. Acesso em: 24 abr. 2024.

POUDEL, Pralabh. World Energy Consumption. Disponível em:

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

24 abr. 2024.

RBE ENERGIA. Além das fronteiras: entenda como funciona o mercado de energia em diferentes partes do mundo, 2022. Disponível em:

<https://www.rbenergia.com.br/blog/producao-mundial-de-energia-eletrica/>. Acesso em: 25 abr. 2024.

SMIL, Vaclav. Energy and Civilization: A History. MIT Press, 2017.