

**UNIVERSIDADE PRESBITERIANA MACKENZIE**  
**FACULDADE DE COMPUTAÇÃO E INFORMÁTICA (FCI)**

THAINA VIEIRA DOS SANTOS

VINICIUS CAUMO SEGATTO

PREVISÃO DE EMISSÕES DE CO<sub>2</sub> E OTIMIZAÇÃO DO USO DE ENERGIAS  
RENOVÁVEIS: UMA ABORDAGEM COM MACHINE LEARNING

São Paulo

2024

THAINA VIEIRA DOS SANTOS

VINICIUS CAUMO SEGATTO

PREVISÃO DE EMISSÕES DE CO<sub>2</sub> E OTIMIZAÇÃO DO USO DE ENERGIAS  
RENOVÁVEIS: UMA ABORDAGEM COM MACHINE LEARNING

Trabalho Aplicando conhecimento para entrega no Moodle referente ao conteúdo 1 de aprendizagem do componente curricular Ciência, tecnologia e sociedade;

ORIENTADOR: Prof. GUSTAVO SCALABRINI SAMPAIO

São Paulo

2024

THAINA VIEIRA DOS SANTOS

VINICIUS CAUMO SEGATTO

PREVISÃO DE EMISSÕES DE CO<sub>2</sub> E OTIMIZAÇÃO DO USO DE ENERGIAS  
RENOVÁVEIS: UMA ABORDAGEM COM MACHINE LEARNING

Trabalho Aplicando conhecimento para entrega no Moodle referente ao conteúdo 1 de aprendizagem do componente curricular Ciência, tecnologia e sociedade;

Aprovado em

BANCA EXAMINADORA

---

Prof. GUSTAVO SCALABRINI SAMPAIO

Universidade Presbiteriana Mackenzie

## RESUMO

Este trabalho apresenta uma análise preditiva sobre emissões de CO<sub>2</sub> per capita e geração de energia renovável no Brasil, utilizando modelos de séries temporais SARIMA. A motivação do estudo surge da crescente necessidade de compreender padrões históricos e projetar tendências futuras para apoiar políticas ambientais e energéticas sustentáveis. O objetivo principal foi realizar previsões para o período de 2021 a 2030, avaliando a estabilidade das emissões e a produção de energia renovável. A justificativa baseia-se na relevância dessas informações para o planejamento estratégico e a mitigação de impactos ambientais. A metodologia adotada incluiu a aplicação de modelos SARIMA para séries temporais, com ajustes e validações baseados em testes diagnósticos como Ljung-Box e Jarque-Bera. Os resultados mostraram que, enquanto as emissões de CO<sub>2</sub> per capita apresentaram tendência de leve declínio, a geração de energia renovável manteve-se estável, ambas com intervalos de confiança que refletem a incerteza das previsões. Conclui-se que o modelo é eficaz para captar padrões históricos e projetar tendências, embora melhorias possam ser alcançadas com a inclusão de variáveis exógenas e séries temporais mais longas. A análise fornece subsídios relevantes para a formulação de políticas públicas voltadas à sustentabilidade no Brasil.

**Palavras-chave:** emissões de CO<sub>2</sub>; energia renovável; séries temporais; SARIMA; sustentabilidade.

## **ABSTRACT**

This study presents a predictive analysis of per capita CO<sub>2</sub> emissions and renewable energy generation in Brazil, using SARIMA time series models. The motivation for this study arises from the growing need to understand historical patterns and project future trends to support sustainable environmental and energy policies. The main objective was to make forecasts for the period from 2021 to 2030, assessing the stability of emissions and renewable energy production. The justification is based on the relevance of this information for strategic planning and mitigation of environmental impacts. The methodology adopted included the application of SARIMA models for time series, with adjustments and validations based on diagnostic tests such as Ljung-Box and Jarque-Bera. The results showed that while per capita CO<sub>2</sub> emissions showed a trend of slight decline, renewable energy generation remained stable, both with confidence intervals that reflect the uncertainty of the forecasts. It is concluded that the model is effective in capturing historical patterns and projecting trends, although improvements can be achieved by including exogenous variables and longer time series. The analysis provides relevant subsidies for the formulation of public policies aimed at sustainability in Brazil.

**Keywords:** CO<sub>2</sub> emissions; renewable energy; time series; SARIMA; sustainability.

## **SUMÁRIO**

### **1. INTRODUÇÃO**

#### **1.1 CONTEXTO**

#### **1.2 MOTIVAÇÃO E JUSTIFICATIVA**

#### **1.3 OBJETIVO**

### **2. REFERENCIAL TEÓRICO**

### **3. DIAGRAMA DE SOLUÇÃO**

### **4. EDA E PRÉ-PROCESSAMENTO DOS DADOS**

### **5. INTRODUÇÃO À MODELAGEM PREDITIVA: UTILIZANDO O MODELO SARIMAX**

### **6. RESULTADOS**

### **7. DISCUSSÃO E CONCLUSÃO**

### **8. APRESENTAÇÃO**

### **9. REFERÊNCIAS BIBLIOGRÁFICAS**

### **10. APÊNDICES**

## **LISTA DE FIGURAS**

Figura 1 - Evolução das emissões totais por ano

Figura 2 - Emissões de CO<sub>2</sub> per capita do Brasil

Figura 3 - Emissões de CO<sub>2</sub> per capita de Argentina

Figura 4 - Trecho do código

Figura 5 - Função dfSummary dos dados para os anos de 2010 a 2020

Figura 6 - Função dfSummary sobre emissões de energia renovável ao longo dos anos

Figura 7 - Evolução das produções de energias renováveis por ano

Figura 8 - Análise Comparativa da Produção de Energia Renovável

Figura 9 - Previsão de Criação de Energia Renovável no Brasil com SARIMA(2021-2030)

## 1. INTRODUÇÃO

### 1.1 CONTEXTO

A transição para uma matriz energética mais limpa e sustentável é um dos maiores desafios do século XXI. De acordo com uma matéria da *BBC News*, a crise climática tem se manifestado de diversas formas, a seca na Amazônia é um exemplo alarmante de seus impactos<sup>(1)</sup>. A região amazônica, conhecida como o "pulmão do mundo", desempenha um papel crucial na regulação do clima global. A perda de sua biodiversidade e a redução de sua capacidade de absorver carbono contribuem significativamente para o aumento das temperaturas e a intensificação de eventos climáticos extremos. A recente revisão da meta climática brasileira, conforme anunciado pelo governo federal, demonstra um compromisso renovado com a luta contra as mudanças climáticas. Ao retomar a meta estabelecida no Acordo de Paris<sup>(1)</sup>, o país sinaliza sua intenção de reduzir significativamente as emissões de gases de efeito estufa até 2025 e 2030.

A relação entre as emissões de CO<sub>2</sub> e a seca na Amazônia é complexa e multifacetada. O aumento das concentrações de gases de efeito estufa na atmosfera, provenientes principalmente da queima de combustíveis fósseis, intensifica o efeito estufa, elevando as temperaturas globais e alterando os padrões de precipitação. A seca



prolongada na Amazônia, por sua vez, reduz a capacidade da floresta de absorver CO<sub>2</sub>, gerando um ciclo vicioso que acelera o aquecimento global.

A transição para uma matriz energética mais limpa, baseada em fontes renováveis, é fundamental para mitigar os efeitos das mudanças climáticas e reduzir a frequência e a intensidade de eventos extremos como a seca na Amazônia. Ao substituir os combustíveis fósseis por fontes renováveis, como a energia solar, eólica e hidrelétrica, reduzimos as emissões de gases de efeito estufa e diminuimos a pressão sobre os ecossistemas.

O entendimento da relação entre as emissões de CO<sub>2</sub>, o uso de energias renováveis e os eventos climáticos extremos é o foco deste estudo. Busca-se quantificar o impacto das energias renováveis na redução das emissões e identificar os fatores que

influenciam a vulnerabilidade dos ecossistemas às mudanças climáticas, permitindo o desenvolvimento de políticas públicas e estratégias de mitigação mais eficazes.

## **1.2 MOTIVAÇÃO E JUSTIFICATIVA**

A previsão de emissões de CO<sub>2</sub> e a otimização do uso de energias renováveis, com foco no Brasil, representa uma escolha de tema extremamente relevante e estratégica por diversos motivos. O Brasil possui uma matriz energética diversificada, com um grande potencial para o desenvolvimento de fontes renováveis como a hidrelétrica, eólica e solar. O projeto tem o potencial de gerar novas soluções concretas para os desafios relacionados à energia e ao clima. Ao desenvolver modelos precisos de previsão e otimização, poderemos auxiliar na tomada de decisões mais informadas por parte de governos, empresas e sociedade em geral. O tema também abrange diversas áreas do conhecimento, como ciência da computação, engenharia, economia e ciências ambientais. Essa característica multidisciplinar permite a formação de equipes de trabalho mais diversificadas e a geração de soluções mais completas e inovadoras.

## **1.3 OBJETIVO**

Um modelo preditivo robusto e preciso será desenvolvido para estimar as emissões de CO<sub>2</sub> no Brasil, com o objetivo de otimizar o uso de fontes de energia renovável e apoiar a tomada de decisões estratégicas para a transição energética.

## **Objetivos Específicos Relacionados ao Desenvolvimento do Projeto:**

### **Coleta e Preparação de Dados**

- Dados relevantes serão identificados e coletados a partir de diversas fontes, como bases de dados governamentais, agências internacionais e estudos acadêmicos.
- Os dados serão limpos e tratados, assegurando-se a qualidade e a consistência das informações.
- Um banco de dados integrado e estruturado será criado para facilitar a análise e o desenvolvimento dos modelos.

### **Engenharia de Features**

- Variáveis relevantes para a previsão das emissões de CO<sub>2</sub> serão selecionadas e criadas, considerando fatores como produção industrial, consumo energético, atividades agrícolas, desmatamento e variáveis climáticas.
- A engenharia de features será realizada para extrair informações úteis dos dados e melhorar o desempenho dos modelos.

### **Modelagem Preditiva**

- Diferentes algoritmos de machine learning serão experimentados para a construção dos modelos preditivos.
- A performance dos modelos será avaliada com o uso de métricas apropriadas, como RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) e R<sup>2</sup>.
- O modelo de melhor desempenho será selecionado para a previsão das emissões de CO<sub>2</sub>.

### **Análise de Cenários**

- Os modelos serão validados utilizando dados históricos e comparando os resultados com previsões anteriores.
- A precisão das previsões será continuamente aprimorada por meio de refinamentos nos modelos e incorporação de novos dados.

## **2. REFERENCIAL TEÓRICO**

A produção e o consumo de energia, desde a era industrial, têm estado intrinsecamente ligados às emissões de gases de efeito estufa (GEE). O dióxido de carbono (CO<sub>2</sub>), principal GEE, é gerado como subproduto da queima de combustíveis fósseis, amplamente utilizados para a geração de energia. A permanência do CO<sub>2</sub> na atmosfera por longos períodos contribui para o aumento do efeito estufa e para as mudanças climáticas subsequentes. A compreensão dessa relação é essencial para que políticas e estratégias possam ser elaboradas com o objetivo de mitigar os impactos ambientais da atividade humana <sup>(4)</sup>.

No Brasil, que possui uma matriz energética diversificada e rica em fontes renováveis, identifica-se um grande potencial para a geração de créditos de carbono e participação ativa no mercado internacional. Segundo o The World Bank (2021) <sup>(4)</sup>, o país pode beneficiar-se consideravelmente da comercialização desses créditos. O setor de transportes, especialmente com o uso crescente de etanol hidratado, revela uma oportunidade promissora para a geração de créditos de carbono. No entanto, a falta de regulamentação clara e a ausência de um mercado de carbono consolidado ainda são desafios a serem superados.

Paralelamente, a previsão da demanda por combustíveis, como a gasolina, é considerada essencial para a tomada de decisões estratégicas no setor energético.

Estudos como o de Santos Jr. (2018) <sup>(4)</sup> demonstram a complexidade dessa tarefa, na qual diversos modelos estatísticos são utilizados para captar as nuances do comportamento dos consumidores. A escolha do modelo mais adequado depende de fatores como a disponibilidade de dados, o período de análise e os objetivos da pesquisa.

De acordo com o artigo "Aplicação de Machine Learning para Previsão das Emissões de CO<sub>2</sub> Produzidas pela Produção de" <sup>(5)</sup>, de Soares, o uso de técnicas de Machine Learning (ML) para prever as emissões de CO<sub>2</sub> é uma área em constante evolução. No artigo, apresenta-se uma revisão da literatura sobre o tema e uma aplicação prática de ML para a previsão das emissões de CO<sub>2</sub> em uma empresa de cimento. Um conjunto de dados de 10 anos foi utilizado para treinar um modelo de regressão linear, que foi então empregado para prever as emissões de CO<sub>2</sub> dos próximos 5 anos. Os resultados mostram que o modelo prevê as emissões de CO<sub>2</sub> com boa precisão. Conclui-se que o uso de ML para a previsão das emissões de CO<sub>2</sub> é promissor, embora desafios como a falta de dados históricos e a complexidade dos processos de produção ainda devam ser superados.

### **3. DIAGRAMA DE SOLUÇÃO**

O diagrama da solução, representado pelo pipeline proposto, estabelece um fluxo de trabalho sequencial para o desenvolvimento do modelo de previsão das emissões de CO<sub>2</sub>, desde a concepção até a implementação final. Essa estrutura serve como um guia orientador e facilita o acompanhamento do progresso e a adaptação do projeto conforme necessário. A seguir, discute-se cada etapa do pipeline:

#### **Definição do Problema e Objetivos**

Clareza: A primeira fase visa delimitar com precisão o problema específico das emissões de CO<sub>2</sub>, esclarecendo o objetivo central da previsão e os impactos ambientais e econômicos que se buscam mitigar.

**Alcance:** Essa etapa define o escopo do projeto, incluindo as áreas e setores abrangidos na análise e os limites temporais e espaciais dos dados coletados. Essa delimitação é essencial para garantir a viabilidade do projeto e o alinhamento com os recursos disponíveis.

**Metas:** Define-se aqui o sucesso do modelo, estabelecendo métricas para avaliar o desempenho preditivo. A escolha das métricas é crucial para garantir a qualidade do modelo e orientar ajustes ao longo do desenvolvimento.

## **Coleta e Preparação dos Dados**

**Identificação das Fontes e Coleta:** As fontes de dados (ex.: bases governamentais, setores energéticos) são identificadas e verificadas para garantir a integridade dos dados históricos sobre produção de energia, consumo e emissões. A coleta deve assegurar que os dados sejam representativos e atualizados.

**Limpeza e Transformação:** Após a coleta, os dados passam por um processo de limpeza e remoção de outliers, correção de inconsistências e tratamento de valores ausentes. Em seguida, são transformados e normalizados, e novas variáveis são criadas conforme necessário, visando um formato adequado para a análise preditiva.

**Exploração:** A análise exploratória permite identificar padrões iniciais, correlações entre variáveis e relações causais potenciais, que servirão de base para a seleção de features.

## **Seleção e Engenharia de Features**

**Seleção de Variáveis Relevantes:** A seleção de features foca na escolha de variáveis com maior impacto nas emissões de CO<sub>2</sub>. Fatores como consumo energético, atividades econômicas e variáveis climáticas são priorizados.

**Engenharia de Features:** Caso necessário, cria-se novas variáveis a partir das existentes para aumentar a capacidade explicativa do modelo. Essa etapa pode envolver a criação de índices compostos ou de métricas derivadas que reforcem a compreensão dos padrões de emissão.

## **Escolha do Modelo de Machine Learning**

**Revisão da Literatura e Experimentação:** A escolha do modelo baseia-se em uma revisão de abordagens de machine learning adequadas para séries temporais e previsões de emissões. Modelos como regressão linear, árvores de decisão e redes neurais são testados para comparação de desempenho.

**Experimentação:** A experimentação abrange a aplicação prática de diferentes modelos nos dados coletados. Compara-se o desempenho de cada modelo para determinar a abordagem que melhor atende aos objetivos de previsão.

### **Treinamento e Validação do Modelo**

**Divisão dos Dados:** A divisão em conjuntos de treinamento, validação e teste possibilita o ajuste e a avaliação do modelo de forma equilibrada. Essa estrutura reduz o risco de overfitting e aumenta a confiabilidade do modelo final.

**Treinamento e Validação:** O treinamento ocorre no conjunto de dados específico e ajusta-se os hiperparâmetros no conjunto de validação. Esse ajuste visa melhorar a precisão e a generalização do modelo sem comprometê-lo em dados novos.

### **Avaliação do Modelo**

**Métricas de Desempenho:** A fase de avaliação envolve métricas como MAE, RMSE e MAPE, que permitem uma visão quantitativa do erro preditivo. Cada métrica fornece insights específicos sobre a precisão e robustez do modelo.

**Interpretação dos Resultados:** A interpretação dos resultados e da importância das variáveis contribui para entender quais fatores influenciam mais as emissões. Esses insights são fundamentais para a tomada de decisões informadas e para definir estratégias de mitigação direcionadas.

Cada etapa do pipeline, conectada às demais, permite um fluxo de desenvolvimento dinâmico e iterativo. A solução é continuamente aprimorada, e ajustes podem ser incorporados à medida que novos dados ou desafios são identificados. Esse roadmap confere agilidade e precisão ao projeto, garantindo que as metas sejam alcançadas com rigor e adaptabilidade.

## **4. EDA E PRÉ-PROCESSAMENTO DOS DADOS**

Este capítulo descreve o processo de análise exploratória e pré-processamento dos dados de emissões de CO<sub>2</sub> per capita em toneladas métricas, para países da América Latina e Caribe, com o objetivo de preparar as informações para análises mais aprofundadas. As etapas foram realizadas em Python, utilizando a biblioteca *pandas*, e focaram na organização e limpeza dos dados, para que fossem consistentes e prontos para uso em modelos analíticos e visualizações.

## **4.1 Análise Exploratória de Dados (EDA) sobre Emissões de CO<sub>2</sub> Per Capita por País/Região**

### **Contextualização do Conjunto de Dados**

O conjunto de dados contém informações históricas sobre as emissões per capita de CO<sub>2</sub> (em toneladas métricas) de diversos países e regiões, abrangendo o período de 1990 a 2020. Esse dataset foi elaborado com o objetivo de analisar padrões de emissões ao longo do tempo, identificar diferenças regionais e contribuir para a formulação de estratégias de redução de emissões globais.

### **Estrutura do Conjunto de Dados**

O dataset possui 266 observações (países ou regiões) e 33 atributos (colunas), sendo composto por:

- Categórico: Country Name, que identifica os países ou regiões.
- Numérico: Colunas de 1990 a 2020 contendo valores contínuos que representam as emissões per capita de CO<sub>2</sub>. Foi criada uma nova coluna, Total Emissions, que calcula a soma total das emissões de cada país/região ao longo do período disponível.

### **Informações dos Atributos**

Nomes das Colunas: Cada coluna representa um ano específico (1990-2020), enquanto a coluna Country Name identifica o país ou a região correspondente.

### **Tipos de Dados:**

- Country Name: Dados categóricos.
- Colunas de 1990 a 2020: Dados numéricos contínuos.
- Total Emissions: Também numérica contínua, criada para consolidar as emissões acumuladas.

- Estatísticas de Valores Faltantes: Algumas colunas referentes a anos específicos contêm valores ausentes (NaN), indicando a ausência de dados históricos para determinados países.

A análise descritiva dos dados revelou informações importantes sobre as emissões per capita globais, destacando uma variabilidade significativa entre países e regiões. Enquanto algumas regiões apresentam emissões muito baixas ou até mesmo nulas, outras exibem níveis substancialmente elevados. Esse contraste ressalta disparidades regionais nas contribuições para as emissões globais, sendo que os dados categóricos da coluna Country Name confirmam a unicidade de cada entrada, com 266 valores distintos. Adicionalmente, os valores médios, medianos, mínimos, máximos e os desvios padrão calculados para os anos de 1990 a 2020 oferecem insights preliminares sobre a dispersão e a concentração das emissões, auxiliando na identificação de tendências e outliers.

No processo de limpeza e transformação de dados, etapas fundamentais foram realizadas para garantir a consistência e relevância da análise. A remoção de atributos irrelevantes, como as colunas Country Code, Indicator Name e Indicator Code, foi essencial, pois essas informações não contribuíam para os objetivos analíticos. De forma semelhante, colunas de anos fora do período de interesse (1960-1989 e 2021-2023) foram descartadas devido à ausência de dados consistentes. Além disso, uma nova métrica consolidada, a coluna Total Emissions, foi criada para representar a soma das emissões per capita de cada país ou região entre 1990 e 2020, permitindo comparações simplificadas e diretas. Outros ajustes, como a conversão dos nomes das colunas de anos para o formato numérico, facilitaram a manipulação dos dados para análises futuras.

Em relação ao tratamento de valores ausentes, foi identificado que algumas lacunas estavam concentradas em anos específicos ou em países com registros incompletos. Esse aspecto será tratado cuidadosamente para minimizar interferências nos resultados. Adicionalmente, a análise exploratória revelou disparidades marcantes na distribuição das emissões ao longo dos anos e entre os países. Para variáveis numéricas, foram identificados valores extremos que possivelmente refletem nações altamente industrializadas ou com grande dependência de combustíveis fósseis. Regiões como "Africa Eastern and Southern" foram destacadas pela sua significativa contribuição para



as emissões totais, enquanto países menores, como Aruba, registraram emissões desprezíveis ou inexistentes.

Esses insights iniciais, complementados pelas estatísticas detalhadas geradas com a função *dfSummary*, apontam para a necessidade de uma análise mais aprofundada. A visualização e o estudo das disparidades regionais serão cruciais para entender melhor as dinâmicas subjacentes às emissões globais e para informar políticas públicas mais direcionadas e eficazes.

### Estrutura Final do Dataset

Após o tratamento, o dataset apresenta a seguinte estrutura:

- Atributo categórico: Country Name: Nome do país ou região.
- Atributos numéricos (1990 a 2020): Representam emissões de CO<sub>2</sub> per capita em toneladas métricas, variando em precisão decimal.
- Dimensões do dataset: 264 observações (países/regiões) e 32 atributos.

Um exemplo da estrutura final dos dados é apresentado abaixo:

Country Name	1990	1991	1991	...	2020
Aruba	NaN	NaN	NaN		NaN
Africa Eastern and Southern	0.982975	0.942212	0.907936		0.795420
Afghanistan	0.191389	0.180674	0.126157		0.223479
Africa Western and Central	0.470111	0.521084	0.558013		0.463150
Angola	0.554941	0.545807	0.544413		0.592743

A função *dfSummary()* da biblioteca *summarytools* em Python oferece uma ferramenta poderosa para a análise exploratória de dados. Ao aplicar essa função a um data frame, obtém-se um resumo detalhado de cada variável, incluindo informações como

tipo de dado, valores ausentes, distribuição de frequência e estatísticas descritivas. Essa funcionalidade permite uma rápida compreensão da natureza dos dados e facilita a identificação de padrões, outliers e outras características relevantes para a análise subsequente.

Durante a década de 2010 a 2020, os dados relacionados às emissões globais de gases de efeito estufa revelam tendências importantes que merecem análise detalhada. Primeiramente, observa-se uma redução progressiva na média das emissões globais, que passou de 4,4 em 2010 para 3,9 em 2020. Essa diminuição de aproximadamente 11,4% ao longo da década sugere que políticas ambientais implementadas em diversos países tiveram impacto, ainda que moderado, na redução geral das emissões. Contudo, o ritmo dessa redução aponta para a necessidade de intensificar esforços globais para atender às metas climáticas internacionais.

A mediana das emissões, por outro lado, manteve-se relativamente estável durante o período, variando de 2,7 em 2010 para 2,5 em 2020. Essa pequena variação, de apenas 0,2 unidades (7,4%), sugere que a maior parte dos países não experimentou mudanças significativas em seus níveis de emissões. A persistência de uma mediana consideravelmente inferior à média indica a presença de outliers, ou seja, países com emissões extremamente altas que distorcem a média para cima. Isso demonstra que, embora esforços tenham sido direcionados aos grandes emissores, o impacto nas emissões típicas da maioria dos países foi menos expressivo.

A análise do desvio padrão (SD) reforça a tendência de redução nas diferenças extremas entre países emissores. O SD caiu de 5,2 em 2010 para 4,4 em 2020, refletindo uma menor dispersão das emissões globais e, possivelmente, uma convergência nos esforços de mitigação climática. Essa menor variabilidade é um indicativo de que países com emissões elevadas estão começando a alinhar-se com os níveis globais médios, embora ainda haja um caminho a percorrer.

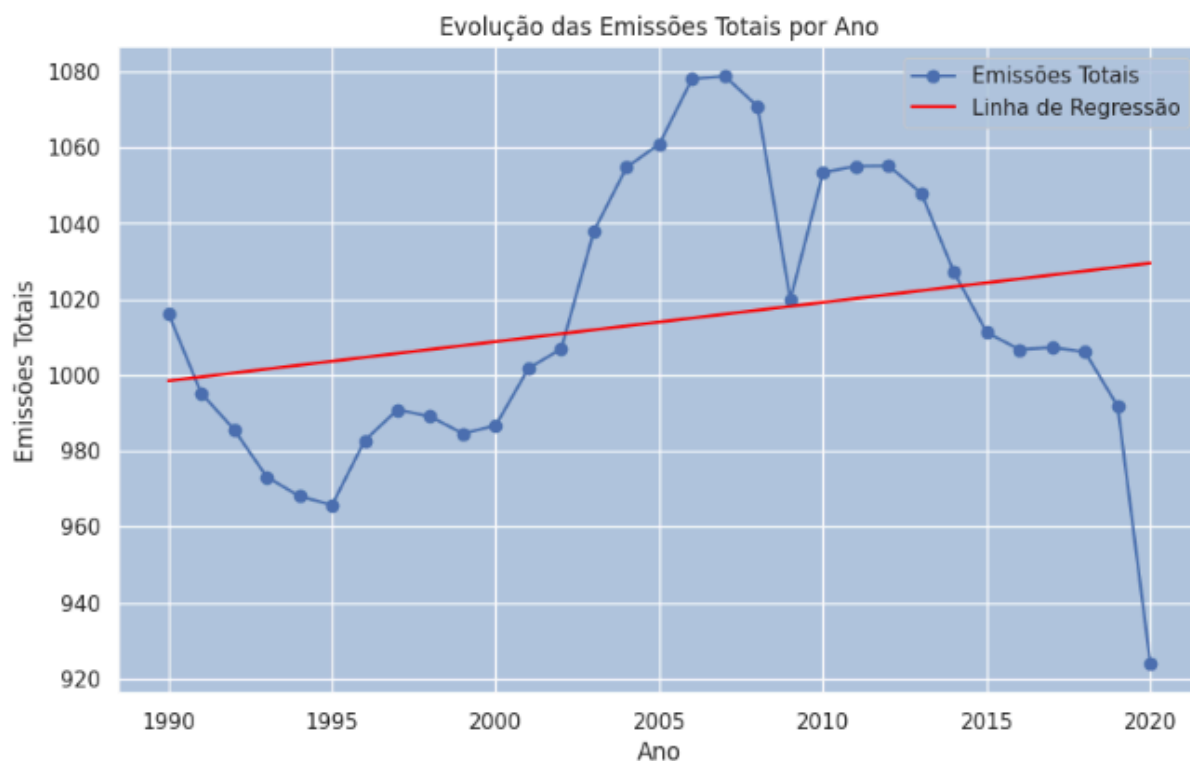
Um ponto importante a ser destacado é a redução nos valores máximos de emissões anuais. Enquanto em 2010 o valor máximo era de 35,5, ele caiu para 31,7 em 2020, refletindo uma redução significativa nos maiores emissores globais. Essa mudança é um sinal positivo de que os esforços de mitigação estão surtindo efeito nos países mais industrializados, que tradicionalmente lideram os níveis de emissões. Entretanto, o valor

mínimo manteve-se constante em 0,0 ao longo da década, evidenciando que muitos países ou regiões continuam apresentando emissões muito baixas ou inexistentes, como é o caso de nações menos industrializadas.

Ao longo da década, o intervalo interquartil (IQR), que mede a dispersão entre os valores do primeiro e terceiro quartis, manteve-se estável em torno de 5,0 a 5,5. Essa estabilidade demonstra que as emissões típicas da maioria dos países pouco variaram ao longo do período, reforçando a interpretação de que os grandes emissores são os principais responsáveis pelas mudanças nas métricas globais.

Nos anos finais da década, entre 2018 e 2020, percebe-se uma aceleração na redução das emissões médias, que passaram de 4,2 para 3,9. Além disso, a mediana também reduziu de 3,0 para 2,5, indicando que, embora lentamente, os esforços começaram a impactar um número maior de países, inclusive aqueles com emissões centrais. Esse movimento sugere uma maior adesão às políticas climáticas globais nos anos recentes.

Conclui-se, portanto, que a década de 2010 a 2020 foi marcada por avanços significativos na redução das emissões globais, especialmente entre os maiores emissores. Contudo, a estabilidade dos valores medianos e do intervalo interquartil demonstra que, para a maioria dos países, as mudanças ainda foram insuficientes. A persistência de outliers e a discrepância entre média e mediana destacam a importância de direcionar esforços para grandes economias emissoras e para o fortalecimento das políticas globais de mitigação. Esses resultados reforçam a necessidade de ações coordenadas e abrangentes para alcançar reduções mais expressivas nos próximos anos.



*Figura 1 - Evolução das emissões totais por ano*

### **Evolução Temporal das Emissões Totais e Análise Regional na América do Sul**

O gráfico acima apresenta a evolução das emissões totais de CO<sub>2</sub> ao longo do período de 1990 a 2020, demonstrando uma análise descritiva e a modelagem de tendência por meio de regressão linear simples. A curva azul representa os valores observados das emissões ao longo dos anos, enquanto a linha vermelha evidencia a tendência ajustada pelo modelo de Mínimos Quadrados Ordinários (OLS). Essa abordagem permitiu identificar a direção geral das emissões no tempo, destacando padrões que sugerem estabilidade ou alterações significativas no período analisado.

A análise das emissões totais é sustentada por uma etapa crucial de agregação dos dados. Por meio do cálculo da soma dos valores numéricos em cada ano, foi criada uma linha adicional no conjunto de dados, denominada "*Total Emissions*". Essa linha representa a soma consolidada das emissões per capita de todos os países e regiões para cada ano do período analisado. A introdução dessa métrica fornece uma visão macro das emissões globais, eliminando as variabilidades individuais de países e

permitindo a observação de tendências gerais. Além disso, o cálculo e inclusão da linha "Total" facilitam análises comparativas temporais e modelagens preditivas, como a regressão linear apresentada.

Além dessa análise global, os dados foram segmentados para focar nos países da América do Sul. Utilizando uma lista pré-definida de nações sul-americanas, foi possível criar um subconjunto de dados específicos para essa região. A filtragem produziu um novo *DataFrame* contendo apenas os países de interesse, permitindo comparações regionais e análises específicas. Essa abordagem não apenas facilita a visualização das tendências locais, mas também revela particularidades associadas à América do Sul, como o impacto de políticas ambientais regionais, padrões de industrialização e o uso de recursos naturais.

A separação por regiões oferece insights valiosos, como a capacidade de contrastar a tendência global com as emissões específicas da América do Sul. Por exemplo, enquanto as emissões globais podem mostrar estabilidade ou crescimento, os países sul-americanos podem exibir padrões diferentes, refletindo fatores locais como desmatamento, dependência de combustíveis fósseis ou a adoção de energias renováveis. Essa análise regional, portanto, complementa a visão global, permitindo uma interpretação mais detalhada das dinâmicas subjacentes às emissões.

Esse estudo, ao articular tendências globais e regionais, não apenas aprofunda a compreensão das emissões de CO<sub>2</sub>, mas também fornece uma base sólida para futuras pesquisas e para a formulação de políticas públicas específicas. O gráfico apresentado, combinado com a métrica "Total Emissions", exemplifica a relevância de explorar as emissões por meio de perspectivas integradas, enriquecendo a narrativa sobre a evolução das emissões e os desafios associados ao enfrentamento das mudanças climáticas.

**Foco na América do Sul e Comparação Regional**

A tabela abaixo é um recorte da análise presente no código:

Country Name	2012	2013	2014	2015	(...)	2020	Total Emissions
Argentina	4.264111	4.342250	4.209112	4.301914		3.405618	105.670222
Bolivia	1.604682	1.686988	1.784219	1.803674		1.539465	37.644109

Brazil	2.271418	2.413447	2.514592	2.365361		1.942523	52.962273
Chile	4.521603	4.740083	4.328593	4.603236		4.395151	106.234779
Colombia	1.533846	1.680988	1.715821	1.719390		1.552259	43.203476
Ecuador	2.395078	2.525035	2.619452	2.555603		1.957575	59.356498
Guyana	2.689370	2.643226	2.747101	2.754761		3.472771	67.704955
Peru	1.617522	1.656254	1.751594	1.776301		1.398566	36.957167
Paraguay	0.874070	0.882702	0.924029	1.030682		1.144591	24.306140
Suriname	4.374811	4.060328	4.599057	4.733655		4.285077	107.161950
Uruguay	2.526658	2.176045	1.916582	1.956379		1.899719	51.100587
Venezuela, RB	6.146121	6.046735	5.936295	5.343217		2.545028	145.526692

Para compreender melhor as emissões de CO<sub>2</sub> per capita e possibilitar uma análise comparativa detalhada, os dados foram segmentados em dois grandes grupos: países da América do Sul e outras nações. Essa segmentação foi realizada a partir de uma lista que inclui 12 países sul-americanos, como Argentina, Brasil, Chile, e Venezuela, permitindo isolar e observar os padrões específicos da região. A partir dessa filtragem, foi criado um novo conjunto de dados denominado `df_america_do_sul`, contendo exclusivamente os registros referentes a esses países, enquanto os dados das demais regiões foram alocados no `DataFrame` `df_outros_paises`.

### Dados da América do Sul: Um Olhar Regional

Na tabela acima, é possível observar os dados de emissões per capita para três países sul-americanos: Argentina, Bolívia e Brasil. Cada linha do *DataFrame* representa as emissões per capita de um país ao longo do período de 1990 a 2020, além do valor acumulado, armazenado na coluna *Total Emissions*. Esses valores consolidados fornecem uma visão clara da soma das emissões per capita ao longo de três décadas, facilitando análises comparativas e destacando tendências regionais.

- Argentina: Apresenta um total de emissões de aproximadamente 105.67, sendo uma das maiores na região, com uma média per capita relativamente constante ao longo dos anos. Notam-se valores elevados, especialmente a partir da década de 2010.

- Bolívia: Apresenta emissões totais acumuladas de 37.64, com valores per capita significativamente menores do que os observados na Argentina e no Brasil. No entanto, observa-se um aumento gradual a partir do início dos anos 2000, provavelmente relacionado ao crescimento industrial e ao consumo energético.
- Brasil: Com emissões acumuladas de 52.96, o Brasil ocupa uma posição intermediária entre os países analisados. Observa-se um crescimento lento mas consistente nas emissões per capita até 2014, seguido de uma ligeira redução nos anos subsequentes. Isso pode estar associado a esforços para diversificar a matriz energética, com maior uso de energias renováveis.

### **Importância do Recorte Regional**

A análise isolada dos países da América do Sul é essencial para identificar particularidades locais que podem ser mascaradas em uma abordagem global. Por exemplo, enquanto países como a Argentina destacam-se pelo alto volume de emissões, a Bolívia apresenta uma dinâmica distinta, com menores contribuições. O Brasil, foco desta análise, apresenta um comportamento intermediário, mas com peculiaridades como uma dependência histórica de hidrocarbonetos e esforços recentes de mitigação de emissões.

Essa segmentação permitirá comparações diretas e insights detalhados sobre o papel do Brasil em relação aos seus vizinhos sul-americanos. Além disso, os dados regionais são fundamentais para explorar relações entre políticas locais e padrões de emissão, auxiliando na formulação de estratégias mais eficazes de redução de emissões.

A análise a seguir aborda a avaliação das emissões de CO<sub>2</sub> per capita de países da América do Sul ao longo de um período de 30 anos, entre 1990 e 2020. O objetivo principal é entender as tendências dessas emissões, utilizando uma abordagem visual que permita uma comparação clara entre as diferentes nações da região. Inicialmente, foi realizada uma limpeza nos dados, removendo informações irrelevantes que poderiam interferir na análise, como os dados de emissões totais, focando exclusivamente nas emissões per capita, o que proporciona uma visão mais precisa das emissões por indivíduo em cada país.

A partir dessa limpeza, foi desenvolvida uma metodologia para gerar gráficos individuais para cada país, representando as emissões de CO<sub>2</sub> per capita ao longo dos anos. Para isso, foi considerado o comportamento de cada nação de forma isolada, permitindo uma análise detalhada das emissões em cada país. A construção desses gráficos foi complementada com a aplicação de uma regressão linear, que ajuda a visualizar a tendência das emissões de CO<sub>2</sub> ao longo do tempo. A linha de regressão, ajustada com base nos dados de emissões per capita, permite identificar se as emissões estão aumentando, diminuindo ou permanecendo estáveis ao longo dos anos, proporcionando uma visão clara da evolução das emissões de CO<sub>2</sub>.

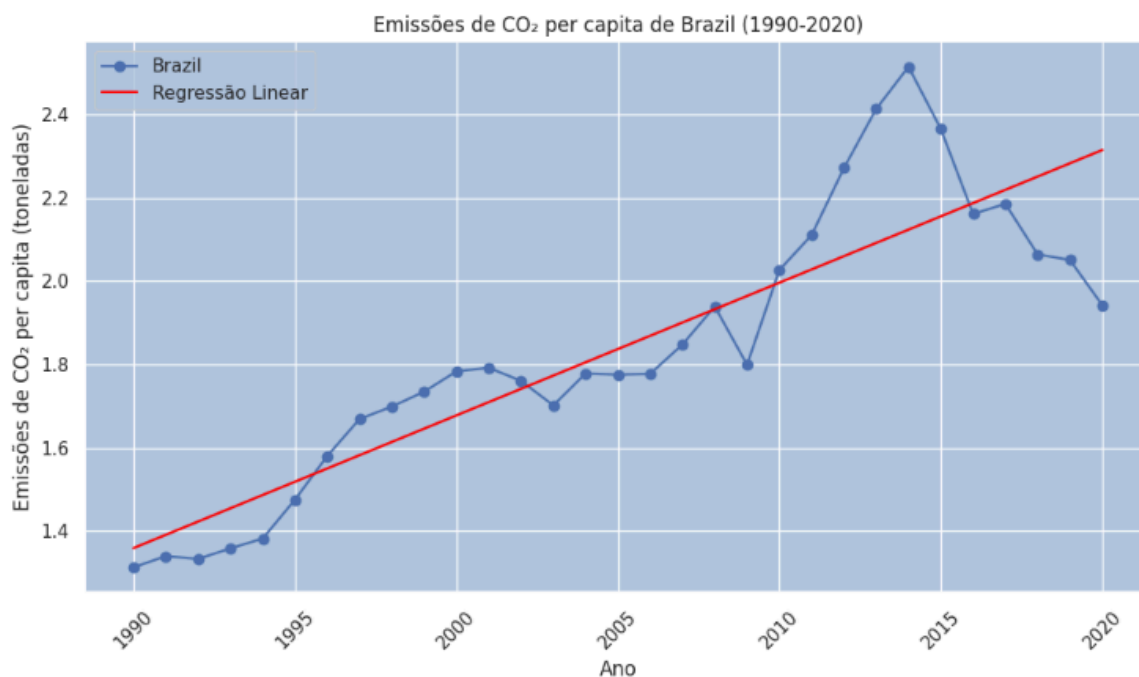
A análise considera o impacto do tempo nas emissões de CO<sub>2</sub>, observando se os países apresentaram um aumento ou uma diminuição nas emissões per capita, o que é crucial para entender o comportamento desses países no contexto da mudança climática. A utilização da regressão linear oferece uma representação estatística robusta das tendências, permitindo prever como as emissões podem se comportar no futuro, com base nas trajetórias observadas. A visualização das emissões ao longo do tempo, combinada com a linha de regressão, facilita a interpretação dos dados, tornando evidente a evolução das emissões em cada país e destacando os padrões regionais.

Além disso, o processo de salvar os gráficos gerados como imagens individuais facilita o armazenamento e a comparação entre os países, permitindo que os resultados sejam acessados e analisados separadamente. Isso oferece uma maneira prática de documentar a evolução das emissões de CO<sub>2</sub> per capita ao longo do período estudado, sendo uma ferramenta valiosa para a análise de políticas ambientais. A metodologia aplicada, ao gerar gráficos claros e ao mesmo tempo fornecer uma análise estatística das tendências, proporciona um recurso eficaz para pesquisadores e formuladores de políticas públicas na área de meio ambiente.

Essa abordagem de análise é essencial para compreender a dinâmica das emissões de CO<sub>2</sub> na América do Sul e para avaliar os esforços dos países em mitigar os impactos das mudanças climáticas. A visualização das tendências ao longo do tempo, combinada com modelos de regressão, oferece uma forma intuitiva e eficaz de identificar padrões, comparar nações e, finalmente, tomar decisões informadas com base nas evidências empíricas geradas pela análise. Em um contexto global de crescente



preocupação com as mudanças climáticas, estudos como este são fundamentais para a compreensão do papel de cada país nas emissões de gases de efeito estufa e no desenvolvimento de estratégias para a redução das emissões.



*Figura 2 - Emissões de CO<sub>2</sub> per capita do Brasil*

O gráfico apresentado reflete a evolução das emissões de CO<sub>2</sub> per capita no Brasil ao longo de um período de 30 anos, entre 1990 e 2020. A análise visual do gráfico revela informações importantes sobre a tendência das emissões, que, em termos gerais, mostram uma trajetória de crescimento durante o período analisado. A linha de regressão linear, representada pela linha vermelha, sugere que as emissões de CO<sub>2</sub> per capita aumentaram ao longo dos anos, indicando que, em média, cada brasileiro emitiu mais CO<sub>2</sub> à medida que o tempo passou. Este padrão de crescimento pode refletir uma série de fatores interligados, como o aumento das atividades industriais e o consumo de energia no país.

Embora o gráfico mostre uma tendência geral de crescimento, é importante observar que as emissões não aumentaram de forma linear e constante ao longo dos anos. As variações anuais nas emissões são visíveis, com alguns anos apresentando aumentos mais acentuados, enquanto outros mostraram um crescimento mais modesto. Isso sugere que as emissões de CO<sub>2</sub> per capita no Brasil foram influenciadas por uma

série de variáveis que variaram de ano para ano, como mudanças na economia, políticas públicas e fatores externos, como crises econômicas e alterações no consumo de energia.

Um ponto notável no gráfico é a presença de um ponto de inflexão por volta de 2010. A partir desse ano, o crescimento das emissões parece ter se acelerado, indicando uma possível mudança no padrão de emissão do Brasil. Esse período de aceleração no crescimento das emissões pode estar relacionado a uma série de fatores, incluindo a expansão de setores industriais de alto impacto ambiental e um aumento no consumo de combustíveis fósseis, particularmente petróleo e carvão. O pico máximo das emissões foi registrado em torno de 2015, representando o maior valor de emissões de CO<sub>2</sub> per capita durante o período estudado, o que destaca um ponto crítico para a política ambiental do país.

A partir de 2015, o gráfico sugere uma leve tendência de estabilização nas emissões, e em alguns casos, um pequeno declínio nas emissões per capita. No entanto, é importante ressaltar que ainda é cedo para afirmar com certeza que essa tendência será mantida nos próximos anos. A estabilização ou redução das emissões pode ser influenciada por diversas políticas públicas, como o incentivo a fontes de energia mais limpas, além de fatores econômicos e sociais que podem ter desencorajado a expansão de atividades de grande impacto ambiental.

Diversos fatores podem ter influenciado a evolução das emissões de CO<sub>2</sub> per capita no Brasil ao longo desse período. O crescimento econômico, particularmente nos anos 2000, foi um dos principais motores do aumento das emissões, uma vez que o desenvolvimento econômico geralmente está associado a um aumento no consumo de energia e à intensificação de atividades industriais, ambas fontes significativas de emissões de CO<sub>2</sub>. Além disso, a dependência do Brasil de fontes de energia fósseis, como petróleo e carvão mineral, tem sido um fator crítico, já que a queima desses combustíveis gera grandes quantidades de emissões de CO<sub>2</sub>.

O desmatamento também desempenha um papel crucial nas emissões de CO<sub>2</sub> do Brasil. A destruição de florestas, especialmente na Amazônia, reduz a capacidade do país de absorver carbono, resultando em uma liberação significativa de CO<sub>2</sub> na atmosfera.

Assim, o desmatamento é um fator-chave que contribui diretamente para o aumento das emissões de gases de efeito estufa no Brasil. Além disso, as políticas públicas implementadas ao longo do período estudado tiveram um impacto direto nas emissões, seja para aumentar ou reduzir as emissões de CO<sub>2</sub>. A adoção de políticas ambientais mais rigorosas ou de incentivos a energias renováveis pode ter ajudado a mitigar o crescimento das emissões em alguns momentos, enquanto políticas menos eficazes podem ter agravado a situação.

Por fim, crises econômicas, como a crise financeira de 2008<sup>(6)</sup>, também podem ter influenciado as emissões de CO<sub>2</sub> no Brasil. A desaceleração econômica durante esses períodos pode ter levado a uma redução temporária nas emissões, uma vez que a diminuição das atividades econômicas geralmente resulta em menor demanda por energia e menos atividades industriais intensivas em carbono. Assim, os dados refletem a interação de uma série de fatores econômicos, sociais, políticos e ambientais que moldaram as emissões de CO<sub>2</sub> per capita no Brasil ao longo das três décadas analisadas.

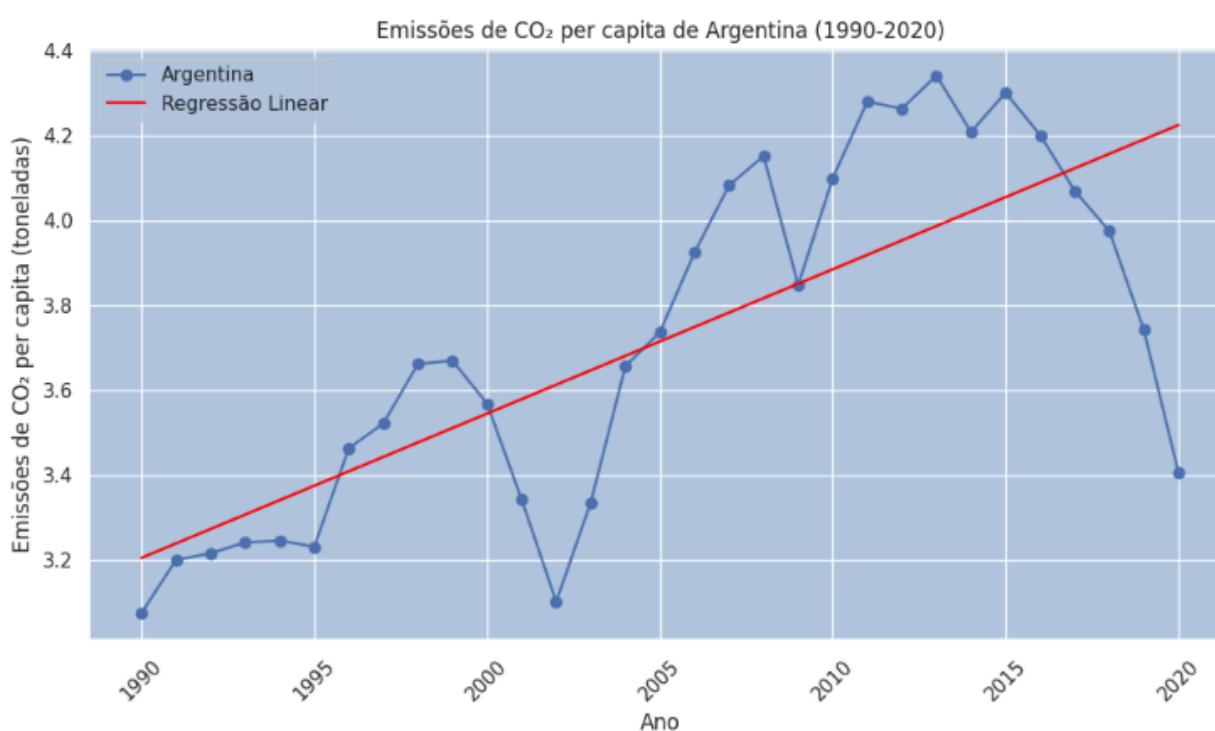


Figura 3 - Emissões de CO<sub>2</sub> per capita de Argentina

A análise comparativa das emissões de CO<sub>2</sub> per capita entre Brasil e Argentina revela tanto semelhanças quanto diferenças significativas ao longo do período de 1990 a 2020. Em termos gerais, ambos os países apresentam uma tendência de crescimento das emissões de CO<sub>2</sub> per capita, embora com variações anuais notáveis. Essa tendência de crescimento é claramente evidenciada pela linha de regressão linear nos gráficos de ambos os países, que sugere um aumento contínuo das emissões ao longo do tempo, refletindo uma expansão das atividades econômicas e, possivelmente, um aumento no consumo de combustíveis fósseis. No entanto, apesar de essa ser a tendência predominante, tanto no Brasil quanto na Argentina, as emissões flutuam de ano para ano, com picos e quedas que podem ser atribuídos a fatores como crises econômicas, mudanças climáticas e a implementação de políticas públicas de controle ambiental.

No entanto, as duas nações também apresentam algumas diferenças notáveis. Uma das principais diferenças observadas refere-se aos níveis absolutos das emissões de CO<sub>2</sub> per capita. A Argentina, em geral, apresenta níveis ligeiramente mais altos de emissões, especialmente nos últimos anos do período analisado. Essa diferença pode ser explicada por vários fatores, incluindo a matriz energética de cada país, com a Argentina possivelmente dependendo mais de fontes fósseis ou com uma diversificação energética menor, o que contribui para um aumento maior das emissões. Além disso, o ponto de inflexão, que marca a aceleração do crescimento das emissões, ocorre um pouco antes na Argentina do que no Brasil, sugerindo que a dinâmica de crescimento das emissões no país vizinho começou a se intensificar mais cedo. De forma semelhante, o pico das emissões na Argentina também parece ocorrer em um período ligeiramente anterior ao do Brasil, o que pode indicar um padrão de emissão distinto entre os dois países.

A tendência recente das emissões de CO<sub>2</sub> per capita também difere entre as duas nações. Enquanto o Brasil parece ter experimentado uma leve estabilização ou até um pequeno declínio nas emissões após 2015, a Argentina apresenta uma possível desaceleração mais acentuada ou até um declínio mais expressivo nas emissões nos últimos anos do período. Esse comportamento pode refletir diferenças nas políticas públicas adotadas em cada país ou variações na dinâmica econômica e industrial. A Argentina, por exemplo, pode ter implementado políticas de mitigação mais eficazes ou

pode estar enfrentando mudanças estruturais em sua economia que reduziram a intensidade das emissões.

As diferenças observadas entre os dois países podem ser explicadas por uma série de fatores, como a matriz energética, a estrutura do setor industrial e as políticas climáticas implementadas. A matriz energética dos dois países pode influenciar de maneira significativa as emissões de CO<sub>2</sub>. A dependência de fontes de energia fósseis, como petróleo e gás natural, na Argentina, pode ser um fator determinante para as emissões mais altas em comparação com o Brasil, que possui uma maior participação de fontes renováveis em sua matriz energética. Além disso, a intensidade energética dos processos industriais nos dois países pode variar, afetando diretamente as emissões. A estrutura industrial de cada país, incluindo o uso de tecnologias mais limpas ou mais poluentes, também pode explicar as disparidades nas emissões de CO<sub>2</sub> per capita.

Outros fatores que podem ter impactado as emissões incluem as políticas públicas de cada país voltadas à mitigação das mudanças climáticas. A implementação de políticas ambientais mais rígidas ou a adoção de tecnologias de baixo carbono pode ter contribuído para a redução das emissões na Argentina nos últimos anos, enquanto no Brasil o impacto dessas políticas pode ter sido mais limitado ou demorado para se refletir nos dados. Fatores econômicos, como as crises financeiras e os ciclos de crescimento de cada país, também desempenham um papel importante nas emissões, com a desaceleração econômica possivelmente resultando em uma redução temporária da demanda por energia e, conseqüentemente, nas emissões de CO<sub>2</sub>.

#### **4.1 Análise Exploratória de Dados (EDA) relacionada à energia renovável**

Dando continuidade à nossa análise, agora abordaremos os dados relacionados à energia renovável, os quais foram obtidos no site do World Bank Group, especificamente na seção "Renewable energy consumption (% of total final energy consumption)" para a região da América Latina e Caribe. Esses dados estão disponíveis no arquivo Excel localizado em nosso diretório com o caminho `"/content/API_EN.ATM.CO2E.PC_DS2_en_excel_v2_112.xls"`, que, assim como o dataset anterior, faz parte dos dados utilizados nesta análise.

O processo de análise dos dados de energia renovável segue etapas semelhantes ao tratamento dos dados de emissões de CO<sub>2</sub>. Primeiramente, o código carrega os dados a partir do arquivo Excel, garantindo que a informação necessária seja importada para o ambiente de trabalho. Após o carregamento dos dados, procede-se com a limpeza do conjunto de dados. Nessa fase, são removidas linhas e colunas que são consideradas desnecessárias para a análise, como aquelas que não contêm dados relevantes ou que apenas adicionam complexidade sem oferecer informações significativas.

Em seguida, são realizados ajustes nos nomes das colunas para torná-los mais legíveis e compreensíveis, facilitando o trabalho nas etapas subsequentes de análise. A renomeação de colunas é uma etapa fundamental, pois garante que os nomes dos atributos estejam em conformidade com a terminologia utilizada na análise e a clareza necessária para interpretação. Além disso, em alguns casos, o código pode também envolver a conversão do tipo de dados das colunas para garantir que os valores sejam tratados corretamente, como a conversão de dados de texto em valores numéricos ou de datas, conforme a necessidade da análise.

Ao final desse processo de preparação de dados, o código exibe uma prévia das primeiras linhas do conjunto de dados limpo, oferecendo uma visão inicial das informações que serão analisadas. Esse processo de limpeza e preparação dos dados é crucial, pois assegura que a análise posterior seja realizada com dados precisos, consistentes e organizados, evitando erros ou interpretações equivocadas que poderiam surgir de dados desordenados ou mal estruturados. Além disso, a qualidade dos dados limpos influencia diretamente a qualidade das visualizações e conclusões que podem ser extraídas a partir da análise, tornando essa etapa essencial para o sucesso da investigação.

```
[ ] 1 data_reno = data_reno.drop(columns=colunas_a_remover)
    2
    3 def soma_emissoes_por_pais(data_reno):
    4     data_reno['Total Energia Renovavel'] = data_reno.iloc[:, 4:].sum(axis=1)
    5     return data_reno
    6
    7 data_reno = soma_emissoes_por_pais(data_reno)
    8 data_reno.head(5)
```

#### *Figura 4 - Trecho do código*

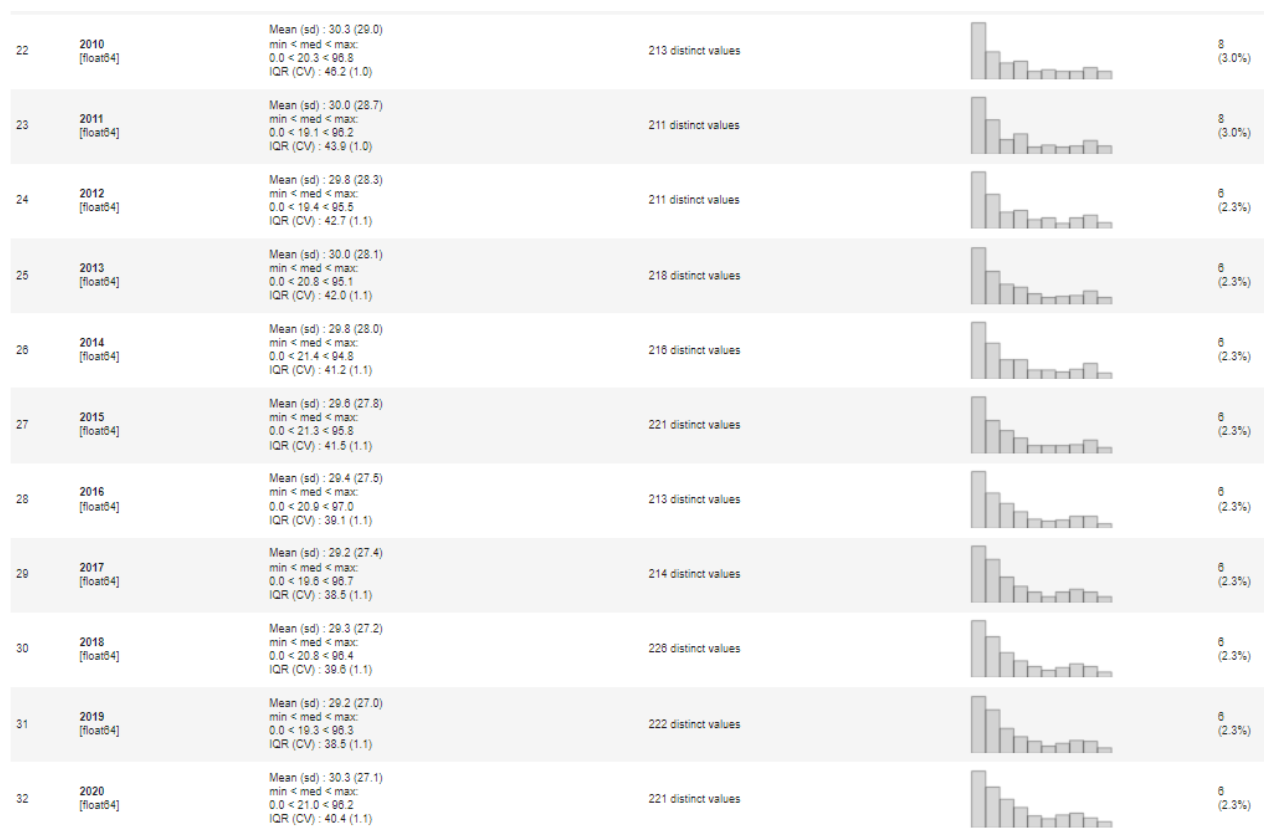
O código apresentado realiza a transformação de um DataFrame contendo dados sobre energia renovável por país e ano, com o objetivo de facilitar a análise e a comparação dos níveis de energia renovável entre diferentes países. A transformação envolve a criação de uma nova coluna no DataFrame, chamada "Total Energia Renovável", que agrega o total de energia renovável consumida por cada país ao longo dos anos analisados. Essa soma é crucial, pois oferece uma visão geral do consumo de energia renovável por país, facilitando comparações entre os mesmos.

Inicialmente, o código remove colunas indesejadas, que não são necessárias para o cálculo ou análise dos dados de energia renovável. A remoção dessas colunas é realizada através da função `drop()`, com a especificação de quais colunas devem ser excluídas com base na lista fornecida, chamada `colunas_a_remover`.

Na sequência, é definida a função `soma_emissoes_por_pais`, que calcula o total de energia renovável consumida por cada país. Essa função seleciona todas as colunas a partir da quinta (que representam os anos de 1960 a 2020, por exemplo), e usa a função `sum(axis=1)` para somar os valores de cada linha (relativa a um país) ao longo dos anos. O resultado é armazenado na nova coluna "Total Energia Renovável".

Após a aplicação dessa função, o DataFrame resultante é armazenado na variável `data_reno`, e os primeiros cinco registros são exibidos para que seja possível verificar a transformação realizada. A nova coluna "Total Energia Renovável" fornece uma visão mais clara do total de energia renovável consumida por cada país durante o período considerado, permitindo análises comparativas e facilitando a visualização de tendências gerais de consumo de energia renovável.

Com essa abordagem, o DataFrame resultante pode ser utilizado em análises subsequentes, como a avaliação de padrões de consumo de energia renovável, comparações entre países e identificações de possíveis correlações entre o consumo de energia renovável e outros indicadores econômicos ou ambientais.



*Figura 5 - Função dfSummary dos dados para os anos de 2010 a 2020*

A análise dos dados apresentados para os anos de 2010 a 2020, assim como para a variável "Total Energia Renovável", revela uma série de características importantes sobre a distribuição e variação dos dados. A média das emissões de energia renovável por ano se manteve relativamente constante, variando de 29,0 a 30,3, com um desvio padrão variando de 27,0 a 29,0. Esses valores indicam que, embora a média anual tenha sido estável, houve uma considerável dispersão nas emissões, refletida pelo desvio padrão elevado. Esse comportamento sugere que, apesar de um padrão geral, existiram flutuações notáveis nos dados de energia renovável ao longo do período analisado, com algumas observações muito distantes da média.

A distribuição dos dados, observada através dos valores mínimo, mediano e máximo, demonstra que, em todos os anos, o valor mínimo foi 0.0, o que pode indicar a presença de valores ausentes ou a ausência de geração de energia renovável em alguns países. A mediana dos valores foi consistentemente inferior ao valor máximo, indicando que, para a maioria dos países e anos, as emissões de energia renovável se concentraram em faixas mais baixas, mas com picos significativos em algumas



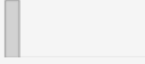
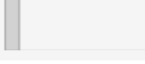

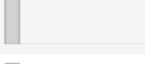

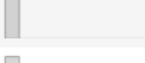
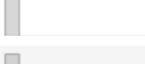
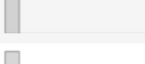

observações. O valor máximo de energia renovável em certos anos ultrapassou 90, evidenciando que países específicos ou anos isolados apresentaram quantidades notavelmente altas de geração de energia renovável.

Em relação ao índice de dispersão, medido pelo intervalo interquartil (IQR), observou-se uma variação entre 38,5 e 46,2 ao longo dos anos, o que denota uma amplitude considerável na dispersão dos dados centrais, com uma relativa consistência ao longo do tempo. O coeficiente de variação (CV) permaneceu constante em torno de 1.0, indicando que, embora os valores absolutos de energia renovável apresentassem variações de ano para ano, a dispersão relativa das emissões ao longo do período foi estável, sugerindo uma estrutura similar de variação entre os países.

Além disso, a quantidade de valores distintos nos dados foi bastante consistente, variando de 211 a 226 valores por ano, o que sugere uma boa diversidade nas observações. A percentagem de valores ausentes foi excepcionalmente baixa, em torno de 2,3% por ano, o que indica que o conjunto de dados é substancialmente completo e confiável para análises subsequentes. Este baixo nível de dados ausentes contribui para a qualidade da base de dados, permitindo análises mais robustas e seguras.

Por fim, a análise do "Total Energia Renovável", que inclui o somatório das emissões ao longo do período analisado, revela uma média de 842,3, com um desvio padrão de 813,0. Este alto desvio padrão indica uma grande variação nas emissões totais entre os países, com alguns apresentando valores elevados, como o máximo de 2707,5, enquanto outros apresentaram valores muito próximos de zero. Esses extremos destacam as desigualdades na produção e consumo de energia renovável entre diferentes nações, sendo um reflexo das distintas políticas energéticas e capacidades de geração de energia renovável nos países analisados.

Em suma, os dados demonstram uma variação considerável na geração de energia renovável entre os países, com uma concentração maior de valores baixos, mas com algumas nações apresentando contribuições significativamente altas. O comportamento consistente dos coeficientes de variação e a pequena quantidade de valores ausentes sugerem que os dados são de boa qualidade, permitindo uma análise detalhada e confiável da evolução das emissões de energia renovável ao longo do tempo.

21	2010 [float64]	Mean (sd) : 58.5 (476.8) min < med < max: 0.0 < 19.6 < 7808.5 IQR (CV) : 48.0 (0.1)	214 distinct values		0 (0.0%)
22	2011 [float64]	Mean (sd) : 58.0 (473.1) min < med < max: 0.0 < 18.8 < 7745.3 IQR (CV) : 44.0 (0.1)	212 distinct values		0 (0.0%)
23	2012 [float64]	Mean (sd) : 58.1 (473.4) min < med < max: 0.0 < 18.3 < 7750.8 IQR (CV) : 43.7 (0.1)	212 distinct values		0 (0.0%)
24	2013 [float64]	Mean (sd) : 58.4 (475.3) min < med < max: 0.0 < 19.5 < 7798.2 IQR (CV) : 42.5 (0.1)	219 distinct values		0 (0.0%)
25	2014 [float64]	Mean (sd) : 58.1 (473.4) min < med < max: 0.0 < 19.4 < 7750.9 IQR (CV) : 42.1 (0.1)	217 distinct values		0 (0.0%)
26	2015 [float64]	Mean (sd) : 57.7 (470.4) min < med < max: 0.0 < 21.0 < 7702.3 IQR (CV) : 42.1 (0.1)	222 distinct values		0 (0.0%)
27	2016 [float64]	Mean (sd) : 57.3 (467.2) min < med < max: 0.0 < 20.2 < 7648.9 IQR (CV) : 39.4 (0.1)	214 distinct values		0 (0.0%)
28	2017 [float64]	Mean (sd) : 56.9 (464.3) min < med < max: 0.0 < 18.5 < 7602.7 IQR (CV) : 38.9 (0.1)	215 distinct values		0 (0.0%)
29	2018 [float64]	Mean (sd) : 57.1 (465.6) min < med < max: 0.0 < 18.7 < 7623.7 IQR (CV) : 38.3 (0.1)	227 distinct values		0 (0.0%)
30	2019 [float64]	Mean (sd) : 56.9 (464.1) min < med < max: 0.0 < 18.9 < 7599.6 IQR (CV) : 38.7 (0.1)	223 distinct values		0 (0.0%)
31	2020 [float64]	Mean (sd) : 59.0 (480.7) min < med < max: 0.0 < 20.9 < 7872.1 IQR (CV) : 39.8 (0.1)	222 distinct values		0 (0.0%)
32	Total Energia Renovavel [float64]	Mean (sd) : 1078.2 (13883.5) min < med < max: 0.0 < 547.5 < 224038.7 IQR (CV) : 1281.8 (0.1)	254 distinct values		0 (0.0%)

*Figura 6 - Função dfSummary sobre emissões de energia renovável ao longo dos anos*

A análise dos dados apresentados revela uma série de informações importantes sobre as emissões de energia renovável ao longo dos anos. A base de dados foi tratada com a substituição dos valores ausentes por zero, o que permitiu calcular a soma total da energia renovável gerada a cada ano, e uma nova linha com os totais foi adicionada ao DataFrame. Com isso, foi possível observar a evolução anual da energia renovável gerada em diferentes países.

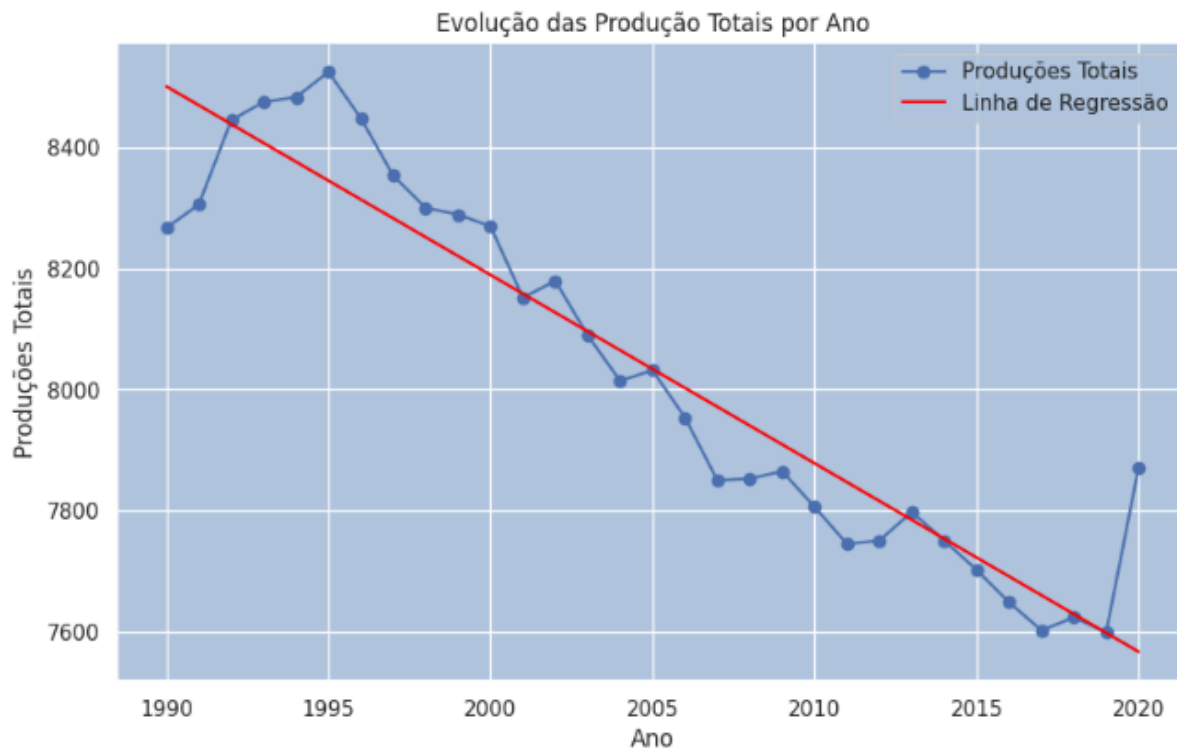
Para cada ano de 2010 a 2020, a média da energia renovável gerada por país variou entre 56,9 e 59,0, com um desvio padrão considerável, variando de 464,1 a 480,7. Esse alto desvio padrão sugere que as emissões de energia renovável variaram consideravelmente entre os países, com valores mínimos frequentemente próximos de zero e máximos que ultrapassaram 7.800 GWh em alguns anos. A mediana foi consistentemente abaixo do valor máximo, o que indica que a maior parte dos países gerou quantidades relativamente pequenas de energia renovável, com alguns casos

extremos de países com grandes contribuições. O coeficiente de variação (CV) permaneceu em torno de 0.1, refletindo uma dispersão moderada em relação à média das emissões por ano.

Além disso, foi observada uma boa diversidade nas emissões de energia renovável ao longo dos anos, com o número de valores distintos variando entre 211 e 227. A quantidade de dados ausentes foi praticamente inexistente, com 0% de valores ausentes em cada ano, o que indica que a base de dados está bem completa.

O total de energia renovável gerada ao longo de todo o período (2010-2020) apresentou uma média de 1.678,2 GWh, com um desvio padrão significativamente alto de 13.683,5 GWh, refletindo uma enorme variação nas contribuições individuais dos países. A mediana e os valores mínimos também apontam para a grande disparidade nas contribuições, com alguns países gerando apenas pequenas quantidades de energia renovável, enquanto outros apresentaram totais muito elevados, com o valor máximo atingindo mais de 224.000 GWh.

Dado esse panorama geral, nossa análise a seguir se concentrará nas informações relativas ao Brasil, com o intuito de compreender de forma mais detalhada a evolução das emissões de energia renovável no país e explorar as possíveis tendências e peculiaridades desse contexto, de maneira similar ao procedimento adotado na análise anterior.



*Figura 7 - Evolução das produções de energias renováveis por ano*

### **Análise da Evolução da Produção Total de Energia Renovável: Uma Perspectiva Histórica**

A análise da série temporal da produção total de energia renovável revela uma tendência de declínio ao longo do período analisado, com algumas flutuações anuais. A linha de regressão, que representa a tendência central dos dados, apresenta uma inclinação negativa, corroborando a observação visual do gráfico. Diversos fatores podem ter contribuído para essa trajetória decrescente, como a priorização de outras fontes energéticas, a falta de incentivos governamentais, instabilidades econômicas e condições climáticas adversas. No entanto, a partir de meados da primeira década do século XXI, observa-se uma aparente estabilização da produção em um patamar inferior, sugerindo uma possível adaptação do setor a novas condições ou a implementação de medidas para reverter a tendência de queda. Projeções futuras, baseadas na tendência histórica, indicam a persistência do declínio caso as condições atuais se mantenham. Contudo, é fundamental ressaltar que o futuro da produção de energia renovável é incerto e pode ser significativamente influenciado por mudanças nas políticas energéticas, avanços tecnológicos, e a crescente conscientização ambiental. Para reverter a tendência de

declínio e estimular o crescimento do setor, é crucial a implementação de políticas públicas que promovam o desenvolvimento das fontes renováveis, o investimento em pesquisa e desenvolvimento, e a conscientização da sociedade sobre a importância da transição para uma matriz energética mais limpa e sustentável.

	Country Name	1990	1991	1992	1993	1994	1995	1996	1997	1998	...	2012	2013	2014	2015	2016	2017	2018	2019	2020	Total Energia Renovavel
9	Argentina	8.6	8.4	9.0	9.7	10.0	10.2	8.6	9.7	9.6	...	8.6	8.9	9.8	9.4	9.4	10.4	10.5	10.7	9.8	267.5
28	Bolivia	37.0	38.1	36.9	36.9	35.2	33.2	32.4	32.0	30.6	...	17.4	16.0	15.0	14.2	13.6	14.3	14.3	13.5	16.9	655.7
29	Brazil	49.8	48.9	48.7	48.2	48.3	46.1	44.5	43.7	43.8	...	43.5	42.3	41.7	43.7	45.4	45.3	46.9	47.5	50.0	1270.0
39	Chile	34.0	37.3	38.6	35.8	34.8	34.2	32.2	30.0	29.8	...	30.3	30.2	27.0	25.1	24.6	24.1	25.5	25.3	26.7	838.1
45	Colombia	38.1	36.8	33.1	33.0	33.8	33.5	33.6	25.7	25.8	...	29.6	32.0	32.1	31.1	30.5	32.2	30.4	31.5	32.0	851.9
66	Ecuador	24.2	24.2	20.8	21.6	20.9	20.4	19.3	16.6	18.2	...	13.2	11.8	12.2	13.1	14.8	17.0	16.3	17.7	20.2	459.3
94	Guyana	55.1	41.6	43.2	43.1	32.3	31.0	29.0	27.4	27.6	...	28.7	26.9	24.0	25.0	21.3	22.1	16.8	11.3	12.0	802.7
186	Peru	39.4	39.9	37.3	37.1	35.9	33.3	31.7	31.7	39.4	...	30.8	28.4	28.6	27.4	27.2	27.6	27.9	27.1	31.6	943.6
195	Paraguay	78.5	79.2	75.7	72.6	69.9	70.2	70.8	69.5	67.4	...	63.4	62.5	62.5	60.5	62.5	60.1	59.2	60.1	61.5	1841.3
220	Suriname	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	15.8	17.6	12.8	11.6	11.1	15.5	18.3	14.9	14.6	416.5
250	Uruguay	44.8	43.7	43.3	42.7	42.2	41.0	37.5	37.7	36.9	...	47.9	51.9	56.7	59.4	60.3	60.8	60.8	59.3	61.2	1310.0
254	Venezuela, RB	12.0	13.2	14.2	13.2	13.8	13.2	12.9	13.9	13.7	...	13.8	15.4	12.1	15.4	14.6	14.3	14.6	15.8	31.2	420.6

12 rows x 33 columns

Figura 8 - Análise Comparativa da Produção de Energia Renovável

**Análise Comparativa da Produção de Energia Renovável: Brasil x Demais Países da América do Sul**

A análise da tabela de produção de energia renovável na América do Sul revela um cenário diversificado, com o Brasil se destacando como o principal produtor da região. A matriz energética brasileira, historicamente marcada pela predominância da hidroeletricidade, tem demonstrado uma capacidade significativa de geração de energia limpa. A comparação com os demais países da América do Sul evidencia a posição de destaque do Brasil nesse contexto.

Países como o Chile e o Uruguai, embora com produção menor, têm apresentado um crescimento expressivo na geração de energia eólica e solar, impulsionados por políticas públicas favoráveis e pelo potencial energético de seus territórios. A Argentina, por sua vez, possui uma matriz energética diversificada, com a contribuição de diversas fontes renováveis, incluindo biomassa e biogás.

No entanto, é importante ressaltar que a produção de energia renovável nos demais países da América do Sul ainda apresenta desafios, como a falta de infraestrutura, a instabilidade política e econômica, e a dependência de fontes fósseis.

Em resumo, a análise comparativa permite concluir que:

- O Brasil é líder na produção de energia renovável na América do Sul, impulsionado principalmente pela hidroeletricidade.
- Países como o Chile e o Uruguai têm apresentado um crescimento significativo em fontes eólicas e solares.
- A produção de energia renovável nos demais países da América do Sul ainda apresenta desafios e variações significativas.

A comparação entre o Brasil e os demais países da América do Sul evidencia a importância de políticas públicas consistentes e investimentos em infraestrutura para o desenvolvimento das fontes renováveis. O Brasil, com sua vasta experiência na área, pode servir como referência para outros países da região, compartilhando conhecimentos e tecnologias.

## **5. INTRODUÇÃO À MODELAGEM PREDITIVA: UTILIZANDO O MODELO SARIMAX**

Para projetar as emissões de CO<sub>2</sub> e a geração de energia renovável até 2030, utilizaremos um modelo avançado de séries temporais, o *SARIMAX* (Seasonal Auto-Regressive Integrated Moving Average with Exogenous Factors). Este modelo é uma extensão do ARIMA tradicional, projetado para lidar com dados que apresentam padrões sazonais e são influenciados por variáveis externas. A escolha do *SARIMAX* está fundamentada em sua capacidade de capturar a complexidade intrínseca dos dados temporais, possibilitando previsões robustas e alinhadas às necessidades do projeto.

Um dos principais diferenciais do *SARIMAX* é sua habilidade de incorporar padrões sazonais. Tanto as emissões de CO<sub>2</sub> quanto a geração de energia renovável tendem a exibir comportamentos cíclicos, influenciados por fatores como sazonalidade climática e flutuações na demanda energética. Ao modelar explicitamente esses padrões, o

SARIMAX permite identificar e prever essas variações com maior precisão. Além disso, o modelo permite a inclusão de variáveis exógenas, como políticas públicas, preços de combustíveis fósseis e avanços tecnológicos. Essas variáveis externas enriquecem o modelo ao adicionar informações contextuais que são fundamentais para capturar as dinâmicas subjacentes dos dados.

Outro ponto forte do SARIMAX é sua capacidade de lidar com séries temporais que possuem tanto tendências de longo prazo quanto ruídos. Esse equilíbrio é essencial para garantir que as previsões reflitam os comportamentos reais dos dados, filtrando interferências pontuais sem perder de vista as tendências mais amplas. Essa característica torna o modelo uma escolha confiável para projeções de longo prazo, como as requeridas para 2030, onde é necessário lidar com incertezas e dinâmicas complexas.

Além de sua precisão, o SARIMAX também oferece uma estrutura que facilita a interpretação de seus parâmetros. Isso possibilita uma compreensão clara de como diferentes fatores, internos e externos, influenciam os resultados. A capacidade de validar os parâmetros estatisticamente adiciona uma camada de confiabilidade ao modelo, assegurando que as previsões sejam fundamentadas em bases matemáticas robustas e teoricamente consistentes.

Portanto, o modelo SARIMAX é particularmente adequado ao objetivo deste projeto. Ele combina flexibilidade, precisão e interpretabilidade para fornecer projeções que são não apenas quantitativamente confiáveis, mas também qualitativamente informativas. Essas características fazem do SARIMAX uma ferramenta poderosa para entender e prever as emissões de CO<sub>2</sub> e a produção de energia renovável em um horizonte de longo prazo.

### **5.1. Análise dos Dados de Emissões de CO<sub>2</sub> no Brasil**

Para compreender o comportamento histórico das emissões de CO<sub>2</sub> no Brasil e desenvolver modelos preditivos robustos, foi necessário realizar uma série de procedimentos exploratórios e análises iniciais. Os resultados fornecem uma base sólida para ajustar e aplicar o modelo SARIMAX, mas também destacam características importantes da série temporal que devem ser tratadas.

## **1. Filtragem, Seleção e Visualização de Dados**

Inicialmente, os dados referentes ao Brasil foram extraídos do conjunto de dados global, resultando em uma série temporal com informações anuais desde 1990. Esse processo garantiu que nossa análise fosse focada exclusivamente no Brasil, permitindo uma visão detalhada do comportamento histórico de suas emissões. Ao visualizar os primeiros anos da série, verificou-se que os valores anuais das emissões apresentam variações moderadas, sugerindo um possível padrão de tendência.

## **2. Limpeza dos Dados**

Etapas adicionais de pré-processamento foram realizadas para preparar os dados. Primeiramente, a coluna "Country Name" foi removida, pois não seria necessária para as análises subsequentes. Em seguida, a linha que representava o "Total Emissions" foi excluída, mantendo apenas os dados históricos do Brasil. Isso garantiu que a análise estivesse focada unicamente nos valores específicos de emissões anuais, eliminando interferências de valores agregados que poderiam distorcer os resultados.

## **3. Teste de Raiz Unitária (ADF)**

Com os dados limpos, foi realizado o Teste de Dickey-Fuller Aumentado (ADF) para avaliar a estacionariedade da série temporal. O teste é uma ferramenta estatística essencial para identificar se os dados possuem raiz unitária, indicando ausência de estacionariedade. A estatística calculada foi de -1.610452, enquanto o *p-valor* foi de 0.477739.

O resultado sugere que a série de emissões do Brasil não é estacionária, uma vez que o *p-valor* excede o nível de significância de 0,05. Essa conclusão implica que os dados provavelmente contêm uma tendência subjacente ou um comportamento aleatório ao longo do tempo, impedindo o retorno a uma média constante. Isso é consistente com as dinâmicas históricas de emissões, que frequentemente são influenciadas por fatores econômicos, tecnológicos e políticos de longo prazo.

## **5.2. PREPARAÇÃO DOS DADOS PARA APLICAÇÃO DO MODELO SARIMAX**



Com o objetivo de viabilizar a aplicação de manipulações numéricas e ajustar o modelo SARIMAX de forma eficiente, os dados da série temporal de emissões de CO<sub>2</sub> do Brasil foram convertidos para um array NumPy do tipo float. Essa conversão assegura que os dados estejam no formato ideal para cálculos matemáticos e manipulações durante as etapas de modelagem. O código utilizado realizou a conversão diretamente e verificou o tipo de dado resultante, confirmando que os valores agora são do tipo float. Para compreender melhor a estrutura dos dados e garantir sua integridade, foi realizada uma verificação dos índices e dos tipos de dados no DataFrame. Os índices correspondem aos anos da série temporal, iniciando em 1990 e terminando em 2020, com um total de 31 entradas. No entanto, a análise revelou que o índice está no formato *object*, o que pode dificultar algumas operações numéricas ou temporais.

Além disso, foi realizada uma revisão da estrutura do DataFrame por meio do método `.head()`, o que confirmou a presença dos valores esperados para os primeiros anos da série. O tipo de dado *object* dos índices será corrigido posteriormente, se necessário, para facilitar a manipulação temporal e atender aos requisitos do modelo SARIMAX.

### **5.3. Determinação dos Parâmetros para o Modelo SARIMA**

O modelo SARIMA (Seasonal AutoRegressive Integrated Moving Average) foi selecionado devido à sua capacidade de capturar padrões sazonais e tendências em séries temporais. Para determinar os parâmetros ideais, utilizou-se a função `auto_arima` da biblioteca `pmdarima`, que realiza uma busca automatizada com o objetivo de minimizar o Critério de Informação de Akaike (AIC). Esse critério avalia a qualidade do modelo ajustado, penalizando sua complexidade para evitar o superajuste. O modelo identificado como mais adequado foi o SARIMAX(0, 1, 0)x(0, 1, [1], 2), que apresentou o menor AIC (-41.583). Este resultado indica que o modelo incorporou corretamente a estrutura sazonal de dois períodos, bem como o componente de média móvel sazonal, para descrever adequadamente a dinâmica da série.

A escolha automatizada dos parâmetros é fundamental em análises preditivas, pois garante que o modelo capture de forma eficiente as características estruturais dos dados. No caso específico das emissões de CO<sub>2</sub> no Brasil, o processo de seleção considerou

diferentes combinações de ordens autoregressivas, integrações e médias móveis, tanto para os componentes sazonais quanto para os não sazonais. Esse rigor metodológico é essencial para desenvolver modelos preditivos confiáveis.

#### **5.4. Ajuste do Modelo SARIMA com os Parâmetros Otimizados**

Após a determinação dos parâmetros, o modelo SARIMA foi ajustado utilizando a biblioteca statsmodels. O ajuste seguiu as configurações propostas pelo processo automatizado, incorporando ordens específicas para os componentes autorregressivo (AR), de integração (I) e de média móvel (MA), bem como suas correspondências sazonais. Durante o ajuste, as restrições de estacionariedade e invertibilidade foram flexibilizadas, permitindo explorar um espaço de soluções mais amplo. Essa abordagem é particularmente útil em séries temporais não estacionárias, como a série analisada, que apresenta tanto tendências quanto sazonalidades.

O ajuste bem-sucedido do modelo reflete a compatibilidade entre os dados e a estrutura paramétrica definida. Ao considerar tanto os dados históricos quanto as incertezas inerentes, o modelo é capaz de fornecer previsões mais precisas e alinhadas com o comportamento observado.

#### **5.5. Previsão de Emissões de CO<sub>2</sub> para o Período de 2021 a 2030**

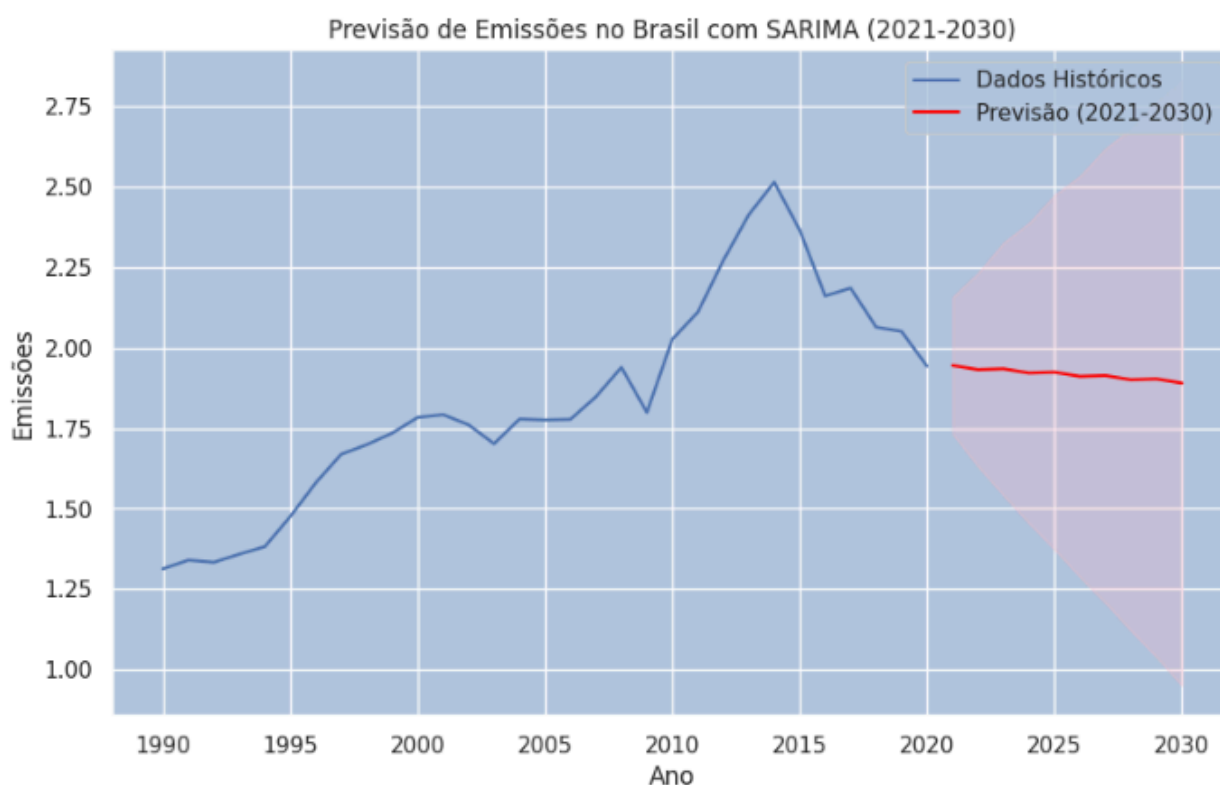
Com o modelo SARIMA ajustado, foram geradas previsões para as emissões de CO<sub>2</sub> no Brasil no horizonte de 2021 a 2030. Essas previsões incluem valores pontuais para cada ano, acompanhados de intervalos de confiança que refletem a incerteza associada às estimativas. Os intervalos de confiança são particularmente importantes, pois permitem avaliar a variabilidade esperada nas previsões, fornecendo subsídios para decisões estratégicas mesmo diante de cenários incertos.

Para organizar as projeções, foi criado um índice temporal que cobre os anos previstos. Isso facilita a interpretação e a integração dos resultados com outros estudos ou modelos. As previsões indicam a tendência esperada das emissões de CO<sub>2</sub> no Brasil, considerando os padrões históricos e os componentes sazonais identificados pelo modelo.

## 5.6. Visualização dos Resultados

Os resultados das previsões foram apresentados graficamente para facilitar a análise e interpretação. O gráfico combina a série histórica das emissões de CO<sub>2</sub> (de 1990 a 2020) com as previsões do modelo SARIMA para o período de 2021 a 2030. Os dados históricos foram representados como uma linha contínua, enquanto as previsões foram destacadas em vermelho. Além disso, os intervalos de confiança foram visualizados como uma área sombreada, indicando a faixa de variabilidade esperada.

Essa visualização permite uma análise detalhada das tendências projetadas e destaca as incertezas associadas às estimativas. Com isso, é possível avaliar a consistência do modelo, identificar padrões de crescimento ou estabilização nas emissões e informar estratégias de mitigação e políticas públicas voltadas para a redução de emissões no Brasil.



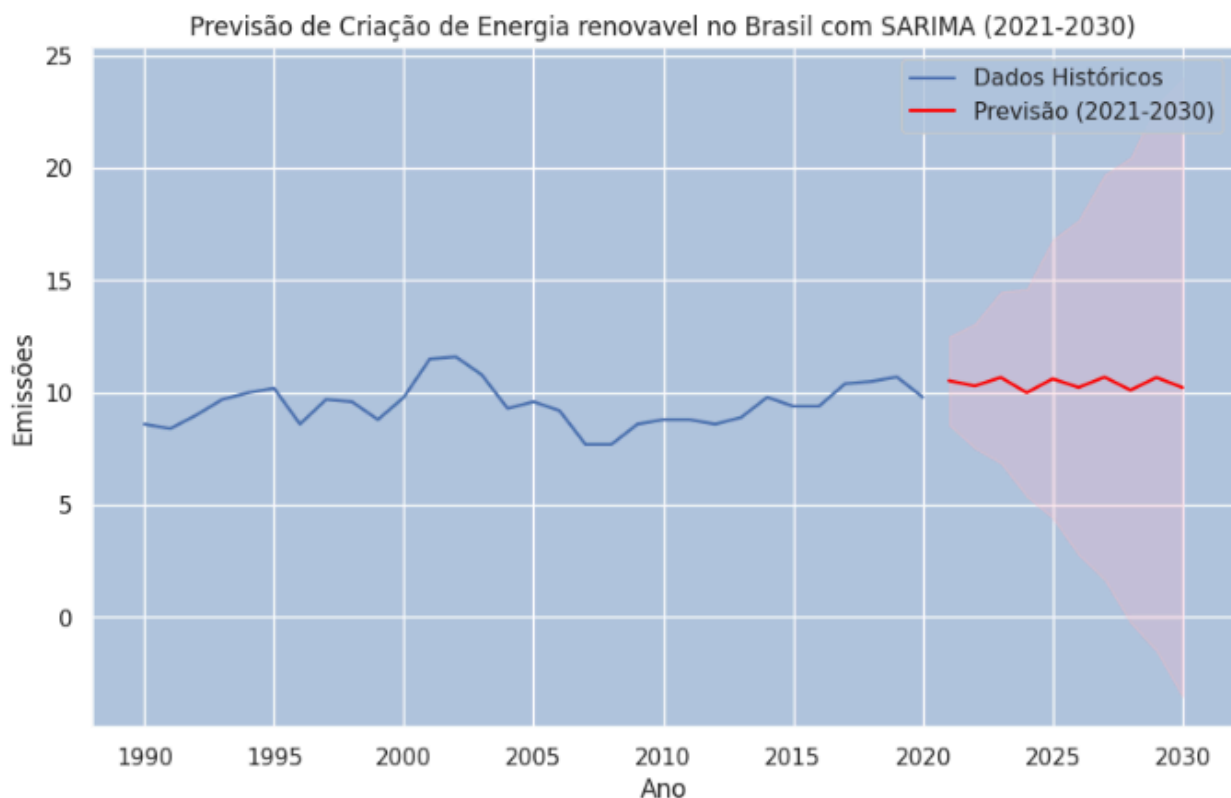
*Figura 8 - Previsão de Emissões no Brasil com SARIMA (2021-2030)*

O gráfico apresentado oferece uma visão geral da evolução histórica e da projeção futura das emissões no Brasil, utilizando o modelo estatístico SARIMA. A linha azul ilustra

um crescimento nas emissões até meados da década de 2010, seguido de uma tendência de estabilização ou leve declínio. A linha vermelha, por sua vez, representa a previsão para os próximos anos, indicando uma continuação desse patamar mais baixo, com uma ligeira tendência de queda a partir de 2025. A área sombreada, que representa o intervalo de confiança da previsão, demonstra a incerteza inerente a qualquer projeção futura.

O modelo SARIMA sugere que as medidas e políticas implementadas para mitigar as emissões no Brasil podem estar surtindo efeito. No entanto, é crucial ressaltar que essa é uma simplificação da realidade. Fatores como mudanças climáticas extremas, políticas governamentais futuras e avanços tecnológicos podem influenciar significativamente os resultados reais. Além disso, a previsão se baseia em dados históricos e nas premissas do modelo, sendo sensível a alterações nessas condições.

Para uma análise mais completa, seria fundamental aprofundar a investigação em diversos aspectos. A desagregação dos dados por setor (energia, indústria, agropecuária, etc.) permitiria identificar as principais fontes de emissões e direcionar as políticas de forma mais eficaz. A análise das políticas públicas implementadas, como programas de incentivo à energia renovável e desmatamento zero, seria crucial para avaliar sua efetividade. Além disso, a consideração de diferentes cenários futuros, levando em conta variáveis como o crescimento econômico, a transição energética e as metas climáticas globais, proporcionaria uma visão mais abrangente e robusta das perspectivas para as emissões brasileiras.



*Figura 9 - Previsão de Criação de Energia Renovável no Brasil com SARIMA(2021-2030)*

O gráfico apresentado projeta a geração de energia renovável no Brasil para o período de 2021 a 2030, utilizando o modelo SARIMA com base em dados históricos. A área sombreada no gráfico representa o intervalo de confiança das previsões, oferecendo uma medida da incerteza associada às estimativas realizadas. Essa representação gráfica combina informações históricas e futuras, fornecendo uma visão clara da tendência esperada.

A análise dos dados revela uma estabilidade na criação de energia renovável no Brasil nos últimos anos, com pequenas flutuações em torno de uma média estável. Essa tendência é mantida nas previsões do modelo SARIMA, que sugere que a geração de energia renovável deve permanecer consistente, sem variações significativas, ao longo da próxima década. Além disso, a faixa relativamente estreita do intervalo de confiança reforça a robustez das estimativas, indicando baixa variabilidade esperada na produção de energia renovável no futuro. Esses resultados destacam a previsibilidade do setor e

podem auxiliar no planejamento estratégico para políticas energéticas sustentáveis no Brasil.

## 6. RESULTADOS

O modelo  $ARIMA(0,1,0)(0,1,1)[2]$  foi identificado como o mais adequado para modelar as emissões de  $CO_2$  per capita no Brasil. Esse modelo utiliza uma diferenciação de primeira ordem para lidar com a não-estacionariedade da série, além de incluir uma componente sazonal de médias móveis de ordem 1, com sazonalidade bienal. Os indicadores de qualidade do ajuste, como o Log-Likelihood positivo (22.791) e os valores negativos de AIC (-41.583) e BIC (-38.918), demonstram a adequação do modelo em capturar os padrões da série temporal com simplicidade e precisão.

Os coeficientes estimados reforçam as características do modelo. O termo sazonal de médias móveis (-0.6770) indica um impacto inverso de choques ao longo do tempo sazonal. O valor reduzido de  $\sigma^2$  (0.0110) evidencia baixos erros residuais. Além disso, testes diagnósticos indicam normalidade nos resíduos e ausência de autocorrelação significativa, embora tenha sido observada alguma heterocedasticidade.

A análise gráfica revela uma tendência histórica de crescimento nas emissões per capita até 2015, seguida por uma leve queda até 2020. A projeção para 2021 a 2030 sugere um declínio gradual e estável nas emissões, com baixa variabilidade inicial e aumento da incerteza em longo prazo. Contudo, a faixa de confiança indica que as previsões mantêm-se confiáveis, sugerindo uma leve redução nas emissões per capita ao longo da década. Essas descobertas são relevantes para embasar políticas públicas ambientais e iniciativas de redução de impacto ambiental no Brasil.

A análise das emissões relacionadas à criação de energia renovável no Brasil, utilizando o modelo SARIMA, apresenta uma tendência de estabilidade para o período de 2021 a 2030. Historicamente, os dados mostram pequenas oscilações, sem grandes variações ou tendências marcantes de crescimento ou declínio.

A previsão para o período analisado sugere continuidade da estabilidade observada, com os valores previstos alinhados aos dados históricos mais recentes. A incerteza das previsões, representada pela faixa de confiança, aumenta ao longo do tempo, especialmente após 2025. No entanto, o intervalo permanece relativamente

estreito no início do período de projeção, indicando maior confiança nas previsões de curto prazo.

O modelo SARIMA, portanto, sugere que as emissões associadas à energia renovável tendem a se manter constantes, sem grandes alterações esperadas até 2030. Essa perspectiva pode ajudar a embasar decisões estratégicas relacionadas ao planejamento energético e às metas de sustentabilidade no Brasil, permitindo um melhor gerenciamento das incertezas inerentes ao futuro da geração renovável e suas implicações ambientais.

## **7. DISCUSSÃO E CONCLUSÃO**

O presente estudo analisou as emissões de CO<sub>2</sub> per capita e a previsão de geração de energia renovável no Brasil utilizando modelos de séries temporais SARIMA, oferecendo insights relevantes sobre a dinâmica ambiental e energética do país. Os resultados indicam uma relativa estabilidade para ambos os indicadores nas próximas décadas, confirmando o alcance do objetivo inicial de explorar tendências futuras e contribuir para o planejamento sustentável. Apesar das variações observadas nos dados históricos, as previsões sugerem que o Brasil tende a manter níveis constantes de emissões per capita e produção de energia renovável, o que reforça a importância de iniciativas estratégicas para consolidar essas tendências.

Os modelos aplicados, SARIMA(0,1,0)(1,1,0)[2] para energia renovável e SARIMA(0,1,0)(0,1,1)[2] para emissões de CO<sub>2</sub>, destacaram-se por sua capacidade de capturar a dinâmica temporal das séries e gerar projeções confiáveis. Testes diagnósticos robustos, como a análise dos resíduos e a verificação da normalidade, corroboraram a adequação dos modelos. No entanto, desafios como a heterocedasticidade observada em alguns resíduos e a limitada abrangência de dados históricos destacam áreas para aprimoramento metodológico.

### **Análise Crítica do Trabalho**

#### **Aspectos Positivos:**

**Metodologia Sólida:** A aplicação dos modelos SARIMA possibilitou a análise sazonal e de longo prazo, garantindo previsões bem fundamentadas.

**Resultados Relevantes:** A inclusão de intervalos de confiança reforçou a transparência e a confiabilidade das previsões, contribuindo para uma análise prática das tendências ambientais e energéticas no Brasil.

**Contribuição ao Planejamento:** Os insights gerados são de grande utilidade para políticas públicas e estratégias de sustentabilidade.

### **Limitações:**

**Dados Limitados:** A disponibilidade restrita de séries temporais mais longas pode ter afetado a precisão das previsões.

**Modelo Simples:** Apesar de eficiente, os modelos SARIMA desconsideram variáveis exógenas que podem impactar as emissões e a produção de energia, como mudanças políticas e tecnológicas.

**Heterocedasticidade:** A variabilidade não constante dos resíduos pode ter influenciado a precisão das estimativas para certos períodos.

### **Propostas de Melhoria e Perspectivas Futuras**

Para superar as limitações identificadas, propõe-se a adoção de abordagens mais avançadas, como modelos SARIMAX ou métodos de aprendizado de máquina, que incorporam variáveis externas relevantes. A expansão e atualização das séries temporais, assim como a inclusão de fatores qualitativos como políticas governamentais e eventos econômicos, poderiam enriquecer as análises. Além disso, a aplicação de modelos GARCH seria útil para lidar com a heterocedasticidade e melhorar a precisão das previsões em cenários de maior variabilidade.

### **Conclusão**

O estudo demonstrou a aplicabilidade dos modelos SARIMA na análise de emissões de CO<sub>2</sub> per capita e geração de energia renovável no Brasil, apresentando previsões consistentes para o período de 2021 a 2030. Embora as projeções apontem para estabilidade, é fundamental considerar as incertezas e explorar ferramentas mais avançadas para capturar as complexidades do sistema energético e ambiental. Este trabalho fornece uma base sólida para o planejamento de políticas públicas e estratégias de mitigação de impacto ambiental, ao mesmo tempo que destaca a necessidade de abordagens integradas para enfrentar os desafios futuros.



## 8. APRESENTAÇÃO

Este capítulo tem como objetivo fornecer uma visão geral do projeto, abordando tanto os aspectos conceituais quanto os detalhes técnicos da implementação. Devido à extensão do projeto, a parte contextual foi focada, neste caso, na explicação de alguns dos gráficos gerados, que ilustram os principais resultados e análises realizadas. Para complementar a descrição, disponibilizamos dois vídeos explicativos: o primeiro apresenta a parte conceitual do projeto, com ênfase na explicação dos gráficos selecionados, enquanto o segundo vídeo detalha a explicação do código desenvolvido, abordando as funcionalidades implementadas e o fluxo de execução. Os vídeos foram elaborados para oferecer uma compreensão mais completa dos aspectos teóricos e práticos do projeto. A seguir, encontram-se os links para acesso a ambas as apresentações.

Video Conceitual	<a href="https://youtu.be/Q2VROtDg-aY">https://youtu.be/Q2VROtDg-aY</a>
Video Explicação Código	<a href="https://youtu.be/3HjOLIO7wqQ">https://youtu.be/3HjOLIO7wqQ</a>

#### 4. REFERÊNCIAS BIBLIOGRÁFICAS

- (1) POYTING, Mark; PRAZERES, Leandro. Mudanças climáticas foram 'principal' fator para seca recorde na Amazônia, diz estudo: o que isso significa para o futuro da floresta?. [S. l.], 24 jan. 2024. Disponível em:  
<https://www.bbc.com/portuguese/articles/c88nr0940j8o>. Acesso em: 11 set. 2024.
- (2) CO2 emissions (metric tons per capita) - Latin America & Caribbean. [S. l.], 14 fev. 2023. Disponível em:  
<https://data.worldbank.org/indicator/EN.ATM.CO2E.PC?end=2020&locations=ZJ&start=1990&view=chart>. Acesso em: 11 set. 2024.
- (3) RENEWABLE energy consumption (% of total final energy consumption) - Latin America & Caribbean. [S. l.], 11 abr. 2023. Disponível em:  
<https://data.worldbank.org/indicator/EG.FEC.RNEW.ZS?locations=ZJ>. Acesso em: 11 set. 2024.
- (4) ARAUJO, Arney Rayol Moura de; LOBATO, Tarcísio da Costa. Análise e previsão das emissões de CO2 provenientes dos Combustíveis do Setor de Transporte Rodoviário no Brasil 1990- 2020. [S. l.], 1 dez. 2020. Disponível em:  
<http://repositorioinstitucional.uea.edu.br/bitstream/riuea/5280/1/Análise%20e%20previsão%20das%20emissões%20de%20CO2%20provenientes%20dos%20Combustíveis%20do%20Setor%20de%20Transporte%20Rodoviário%20no%20Brasil%201990-2020.pdf>. Acesso em: 30 set. 2024.
- (5) SOARES, Francisco. Aplicação de Machine Learning para Previsão das Emissões de CO2e Produzidas pela Produção de Energia no Estado do Amazonas: Um Estudo de Caso Usando R. [S. l.], 12 fev. 2024. Disponível em:  
<https://medium.com/@fanciscosoares/aplicação-de-machine-learning-para-previsão-das-emissões-de-co2e-produzidas-pela-produção-de-1817cdd34729>. Acesso em: 30 set. 2024.
- (6) FONSECA, Neide de. A CRISE DE 2008 E POLÍTICAS ECONÔMICAS NO BRASIL. In: FONSECA, Neide de. A CRISE DE 2008 E POLÍTICAS ECONÔMICAS NO BRASIL. 2013. Monografia de Graduação (Bacharel. Ciências Econômicas) - UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO, [S. l.], 2013. f. 37. Disponível em:

<https://itr.ufrj.br/portal/wp-content/uploads/2017/10/t74.pdf>. Acesso em: 17 nov. 2024.

## 5. APÊNDICES

(1) Link do repositório do projeto:

<https://github.com/ViniSegatto/ProjetoAplicadoIV/tree/main/Projeto>

(2) Código no Google Colaboratory:

 Projeto Etapa 4.ipynb