



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

MARIA ILANA DOS SANTOS GUEDES

**ANÁLISE DE *CLUSTER* PARA A IDENTIFICAÇÃO DE MUNICÍPIOS COM MAIOR
POTENCIAL PARA GERAÇÃO EÓLICA E FOTOVOLTAICA NO CEARÁ**

FORTALEZA

2023

MARIA ILANA DOS SANTOS GUEDES

ANÁLISE DE *CLUSTER* PARA A IDENTIFICAÇÃO DE MUNICÍPIOS PROMISSORES
PARA GERAÇÃO EÓLICA E FOTOVOLTAICA NO CEARÁ

Trabalho de conclusão de curso apresentada a coordenação do curso de engenharia elétrica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de bacharel em engenharia elétrica.

Orientador: Prof. Dr. Dalton de Araújo Honório
Coorientador: Rebeca Dieb Holanda Silva

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

G958a Guedes, Maria Ilana dos Santos.
Análise de Cluster para a Identificação de Municípios com Maior Potencial para Geração Eólica e Fotovoltaica no Ceará / Maria Ilana dos Santos Guedes. – 2023.
49 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia, Curso de Engenharia Elétrica, Fortaleza, 2023.

Orientação: Prof. Dr. Dalton de Araújo Honório.

Coorientação: Profa. Rebeca Dieb Holanda Silva.

1. Energia Renovável. 2. Agrupamento. 3. Energia Eólica. 4. Energia Fotovoltaica. 5. Infraestrutura de Transmissão. I. Título.

CDD 621.3

MARIA ILANA DOS SANTOS GUEDES

ANÁLISE DE *CLUSTER* PARA A IDENTIFICAÇÃO DE MUNICÍPIOS PROMISSORES
PARA GERAÇÃO EÓLICA E FOTOVOLTAICA NO CEARÁ

Trabalho de conclusão de curso apresentada a coordenação do curso de engenharia elétrica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de bacharel em engenharia elétrica.

Aprovada em: 12/07/2023.

BANCA EXAMINADORA

Prof. Dr. Dalton de Araújo Honório (Orientador)
Universidade Federal do Ceará (UFC)

Dr. José Willamy Medeiros de Araújo
Universidade Federal do Ceará (UFC)

Eng. Me. Paulo Honório Filho
Universidade Federal do Ceará (UFC)

Est. Rebeca Dieb Holanda Silva
Universidade Federal do Ceará (UFC)

Aos meus pais, Maria Zenilda Freitas dos
Santos e Raimundo Ivan Guedes.

AGRADECIMENTOS

Espero não esquecer nenhum nome, mas gostaria de aproveitar este espaço para expressar meus sinceros agradecimentos a todas as pessoas que foram fundamentais para a conclusão deste trabalho e estiveram ao meu lado ao longo dessa jornada no curso.

Primeiramente, sou imensamente grata à minha família, em especial à minha mãe Zenilda e à minha irmã Isabel, pelo amor incondicional, paciência e constante ajuda ao longo de toda a minha vida. Vocês foram minha base e fonte de inspiração, sou grata por todo o suporte que me proporcionaram.

Aos meus amigos de graduação, Moacir, Sarah, Marcos, Squarts, Jamile, Francimar, Rick, Comberstais e Davi, quero expressar minha gratidão pelo apoio incondicional tanto nos momentos difíceis quanto nos momentos de conquista. Vocês estiveram ao meu lado, compartilhando experiências, me auxiliando e aguentando todos os obstáculos do curso, sem vocês eu não conseguiria finalizar a graduação.

Agradeço também ao Josué, Suyane, Darlan, Gabriele e Joel por tantos momentos de leveza e conversas edificantes que me ajudaram muito ao longo dessa jornada. Suas amizades foram um verdadeiro suporte e fonte de motivação.

A Juliana e o Gladson que estiveram presentes, em grande parte, de um momento difícil durante a pandemia, no qual formávamos grupos de trabalho e nos ajudávamos mutuamente. Ambos foram essenciais para que eu pudesse continuar no curso e o período remoto fortaleceu bastante nossos laços de amizade. Sou grata por ter vocês ao meu lado.

À minha coorientadora Rebeca Dieb, gostaria de expressar minha profunda gratidão por todos os ensinamentos técnicos e não técnicos que compartilhou comigo desde 2020. Sua orientação, amizade e suporte foram fundamentais para o sucesso deste projeto, sou grata por ter a tido como parte dessa jornada.

Agradeço também ao Lucas Lima, meu amigo de longa data, que inesperadamente reapareceu e me ajudou bastante na execução do projeto. Sua colaboração foi valiosa e estou feliz por contar com você.

Aos amigos do trabalho, Cacau, Natã, Suellen, Victor, Gabi, Fabiana, Letícia, Ygor, André, Carol, Felipe e Jeffin, quero expressar minha gratidão pelas risadas, momentos de

descontração e pelo apoio ao conciliar trabalho e graduação. Suas amizades e companheirismo tornaram essa jornada mais leve e prazerosa.

Gostaria de agradecer também ao grupo "Enjuados" e ao canal no Discord, que desempenharam um papel fundamental no fortalecimento da nossa amizade durante a pandemia. Os momentos divertidos que passamos juntos tomando café, ouvindo música ou jogando *League of Legends* foram uma verdadeira válvula de escape diante das dificuldades enfrentadas no trabalho e na graduação. Agradeço de coração por todas as risadas compartilhadas e pelo apoio mútuo. Vocês foram e continuam sendo uma parte indispensável da minha vida.

Por fim, agradeço ao professor Dalton de Araújo Honório por toda a ajuda durante a execução deste trabalho, sua orientação e apoio foram essenciais simplificar o que eu considerava difícil. Sou grata por sua contribuição significativa para este projeto.

“Several things we had to learn, countless mistakes from times of yore.”
(BITTENCOURT, 2004)

RESUMO

Este trabalho teve como objetivo identificar, por meio de um método de agrupamento, os municípios no estado do Ceará com maior potencial para geração de energia eólica e fotovoltaica. Para esse fim, foram consideradas variáveis como altitude, velocidade do vento, irradiação solar e temperatura média e, com base nesses critérios, as cidades foram distribuídas em *clusters* distintos utilizando o algoritmo *K-means*, na qual os resultados revelaram a existência de quatro grupos com características específicas.

Em seguida, foi empregada metodologia de pontuação, na qual constatou-se que o grupo 3 demonstrou ser favorável para a geração eólica, devido à sua alta temperatura e velocidade do vento. Por outro lado, o grupo 1 foi identificado como mais adequado para a geração fotovoltaica, devido à sua alta altitude e baixa temperatura.

Além disso, realizou-se uma análise do escoamento de energia, que incluiu o mapeamento das subestações instaladas, na qual pôde-se observar limitações na infraestrutura de transmissão, principalmente no grupo 1, que foi identificado como o mais adequado para a geração fotovoltaica. Por fim, essas constatações reforçam, portanto, a necessidade de expansão da infraestrutura de distribuição de energia no estado e, dessa forma, será possível aproveitar de maneira mais eficiente os recursos naturais disponíveis em cada região, impulsionando o potencial da geração de energia renovável no estado do Ceará.

Palavras-chave: Energia Renovável, Agrupamento, Energia Eólica, Energia Fotovoltaica, Infraestrutura de Transmissão.

ABSTRACT

This study aimed to identify, through a clustering method, the municipalities in the state of Ceará with the greatest potential for wind and photovoltaic energy generation. For this purpose, variables such as altitude, wind speed, solar radiation, and average temperature were considered. Based on these criteria, the cities were divided into distinct clusters using the K-means algorithm, which revealed the existence of four groups with specific characteristics.

Subsequently, a scoring methodology was employed, which found that group 3 demonstrated suitability for wind generation due to its high temperature and wind speed. On the other hand, group 1 was identified as more suitable for photovoltaic generation due to its high altitude and low temperature.

Additionally, an analysis of energy flow was conducted, including the mapping of installed substations, which revealed limitations in the transmission infrastructure, particularly in group 1, which was identified as the most suitable for photovoltaic generation. Finally, these findings reinforce the need to expand the energy distribution infrastructure in the state, enabling a more efficient utilization of the natural resources available in each region and boosting the potential of renewable energy generation in the state of Ceará.

Keywords: Renewable Energy, Clustering, Wind Energy, Photovoltaic Energy, Transmission Infrastructure.

LISTA DE FIGURAS

Figura 1 - Oferta interna de energia elétrica por fonte.....	16
Figura 2 - Consumo de energia por setor em 2021	16
Figura 3 - Evolução matriz energética no Ceará.....	17
Figura 4 - Principais componentes de um aerogerador.....	19
Figura 5 - Infraestrutura comum de conexão dos parques eólicos à rede de distribuição elétrica.....	20
Figura 6 - Esquemático do processo de geração por parques fotovoltaicos até a rede de distribuição elétrica.....	22
Figura 7 - Representação dos dados: antes e após o método de agrupamento.....	24
Figura 8 - Passo a passo algoritmo <i>k-means</i>	26
Figura 9 - Representação da curva do cotovelo	27
Figura 10 - Pré-visualização do <i>dataframe</i> de irradiação anual tratado	30
Figura 11 - Mapa dos municípios com informações históricas no INMET	31
Figura 12 - Esquemático de tratamento de dados	32
Figura 13 - Gráfico do cotovelo para a clusterização de municípios.....	33
Figura 14 - Pré-visualização da base de dados clusterizada	33
Figura 15 - Mapa de municípios clusterizados	35
Figura 16 - Mapa topográfico do Ceará.....	37
Figura 17 - Mapa 3D das variáveis mais relevantes para geração eólica	40
Figura 18 - Empreendimentos de geração eólica no Ceará.....	41
Figura 19 - Mapa 3D das variáveis mais relevantes para geração solar	42
Figura 20 - Empreendimentos de geração fotovoltaica no Ceará	43
Figura 21 - Mapa das subestações do Ceará	44
Figura 22 - Mapa das subestações com os clusters por município	45
Figura 23 - Algoritmo para o cálculo haversine.....	45
Figura 24 - Algoritmo média da distância dos municípios até subestações mais próximas.....	46
Figura 25 - Algoritmo para calcular cidades abaixo da média da distância até as subestações.....	46

LISTA DE TABELAS

Tabela 1 - Média das variáveis por grupo	36
Tabela 2 - Estatísticas de altitude por cluster	37
Tabela 3 - Estatísticas de temperatura por cluster	38
Tabela 4 - Estatísticas de irradiação por cluster	38
Tabela 5 - Estatísticas de velocidade média do vento por cluster	38
Tabela 6 - Grupos classificados para geração eólica.....	39
Tabela 7 - Grupos classificados para geração fotovoltaica	42

LISTA DE ABREVIATURAS E SIGLAS

IPECE	Instituto de Pesquisa e Estratégia Econômica do Ceará
EPE	Empresa de Pesquisa Energética
ANEEL	Agência Nacional de Energia Elétrica
CEPEL	Centro de Pesquisas de Energia do Ceará
CBIE	Centro Brasileiro de Infraestrutura
SSE	<i>Sum of Squared Errors</i>
WSS	<i>Within Sum of Squares</i>
INMET	Instituto Nacional de Meteorologia
API	<i>Application Programming Interface</i>
DBSCAN	<i>Density-Based Spatial Clustering of Applications with Noise</i>
SIN	Sistema Interligado Nacional
BIC	Critério de Informação Bayesiana

SUMÁRIO

1 INTRODUÇÃO.....	14
2 REVISÃO BIBLIOGRÁFICA.....	15
2.1 Matriz Elétrica no Brasil	15
2.1.1 Contexto Energético no Ceará.....	17
2.2 Fatores que influenciam as gerações de Energia Eólica e Fotovoltaica	18
2.2.1 Energia Eólica.....	18
2.2.2 Energia Fotovoltaica.....	21
2.3 Método de agrupamento (clustering).....	24
2.3.1 Método <i>K-means</i>	25
3 APLICAÇÃO DA PROPOSTA DE PROJETO.....	28
3.1 Conjuntos de dados avaliados.....	28
3.1.1 Irradiação Solar:	28
3.1.2 Limites Municipais do Ceará:.....	28
3.1.3 Temperatura e Velocidade do Vento:	29
3.1.4 Dados das Subestações no Brasil:	29
3.1.5 Altitudes dos municípios do Ceará.....	29
3.2 Linguagem de programação, bibliotecas e softwares utilizados	29
3.3 Transformações nas bases de dados	30
3.4 Aplicação do <i>K-means</i>	32
4 RESULTADOS E DISCUSSÕES	34
4.1 Análise de cluster.....	34
4.2 Classificação dos melhores grupos para Geração Eólica e Fotovoltaica.....	39
4.2.1 Geração Eólica.....	39
4.2.2 Geração Fotovoltaica	41
4.3 Análise da Geolocalização das Subestações Elétricas no Ceará em Relação às Áreas Propícias para Geração	44
5 CONCLUSÃO.....	47

1 INTRODUÇÃO

A crescente demanda por energia sustentável tem impulsionado a busca por fontes renováveis, apoiada pelos avanços tecnológicos (Brummer, 2018). O estado do Ceará, localizado no nordeste do Brasil, apresenta um enorme potencial para a geração de energia eólica e fotovoltaica devido às suas características geográficas e climáticas favoráveis. A abundância de recursos naturais, como ventos constantes e intensidade solar elevada, possibilita o estado a ter uma região propícia para o desenvolvimento dessas fontes limpas de energia (IPECE, 2018)

Diante disso, para aproveitar plenamente essa potencialidade, é necessário realizar uma análise cuidadosa das áreas mais promissoras e identificar os municípios com condições ideais para otimizar a produção energética. Desse modo, as características que impactam as gerações e que serão utilizadas neste estudo compreendem a irradiação solar, a velocidade do vento, a altitude e a temperatura, baseado nisso, será possível agrupar as cidades de acordo com suas características semelhantes. Esses parâmetros climáticos e geográficos desempenham um papel fundamental na análise e avaliação da viabilidade da geração de energia renovável (C. Marimuthu, V. Kirubakaran, 2014; BARROS ZANUT, 2022; FEIJÓO COELHO, 2017).

Ademais, o método de agrupamento (*clustering*) escolhido para este estudo é o algoritmo *K-means*, reconhecido por sua eficácia na identificação de grupos em conjuntos de dados. Embora existam diversos métodos de clusterização disponíveis na literatura, como o *Hierarchical Clustering*, o *DBSCAN* e o *Gaussian Mixture Model*, optou-se pelo algoritmo *K-means* devido à sua simplicidade e eficiência na análise de grandes volumes de dados (BERKHIN, 2002). Esse algoritmo promove uma redução significativa do volume de informações a serem analisadas, permitindo a classificação e priorização do conjunto de cidades com as condições mais favoráveis para a geração de energia.

Além disso, é imprescindível destacar que a possível incorporação de novos empreendimentos na estrutura energética requer a superação dos desafios inerentes à infraestrutura de transmissão (SOARES ACCON, 2023). Entre os principais desafios, destacam-se a necessidade de investimentos significativos para ampliar e modernizar as redes de distribuição e a mitigação de perdas energéticas ao longo das linhas de transmissão, esses obstáculos demandam abordagens estratégicas e soluções eficientes para garantir a efetiva integração dos novos empreendimentos e o fornecimento confiável de energia elétrica.

Nesse contexto, a observação integrada que envolve a clusterização de municípios e o estudo do escoamento de energia é imprescindível no processo de desenvolvimento e implementação efetiva de projetos de geração renovável no estado do Ceará, proporcionando uma base sólida para impulsionar um aumento expressivo em sua contribuição na participação das fontes de energia limpa no sistema do país.

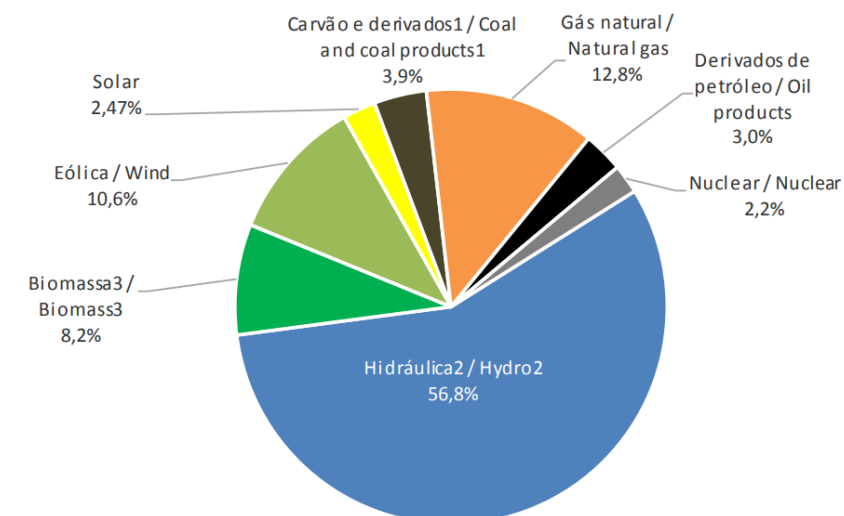
2 REVISÃO BIBLIOGRÁFICA

Neste capítulo, será fornecida a base teórica essencial para embasar as questões abordadas neste trabalho. Nele será abordado o contexto atual da matriz elétrica no Brasil e os conceitos fundamentais de produção e distribuição de energias renováveis, com ênfase nas variáveis climáticas e geográficas que exercem um papel determinante na geração de energia eólica e fotovoltaica. Além disso, será apresentada em detalhes a técnica de agrupamento (*clustering*) com foco no algoritmo *K-means* utilizado neste estudo e sua aplicação na formação de grupos com características similares.

2.1 Matriz Elétrica no Brasil

A matriz energética brasileira é amplamente sustentada por fontes de energia renovável, fornecendo a maior parte da eletricidade consumida no país (EPE, 2022). Conforme demonstrado na Figura 1, extraída do Balanço Energético Nacional de 2021 publicado pela Empresa de Pesquisa Energética (EPE) em 2022, é possível observar as proporções de geração de energia provenientes de cada fonte.

Figura 1 - Oferta interna de energia elétrica por fonte



Fonte: EPE, 2022. Balanço Energético Nacional, Ano-Base 2021.

Considerando as informações apresentadas é evidente que a fonte hídrica desempenha um papel fundamental na matriz energética do Brasil, contribuindo com 56,8% da oferta interna de eletricidade. Além disso, é importante ressaltar que as fontes renováveis, de maneira geral, têm uma significativa participação, representando 78,1% da matriz energética do país. Esse cenário reforça o compromisso brasileiro com a sustentabilidade e o aproveitamento de recursos limpos e renováveis para suprir a demanda energética (EPE, 2022).

Consonantemente verificou-se um aumento significativo de 3,4% no consumo de energia no ano de 2021 em comparação com o ano anterior, um reflexo da crescente demanda por energia em diferentes setores da sociedade, impulsionada pelo desenvolvimento econômico e pela expansão das atividades produtivas (EPE, 2022).

Figura 2 - Consumo de energia por setor em 2021



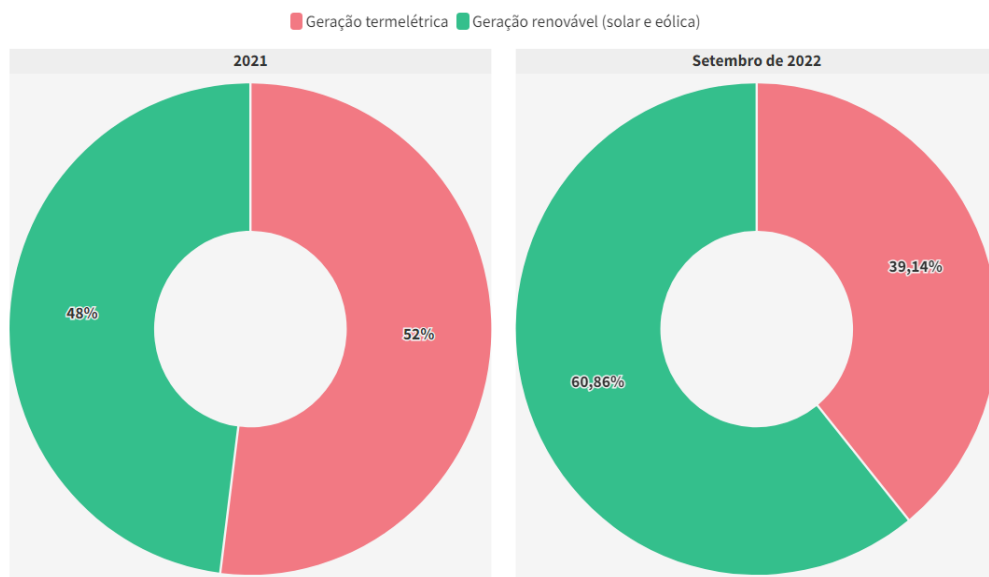
Fonte: EPE, 2022. Balanço Energético Nacional, Ano-Base 2021.

Esse aumento no consumo energético traz consigo desafios e oportunidades para o setor energético, uma vez que é necessário buscar soluções sustentáveis e eficientes para suprir essa demanda crescente, ao mesmo tempo em que se busca reduzir os impactos ambientais associados à geração e ao uso de energia.

2.1.1 Contexto Energético no Ceará

No contexto do estado do Ceará, a matriz de energia elétrica também apresenta características distintas. Atualmente, a composição da matriz energética cearense é composta por 39% de geração termelétrica e 60% de fontes renováveis, com destaque para a energia solar e eólica (ANEEL, 2022).

Figura 3 - Evolução matriz energética no Ceará



Fonte: ANEEL, 2022.

O estado se destaca como uma região propícia para o desenvolvimento de energias renováveis, contando com 159 empreendimentos em plena operação. Esses empreendimentos totalizam uma potência instalada de mais de 5,2 gigawatts (GW). Desse total, cerca de 3,2 GW correspondem a fontes de energia renovável, como a solar e eólica, enquanto 2 GW são provenientes de fontes térmicas (ANEEL, 2022).

Ademais, a implementação da Resolução Normativa ANEEL nº 482/2012 e a posterior instituição do Marco Legal da Geração Distribuída (GD) fortaleceram significativamente a geração de energia limpa no estado do Ceará, pois os consumidores passaram a ter o direito de

gerar sua própria energia elétrica a partir de fontes renováveis. Essas medidas promoveram a sustentabilidade energética, impulsionando o Ceará a se consolidar como um importante protagonista na transição para uma matriz energética mais sustentável (ANEEL, 2022; NeoSolar, [s.d]).

Outro fator importante foi o Plano Estadual de Transição Energética Justa - Ceará Verde instituído no Ceará, por meio de um decreto publicado em maio de 2022 (DECRETO Nº 34.733, 2022). Esse documento desempenha um papel crucial ao estabelecer uma estrutura propícia para atrair investimentos e financiamentos na área de energia limpa, demonstrando o objetivo de fortalecer parcerias e posicionar-se como um líder em energia renovável. Ademais, proporciona o surgimento de novas oportunidades de emprego, gerando benefícios tanto para a preservação do meio ambiente como para o desenvolvimento econômico sustentável (UNEP, I., & IOE, I., 2008).

Dessa forma, essa transição para fontes alternativas no Ceará traz consigo uma série de vantagens, tanto em termos de sustentabilidade ambiental, ao reduzir as emissões de gases de efeito estufa, quanto em termos econômicos, ao impulsionar a geração empregos na cadeia produtiva (UNEP, I., & IOE, I., 2008). Por fim, é importante ressaltar que o sucesso na transição energética requer um planejamento estratégico eficiente, identificando e selecionando as melhores áreas para a geração de energia limpa, garantindo assim o máximo aproveitamento dos recursos disponíveis.

2.2 Fatores que influenciam as gerações de Energia Eólica e Fotovoltaica

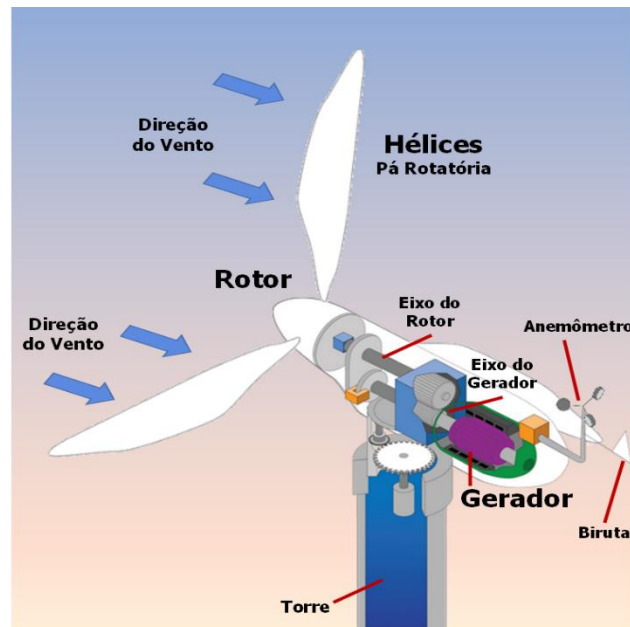
2.2.1 Energia Eólica

A energia eólica é uma forma sustentável de produzir eletricidade, aproveitando a força dos ventos para gerar energia cinética e convertê-la em energia elétrica. Esse processo é realizado por meio de turbinas eólicas ou aerogeradores, que se assemelham aos tradicionais moinhos de vento, formando assim os parques eólicos (CBIE, 2020).

Embora existam diferentes tipos de geradores eólicos, todos seguem a mesma ideia básica de aproveitar a força dos ventos para gerar eletricidade. O modelo mais comum de aerogerador é o constituído por rotores de três pás instalado sobre uma torre em forma de cata-vento. A Figura 4 fornece uma representação simplificada das principais partes desses

aerogeradores, sendo a configuração mais utilizada devido à sua eficiência na captação da energia dos ventos (CBIE, 2020).

Figura 4 - Principais componentes de um aerogerador

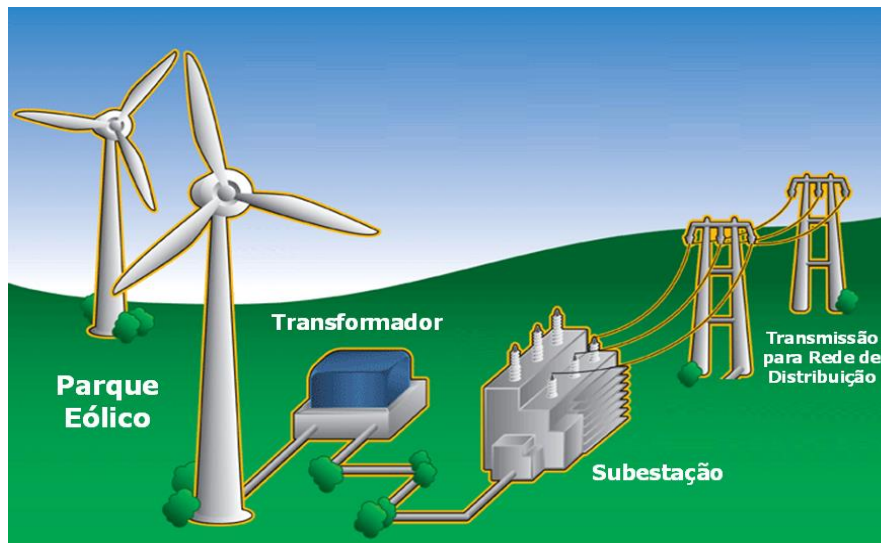


Fonte: CBIE, 2020

Em relação ao funcionamento, os aerogeradores podem ser instalados tanto de forma individual para aplicações em pequena escala, quanto agrupados em parques eólicos, que consistem em espaços terrestres ou marítimos onde estão concentradas várias unidades. Ademais, podem ser conectados a um sistema independente, porém atualmente é mais comum que estejam integrados a uma rede de distribuição elétrica.

Ademais, Sistema Interligado Nacional (SIN) desempenha um papel fundamental na integração e distribuição da energia produzida por parques eólicos em diferentes regiões do país, pois o SIN permite que a energia gerada pelos parques eólicos, localizados em áreas com boas condições de vento, seja transmitida para outras regiões onde a demanda por energia é maior (ONS, 2023).

Figura 5 - Infraestrutura comum de conexão dos parques eólicos à rede de distribuição elétrica.



Fonte: CBIE, 2020

Na visão de benefícios, é uma fonte de energia limpa e renovável que não emite gases de efeito estufa nem resíduos tóxicos durante sua operação, contribuindo diretamente para o combate às mudanças climáticas e ajuda a reduzir a emissão de poluentes na atmosfera. Além disso, os parques eólicos são projetados para operar por mais de 20 anos, o que representa uma vida útil prolongada. Ao final desse período, a área ocupada por esse empreendimento pode ser restaurada com baixos custos financeiros e ambientais, tornando o desenvolvimento eólico uma opção essencialmente reversível e sustentável (CEPEL, 2013).

No entanto é importante destacar que a eficiência de geração eólica pode ser influenciada por diversas variáveis climáticas e topográficas, dentre as principais:

- **Velocidade do vento:** a velocidade do vento é um fator crucial na geração de energia eólica, pois a quantidade de energia que pode ser extraída de um gerador é diretamente proporcional ao cubo da velocidade do vento. Isso significa que um pequeno aumento na velocidade do vento pode resultar em um aumento significativo na energia gerada (C. Marimuthu, V. Kirubakaran, 2014).
- **Altitude:** quanto mais densa for a atmosfera, maior será a energia recebida pela turbina. A densidade do ar varia com a altitude, pois o ar é menos denso em altitudes elevadas em comparação ao nível do mar e o ar quente é menos pesado do que o ar frio. Dessa forma, as turbinas produzirão mais energia em altitudes mais baixas (C. Marimuthu, V. Kirubakaran, 2014).

- Temperatura: a temperatura também influencia na densidade do ar, pois o ar quente é menos denso que o ar frio. Assim, as turbinas produzirão em locais com temperaturas médias mais frias (C. Marimuthu, V. Kirubakaran, 2014)

Além dos fatores climáticos e topográficos, o tamanho da área varrida pela turbina eólica desempenha um papel importante. Quanto maior a área varrida (ou seja, o tamanho da área através da qual o rotor gira), maior é a quantidade de energia que a turbina pode capturar do vento (C. Marimuthu, V. Kirubakaran, 2014). No entanto, vale ressaltar que esse fator não será abordado no método de clusterização utilizado neste trabalho.

2.2.2 Energia Fotovoltaica

O sistema fotovoltaico é uma forma eficiente de geração de energia elétrica pois se baseia na captura dos raios solares. Os painéis solares, por sua vez, desempenham um papel fundamental nesse processo, visto que são responsáveis por absorver a radiação solar e convertê-la em eletricidade utilizável, podendo ser empregada para alimentar residências, estabelecimentos comerciais, indústrias e propriedades rurais, oferecendo assim uma fonte sustentável e renovável de eletricidade (HCC Energia Solar, 2022).

Vale ressaltar que as placas solares geram energia elétrica em corrente contínua e, para ser utilizada em residências e empresas conectadas à rede no Brasil, é necessária a conversão para corrente alternada. Dessa maneira, o componente responsável por efetuar essa conversão é denominado inversor, o qual converte a energia gerada pelos painéis para uma voltagem de até 380 Volts. Se tratando de transmissão em linhas de alta tensão são utilizados transformadores para elevar a voltagem para valores como 13,8, 69, 138 e até acima de 230 KVolts. Essa energia é enviada das usinas fotovoltaicas, geralmente localizadas em áreas remotas, através das linhas de transmissão até os centros urbanos, onde é distribuída para o consumo de casas, empresas e outros empreendimentos conectados à rede da distribuidora local como representa a Figura 6 (PORTAL SOLAR, c.2014-2023).

Figura 6 - Esquemático do processo de geração por parques fotovoltaicos até a rede de distribuição elétrica.



Fonte: PORTAL SOLAR, c.2014-2023

A seguir, apresentam-se os fatores climáticos e topográficos que podem influenciar a geração fotovoltaica:

- Irradiação solar: a irradiação, que representa a intensidade luminosa do sol, desempenha um papel crucial na geração fotovoltaica. Em dias ensolarados, a capacidade de geração é maior, dessa forma, é fundamental posicionar as placas solares de modo a captar o máximo de radiação solar disponível. No Brasil, a orientação ideal para os painéis solares é voltada para o norte, proporcionando condições mais favoráveis de exposição aos raios solares. (FEIJÓO COELHO, 2017; HCC Energia Solar, 2022).
- Temperatura: células solares são bastante suscetíveis a danos por altas temperaturas devido às suas grandes superfícies. A degradação das características elétricas e ópticas das células solares, como detectores de luz em condições de aumento de temperatura, é um dos fatores limitantes mais importantes para sua aplicação. Uma vez que a maioria dos processos elétricos em dispositivos semicondutores depende da temperatura, investigações em temperaturas mais altas que ambiente podem revelar possíveis alterações nas características de saída do dispositivo. Além disso, a vida útil da célula solar é restrita pelo grau de danos por radiação que a célula recebe (C. Marimuthu, V. Kirubakaran, 2014).
- Altitude: em altitudes mais elevadas, a densidade do ar é menor, resultando em uma menor obstrução atmosférica para a passagem dos raios solares. Isso significa que a

radiação solar incidente é menos dispersa e mais direta, aumentando a quantidade de energia solar disponível para conversão em eletricidade. (LUCAS, 2021)

É importante ressaltar outros fatores que influenciam nas perdas e eficiência dos equipamentos, mas que não estarão presentes neste projeto devido à metodologia a ser abordada:

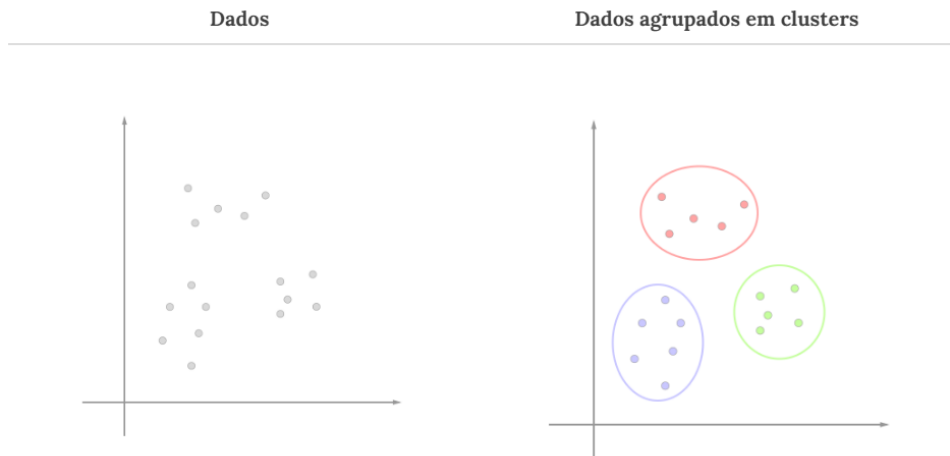
- Sombras: presença de sombras impactam negativamente, caso ocorra bloqueio nas células solares, elas deixam de gerar corrente elétrica e passam a funcionar como uma resistência, limitando toda a corrente em série de todas as células fotovoltaicas. Conseqüentemente, se esse problema persistir, é possível que as células afetadas pelo sombreamento sofram danos, resultando em uma redução na geração de energia do sistema fotovoltaico (HCC Energia Solar, 2022).
- Incompatibilidade elétrica entre os módulos: a combinação de módulos com diferentes características elétricas resulta em níveis de desempenho mais baixos, além das perdas associadas ao inversor durante os processos de conversão e chaveamento de potência (FEIJÓO COELHO, 2017).
- Orientação e inclinação: as perdas relacionadas a orientação e inclinação devem ser consideradas, pois a posição que resultará na maior geração de energia é aquela em que o feixe de luz incide diretamente no painel solar de forma perpendicular ao plano do módulo. Essa configuração garante uma absorção ótima da radiação solar, maximizando assim a eficiência e a produção de energia dos painéis solares (HCC Energia Solar, 2022).
- Poluição: a manutenção adequada dos painéis solares é essencial para garantir a máxima eficiência na geração de energia solar, pois o acúmulo de sujeira, como poeira, folhas, poluição atmosférica e resíduos, pode se tornar significativo ao longo do tempo e afetar demasiadamente a capacidade de produção energética dos painéis solares. Essa acumulação pode reduzir a quantidade de luz solar que atinge as células fotovoltaicas, diminuindo a eficiência de conversão e, conseqüentemente, a produção de eletricidade. (FEIJÓO COELHO, 2017).

2.3 Método de agrupamento (*clustering*)

A clusterização é um método analisado principalmente nas áreas de estatística e matemática (BERKHIN, 2002). No entanto, na área da computação, esses problemas surgiram com a popularização do conceito de mineração de dados (*Data Mining*). Trata-se de uma técnica não supervisionada de aprendizagem de máquina, ou seja, é uma abordagem da inteligência artificial em que algoritmos são utilizados para descobrir padrões ou estruturas em conjuntos de dados não rotulados. Diferentemente do aprendizado supervisionado, em que já existem exemplos rotulados, o método não supervisionado requer que os algoritmos identifiquem por si mesmos a estrutura subjacente nos dados sem rótulos ou categorias predefinidas (HONDA; FACURE; YAOHAO, 2017).

De uma forma geral, busca-se encontrar uma representação mais informativa de uma grande quantidade de dados disponíveis, sendo possível analisar padrões grupais por meio de técnicas que medem a similaridade entre as informações e os grupos gerados por esse tipo de classificação são chamados de *clusters* (BERKHIN, 2002). Sob essa perspectiva, o principal objetivo é obter *insights* e conhecimentos relevantes, que possam ser aplicados em diferentes áreas de estudo e pesquisa.

Figura 7 - Representação dos dados: antes e após o método de agrupamento



Fonte: HONDA; FACURE; YAOHAO, 2017

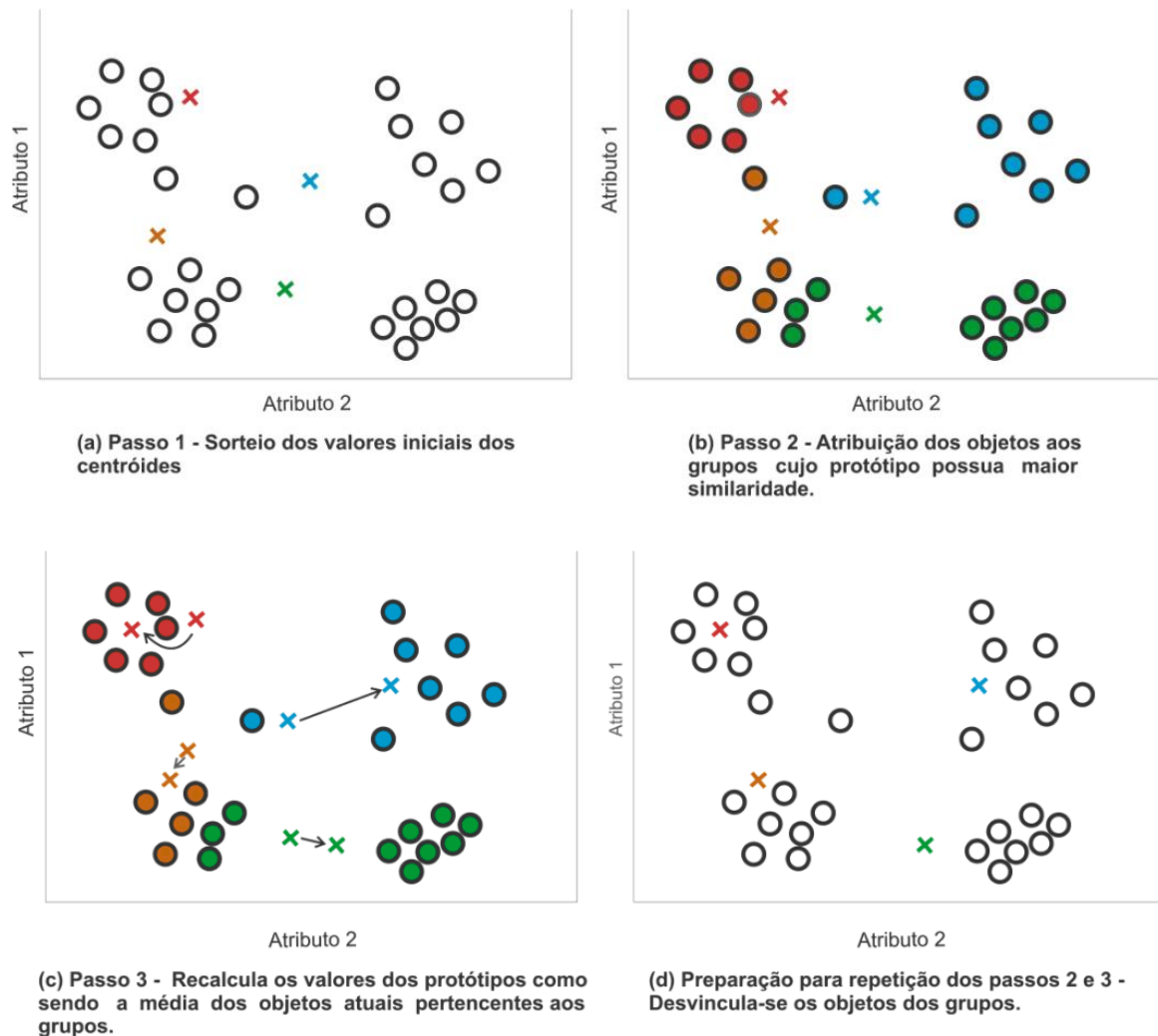
2.3.1 Método *K-means*

O algoritmo *K-means* ou K-médias tem essa nomenclatura porque sua principal tarefa é encontrar K grupos distintos dentro de um conjunto de dados. Para cada agrupamento identificado é atribuído um centroide que representa o centro do *cluster* e, a partir disso, é calculado como a média dos valores dos elementos pertencentes a esse *cluster*. Diante disso, a abordagem do *K-means* permite uma organização eficiente dos dados em grupos coerentes, além de facilitar a análise e interpretação dos padrões presentes nos conjuntos de dados (HONDA, 2017).

Nessa metodologia é comumente utilizada a métrica de soma dos erros quadrados, conhecida como *Sum of Squared Errors* (SSE), para minimizar a distância entre os pontos e seus centroides no algoritmo *K-means*. Após a criação do modelo são atribuídas as médias de cada *cluster* e assim é identificado o centroide mais próximo para cada novo dado, representando o centro de cada *cluster* para determinar organização e categorização dos dados (PIRES GUEDES, 2019).

De forma simplificada, o algoritmo *K-means* pode ser descrito pelos seguintes passos, de acordo com Fontana e Naldi (2009):

1. Inicialização: atribuição de valores iniciais para os centroides dos *clusters*. Podendo esses serem escolhidos de forma aleatória dentro dos limites de domínio de cada atributo ou por métodos.
2. Atribuição dos objetos aos *clusters*: cada objeto é atribuído ao grupo cujo centroide possui a maior similaridade. Essa semelhança pode ser medida utilizando distâncias, como a distância euclidiana, que foi utilizada no presente trabalho.
3. Atualização dos centroides: o valor do centroide de cada *cluster* é recalculado como a média dos objetos pertencentes a esse *cluster*.
4. Iteração: os passos 2 e 3 são repetidos até que os grupos se estabilizem, ou seja, quando não ocorrerem mais alterações nas atribuições dos objetos aos *clusters* ou nos valores dos centroides.

Figura 8 - Passo a passo algoritmo *k-means*

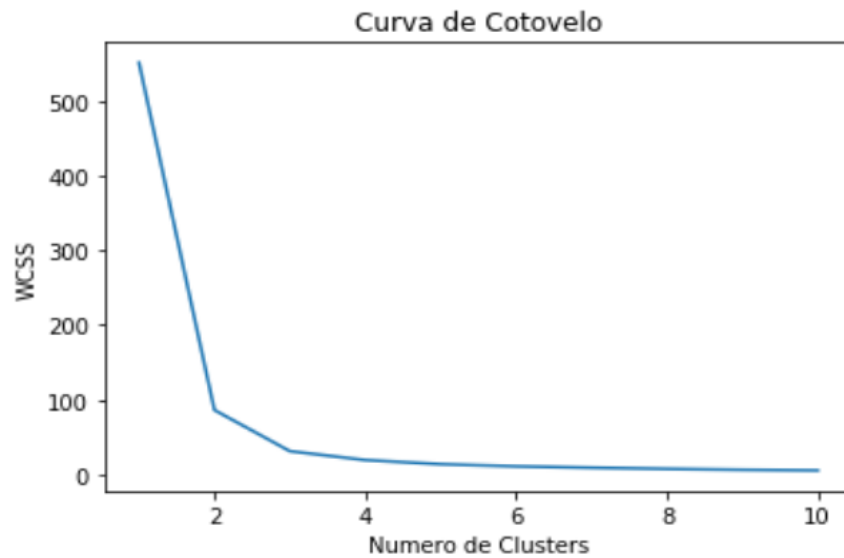
Fonte: (FONTANA; NALDI 2009)

Os passos descritos na Figura 8 são essenciais para a convergência do algoritmo e para a formação dos *clusters* finais.

Outro fator importante para determinar a quantidade ideal de *clusters* ao realizar um algoritmo, é a utilização do método *Elbow Curve* (também conhecido como Curva de Cotovelo) que consiste em uma técnica amplamente utilizada para determinar a quantidade ideal de grupos (K) em um conjunto de dados. Essa abordagem avalia a variância dos dados em relação ao número de *clusters*, identificando que o valor ideal de K é aquele que resulta em uma menor soma dos quadrados dentro dos *clusters* (*Within Sum of Squares* - WSS), ao mesmo tempo em que busca minimizar o número de *clusters* utilizados. O nome "Curva de Cotovelo" é atribuído a esse método porque, a partir do ponto em que ocorreria a curvatura acentuada na curva, não é observada uma redução significativa na variância (PIRES GUEDES, 2019). Assim, o número

de *clusters* é definido pelo ponto que existe uma mudança abrupta no gráfico como, por exemplo, a Figura 9 demonstra que é no ponto 3 do eixo x do gráfico.

Figura 9 - Representação da curva do cotovelo



Fonte: PIRES GUEDES, 2019

Existem vários métodos para determinar a quantidade ideal de *clusters* em uma clusterização, incluindo o Método de Cotovelo, *Silhouette*, Critério de Informação Bayesiana (BIC) e *Gap Statistics*. O método de Cotovelo foi escolhido pois proporcionou uma identificação clara e objetiva do número adequado de *clusters* para o conjunto de dados analisado, tornando a interpretação mais direta e facilitando a tomada de decisões mais apropriadas.

3 APLICAÇÃO DA PROPOSTA DE PROJETO

Neste capítulo, descreve-se a metodologia empregada nesta pesquisa, abordando os seguintes aspectos: conjunto de dados analisados, linguagem de programação, *softwares* e bibliotecas adotadas, tratamentos realizados nas bases de dados e aplicação da técnica de clusterização.

3.1 Conjuntos de dados avaliados

Para a obtenção das informações necessárias neste estudo, foram consideradas as variáveis-chave que exercem influência na geração de energia eólica e fotovoltaica e a disponibilidade de dados públicos. Além disso, foram coletados alguns dados complementares de acordo com a necessidade na fase de transformação de dados.

3.1.1 Irradiação Solar:

Para essa variável foi extraída uma base de dados (*dataset*) que contém informações sobre a irradiação solar a nível de média anual e média mensal, ambas do total diário da irradiação Global Horizontal no Ceará em (Wh/m²). Além disso, as informações foram coletadas em formato *shapefile* que consiste em um armazenamento de dados vetoriais da *Esri* e é amplamente utilizado para armazenar informações sobre a posição, forma e atributos de feições geográficas.

A fonte desse conjunto de dados é o ATLAS brasileiro de Energia Solar (ATLAS, 2017), na qual fornece informações sobre a irradiação solar em várias regiões do Brasil, permitindo avaliar o potencial de geração de energia solar em diferentes estados. É importante destacar que não há limitações no uso deste dado como entrada para diversas análises, desde que a fonte seja devidamente mencionada.

3.1.2 Limites Municipais do Ceará:

Esse conjunto de dados em formato *shapefile* contém os limites municipais do estado do Ceará. O Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE) disponibilizou essas informações em 2021 que são essenciais para delimitar geograficamente os municípios do estado e o arquivo foi utilizado para associar os dados de interesse aos seus respectivos municípios junto ao *dataset* de irradiação solar.

3.1.3 Temperatura e Velocidade do Vento:

Os dados históricos dessas variáveis por município foram obtidos a partir do Instituto Nacional de Meteorologia (INMET), na qual se refere ao período compreendido entre janeiro e junho do ano de 2023 e têm-se informações a nível de hora. Deste modo, podemos ressaltar que essas informações são relevantes para a análise de clusterização pois contém informações acerca da temperatura (em graus Celsius) e a velocidade do vento (em m/s). Contudo, para o Estado do Ceará, esse *dataset* só contém informações de uma lista de 16 municípios, sendo eles: Fortaleza, Sobral, Guaramiranga, Barbalha, Iguatu, Tauá, Quixeramobim, Morada Nova, Jaguaruana, Crateús, Campos Sales, Jaguaribe, Itapipoca, Acaraú, Tianguá e Quixadá, sendo necessária uma designação de valores para as cidades com informações faltantes.

3.1.4 Dados das Subestações no Brasil:

Esses dados são disponibilizados em formato *shapefile* pela Empresa de Pesquisa Energética (EPE) e se referem às subestações instaladas no Brasil. O objetivo de coleta dessas informações é fornecer informações sobre a infraestrutura energética e a possibilidade de identificar lacunas na distribuição de energia renovável produzida no Ceará.

3.1.5 Altitudes dos municípios do Ceará

Através de uma *Application Programming Interface* (API) do Google chamada *Geocoding* e utilizando um arquivo obtido no repositório o "Municípios-Brasileiros" mantido pelo usuário chamado "kelvins" no *GitHub* que continha as informações com latitudes e longitudes de todos os municípios do Brasil, foi possível extrair a informação de *elevation* por cidade do Ceará que, trata-se, do dado de altitude e impacta nas gerações evidenciadas nesse trabalho.

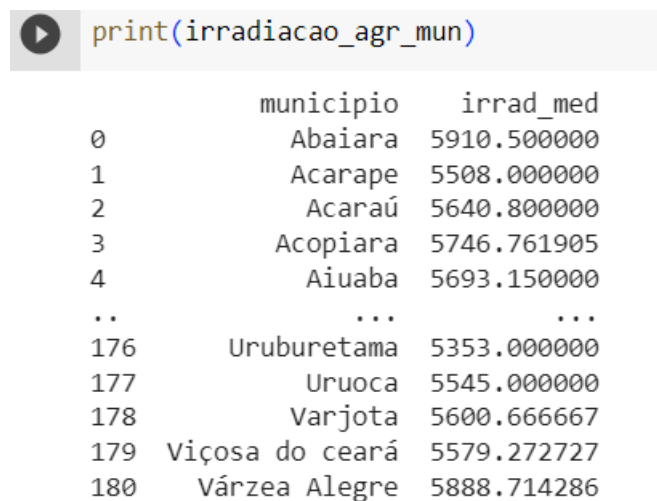
3.2 Linguagem de programação, bibliotecas e softwares utilizados

Para o tratamento dos dados e aplicação do método de clusterização foi utilizada a linguagem Python (versão 3.10.12), com a aplicação de algumas bibliotecas de manipulação, cálculos, plotagem de gráficos e mapas como, por exemplo, *Pandas*, *NumPy*, *Matplotlib*, *Folium*, *Seaborn* e, em particular, a biblioteca *sklearn.cluster* importando *KMeans* para executar o agrupamento final dos dados em *clusters*. Além disso, o software *PowerBi* foi explorado também para plotagem extras de mapas.

3.3 Transformações nas bases de dados

Diante da premissa que o *cluster* será executado a nível de cidades do Ceará, primeiramente foi necessário realizar uma padronização de todos os arquivos para esse formato. Assim, inicialmente, foi realizada associação de municípios ao arquivo *shapefile* de irradiação, pois este continha apenas informações a nível de latitude e longitude. Para essa associação, utilizou-se outro conjunto de dados disponibilizados pelo IPECE, também, em formato *shapefile* contendo as delimitações do estado, com o objetivo de atribuir o nome das cidades a cada coordenada presente na base de dados inicial. Em seguida, utilizando a biblioteca *pandas*, calculou-se a média da irradiação anual por município, resultando em um *dataframe* com 181 municípios como mostra a Figura 10.

Figura 10 - Pré-visualização do *dataframe* de irradiação anual tratado



```
print(irradiacao_agr_mun)
```

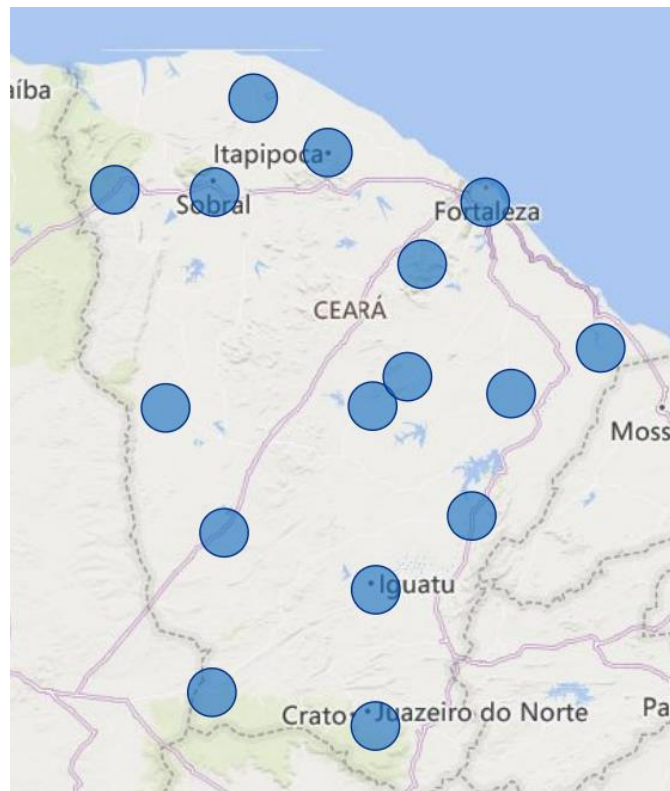
	municipio	irrad_med
0	Abaiara	5910.500000
1	Acarape	5508.000000
2	Acaraú	5640.800000
3	Acopiara	5746.761905
4	Aiuaba	5693.150000
..
176	Uruburetama	5353.000000
177	Uruoca	5545.000000
178	Varjota	5600.666667
179	Viçosa do ceará	5579.272727
180	Várzea Alegre	5888.714286

Fonte: elaborada pelo autor.

Considerando que o estado tem 184 municípios (IPECE, 2021), as três cidades que estavam sem informações tiveram uma designação por meio do método de *haversine* em Python, que determina as três cidades mais próximas de cada um desses municípios e, em seguida, designa a média de irradiação solar de três cidades adjacentes para preencher os valores faltantes das cidades em questão.

Além disso, a base de referências meteorológicas do INMET só contém informações de 16 municípios do Ceará, sendo estes distribuídos de forma que contempla todas as regiões do estado, como está evidenciado na Figura 11.

Figura 11 - Mapa dos municípios com informações históricas no INMET



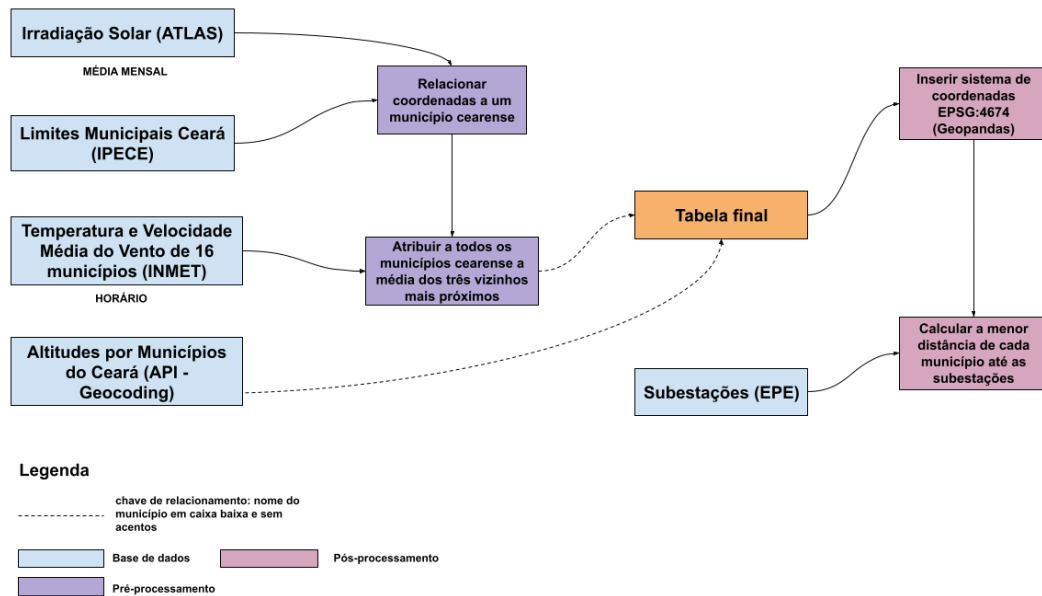
Fonte: elaborada pelo autor.

Em função disso, também foi necessário aplicar o método de *haversine* a fim de designar o valor da média das três cidades mais próximas de todas as coordenadas dos municípios faltantes para atribuir as informações de velocidade média do vento e temperatura média.

Por conseguinte, para obter a tabela final necessária para a primeira etapa da análise, foi realizada a união entre os dois conjuntos de dados tratados. Essa união envolveu a integração através da chave de ligação formada pelo nome minúsculo e sem acento dos 184 municípios, unindo informações de irradiação média anual e altitude, provenientes da primeira base de dados preparada, juntamente com os dados de velocidade média do vento e temperatura média obtidos da base do INMET.

Por fim, houve a necessidade de adicionar o campo *geometry* na base tratada, visto que os dados de subestações vêm em formato *shapefile*. A união desses dados nos possibilita analisar a viabilidade de distribuição de energia.

Figura 12 - Esquemático de tratamento de dados



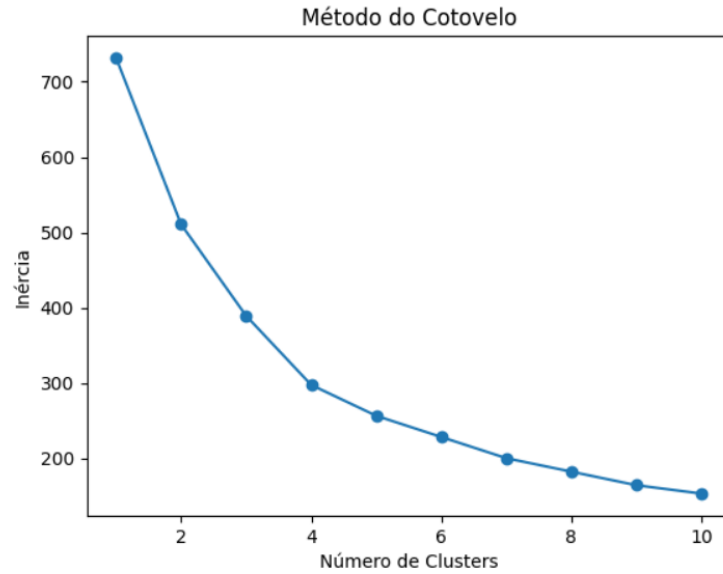
Fonte: elaborada pelo próprio autor.

3.4 Aplicação do *K-means*

A partir da base de dados unificada e tratada, foi possível executar o algoritmo *K-means* utilizando as quatro variáveis disponíveis. Para isso, utilizou-se a biblioteca *sklearn.cluster* e seguiu-se os seguintes passos:

1. Normalização dos dados: inicialmente, os dados foram normalizados com o objetivo de dimensionar os valores dos atributos para uma mesma faixa ou escala. Esse processo é essencial para garantir que todos os atributos tenham a mesma importância durante o cálculo das distâncias entre os pontos, evitando que atributos com escalas maiores dominem o processo de clusterização. Dessa forma, foi utilizada a padronização que converte a média para zero e o desvio padrão para um em todas as variáveis.
2. Determinação do número ideal de grupos: utilizando o método do cotovelo (*Elbow Curve*) foi possível identificar o número ótimo de *clusters*. Esse método envolve a análise do gráfico evidenciado na Figura 12 que mostra a relação entre o número de *clusters* e a variância explicada pelos *clusters*. O ponto de inflexão no gráfico, conhecido como "cotovelo" representa o número ideal de *clusters*.

Figura 13 - Gráfico do cotovelo para a clusterização de municípios



Fonte: elaborada pelo próprio autor.

Ao analisar o gráfico, observou-se que o número ideal de grupos são 4, visto que a partir desse ponto no eixo x existe uma tendência linear na curva.

- Execução do algoritmo de clusterização: com base no número de *clusters* determinado no passo anterior, o algoritmo *K-means* foi executado com o objetivo de agrupar os municípios de acordo com seus atributos parecidos, levando em consideração as variáveis analisadas.
- Adição do rótulo de *cluster*: ao final do processo de clusterização, o rótulo de *cluster_label* foi adicionado à base de dados inicial, indicando a qual grupo cada cidade pertence. Essa informação permite a posterior análise e interpretação dos resultados obtidos.

Figura 14 - Pré-visualização da base de dados clusterizada

```
[66] base_final.head()
```

	municipio	uf	latitude	longitude	altitude	media_velocidade_vento	media_temperatura	chave	irrad_med	cluster_label
0	Abaiara	CE	-7.34588	-39.0416	386.262207	1.605062	25.078958	abaiara	5910.500000	2
1	Acaraú	CE	-2.88769	-40.1183	19.753691	1.783420	26.391786	acarau	5508.000000	0
2	Acarape	CE	-4.22083	-38.7055	95.597076	1.618619	25.025069	acarape	5640.800000	0
3	Acopiara	CE	-6.08911	-39.4480	313.874878	1.871038	26.854507	acopiara	5746.761905	3
4	Aluaba	CE	-6.57122	-40.1178	467.503235	1.882526	25.511495	aluaba	5693.150000	1

Fonte: elaborada pelo próprio autor.

De acordo com esses passos foi possível a criação de grupos distintos de municípios com características similares com base nas variáveis analisadas, facilitando a compreensão e tomada de decisões relacionadas à geração de energia solar e eólica nesses locais.

4 RESULTADOS E DISCUSSÕES

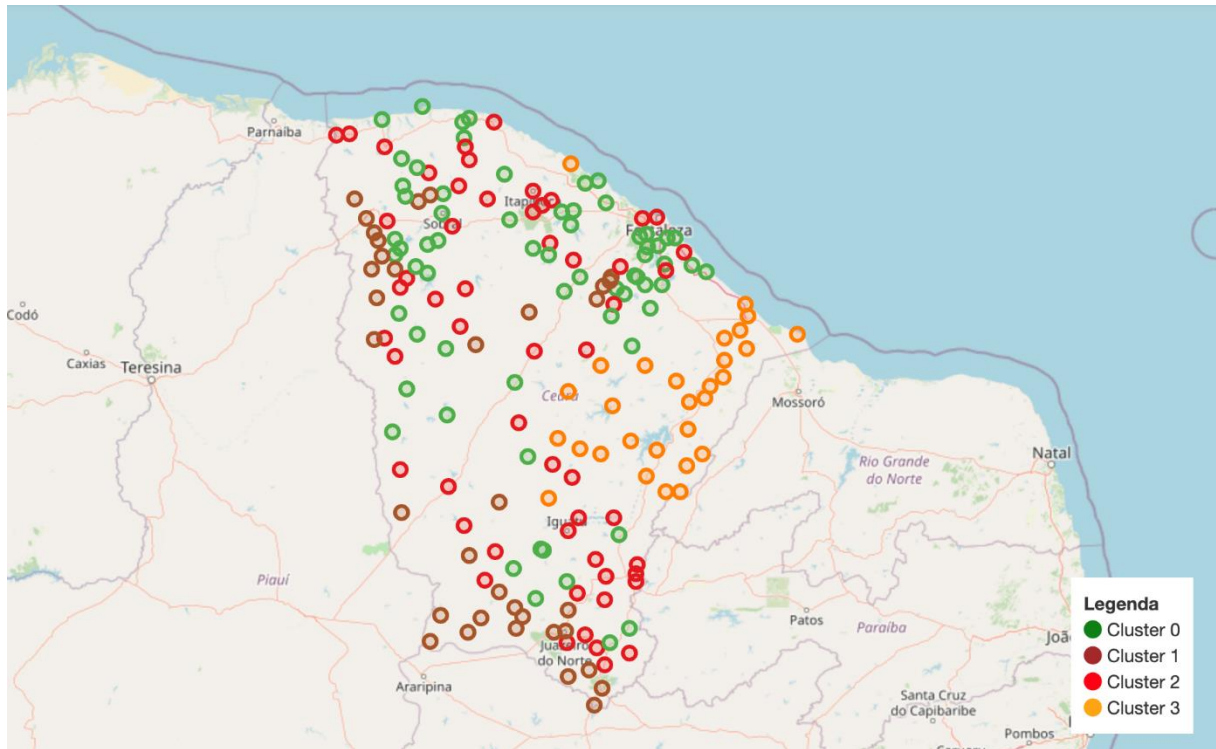
4.1 Análise de *cluster*

Por meio da aplicação do algoritmo de clusterização nos dados geográficos do Ceará foi possível identificar a formação de quatro grupos distintos (0, 1, 2 e 3). Sendo estes constituídos das seguintes formas:

- Grupo 0: 65 municípios
- Grupo 1: 35 municípios
- Grupo 2: 55 municípios
- Grupo 3: 29 municípios

Posteriormente, cada grupo foi associado a um rótulo específico, representando as características similares identificadas nos municípios analisados, conforme demonstrado na Figura 14.

Figura 15 - Mapa de municípios clusterizados



Fonte: elaborado pelo próprio autor

A visualização no mapa proporciona uma compreensão mais clara da distribuição dos grupos no território cearense, permitindo identificar uma dispersão de alguns grupos no estado. Nota-se que o *cluster* 3 apresenta uma uniformidade marcante em relação aos demais no mapa e essa característica pode ser atribuída à influência da variável de altitude, uma vez que os grupos exibem uma variação mais significativa nesse fator topológico em relação aos outros fatores climáticos. Essa hipótese é corroborada pelos dados apresentados na Tabela 1, na qual são exibidas as médias das variáveis para cada *cluster*. Observa-se que os valores mais divergentes entre os *clusters* são relacionados à altitude, indicando que a variável desempenha um papel significativo na diferenciação dos grupos.

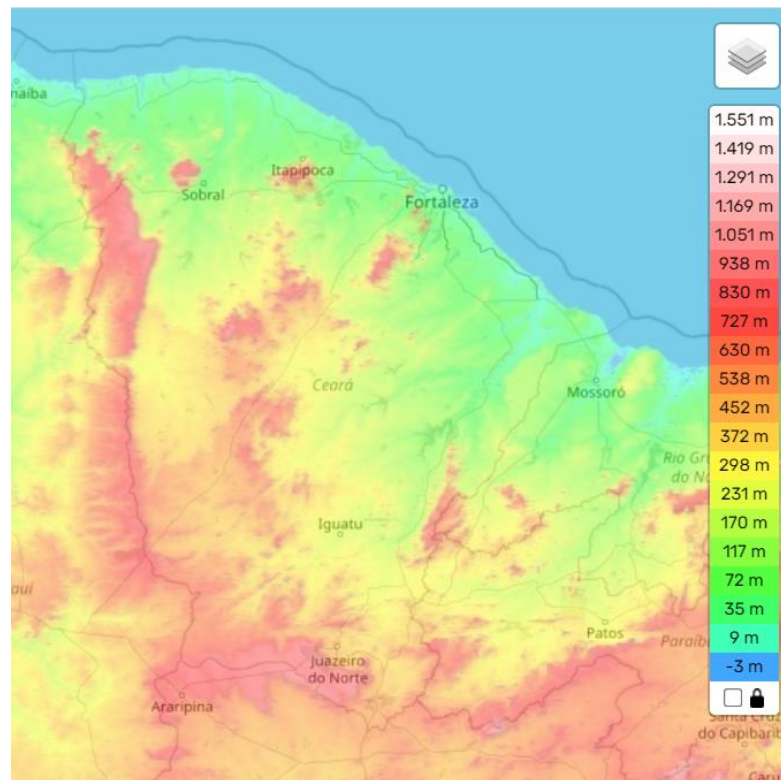
Tabela 1 - Média das variáveis por grupo

Rótulo do cluster	Altitude Média (m)	Irradiação Média (Wh/m²)	Média Velocidade média do Vento (m/s)	Temperatura Média (°C)
0	136,52	5564,69	1,70	25,40
1	653,08	5675,31	1,65	25,14
2	214,39	5822,27	1,72	25,71
3	119,60	5686,03	2,10	27,23

Fonte: elaborada pelo próprio autor.

Além disso, o mapa disponível no site *topographic-map*, evidenciado na Figura 15, ilustra a variação de altitude em todo o estado do Ceará, revelando que a área que o grupo 3 está compreendida não sofre muitas flutuações em suas altitudes.

Figura 16 - Mapa topográfico do Ceará



Fonte: *topographic-map*, [s.d]

Tabela 2 - Estatísticas de altitude por *cluster*

Cluster	Média	Desvio Padrão	Mínimo	Primeiro Quartil (25%)	Mediana	Terceiro Quartil (75%)	Máximo
0	136,52	103,52	8,49	56,36	103,45	183,81	367,85
1	653,08	153,03	404,93	516,19	671,16	776,82	921,45
2	214,39	135,35	7,02	88,40	237,22	302,24	506,29
3	119,60	118,66	8,45	23,54	117,38	180,53	560,88

Fonte: elaborada pelo próprio autor.

Tabela 3 - Estatísticas de temperatura por *cluster*

Cluster	Média	Desvio Padrão	Mínimo	Primeiro Quartil (25%)	Mediana	Terceiro Quartil (75%)	Máximo
0	25,40	0,58	24,88	25,00	25,09	25,73	26,59
1	25,14	0,29	24,93	25,01	25,08	25,11	26,59
2	25,71	0,70	24,79	25,08	25,51	26,42	27,05
3	27,23	0,26	26,61	27,09	27,18	27,43	27,58

Fonte: elaborada pelo próprio autor.

Tabela 4 - Estatísticas de irradiação por *cluster*

Cluster	Média	Desvio Padrão	Mínimo	Primeiro Quartil (25%)	Mediana	Terceiro Quartil (75%)	Máximo
0	5564,69	86,97	5353,00	5508,00	5562,00	5621,50	5715,80
1	5675,31	135,40	5433,00	5600,74	5648,73	5779,40	5935,33
2	5822,27	78,98	5662,00	5756,71	5822,50	5887,56	5995,86
3	5686,03	147,84	5408,00	5589,33	5649,69	5781,63	5980,25

Fonte: elaborada pelo próprio autor.

Tabela 5 - Estatísticas de velocidade média do vento por *cluster*

Cluster	Média	Desvio Padrão	Mínimo	Primeiro Quartil (25%)	Mediana	Terceiro Quartil (75%)	Máximo
0	1,70	0,13	1,19	1,61	1,73	1,77	1,90
1	1,65	0,09	1,55	1,61	1,61	1,75	1,88
2	1,72	0,16	1,19	1,59	1,75	1,78	2,10
3	2,10	0,12	1,78	2,04	2,12	2,21	2,24

Fonte: elaborada pelo próprio autor.

As tabelas 2, 3, 4 e 5 apresentam as estatísticas de cada variável utilizada no agrupamento, complementando as demais análises e contribuindo para uma caracterização mais precisa dos grupos.

4.2 Classificação dos melhores grupos para Geração Eólica e Fotovoltaica

4.2.1 Geração Eólica

Com base no tópico 2.2.1 do presente trabalho e na análise das variáveis durante o processo de clusterização foi identificado que certos aspectos têm um impacto positivo na geração eólica. Essas características incluem altas temperaturas, velocidade do vento mais elevada e altitudes mais elevadas. Diante do exposto, necessitou-se de uma metodologia de pontuação que utiliza esses três fatores como indicadores para determinar o grupo mais adequado para a geração eólica.

Dessa forma, foram adotados os seguintes procedimentos:

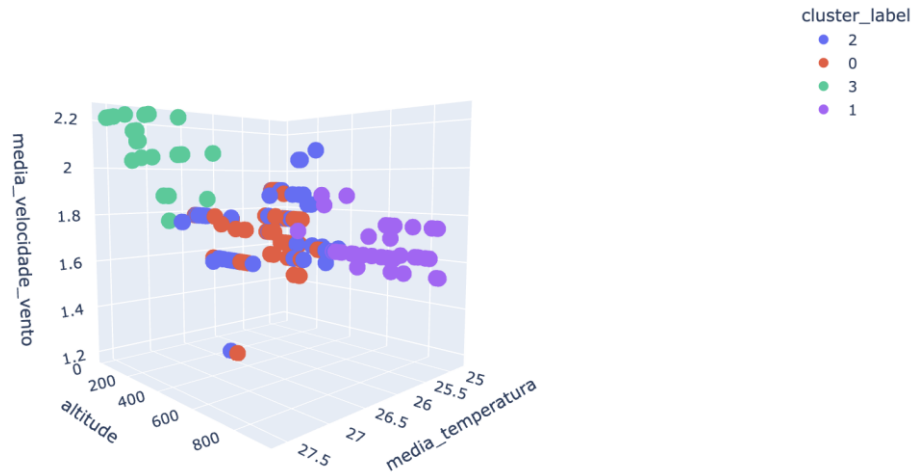
1. Utilizando as médias calculadas para cada variável em cada grupo, atribuiu-se uma pontuação de 1 a 4 para as variáveis de altitude média, média da velocidade média do vento e temperatura média, em que 1 representa o valor mínimo e 4 o valor máximo. Por exemplo, conforme apresentado na Tabela 6, o grupo 1 obteve o maior valor médio de altitude, o que resultou na atribuição do valor 4, indicando sua influência positiva na geração.
2. Por fim, somou-se as três pontuações das variáveis para obter um *score* final, denominado "Total" na Tabela 6.

Tabela 6 - Grupos classificados para geração eólica

Rótulo do cluster	Altitude média (m)	Média velocidade média do vento (m/s)	Temperatura média (°C)	Altitude score	Velocidade vento score	Temperatura score	Total
2	214,39	1,72	25,71	3	3	3	9
3	119,60	2,10	27,23	1	4	4	9
0	136,52	1,70	25,40	2	2	2	6
1	653,08	1,65	25,14	4	1	1	6

Fonte: elaborada pelo próprio autor.

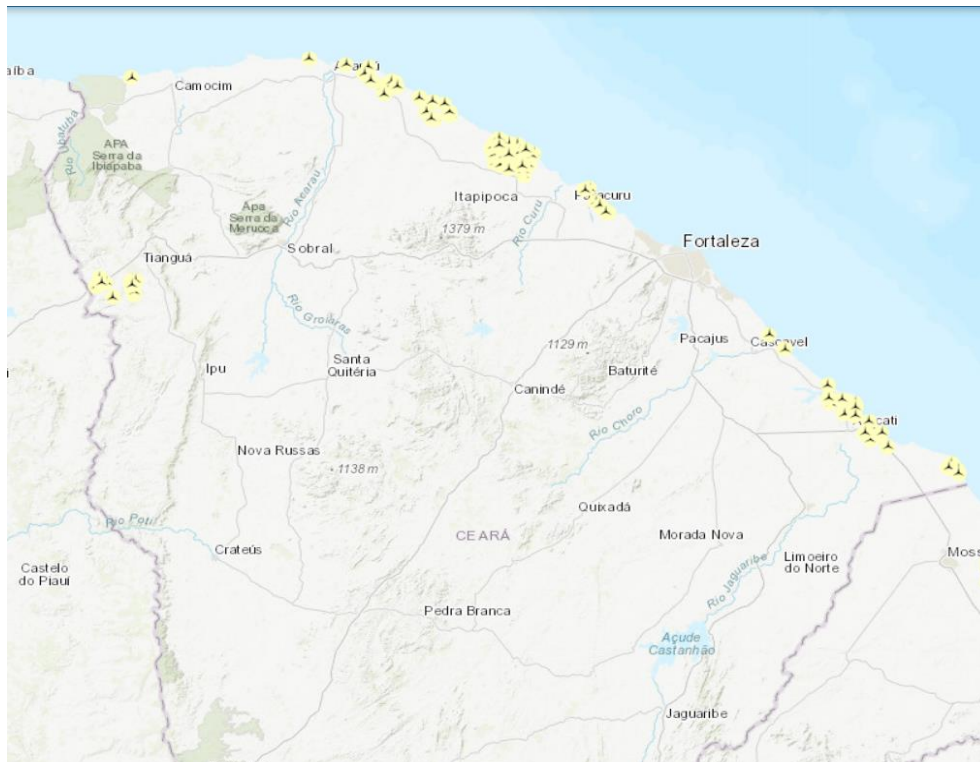
Figura 17 - Mapa 3D das variáveis mais relevantes para geração eólica



Fonte: elaborado pelo próprio autor

Podemos compreender com base na Tabela 6 que houve um empate na pontuação total dos grupos 2 e 3. No entanto, também observando a Figura 16, considerando o grau de impacto na eficiência de geração eólica, é preferível priorizar o grupo 3, pois se destaca pelos valores de velocidade média do vento e temperatura, com todos os membros do grupo apresentando valores acima da média para essas variáveis. Em relação à altitude, o grupo 3 possui uma média mais baixa, com 93% dos seus membros possuindo altitudes inferiores à média geral de 255,39 metros. Sob essa perspectiva, as cidades que formam o *cluster* 3 são: Acopiara, Alto Santo, Aracati, Banabuiú, Fortim, Ibicuitinga, Icapuí, Iracema, Itaiçaba, Jaguaretama, Jaguaribara, Jaguaribe, Jaguaruana, Limoeiro do Norte, Milhã, Morada Nova, Palhano, Pereiro, Potiretama, Quixadá, Quixeramobim, Quixeré, Russas, São João do Jaguaribe, Senador Pompeu, Solonópole, Tabuleiro do Norte, Trairi e Ererê.

Figura 18 - Empreendimentos de geração eólica no Ceará



Fonte: EPE.

Além da classificação observada, observou-se também o *WebMap* EPE de empreendimentos eólicos já existentes no Ceará disponibilizados pela Empresa de Pesquisa Energética (EPE), e atualmente, há uma maior concentração no litoral cearense do que na totalidade de área que abrange o *cluster* três, indicando a necessidade de redirecionar esforços para a alteração ou mais instalações de parques eólicos em uma área mais propícia.

4.2.2 Geração Fotovoltaica

Com base no tópico 2.2.2 do presente trabalho e na análise das variáveis durante o processo de clusterização, foi constatado que certas características exercem um impacto positivo na geração de energia solar. Essas características incluem baixas temperaturas, altos níveis de irradiação solar e altitudes mais elevadas dos municípios. A partir dessas constatações, tornou-se necessário desenvolver uma metodologia de pontuação que utiliza esses três fatores como indicadores para determinar o grupo mais adequado para a produção de energia solar.

Conseqüentemente, as seguintes etapas foram implementadas:

1. Utilizando as médias calculadas para cada variável em cada grupo, atribuiu-se uma pontuação de 1 a 4 para as variáveis de altitude média, irradiação média e temperatura

média, em que 1 representa o valor mínimo e 4 o valor máximo. Por exemplo, conforme apresentado na Tabela 7, o grupo 1 obteve o menor valor médio de temperatura, o que resultou na atribuição do valor 4, indicando que menores temperaturas têm influência positiva na geração solar.

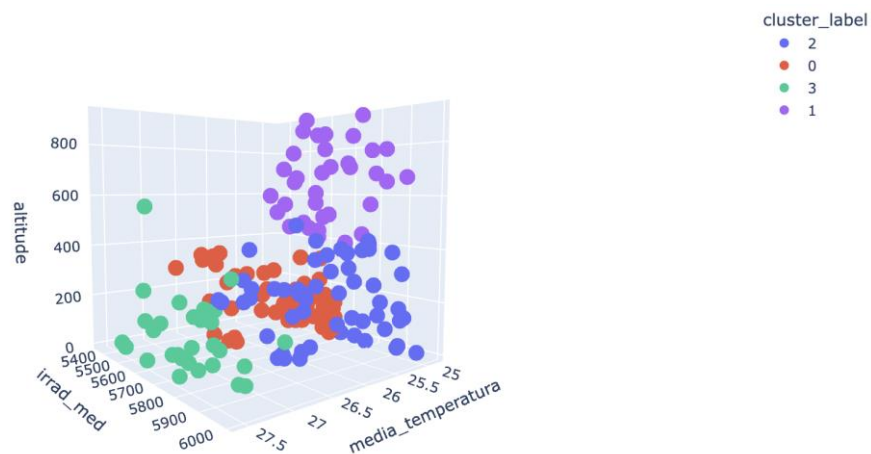
- Por fim, somou-se as três pontuações das variáveis para obter um score final, denominado "Total" na Tabela 7.

Tabela 7 - Grupos classificados para geração fotovoltaica

Rótulo do cluster	Altitude média (m)	Velocidade do vento média (m/s)	Temperatura média (°C)	Altitude score	Velocidade do vento score	Temperatura score	Total
1	653,08	5675,31	25,14	4	2	4	10
2	214,39	5822,27	25,71	3	4	2	9
0	136,52	5564,69	25,40	2	1	3	6
3	119,60	5686,03	27,23	1	3	1	5

Fonte: elaborada pelo próprio autor.

Figura 19 - Mapa 3D das variáveis mais relevantes para geração solar

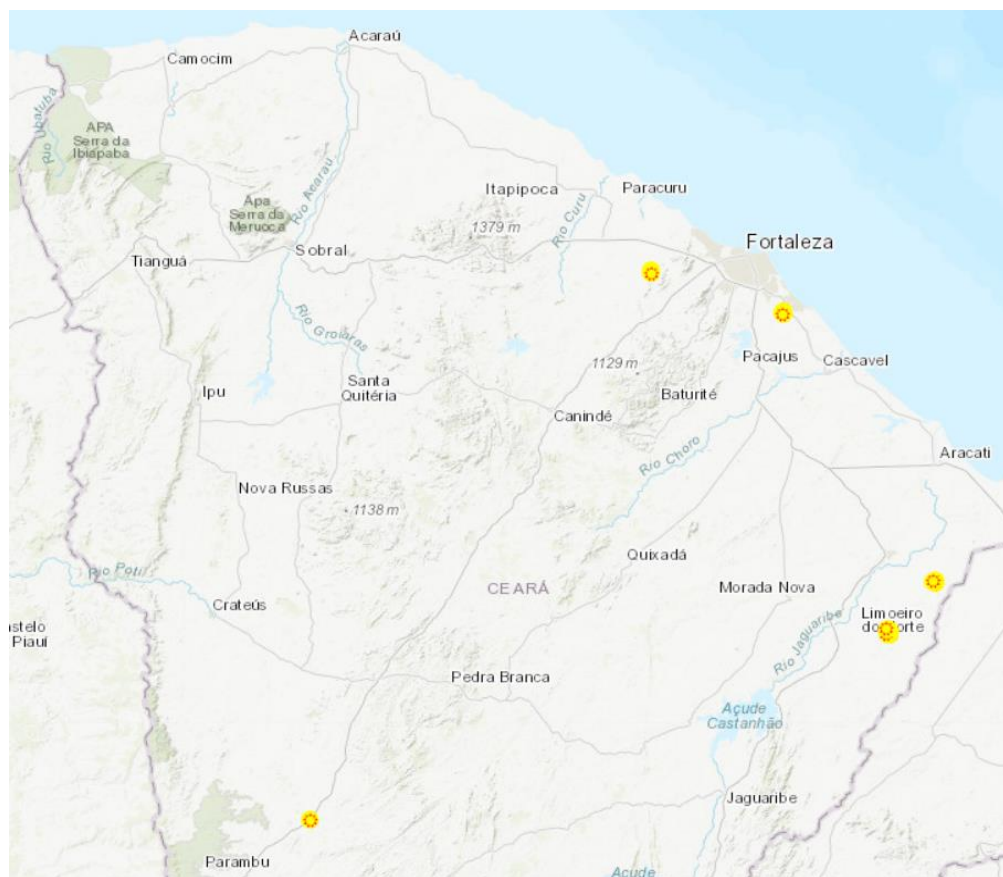


Fonte: elaborado pelo próprio autor

Ao analisar a Tabela 2 e a Figura 17, nota-se que o *cluster* 1 destaca-se pelas altitudes mais elevadas em comparação aos outros grupos, apresentando uma altitude mínima de 404,93m, enquanto nos demais grupos é inferior a 8,5m. Em relação à temperatura, 95% dos municípios deste grupo possuem uma média de temperatura de até 25,52 graus.

Aproximadamente 43% dos municípios nesse grupo possuem uma irradiação solar maior do que a média geral de 5681,85 Wh/m². No entanto, é importante mencionar que há uma maior dispersão entre os municípios desse grupo em relação a essa variável. Se considerarmos exclusivamente a irradiação solar, o *cluster 2* se destaca, uma vez que 98% dos municípios desse grupo estão acima da média. As cidades que formam o *cluster 1* são: Aiuaba, Alcântaras, Altaneira, Araripe, Assaré, Aratuba, Campos Sales, Caririaçu, Carnaubal, Catarina, Crato, Croatá, Guaraciaba do Norte, Ibiapina, Itaitira, Guaramiranga, Jardim, Jati, Juazeiro do Norte, Meruoca, Monsenhor Tabosa, Nova Olinda, Parambu, Penaforte, Poranga, Porteiras, Potengi, Salitre, Mulungu, Santana do Cariri, São Benedito, Pacoti, Tianguá, Ubajara e Viçosa do Ceará.

Figura 20 - Empreendimentos de geração fotovoltaica no Ceará



Fonte: EPE.

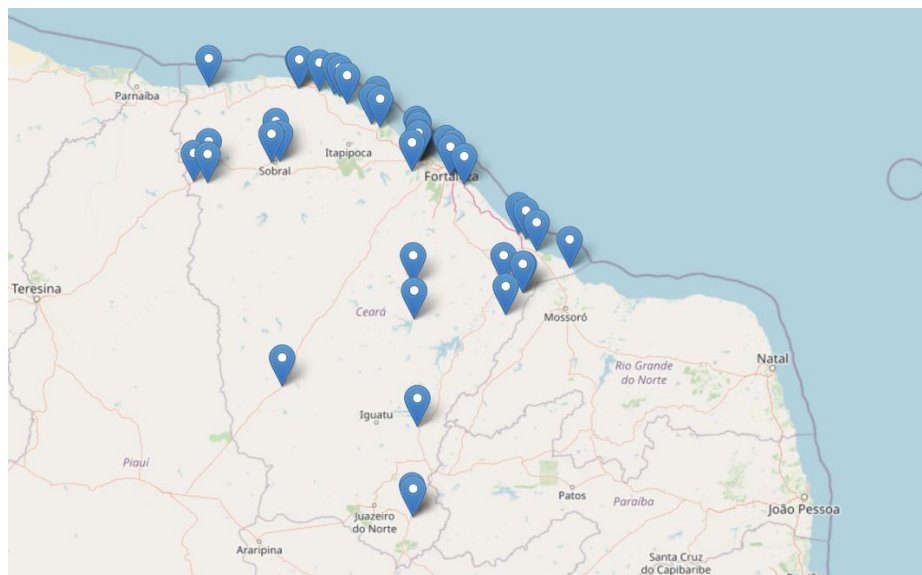
Além da classificação observada, foram analisados os empreendimentos fotovoltaicos existentes no Ceará, disponibilizados pela Empresa de Pesquisa Energética (EPE). Notou-se que ainda não está sendo aproveitado todo o potencial de geração do estado, e os empreendimentos atuais não estão localizados no melhor *cluster* de geração, indicando a necessidade de redirecionar esforços para alteração ou instalação de parques eólicos em áreas mais propícias.

4.3 Análise da Geolocalização das Subestações Elétricas no Ceará em Relação às Áreas Propícias para Geração

Nesse estágio do trabalho, serão aplicados cálculos de distância para avaliar a proximidade dos municípios selecionados em relação às subestações mais próximas. Com isso, poderemos obter uma visão abrangente da estrutura de escoamento de energia nessas áreas e avaliar a eficiência e o potencial de expansão da geração renovável no estado do Ceará.

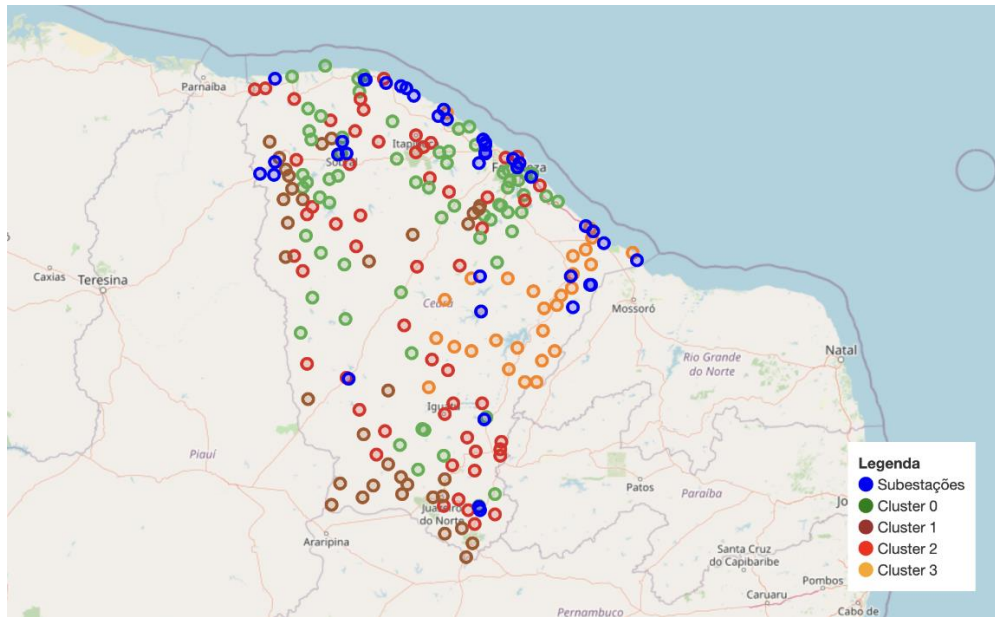
Ao analisarmos a Figura 18, fica evidente que as subestações estão concentradas principalmente na região litorânea do estado do Ceará. Contabilizamos um total de 42 subestações atualmente em atividade no estado. Dessa forma, essa distribuição geográfica das subestações sugere uma maior infraestrutura de escoamento de energia ao longo da costa cearense, proporcionando uma conectividade mais direta com as áreas de geração de energia renovável, como eólica e fotovoltaica, que muitas vezes estão localizadas em regiões próximas ao litoral. Essa concentração de subestações nessa região pode ter implicações na eficiência do sistema de transmissão e no potencial de integração da energia gerada nessas áreas de geração às redes de distribuição.

Figura 21 - Mapa das subestações do Ceará



Fonte: elaborado pelo próprio autor

Figura 22 - Mapa das subestações com os *clusters* por município



Fonte: elaborado pelo próprio autor.

De acordo com a Figura 19, é notório que os *clusters* com a maior quantidade de municípios próximos as subestações são os grupos 0 e 2.

A partir desses dados, foram executados os seguintes passos:

1. Cálculo da distância *haversine* de cada município até o a subestação mais próxima.

Figura 23 - Algoritmo para o cálculo *haversine*

```
[ ] distancias = []
for _, municipio in geo_dados.iterrows():
    ponto_municipio = municipio.geometry
    distancias_municipio = []
    for _, subestacao in subestacoes_ce.iterrows():
        ponto_subestacao = subestacao.geometry
        distancia = geodesic((ponto_municipio.y, ponto_municipio.x), (ponto_subestacao.y, ponto_subestacao.x)).km
        distancias_municipio.append(distancia)
    menor_distancia = min(distancias_municipio)
    distancias.append(menor_distancia)

distancias_km = [distancias for distancias in distancias]

# Adicionar a coluna "distancia" ao DataFrame dos municípios
geo_dados['distancia_km'] = distancias_km

# Exibir o DataFrame dos municípios com a coluna "distancia"
print(geo_dados[['municipio', 'distancia_km']])
```

	municipio	distancia_km
0	Abaiara	12.281274
1	Acaraú	4.444930
2	Acarape	46.502873
3	Acopiara	72.003180
4	Aiuaba	65.375879
...
179	Trairi	4.320092
180	Tururu	32.461551
181	Umirim	38.276280
182	Uruburetama	39.795929
183	Ererê	72.734971

Fonte: elaborado pelo próprio autor.

2. Filtrar os municípios com distância de até 45KM (esse valor foi baseado na média geral da distância entre municípios e subestações calculada no passo 1)

Figura 24 – Algoritmo média da distância dos municípios até subestações mais próximas

```
[ ] media_distancias_km = np.mean(geo_dados['distancia_km'])
    print("Média das distâncias em km:", media_distancias_km)

Média das distâncias em km: 45.09268427938451
```

Fonte: elaborado pelo próprio autor.

Figura 25 - Algoritmo para calcular cidades abaixo da média da distância até as subestações

```
[ ] cidades_abaixo_media = []

#Verificar quais cidades estão abaixo da média
for i in range(len(distancias_km)):
    if distancias_km[i] < media_distancias_km:
        cidades_abaixo_media.append(distancias_km[i])
```

Fonte: elaborado pelo próprio autor.

3. Por fim, agrupar por *cluster* e calcular o percentual de municípios filtrados com relação ao total do *cluster*.

Como resultado da execução dos passos, obtivemos, aproximadamente, os percentuais 57%, 34%, 62% e 66% de cidades acima da média de distância até as subestações para os *clusters* 0, 1, 2 e 3, respectivamente.

5 CONCLUSÃO

Em conclusão, por meio da análise de *cluster* dos municípios do Ceará, foi possível identificar quatro grupos distintos com características específicas relacionadas à altitude, temperatura, irradiação solar e velocidade do vento. O grupo 3 demonstrou ser mais propício para a geração eólica devido à sua alta temperatura e velocidade do vento, enquanto o grupo 1 mostrou-se mais adequado para a geração fotovoltaica, com sua alta altitude e baixa temperatura. Vale ressaltar que, embora essas áreas tenham sido identificadas como mais favoráveis para cada tipo de geração, os empreendimentos eólicos e fotovoltaicos já existentes no Ceará nem sempre estão alocados nessas melhores áreas de geração.

Além disso, ao observar a distância dos municípios até as subestações, constatou-se que os *clusters* 0, 1, 2 e 3 possuem diferentes percentuais de cidades próximas às subestações. Isso indica que os grupos com maior potencial para geração de energia renovável (*cluster* 3 para geração eólica e *cluster* 1 para geração fotovoltaica) não necessariamente estão localizados próximos às subestações de energia. Por outro lado, mesmo que o grupo 3 apresente o maior percentual de cidades localizadas próximas às subestações, é importante destacar que ele possui a menor quantidade de municípios abrangidos. Isso se dá por sua característica mais compacta em comparação aos demais grupos, podendo ser uma vantagem significativa para a expansão da infraestrutura de distribuição. Dessa forma, essas constatações ressaltam a necessidade de expandir a sistema de transmissão e distribuição de energia no estado para potencializar a geração de energia renovável por meio de direcionamentos dos investimentos e políticas públicas visando à expansão das fontes de energia renovável e ao aproveitamento dos recursos naturais disponíveis em cada região.

Por fim, a seleção dos melhores grupos para cada modalidade de geração não deve ser considerada definitiva, especialmente devido ao empate entre *clusters* observado no contexto da geração de energia eólica, pois a escolha pode variar dependendo da análise, permitindo priorizar diferentes variáveis e tornando esse estudo constantemente dinâmico. Dessa forma, pode-se afirmar com êxito que o objetivo principal do presente trabalho foi alcançado, pois foi realizada uma caracterização detalhada dos municípios, proporcionando uma compreensão mais aprofundada das correlações entre os dados relevantes no contexto da geração de energias renováveis. Em trabalhos futuros, é recomendável explorar o cruzamento desses dados com informações climáticas, geográficas e socioeconômicas adicionais para uma análise mais abrangente e, além disso, considerar a densidade populacional é essencial, uma vez que atualmente o Brasil tem como uma das principais premissas esse fator para a geração de energia.

REFERÊNCIAS

ANEEL - Agência Nacional de Energia Elétrica. **Atlas de energia elétrica do Brasil**. Agência Nacional de Energia Elétrica. – Brasília: ANEEL, 2002. Disponível em: <https://www.fisica.net/energia/atlas_de_energia_eletrica_do_brasul_3a-ed.pdf>. Acesso em: 03 de mar. de 2023.

ANEEL - Agência Nacional de Energia Elétrica. **Informações gerenciais das distribuidoras**. Brasília, DF: Aneel. Disponível em: <<https://antigo.aneel.gov.br/documents/656877/14854008/Boletim+de+Informa%C3%A7%C3%B5es+Gerenciais+-+2%C2%BA+trimestre+2018/fa14e464-2b54-bfc8-6bfl-c26b42d00d0a?version=1.0>>. Acesso em: 7 de mar. 2023.

Balanco Energético Nacional 2022: Ano base 2021. Rio de Janeiro, 2022. Disponível em: <<https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2022>>. Acesso em: 04 de mar de 2023.

BARROS ZANUT, Beatriz. **Conheça os fatores que influenciam na produção de energia eólica**. Engenharia 360, 2022. Disponível em: <<https://engenharia360.com/nikola-picape-eletrica-badger/>>. Acesso em: 04 de mar. de 2023.

BERKHIN, Pavel. A survey of clustering data mining techniques. In: **Grouping multidimensional data: Recent advances in clustering**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 25-71.

Brummer, V., 2018. **Of expertise, social capital, and democracy: Assessing the organizational governance and decision-making in German Renewable Energy Cooperatives**. Energy Res. Soc. Sci. 37, 111–121.

CBIE – Centro Brasileiro de Infraestrutura, 2020. **Como funciona a Geração Eólica?**. Disponível em: <<https://cbie.com.br/como-funciona-a-geracao-eolica/>> Acesso em: 03 de mar. de 2022.

CEARÁ. Decreto n.º 34733, 20 de Maio de 2022. **Institui o plano estadual de transição energética justa do Ceará - Ceará verde**. Palacio da Abolição do Governo do Estado do Ceará. Fortaleza, CE, 20 de Maio de 2022.

Centro de Referência para Energia Solar e Eólica Sérgio Brito. **Atlas Solarimétrico do Brasil: banco de dados solarimétricos**. Recife, 2000. Disponível em: <http://www.cresesb.cepel.br/publicacoes/download/Atlas_Solarimetrico_do_Brasil_2000>

Empresa de Pesquisa Energética. **WebMap EPE**. EPE, [s.d]. Disponível em: <<https://gisepeprd2.epe.gov.br/WebMapEPE/>> Acesso em: 18 de jul. de 2023.

ENERGIA EÓLICA: o que é, como funciona, vantagens e desvantagens. Portal Solar, c.2014-2023, Disponível em: <<https://www.portalsolar.com.br/energia-eolica>>. Acesso em: 03 de mar. de 2022.

FEIJÓO COELHO, Guilherme. **Fatores que Influenciam a Geração de Energia Solar (Parte 1)**. Medium, 2017. Disponível em: <<https://guilhermefeijoo.medium.com/fatores-que-influenciam-a-gera%C3%A7%C3%A3o-de-energia-solar-fotovoltaica-parte-1-11ae892ae9af>>. Acesso em: 03 de mar. de 2023.

FEIJÓO COELHO, Guilherme. **Fatores que Influenciam a Geração de Energia Solar (Parte 2)**. Medium, 2017. Disponível em: <<https://guilhermefeijoo.medium.com/fatores-que>>

influenciam-a-gera%C3%A7%C3%A3o-de-energia-solar-parte-2-b0310f02f7c7>. Acesso em: 03 de mar. de 2023.

Fontana, A., Naldi, M. C. **Estudo de Comparação de Métodos para Estimação de Números de Grupos em Problemas de Agrupamento de Dados**. 2009. Universidade de São Paulo. ISSN - 0103-2569

HONDA, Hugo, 2017. **Introdução Básica à Clusterização**. Disponível em: <https://lamfo-unb.github.io/2017/10/05/Introducao_basica_a_clusterizacao/>. Acesso em: 20 de mar. de 2023.

HONDA, Hugo; FACURE, Mateus; YAOHAO, Peng, 2017. **Os Três Tipos de Aprendizado de Máquina**. Disponível em: <<https://lamfo-unb.github.io/2017/07/27/tres-tipos-am/>>. Acesso em: 20 de mar. de 2023.

IPECE - Instituto de Pesquisa e Estratégia Econômica do Ceará. **Ceará ocupa o 13º lugar no ranking energético nacional e tem como principais matrizes de energia a termelétrica e a eólica**. 2018. Disponível em: <<https://www.ipece.ce.gov.br/2018/12/12/ceara-ocupa-o-13o-lugar-no-ranking-energetico-nacional-e-tem-como-principais-matrizes-de-energia-a-termeletrica-e-a-eolica/>> Acesso em: 04 de mar. de 2023.

JAIN, Anil K. **Data clustering: 50 years beyond K-means**. Pattern recognition letters, v. 31, n. 8, p. 651-666, 2010.

kelvins (Organização). **Municípios-Brasileiros**. Disponível em: <<https://github.com/kelvins/Municipios-Brasileiros>>. Acesso em: 21 de mar. de 2023.

LEI 14.300: O Marco Legal da Geração Distribuída. NeoSolar, [s.d]. Disponível em: <<https://www.neosolar.com.br/aprenda/saiba-mais/lei-14300-marco-legal-geracao-distribuida#:~:text=1.-,O%20que%20%C3%A9%20o%20Marco%20Legal%20da%20Gera%C3%A7%C3%A3o%20Distribu%C3%ADa,Resolu%C3%A7%C3%A3o%20Normativa%20482%2F2012>>. Acesso em: 18 de jul. de 2023.

LUCAS, Fabio. **Primeira usina solar com painéis solares flutuantes em grande altitude gera até 50% a mais de energia solar**. Click Petróleo e Gás, 2021. Disponível em: <<https://clickpetroleoegas.com.br/primeira-usina-solar-com-paineis-solares-flutuantes-em-grande-altitude-gera-ate-50-a-mais-de-energia-solar/>>. Acesso em: 21 de maio de 2023.

MARIMUTHU, C.; KIRUBAKARAN, V. **A critical review of factors affecting wind turbine and solar cell system power production**. Int. J. Adv. Engg. Res. Studies/III/II/Jan.-March, v. 143, p. 147, 2014.

Montezano, Bruno. Centro de Pesquisas de Energia Elétrica - CEPEL. **As energias solar e eólicas no Brasil**, 2013. Rio de Janeiro, 2013. PDFs/Greenjobs/UNEP-Green-Jobs-Report.pdf>.

Operador Nacional do Sistema Elétrico (ONS). **O Sistema Interligado Nacional**. ONS, 2023. Disponível em: <<https://www.ons.org.br/paginas/sobre-o-sin/o-que-e-o-sin>>. Acesso em: 18 de jul. de 2023.

PEREIRA, E. B.; MARTINS, F. R.; GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S. V.; SOUZA, J. G. **Atlas brasileiro de energia solar. 2.ed. São José dos Campos: INPE, 2017**. 80p. Disponível em: <<http://doi.org/10.34024/978851700089>>.

PIRES GUEDES, Erivelton. **Clusterização (*k-means*)**, 2019. Disponível em: <<https://www.kaggle.com/code/eriveltonguedes/7-clusteriza-o-k-means-erivelton>>. Acesso em: 20 de mar. de 2023.

SOARES ACCON, Frederico. **Os desafios para o escoamento da geração de energia**. Editora Brasil Energia, 2023. Disponível em: <<https://energiahoje.editorabrasilenergia.com.br/os-desafios-para-o-escoamento-da-geracao-de-energia>>. Acesso em: 03 de jul. de 2023.

SPROUTS of time. Intérprete: Edu Falaschi. Compositor: Rafael Bittencourt e Kiko Loureiro. In: Temple of Shadows. Intérprete: Edu Falaschi. Gera, Alemanha: House Of Audio, 2004. 1 CD, faixa 10.

TRANSFORMAÇÃO ENERGÉTICA a todo vapor. Guia do Ceará, Terra das Oportunidades. Disponível em: <[https://cearaterradasoportunidades.sedet.ce.gov.br/energiasrenovaveis/#:~:text=A%20matriz%20de%20energia%20el%C3%A9trica,El%C3%A9trica%20\(Aneel\)%20de%202022](https://cearaterradasoportunidades.sedet.ce.gov.br/energiasrenovaveis/#:~:text=A%20matriz%20de%20energia%20el%C3%A9trica,El%C3%A9trica%20(Aneel)%20de%202022)> Acesso em: 04 de mar. de 2023.

UNEP, I., & IOE, I. (2008). **Green Jobs: Towards Decent Work in a Sustainable. Low-carbon World, Nairobi**. Disponível em: <http://www.unep.org/labour_environment/>

UNIVERSIDADE FEDERAL DO CEARÁ. Biblioteca Universitária. **Guia de normalização de trabalhos acadêmicos da Universidade Federal do Ceará**. Fortaleza: Biblioteca Universitária, 2013. Disponível em: <<https://biblioteca.ufc.br/wp-content/uploads/2019/10/guia-de-citacao-06.10.2019.pdf>>. Acesso em: 07 de jul. 2023.

USINAR SOLAR: o que é, tipos, como funciona, vantagens e desvantagens. Portal Solar, c.2014-2023. Disponível em: <<https://www.portalsolar.com.br/usina-solar.html>>. Acesso em: 17 de mar. de 2023.

Veja os fatores que influenciam na eficiência do sistema fotovoltaico. HCC Energia Solar, 2022. Disponível em: <<https://hccenergiasolar.com.br/veja-os-fatores-que-influenciam-na-eficiencia-do-sistema-fotovoltaico/>>. Acesso em: 17 de mar. de 2023.