

SQL PARA ANÁLISE DE SÉRIES TEMPORAIS

Um estudo estatístico das emissões de carbono no consumo de energia na Dinamarca (2021 a 2024)

Mateus Monteiro da Costa (Fatec Deputado Ary Fossen)

mateus.costa24@fatec.sp.gov.br

Orientadora

Gisele Souza Moreira (Fatec Deputado Ary Fossen)

gisele.moreira@fatec.sp.gov.br

1. INTRODUÇÃO

Compreender o passado é uma das etapas mais importantes na análise de dados, pois nele está a base que sustenta o presente e o futuro. O SQL (*Structured Query Language*) é uma ferramenta essencial para essa tarefa, pois permite modelar, inserir e analisar dados de maneira explícita e flexível.

Na análise de séries temporais, onde é necessário trabalhar com dados como dias, horas e minutos, o SQL se destaca como uma excelente escolha. Isso ocorre porque, por meio de suas cláusulas de agrupamento (*GROUP BY*) e funções de extração de fragmentos temporais (*YEAR()*, *MONTH()*, *DAY()*, etc.), é possível construir consultas que retornam os valores desejados organizados cronologicamente. Um exemplo disso é a agregação anual da quantidade de carbono emitido no consumo de eletricidade na Dinamarca, conforme ilustrado em algumas figuras deste artigo. De acordo com Tanimura (2022, p. 79), a análise de séries temporais é "uma maneira de entender e quantificar como as coisas mudam com o tempo". Em um cenário onde o consumo de recursos essenciais para o ser humano vem se tornando cada vez mais discutido, a afirmação de Tanimura se torna ainda mais relevante, e analisar as variáveis que giram em torno desta questão se torna uma prática de muito bom gosto.

Os países nórdicos (Dinamarca, Finlândia, Islândia, Noruega, Suécia etc.), são referência quando o assunto é eficiência energética, dados do EESI (Environmental and Energy Study Institute) confirmam esta afirmação, porém há um exemplo que destaca dos demais. A

Dinamarca é um país europeu que se localiza na região da Escandinávia, no norte da Europa, com aproximadamente 42.921 km² de extensão territorial. Sua capital é Copenhague, que possui 1,3 milhão de habitantes, representando cerca de 22,45% da população total do país, que é de aproximadamente 5,79 milhões de pessoas. Este país nórdico possui uma economia que se destaca pelo alto índice de desenvolvimento e pelo seu principal produto, as turbinas eólicas.

No decorrer deste artigo serão apresentados consultas SQL e seus resultados, tabelas de dicionário de termos e expressões, dados tabulados, conceitos estatísticos e representações gráficas para validação dos dados a serem expostos.

2. FUNDAMENTAÇÃO TEÓRICA

Para o desenvolvimento deste artigo, foram utilizados materiais de consulta como artigos acadêmicos e livros didáticos referente às tecnologias utilizadas na construção das análises. Todos os dados apresentados em consultas e gráficos foram disponibilizados pela Electricity Maps, uma empresa dinamarquesa fundada por Olivier Corradi, que surgiu com intuito de mapear a emissão de carbono por consumo de eletricidade. Corradi acreditava que dados em tempo real poderiam servir como um catalisador para a descarbonização do planeta terra.

No artigo Climate, wind energy, and CO₂ emissions from energy production in Denmark, Carlini, Christensen, Gupta e Magistris (2023, p. 1) afirmam que “A energia eólica já é importante na Dinamarca, cobrindo 47% do consumo bruto de eletricidade em 2019”. Esta afirmação já diz muito a respeito dos dados que serão estudados neste artigo, a energia eólica é uma fonte de energia que gera uma quantidade quase nula de carbono, o que acaba reduzindo diretamente a emissão do mesmo durante o consumo de energia elétrica. Tendo em vista o compromisso da Dinamarca por um país mais sustentável no que se diz respeito à geração e consumo de energia elétrica, esta porcentagem aumentou consideravelmente.

O SQL (Structured query language | Linguagem de consulta estruturada), é uma linguagem utilizada para as etapas de tratamento e processamento de dados, auxiliando no armazenamento, manipulação e recuperação de dados em bancos de dados relacionais como SQL Server, PostgreSQL, Oracle, MySQL, etc. As consultas em SQL conseguem extrair informações de dados tabulados, organizados por linhas e colunas em tabelas.

3. DESENVOLVIMENTO DA TEMÁTICA

Uma matéria do DW de 2014 comunica o plano da Dinamarca de até 2050 produzir somente energia e calor limpos, eliminando as emissões de dióxido de carbono no setor. Para o ano de 2020, a estimativa era que cerca de 70% da energia gerada seria a partir de fontes renováveis. Infelizmente a base de dados utilizada neste artigo só continha dados a partir do ano de 2021 até o ano de 2024, mas mesmo assim, através da seção de resultados e discussão, será possível visualizar o sucesso do projeto proposto pelo país.

Figura 1 - Consulta SQL que retorna todos os dados das medições registradas às 04 horas no ano de 2021.

```
SELECT * FROM denmark.year_2021_hourly  
WHERE EXTRACT(HOUR FROM datetime_utc) = 4
```

datetime_utc	country	zone_name	zone_id	carbon_intensity_direct	carbon_intensity_lca	low_carbon_percentage	renewable_percentage	data_source	data_estimated	data_estimation_method
2021-01-01 04:00:00	Denmark	Denmark	DK	119	193.68	82.45	70.54	entsoe.eu	false	[null]
2021-01-02 04:00:00	Denmark	Denmark	DK	120.43	195.99	81.24	68.75	entsoe.eu	false	[null]
2021-01-03 04:00:00	Denmark	Denmark	DK	109.05	171.12	83.96	74.78	entsoe.eu	false	[null]
2021-01-04 04:00:00	Denmark	Denmark	DK	128.71	187.84	82.21	82.06	entsoe.eu	false	[null]
2021-01-05 04:00:00	Denmark	Denmark	DK	113.22	172.89	83.88	83.8	entsoe.eu	false	[null]
2021-01-06 04:00:00	Denmark	Denmark	DK	118.67	181.78	82.25	80.68	entsoe.eu	false	[null]
2021-01-07 04:00:00	Denmark	Denmark	DK	154.48	224.11	78.71	74.08	entsoe.eu	false	[null]
2021-01-08 04:00:00	Denmark	Denmark	DK	185.75	261.58	73.22	62.74	entsoe.eu	false	[null]
2021-01-09 04:00:00	Denmark	Denmark	DK	202.26	276.22	68.5	58.17	entsoe.eu	false	[null]
2021-01-10 04:00:00	Denmark	Denmark	DK	112.85	161.09	83.47	74.89	entsoe.eu	false	[null]
2021-01-11 04:00:00	Denmark	Denmark	DK	199.21	252.59	74.12	70.16	entsoe.eu	false	[null]
2021-01-12 04:00:00	Denmark	Denmark	DK	246.33	310.52	68.41	60.4	entsoe.eu	false	[null]
2021-01-13 04:00:00	Denmark	Denmark	DK	158.32	215.95	79.17	77.85	entsoe.eu	false	[null]
2021-01-14 04:00:00	Denmark	Denmark	DK	180.05	246.67	76.8	73	entsoe.eu	false	[null]
2021-01-15 04:00:00	Denmark	Denmark	DK	252.79	343.77	60.85	57.27	entsoe.eu	false	[null]
2021-01-16 04:00:00	Denmark	Denmark	DK	176.03	259.05	74.13	65.67	entsoe.eu	false	[null]
2021-01-17 04:00:00	Denmark	Denmark	DK	131.08	205.64	80	71.66	entsoe.eu	false	[null]

Fonte: Autoria própria (2025).

Para se trabalhar com dados temporais utilizando o SQL, é necessário conhecer alguns fundamentos. Todo dado criado pertence a uma classificação específica, que define o tipo de informação que está sendo armazenada, em dados temporais, há as seguintes definições de tipo:

Tabela 1 - Tipos de dados temporais.

Tipos de dados temporais	
Tipo	Descrição
DATE	Data (YYYY-MM-DD)
TIME	Hora (HH:MM:SS)
DATETIME	Data e Hora (YYYY-MM-DD HH:MM:SS)
TIMESTAMP	Data e hora (com ou sem fuso horário)
INTERVAL	Diferença entre datas (ex: '2 days', '3 hours')

Fonte: Autoria própria (2025).

Ao entrar na parte de manipulação dos dados, é imprescindível que o analista tenha conhecimento a respeito de quais tipos de dados ele está lidando, segundo Tanimura (2022, p. 34), “Ter um conhecimento sólido das várias formas que os dados podem assumir o ajudará a ser um analista de dados mais eficiente”. É muito comum ter colunas que armazenam valores numéricos mas que estão configuradas como texto, o que impede a aplicação de operações matemáticas nos dados, pois são considerados como valores textuais, mais conhecidos como “strings” ou “texts”, apesar de visualmente parecerem números inteiros ou reais. Para contornar problemas como este, é recomendável que os dados passem por um processo chamado ETL (Extract, Transform and Load | Extrair, Transformar e Carregar), que consiste na obtenção do dado, a transformação do mesmo, trocando tipos, substituindo valores vazios e demais operações, e por fim a exportação para uma plataforma de armazenamento de informações, desde planilhas Excel até bancos de dados.

Durante a obtenção dos dados utilizados nas análises do tema deste trabalho, foi realizado um processo de ETL utilizando a linguagem de programação Python, e sua biblioteca de manipulação de dados, Pandas. O processo foi simples, pois todos os dados estavam formatados de maneira condizente com as informações que armazenavam, com exceção dos valores de dados temporais, que estavam sendo armazenados como textos, o que iria causar a necessidade de aplicar conversões de tipo para cada consulta relacionada a eles, mas contornar esta situação não foi nenhum problema, pois o Pandas possui funções específicas para casos como este, que permitem realizar conversões de valores em qualquer elemento do Data Frame, uma estrutura de dados da biblioteca Pandas, usada para armazenar e manipular tabelas de dados no Python através de um arquivo que contenha estes dados, podendo ser um arquivo CSV, Excel, SQL etc. O Data Frame do Pandas funciona como uma tabela com rótulos, onde cada coluna pode ter um

tipo diferente (números, texto, datas, etc.). Este artigo não visa focar no Python e suas tecnologias, mas ele é uma ferramenta que torna vários processos relacionados a dados, que antes eram considerados braçais, em soluções automatizadas. Considere conhecer mais da linguagem e suas bibliotecas, como o Pandas, Numpy, Seaborn e Matplotlib. As documentações de todas as tecnologias utilizadas neste trabalho estarão na sessão de referências.

Com os dados transformados e carregados no banco de dados, torna-se possível a aplicação das consultas SQL para obter informações relevantes. A escrita de alguns comandos pode variar, pois cada sistema de gerenciamento de banco de dados (SGBD) possui uma nomenclatura específica. Neste trabalho, o SGBD utilizado foi o PostgreSQL, então todos os comandos apresentados estarão seguindo a nomenclatura própria deste sistema, mas a lógica e resultado são os mesmos para os demais SGBDs.

4. RESULTADOS E DISCUSSÃO

Esta seção visa apresentar um aprofundamento do estudo proposto, trazendo consultas SQL e seus resultados, também haverão representações gráficas, para melhor visualização. Alguns conceitos estatísticos serão apresentados, porém serão explicados de forma simples e intuitiva, para melhor compreensão.

A análise descritiva está extremamente vinculada aos conceitos de séries temporais, pois todo fato possui um contexto temporal. A granularidade temporal dos dados está definida em horas, então cada registro representa a média de intensidade de gCO₂/kWh computada no respectivo horário.

Os valores de emissão direta apresentados nas análises são resultantes da soma de todos os valores médios de emissão computados no respectivo horário durante todos os meses do ano. Antes de apresentar os gráficos de análise, é importante contextualizar o uso do SQL para este fim. Pelo fato da consulta ser muito complexa, serão explicados os principais pontos seguindo a tabela abaixo:

Tabela 2 - Funções utilizadas na consulta.

Funções utilizadas nas consultas		
Comando / Função	Descrição	Aplicação na Consulta
WITH	Cria expressões temporárias (CTEs - Common Table)	Define os resumos mensais ou hora por ano (ex:

Funções utilizadas nas consultas		
	Expressions) usadas para organizar a consulta.	dk_2021_summary, dk_2022_summary, etc).
SELECT	Recupera dados das tabelas ou CTEs.	Usado em todas as etapas da consulta para definir os campos de interesse.
EXTRACT(MONTH (OU HOUR) FROM ...)	Extrai o mês ou hora de uma data (formato datetime).	Permite agrupar os dados por mês ou hora.
ROUND(valor, 2)	Arredonda o valor para 2 casas decimais.	Utilizado nas somas, médias e variações percentuais para melhor legibilidade.
SUM(...)	Soma os valores de uma coluna numérica	Calcula o total mensal de emissões de carbono (direta e LCA).
AVG(...)	Calcula a média dos valores de uma coluna.	Média das porcentagens de fontes renováveis e de baixo carbono.
GROUP BY	Agrupa os dados por uma ou mais colunas para uso com funções agregadas.	Agrupamento por país, zona e mês ou hora
LEAD(...) OVER (...)	Função de janela que acessa o valor da próxima linha na partição ordenada.	próxima linha na partição ordenada. Pega o valor de emissão da hora seguinte, para calcular a variação.
ORDER BY	Ordena as linhas dentro da partição da função de janela e também no resultado final.	Garante que a comparação horária esteja na ordem correta (ano → mês ou hora).

Fonte: Autoria própria (2025).

4.1 QUANTIDADE DE EMISSÃO DURANTE OS ANOS

Consultando o total de emissões de cada ano, agrupado por série temporal, torna-se possível visualizar a trajetória dos níveis de emissão ao longo do período analisado. Conforme a figura 2, observa-se uma tendência de queda contínua, com destaque para a redução expressiva entre os anos de 2022 e 2023. A partir dessa linha temporal, surge a necessidade de responder a uma questão fundamental: o que causou essa redução nas emissões? Para isso, utiliza-se a análise diagnóstica, cujo objetivo é identificar as variáveis responsáveis por um determinado fenômeno.

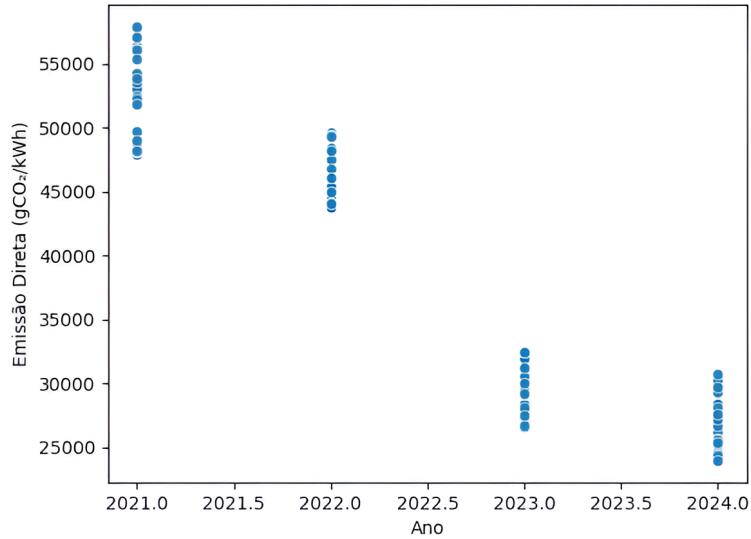
Sabendo da busca constante da Dinamarca por uma produção e consumo de energia mais sustentáveis, uma das possíveis causas da redução significativa entre 2022 e 2023 pode estar relacionada a políticas públicas, inovações tecnológicas ou mudanças no comportamento do consumo energético. Quando se pensa em variáveis causais, é importante manter a mente aberta: hipóteses inicialmente improváveis podem revelar perspectivas relevantes para a compreensão do fenômeno.

Um dado presente na representação gráfica exposta na figura, é o coeficiente de correlação entre ano e emissão direta. Uma das métricas para se estudar correlações entre variáveis é a correlação de Pearson, que não será abordada detalhadamente neste artigo, mas um material de estudo será vinculado na sessão de referências. Na correlação de Pearson, quando duas variáveis possuem um alto nível de correlação, que pode ser apurado através do coeficiente de correlação, que pode ir de -1.0 até 1.0, as dispersões tendem a se concentrar em uma linha reta, representando a linearidade dos dados. Em uma correlação positiva, uma variável tenderia a aumentar em função de outra variável, no caso das amostras utilizadas, percebe-se que o coeficiente tem o valor negativo de -0.94, o que indica que uma variável tende a diminuir em função da outra variável, no caso da figura, a emissão diminui em função do ano.

Nem sempre a correlação expressa o sentido de causalidade, muitas vezes as variáveis se correlacionam por causa de fatores externos, mas não porque uma causa a outra. No sentido de queda de emissão durante os anos, os dados sugerem uma relação causal plausível, tendo em vista o compromisso de sustentabilidade da Dinamarca. Muitas vezes a resposta para esta dúvida é o conhecimento sobre o cenário que envolve os dados.

Figura 2 - Queda de emissão durante os anos.

Granularidade por Horas: Ano | Emissão Direta - Coef de Correlação Ano X Emissão Direta = -0.94 (Negativo Alto)



Fonte: Autoria própria (2025).

4.2 ANÁLISE DA INTENSIDADES DE EMISSÕES POR HORA EM CADA ANO

Originalmente, os dados obtidos estavam agrupados por ano e horário da emissão, para conhecer o pico de emissão em cada horário, foi criada a consulta exposta na figura 3, que calcula a soma total de emissão e a agrupa pelo respectivo horário, por exemplo, para saber a média total de emissão computada às 16:00 horas no ano de 2021, foram somados todos os valores de emissão no horário em todos os dias de todos os meses de 2021.

Figura 3 - Consulta SQL que retorna os picos de emissão em cada horário nos quatro anos.

```
WITH emission_hourly AS (
    SELECT EXTRACT(YEAR FROM datetime_utc) AS year_, EXTRACT(HOUR FROM datetime_utc) AS hour_, ROUND(SUM(carbon_intensity_direct)::NUMERIC, 2) AS max_direct_carbon_emission
    FROM denmark.year_2021_hourly
    GROUP BY EXTRACT(YEAR FROM datetime_utc), EXTRACT(HOUR FROM datetime_utc)

    UNION

    SELECT EXTRACT(YEAR FROM datetime_utc) AS year_, EXTRACT(HOUR FROM datetime_utc) AS hour_, ROUND(SUM(carbon_intensity_direct)::NUMERIC, 2) AS max_direct_carbon_emission
    FROM denmark.year_2022_hourly
    GROUP BY EXTRACT(YEAR FROM datetime_utc), EXTRACT(HOUR FROM datetime_utc)

    UNION

    SELECT EXTRACT(YEAR FROM datetime_utc) AS year_, EXTRACT(HOUR FROM datetime_utc) AS hour_, ROUND(SUM(carbon_intensity_direct)::NUMERIC, 2) AS max_direct_carbon_emission
    FROM denmark.year_2023_hourly
    GROUP BY EXTRACT(YEAR FROM datetime_utc), EXTRACT(HOUR FROM datetime_utc)

    UNION

    SELECT EXTRACT(YEAR FROM datetime_utc) AS year_, EXTRACT(HOUR FROM datetime_utc) AS hour_, ROUND(SUM(carbon_intensity_direct)::NUMERIC, 2) AS max_direct_carbon_emission
    FROM denmark.year_2024_hourly
    GROUP BY EXTRACT(YEAR FROM datetime_utc), EXTRACT(HOUR FROM datetime_utc)
)
SELECT year_, hour_, max_direct_carbon_emission
FROM emission_hourly
ORDER BY hour_, year_
```

year_ numeric 	hour_grouped numeric 	max_direct_carbon_emission numeric 
2021	16	57900.66
2022	16	49359.42
2023	16	31898.13
2024	16	29261.37

Fonte: Autoria própria (2025).

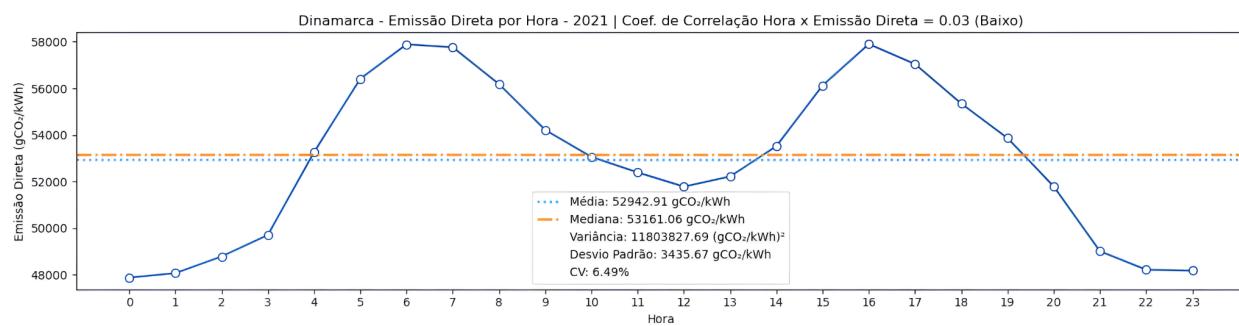
Visualizar estes picos em forma de dados tabulados pode não ser muito claro, então através das bibliotecas Python citadas na seção 3, é possível construir visualizações com os

dados retornados da consulta, como mostra as figuras abaixo, o que acaba facilitando a compreensão da amplitude de emissões e o comportamento geral dos dados.

À primeira vista, os dados parecem formar uma distribuição bimodal, por apresentar dois horários com valores de emissão significativamente superiores aos demais. No entanto, sabendo que na estatística, a moda representa o elemento com maior frequência individual, ou seja, o valor mais presente na base de dados, essa interpretação não é adequada, pois os dados de emissões não refletem a frequência de ocorrências, e sim a média agregada por hora. O eixo Y representa a intensidade média de emissão, e não quantas vezes determinado valor foi registrado.

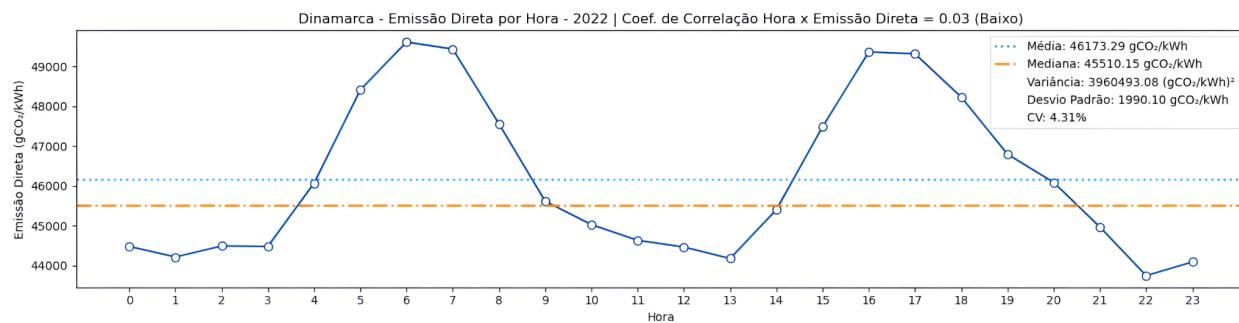
Essa distinção traz um outro ponto importante para a discussão: o eixo x apresenta valores numéricos de 0 a 23, correspondentes às horas do dia. Na ausência de uma legenda ou descrição adequada, o leitor poderia supor que esses números representam classes de frequência, ou até mesmo inverter os eixos, o que comprometeria completamente a interpretação da visualização. Considerando que a análise de dados tem como objetivo fundamental a comunicação clara de um fenômeno, o cuidado com a rotulagem e a explicitação dos eixos é essencial para evitar interpretações equivocadas.

Figura 4 - Gráfico de emissão direta por hora em 2021



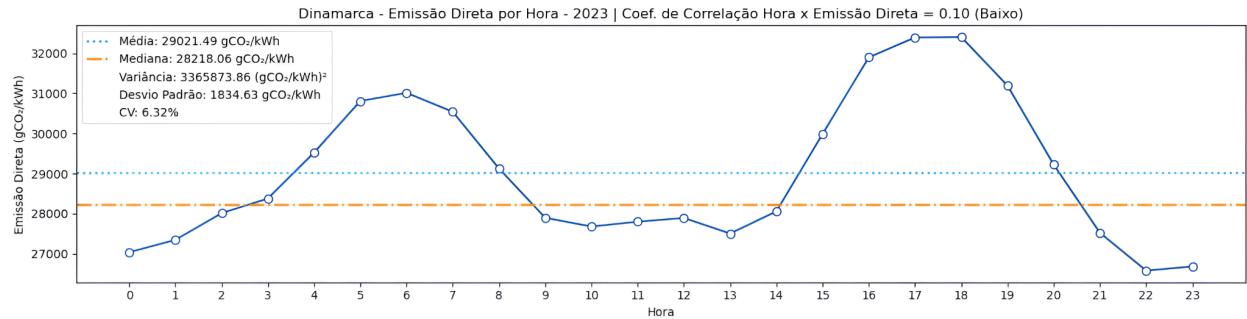
Fonte: Autoria própria (2025)

Figura 5 - Gráfico de emissão direta por hora em 2022



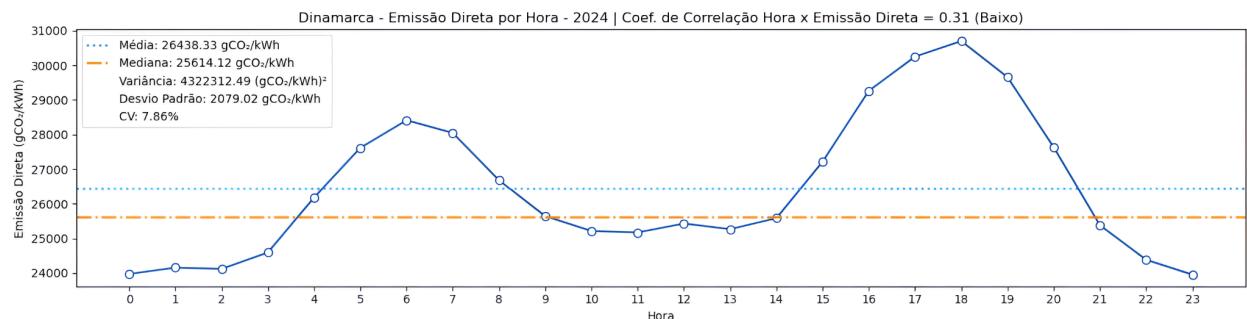
Fonte: Autoria própria (2025)

Figura 6 - Gráfico de emissão direta por hora em 2023



Fonte: Autoria própria (2025)

Figura 7 - Gráfico de emissão direta por hora em 2024



Fonte: Autoria própria (2025)

Analizando os gráficos dos quatro anos estudados, visualiza-se uma constante no pico de emissão no período da manhã. As seis horas sempre é o horário com maior nível de emissão, através de uma rápida pesquisa, têm-se o dado de que normalmente os dinamarqueses costumam começar o dia entre as seis e sete horas da manhã, o que gera um consumo maior de energia elétrica, e consequentemente um nível maior de emissão. No período da tarde, nota-se uma constante nos dois anos iniciais, evidenciada pelas quatro horas como o horário com maior nível de emissão, o que muda nos anos seguintes, tendo o pico de emissão às seis horas.

Sabendo que todos os anos possuem a mesma quantidade de registros, pode-se inferir por meio da média, que o nível de emissões teve uma queda constante, apresentando uma redução de 50,07% do ano inicial até o final, nota-se que a maior queda foi no ano de 2022 para 2023. Todas estas observações do comportamento dos dados reforçam ainda mais a correlação apontada na figura 2 da página 8, porém através de uma análise mais exploratória.

A legenda dos gráficos apresenta alguns valores calculados a partir da estatística descritiva, permitindo uma análise objetiva da distribuição das emissões horárias. A mediana, por exemplo, representa o valor central da amostra. Quando a quantidade de dados é par, a mediana é a média entre os dois valores centrais, se for ímpar, é simplesmente o valor que ocupa a posição central. No caso analisado, como a mediana está próxima da média (obtida pela soma dos valores dividida pelo número total de elementos), pode-se inferir que não há distorções provocadas por valores extremamente altos ou baixos.

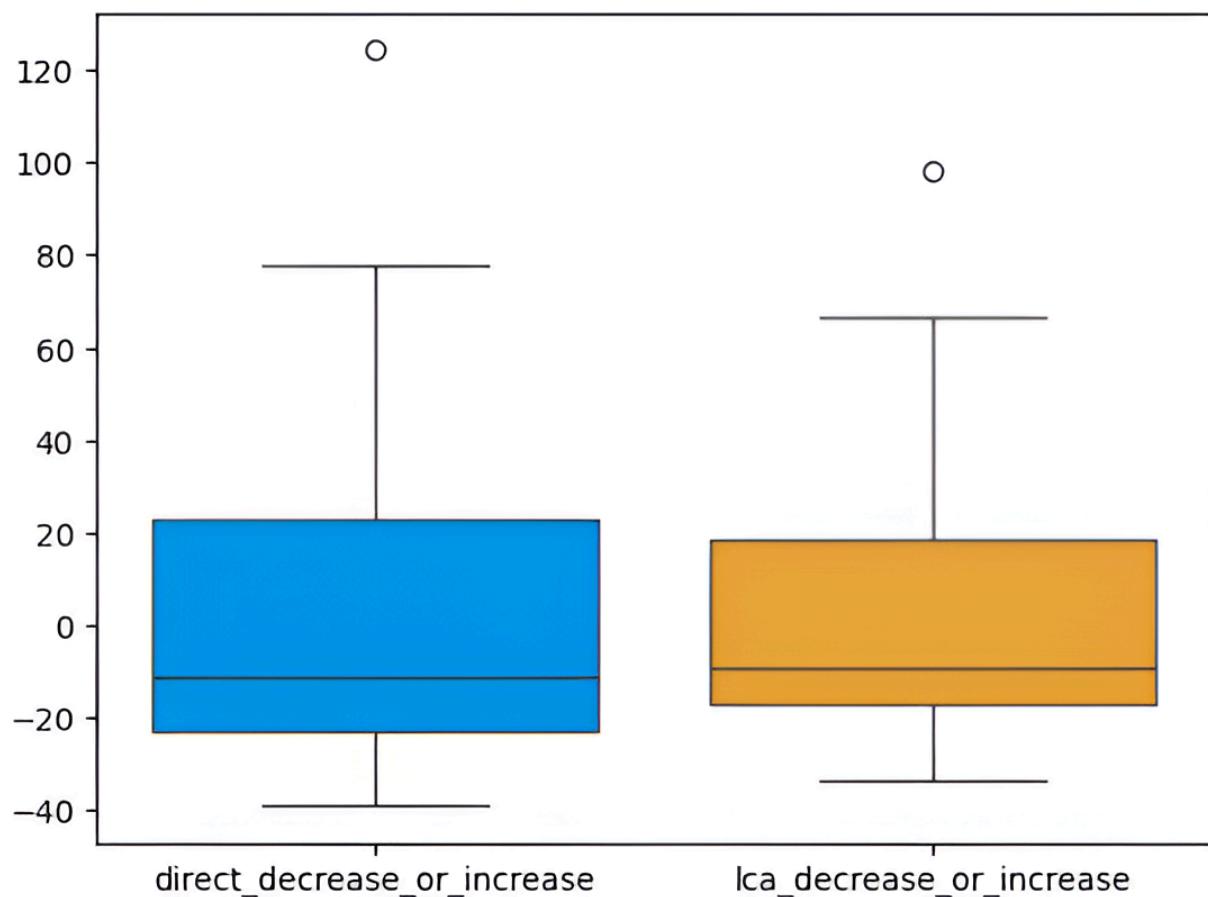
A variância mede o grau de dispersão dos dados em relação à média, porém seu valor é quadrático, o que dificulta a interpretação direta. Por isso, utiliza-se o desvio padrão, que é a raiz quadrada da variância. Ele indica, em média, o quanto os dados se afastam da média, no caso de 2021, cerca de 3.435,67 unidades. Já o coeficiente de variação (CV), dado por $(\text{desvio padrão}/\text{média}) \times 100$, expressa a dispersão em termos percentuais. Um CV inferior a 10%, como observado nos dados, indica baixa variabilidade e alta estabilidade.

No entanto, essa estabilidade pode ser comprometida pela presença de outliers, que são valores que se distanciam significativamente da maioria dos dados. Por exemplo, caso ocorra uma queda total de energia na Dinamarca durante uma hora, a emissão de carbono pode se aproximar de zero, um valor extremamente inferior à média horária, que normalmente está na casa dos milhares. Isso provocaria uma distorção nos indicadores estatísticos, puxando a média para baixo e elevando a variabilidade dos dados.

Uma das formas mais eficazes de identificar outliers é por meio do gráfico boxplot, como ilustrado na figura 8. Nesse tipo de gráfico, a caixa central representa os quartis: a linha inferior da caixa corresponde ao primeiro quartil (Q1, ou 25% dos dados), a linha interna representa a mediana (Q2, 50% dos dados) e a linha superior indica o terceiro quartil (Q3, ou 75%). As linhas horizontais fora da caixa, chamadas de bigodes, definem os limites inferior e superior esperados, geralmente calculados como $Q1 - 1,5 \times \text{IQR}$ e $Q3 + 1,5 \times \text{IQR}$, onde IQR é o intervalo interquartil ($Q3 - Q1$). Os valores fora desses limites são considerados outliers e são representados por pequenos círculos ou pontos.

Nas amostras que comparam a variação percentual da emissão direta entre os meses, observa-se um outlier acima do limite superior, especificamente acima de 120%. Já nas variações mensais das emissões LCA (Emissões referentes a todo o ciclo útil da geração e consumo de energia), o outlier está próximo de 100%. É importante destacar que outliers nem sempre representam um erro ou um problema. Em determinadas análises, eles podem fornecer informações valiosas sobre eventos excepcionais, comportamentos atípicos ou mudanças estruturais. A decisão de mantê-los ou excluí-los deve sempre considerar o contexto da análise e o objetivo do estudo.

Figura 8 - Gráfico boxplot, identificando outliers no aumento ou queda de emissão direta e lca



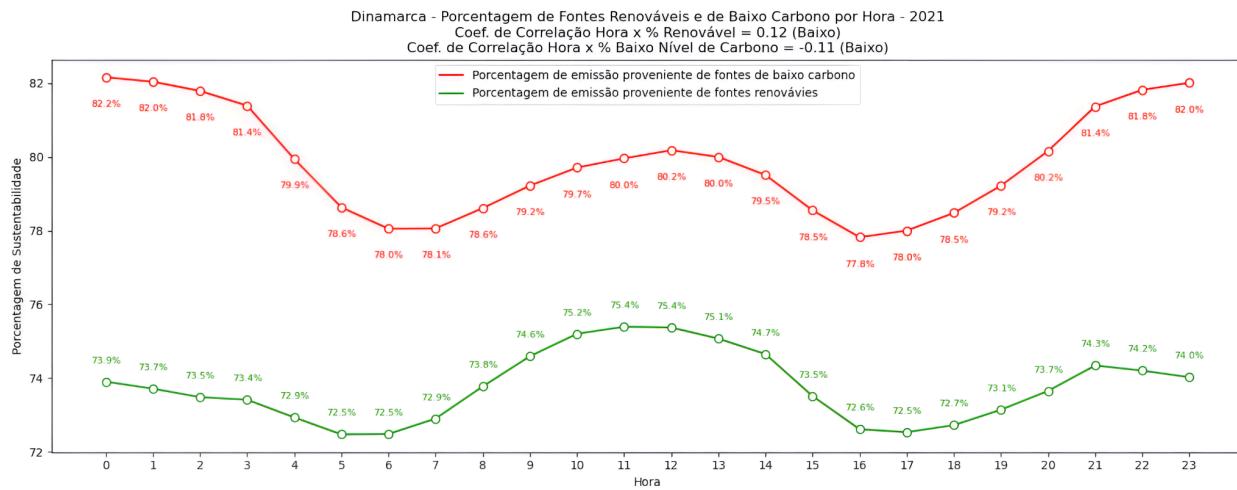
Fonte: Autoria própria (2025).

4.3 ANÁLISE DA SUSTENTABILIDADE NO CONSUMO DE ENERGIA

Retomando a proposta da Dinamarca, de tornar a sua eficiência energética mais sustentável até 2050, é possível visualizar o sucesso deste plano através dos gráficos apresentados abaixo. Estimava-se que, até 2020, cerca de 70% da energia do país seria proveniente de fontes renováveis. Observando os dados entre 2021 e 2024, percebe-se que esta meta foi alcançada, e em determinados horários, até superada, com registros que ultrapassam 80% de geração provenientes de fontes renováveis. Estes resultados evidenciam a eficiência da Dinamarca no compromisso por uma energia mais sustentável.

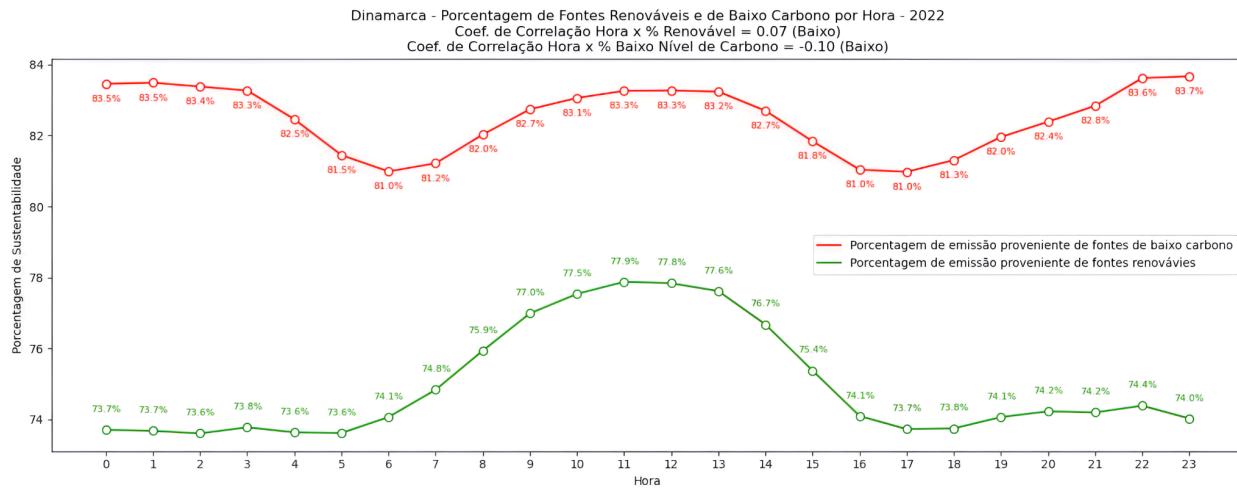
A alta participação de fontes de baixo carbono, como a energia eólica, levanta a hipótese de uma expansão significativa no uso de turbinas eólicas durante esse período. No artigo Climate, wind energy, and CO₂ emissions from energy production in Denmark, os autores apontam que, em 2019, a energia eólica já representava 47% da geração de eletricidade. Considerando os níveis recentes de redução nas emissões diretas e a predominância crescente de energia limpa, é plausível que esse percentual tenha se ampliado em pelo menos 40%, alcançando níveis que ultrapassam 80% em certos momentos, o que reforça o papel central da energia eólica na descarbonização do setor elétrico dinamarquês.

Figura 9 - Gráfico de sustentabilidade por hora em 2021



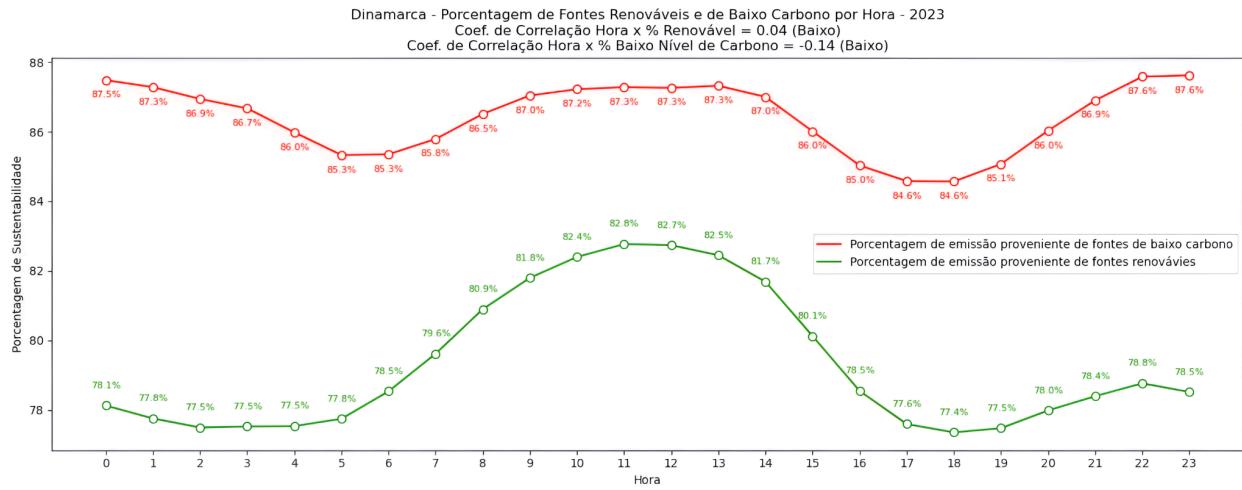
Fonte: Autoria própria (2025).

Figura 10 - Gráfico de sustentabilidade por hora em 2022



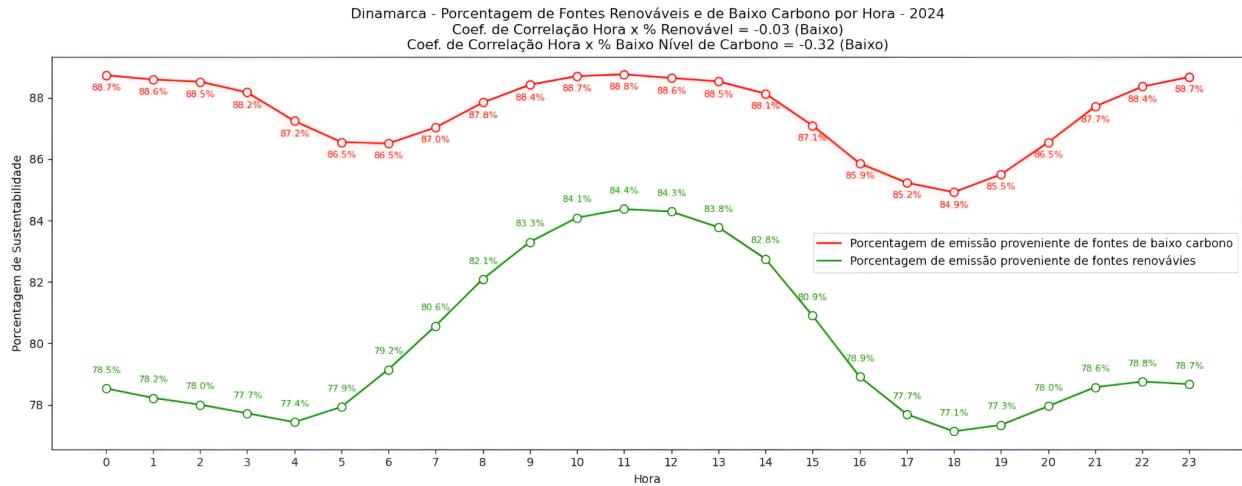
Fonte: Autoria própria (2025).

Figura 11 - Gráfico de sustentabilidade por hora em 2023



Fonte: Autoria própria (2025).

Figura 12 - Gráfico de sustentabilidade por hora em 2024



Fonte: Autoria própria (2025).

5. CONSIDERAÇÕES FINAIS

Por meio das consultas em SQL e da estatística descritiva, foi possível evidenciar o avanço da Dinamarca em sua trajetória rumo a uma matriz energética mais sustentável. Além da redução consistente das emissões diretas de carbono ao longo dos anos analisados, observou-se um aumento expressivo na participação de fontes renováveis e de baixo carbono na geração de energia.

Este artigo também reforça a importância dos dados temporais como instrumento de análise, já que através deles, é possível identificar padrões, tendências e transformações que ocorrem ao longo do tempo, é como se uma história estivesse sendo narrada com base em números. Retomando uma das frases citadas na introdução: “compreender o passado é uma das etapas mais importantes na análise de dados, pois nele está a base que sustenta o presente e o futuro.”

Para além da Dinamarca, este trabalho também buscou valorizar iniciativas como a da equipe da Electricity Maps, cuja dedicação à transparência e à democratização dos dados energéticos contribui significativamente para a conscientização ambiental global. A atuação de organizações com esse propósito é essencial para transformar o presente e construir um futuro mais sustentável para as próximas gerações.

REFERÊNCIAS

CARLINI, Federico; CHRISTENSEN, Bent; GUPTA, Nabanita; MAGISTRIS, Paolo. Climate, wind energy, and CO₂ emissions from energy production in Denmark. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0140988323003195>. Acesso em: 15 maio 2025.

EESI – ENVIRONMENTAL AND ENERGY STUDY INSTITUTE. Nordic countries show potential for low carbon, competitive economies. Disponível em: <https://www.eesi.org/articles/view/nordic-countries-show-potential-for-low-carbon-competitive-economies>. Acesso em: 13 abr. 2025.

ELECTRICITY MAPS. Base de dados de emissões na Dinamarca. Disponível em: <https://portal.electricitymaps.com/datasets/DK>. Acesso em: 10 mar. 2025.

GOMES. Coeficiente de correlação de Pearson: cálculo e aplicação. Data Geeks. Disponível em: <https://www.datageeks.com.br/coeficiente-de-correlacao-de-pearsom/>. Acesso em: 10 maio 2025.

GRUS, Joel. Data science do zero: primeiros princípios com Python. 2. ed. São Paulo: O'Reilly; Alta Books, 2021.

IBM. O que é SQL (Structured Query Language)? Disponível em: <https://www.ibm.com/br-pt/think/topics/structured-query-language>. Acesso em: 13 abr. 2025.

JEPPESEN, Helle. Dinamarca estabelece meta de energia 100% limpa até 2050. Deutsche Welle, 2014. Disponível em: <https://www.dw.com/pt-br/dinamarca-estabelece-meta-de-energia-100-limpa-at%C3%A9-2050/a-17613274>. Acesso em: 17 maio 2025.

MATPLOTLIB. Matplotlib documentation. Disponível em: <https://matplotlib.org/stable/index.html>. Acesso em: 17 maio 2025.

NUMPY. Numpy documentation. Disponível em: <https://numpy.org/doc/>. Acesso em: 14 maio 2025.

PANDAS. Pandas documentation. Disponível em: <https://pandas.pydata.org/>. Acesso em: 17 maio 2025.

POSTGRESQL. PostgreSQL documentation. Disponível em: <https://www.postgresql.org/docs/>. Acesso em: 17 maio 2025.

PYTHON. Python documentation. Disponível em: <https://docs.python.org/3/>. Acesso em: 17 maio 2025.

SEABORN. Seaborn documentation. Disponível em: <https://seaborn.pydata.org/>. Acesso em: 17 maio 2025.

TANIMURA, Cathy. SQL para análise de dados. 1. ed. São Paulo: O'Reilly; Novatec, 2022.

"Os conteúdos expressos no trabalho, bem como sua revisão ortográfica e das normas ABNT são de inteira responsabilidade do(s) autor(es)."

«Declaração de IA generativa e tecnologias assistidas por IA no processo de redação»

“Declara-se pelos autores que durante a preparação deste trabalho foi utilizado o ChatGPT para revisão ortográfica. Após utilizar esta ferramenta/serviço, os autores editaram e revisaram o conteúdo conforme necessário e assumem total responsabilidade pelo conteúdo da publicação.”