

UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO

Guilherme de Souza Dores, RA 2103022
Marcos Vinicius Melo de Oliveira, RA 2100326
Micael Henrique de Jesus Santos, RA 2106909
Nelson Luiz Ferreira, RA 2105813

Análise do Desmatamento Brasileiro

São Paulo - SP
2024

UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO

Análise do Desmatamento Brasileiro

Relatório Técnico-Científico apresentado na disciplina de Projeto Integrador para os cursos: Ciência de Dados e Engenharia da Computação da Universidade Virtual do Estado de São Paulo (UNIVESP).

Cidade - SP
2024

DORES, Guilherme de Souza; OLIVEIRA, Marco Vinicius Melo; SANTOS Micael Henrique de Jesus; FERREIRA, Nelson Luiz. Análise do Desmatamento Brasileiro

Relatório Técnico-Científico. Ciência de Dados e Engenharia da Computação – Universidade Virtual do Estado de São Paulo. Tutor: Glaucia Jardim Pissinelli. Polo UNICEU DRP14-PJI410 / Azul da cor do mar/Curuçá/Parelheiros/Vila Rubi, 2024.

RESUMO

Conforme a proposta da Universidade Virtual do Estado de São Paulo- UNIVESP para o Projeto Integrador IV, “Desenvolver análise de dados em escala utilizando algum conjunto de dados existentes ou capturados por IoT e aprendizagem de máquina. Preparar uma interface para visualização dos resultados”.

Decidimos por consenso, trabalhar com o Banco de Dados aberto fornecido pelo Instituto Nacional de Pesquisas Espaciais – INPE, referentes ao desmatamento da Amazônia fazendo uma Análise da Evolução do desmatamento da Amazônia, sobretudo no território Brasileiro, nos últimos 16 anos, tendo como objetivo: levantar, explorar, identificar, verificar, analisar, descrever e explicar a relação do desmatamento da floresta com os efeitos climáticos no Planeta, entre outros, destacamos o aquecimento global e a piora na qualidade do ar. Para tanto utilizamos: o Design Thinking com a metodologia de Ouvir, Criar e Implementar; algumas tecnologias como, linguagem Python, bibliotecas Pandas e Plotly, GIT, Streamlit, Jupyter e Colab; aplicação dos Conceitos de mineração de dados, utilização dos conceitos de visualização de dados e conhecimentos de estatística descritiva, frutos do aprendizado entre as diversas disciplinas oferecidas durante os cursos de Ciência de dados e Engenharia da Computação da UNIVESP.

PALAVRAS-CHAVE: Análise de dados; Tecnologias; Desmatamento; Resultados

LISTA DE ILUSTRAÇÕES (opcional)

Figura 01	7
Figura 02	7
Figura 03	8
Figura 04 Cerca de 17% do bioma foi devastado nos últimos 50 anos.	8
Figura 1: Gráfico mostrando a distribuição temporal dos focos de queimadas nos últimos anos.	25
Figura 2: Gráfico mostrando a distribuição os meses com maior extensão de queimadas e desmatamento	25
Figura 3: apresentação ano a ano dos meses com maior quantidade de desmatamento	26
Figura 4: Total dos desmatamentos acumulados ano a ano, relacionado ao governante à época	26

1 INTRODUÇÃO	7
2 DESENVOLVIMENTO	10
2.1 Objetivos	10
2.2 JUSTIFICATIVA E DELIMITAÇÃO DO PROBLEMA	11
2.3 Fundamentação teórica	12
2.4 Metodologia	22
2.5 Resultados preliminares: solução inicial	24
3 Referências	28

1 Introdução

A Amazônia é de grande interesse internacional principalmente por sua biodiversidade e papel no combate às mudanças climáticas. Ela é frequentemente chamada de “pulmões do planeta” devido ao seu papel na absorção de dióxido de carbono e produção de oxigênio. Além disso, a Amazônia abriga cerca de três milhões de espécies de plantas e animais e um milhão de pessoas indígenas, sendo um armazém vital de carbono que desacelera o aquecimento global. A região também é protegida por tratados internacionais como o Complexo de Conservação da Amazônia Central, classificado como Patrimônio Mundial pela ONU.



Figura 01



Figura 02

A Amazônia é considerada a região de maior biodiversidade do planeta e o maior bioma do Brasil, compreende um conjunto de ecossistemas que envolve a Floresta Amazônica, bastante diversificada, constituída por árvores, ervas, arbustos, lianas e trepadeiras, tida como a maior

floresta tropical do mundo bem como a bacia do Rio Amazonas, maior bacia hidrográfica do planeta. A fauna é extremamente rica e conta com mais de 30 milhões de espécies.



Figura 03

Não é exclusivamente brasileira, sendo, portanto, encontrada em outros países: " Bolívia, Colômbia, Equador, Guiana, Guiana Francesa, Peru, Suriname e Venezuela".

Há uma preocupação global com o desmatamento e a perda da biodiversidade, o que leva a esforços internacionais para conservar a floresta.



Figura 04 Cerca de 17% do bioma foi devastado nos últimos 50 anos.

Conforme a proposta da UNIVESP para o Projeto Integrador IV, “Desenvolver análise de dados em escala utilizando algum conjunto de dados existentes ou capturados por IoT e aprendizagem de máquina. Preparar uma interface para visualização dos resultados”.

Após algumas reuniões, direcionados pela metodologia do Design Thinking e utilizando a ferramenta Brainstorming, identificamos uma grande oportunidade de desenvolvermos e ampliarmos nosso aprendizado, aplicando na prática, conhecimentos adquiridos em diversas

disciplinas desenvolvidas ao longo de nossos cursos. Decidimos assim, por consenso, trabalhar com o Banco de Dados aberto fornecido pelo Instituto Nacional de Pesquisas Espaciais – INPE, fazendo uma análise da Evolução do desmatamento da Amazônia, sobretudo no território Brasileiro, nos últimos 16 anos relacionando o com:

- A piora da qualidade do ar
- O aquecimento Global

2 Desenvolvimento

2.1 OBJETIVOS

Durante nossa reunião com a orientadora do PI IV, foram abordadas algumas dúvidas e sugestões quanto a como e onde encontrar grandes quantidades de dados referentes a determinados assuntos, onde foi levantado a possibilidade de acessarmos algumas plataformas de “dados abertos”, por exemplo.

Encontramos alguns conjuntos de dados no Instituto Brasileiro de Geografia e Estatística - IBGE, relativos à Habitação, no Ministério da Saúde e na Agência Nacional de Saúde - ANS, dados referentes à Saúde, no Ministério da Justiça e Segurança Pública dados referentes à Segurança.

Através da Agência Nacional de Telecomunicações - ANATEL e do Ministério de Transportes, dados referentes aos Transportes; do Ministério do Esporte, dados referentes ao Esporte; do Ministério da Educação e Cultura, dados referentes à Educação, entre outros.

Analizamos cada um dos temas encontrados. Entretanto, por gerar impacto Internacional, direcionados pela metodologia do Design Thinking e utilizando a ferramenta Brainstorming, chegamos em consenso ao “tema” a ser trabalhado, como o "desmatamento da Amazônia”.

Para o tema escolhido, temos como objetivo geral: Avaliar a evolução do desmatamento da Amazônia ao longo de 16 anos de observações relacionando o com efeitos climáticos como o aquecimento global, a piora da qualidade do ar, a degradação do solo e as perdas de alguns Biomas.

Para tanto, vamos levantar o banco de dados abertos do Instituto Nacional de Pesquisas Espaciais- INPE, referentes ao desmatamento da Amazônia, com o objetivo de explorar e identificar o tamanho da degradação da floresta.

Ao efetuarmos uma análise descritiva, esperamos verificar e avaliar o quanto de floresta já foi perdida ao longo dos anos. Esperamos descobrir, caracterizar e poder descrevermos o quanto o nosso Planeta está sendo afetado.

2.2 JUSTIFICATIVA E DELIMITAÇÃO DO PROBLEMA

Quais as consequências para o nosso planeta, sobretudo para o Brasil com a evolução do desmatamento da Amazônia ?

Durante reunião , nós integrantes do grupo 11, inicialmente escolhemos trabalhar o problema, evolução do desmatamento na Amazônia, analisando a base de dados abertos do Instituto Nacional de Pesquisas Espaciais - INPE relacionando o com os efeitos climáticos, como o aquecimento global, piora da qualidade do ar, degradação do solo, perda de estudos e conhecimentos científicos com a eliminação de alguns Biomas, entre outros.

2.3 FUNDAMENTAÇÃO TEÓRICA

AMAZÔNIA

A Floresta Amazônica leva umidade para toda a América do Sul, influencia o regime de chuvas na região, contribui para estabilizar o clima global e ainda tem a maior biodiversidade do planeta.

O fato é que a Amazônia influencia o equilíbrio ambiental de todo o planeta e tem papel fundamental na economia do Brasil.

A área da Floresta Amazônica, ao contrário de ser improdutiva, produz imensas quantidades de água para o restante do país. Os chamados "rios voadores", formados por massas de ar carregadas de vapor de água gerados pela evapotranspiração na Amazônia, levam umidade da Bacia Amazônica para o Centro-Oeste, Sudeste e Sul do Brasil.

Esses rios voadores também influenciam chuvas na Bolívia, no Paraguai, na Argentina, no Uruguai e até no extremo sul do Chile. A umidade vinda da Amazônia e os rios da região alimentam regiões que geram 70% do PIB da América do Sul.

Segundo estudos do Instituto Nacional de Pesquisas da Amazônia, uma árvore com copa de 10 metros de diâmetro pode bombear para a atmosfera mais de 300 litros de água em forma de vapor por dia – mais que o dobro da água usada diariamente por um brasileiro. Uma árvore maior, com copa de 20 metros de diâmetro, pode evapotranspirar mais de 1.000 litros por dia, bombeando água e levando chuva para irrigar lavouras, encher rios e as represas que alimentam hidrelétricas no resto do país.

O desmatamento prejudica a evapotranspiração e, por consequência, a rota desses rios, podendo afetar assim o regime de chuvas no restante do país e diversas atividades econômicas. Como resultado do desmatamento, até 65% da Amazônia corre o risco de se transformar em savana ao longo dos próximos 50 anos.

Além disso, o Rio Amazonas é responsável por quase um quinto das águas doces levadas aos oceanos no mundo.

MUDANÇAS CLIMÁTICAS

A Amazônia e as florestas tropicais, que armazenam de 90 bilhões a 140 bilhões de toneladas métricas de carbono, ajudam a estabilizar o clima em todo o mundo. Só a Floresta Amazônica representa 10% de toda a biomassa do planeta.

Já as florestas que foram degradadas ou desmatadas são as maiores fontes de emissões de gases do efeito estufa depois da queima de combustíveis fósseis. Isso porque as florestas saudáveis têm uma imensa capacidade de reter e armazenar carbono, mas o desmatamento para o uso agrícola ou extração de madeira libera gases do efeito estufa para a atmosfera e desestabiliza o clima.

Vale lembrar que 2016 foi o terceiro ano consecutivo a ser considerado o mais quente da história. O Acordo de Paris, cujo objetivo é manter o aquecimento da temperatura média do planeta abaixo de 2°C, passa necessariamente pela preservação de florestas.

Dados da ONU de 2015 mostram o Brasil como um dos dez países que mais emitem gases do efeito estufa no mundo, com 2,48% das emissões. Para cumprir suas metas de redução de emissões dentro do acordo internacional, o Brasil se comprometeu a alcançar, na Amazônia brasileira, o desmatamento ilegal zero até 2030 e a compensação das emissões de gases de efeito de estufa provenientes da supressão legal da vegetação até 2030.

EQUILÍBRIO AMBIENTAL

Como a maior floresta tropical do mundo, a Amazônia possui a maior biodiversidade, com uma em cada dez espécies conhecidas. Também há uma grande quantidade de espécies desconhecidas por cientistas, principalmente nas áreas mais remotas. Assegurar a biodiversidade é importante porque ela garante maior sustentabilidade natural para todas as formas de vida, e ecossistemas saudáveis e diversos podem se recuperar melhor de desastres, como queimadas.

Preservar a biodiversidade amazônica, portanto, quer dizer contribuir para estabilizar outros ecossistemas na região. O recife dos corais da Amazônia, por exemplo, um corredor de biodiversidade entre a foz do Amazonas e o Caribe, é um refúgio para corais ameaçados pelo aquecimento global por estar em uma região mais profunda.

Segundo o biólogo Carlos Eduardo Leite Ferreira, da Universidade Federal Fluminense, esse recife poderá ajudar a repovoar áreas degradadas dos oceanos no futuro, mas petroleiras Total e BP têm planos de explorar petróleo perto da região dos corais da Amazônia, ameaçando assim esse ecossistema.



Figura 05 Foto: Arquivo TG G1. globo .com Fonte: Priscila Jordão, Deutsche Welle. 31/08/2017.

Para o desenvolvimento e análise dos problemas gerados ou agravados pelo desmatamento da Amazônia foram utilizadas as seguintes tecnologias:

- Linguagem de programação Python;
- Biblioteca Pandas para tratamento de dados;
- Biblioteca Plotly para exibição de gráficos;

- Jupyter notebooks e google collabs;
- Git para versionamento do código;
- Streamlit para geração de dashboards;
- Estatística descritiva.

LINGUAGEM DE PROGRAMAÇÃO - PYTHON

Linguagem de programação, normalmente composta por regras capaz de gerar códigos a serem compilados e depois decodificados e analisados com a finalidade de enviar e receber instruções computacionais. Os desenvolvedores utilizam determinadas linguagens para elaboração de algoritmos, que nada mais são do que “Programas”. A utilização de uma Linguagem é também de fundamental importância para criação de frameworks. Para o desenvolvimento do Projeto Integrador IV, escolhemos utilizar a Linguagem PYTHON.

Complementarmente, frameworks WEB, por sua vez, são ferramentas práticas, disponíveis em bibliotecas de classes com diversas funções e instruções, em que quando em uso o programador não precisa reescrever todo o código, o que agiliza seu trabalho quanto a complexidade da tecnologia necessária, ajudando-o no desenvolvimento de sistemas, podendo assim dedicar-se às demais etapas da implantação de sua ferramenta ou aplicação.

Python é uma linguagem de programação gratuita, “Open Source” e foi projetada tendo como um dos principais objetivos ser de fácil leitura e utilização, com isso pode ser utilizada para as mais diversas aplicações. Definindo com termos mais técnicos, Python é interpretada, orientada a objetos, funcional, tipada, imperativa e de script.

É comum ouvirmos a expressão de que “programar em Python é como escrever uma carta em inglês para o computador”, pois a linguagem tenta utilizar comandos intuitivos, como “print” para imprimir um texto na tela, “open” para abrir um arquivo, ou “find” para encontrar a posição de uma palavra.

Como Python é uma das linguagens mais ativas em termos de comunidade, a cada dia novas bibliotecas são construídas e aprimoradas. Existem funções e módulos prontos para se executar de tudo, desde manipulações em imagens até algoritmos de inteligência artificial.

Em outras palavras, a programação python é diferenciada pela riqueza de bibliotecas e frameworks prontos para utilização, bem como pelo suporte da comunidade. O fato de existirem bibliotecas robustas também permite que um programador se especialize em uma tarefa específica, por exemplo: “manipulação de tabelas e datasets” para ciência de dados. Nesse caso, bastaria estudar e dominar a biblioteca Pandas.

(DIDATICA.TECH, 2024)

BIBLIOTECA PANDAS

Pandas é uma biblioteca Python que fornece estruturas de dados rápidas, flexíveis e expressivas desenvolvida para tornar o trabalho com dados ‘relacionais’ e ‘rotulados’ fácil e intuitivo. Seu objetivo é ser o bloco de construção fundamental de alto nível para fazer análises de dados práticas, com dados reais, em Python. Adicionalmente possui o objetivo maior de ser a mais poderosa e flexível ferramenta open source para análise de dados e manipulação.

BIBLIOTECA PLOTLY

Plotly Python é uma biblioteca poderosa que permite criar gráficos interativos de maneira fácil e eficiente. Com essa ferramenta, é possível visualizar dados de forma dinâmica, tornando a análise mais intuitiva e atraente.

PLOTLY PYTHON

O Plotly Python é uma extensão da biblioteca Plotly, originalmente desenvolvida para a linguagem R. Com a crescente popularidade do Python, foi criada uma versão específica para essa linguagem.

(AWARI, 2024)

JUPYTER NOTEBOOK

O Jupyter Notebook é uma forma fácil e organizada para “codar” (codificar) ou demonstrar de forma plena algum código, por exemplo um tutorial.

É uma aplicação open source, possibilita a criação de documentos que podem ser facilmente compartilhados, tanto para quem usa o jupyter quanto para quem não usa.

Além disso, também possui vários métodos, funções da própria aplicação, que auxiliam a demonstrar o que está acontecendo no código, ou calcular a performance de duas diferentes operações.

(MEDIUM, 2024)

COLAB NOTEBOOK

Os notebooks do Colab são notebooks do Jupyter hospedados no Colab, permitem combinar código executável e rich text em um só documento, além de imagens, HTML, LaTeX e muito mais. Quando você cria seus próprios notebooks do Colab, eles são armazenados na sua conta do Google Drive. É possível compartilhar os notebooks do Colab facilmente com colegas de trabalho ou amigos e permitir que eles façam comentários ou até editem o documento.

GIT

O Git é um sistema de controle de versões, distribuído e de código aberto.

Definição de cada um dos conceitos:

- **Sistema de controle:** isto significa que o Git é, basicamente, um *tracker* de conteúdo. Portanto, o Git pode ser usado para guardar conteúdo – É usado, principalmente, para armazenar código, devido às outras funcionalidades que disponibiliza.
- **Sistema de controle de versões:** O código que é guardado no Git está em constante mutação à medida que novo código é adicionado. Além disso, muitos programadores

podem adicionar código em paralelo. O sistema de controle de versões, ajuda a gerir tudo isto, mantendo um histórico das alterações que foram realizadas. Adicionalmente, o Git disponibiliza funcionalidades como *branches* e *merges*.

- **Sistema distribuído de controle de versões** : O Git tem um repositório remoto que está alojado em um servidor e um repositório local que está alojado no computador de cada programador. Isto significa que o código não está apenas guardado num servidor central, mas também existe uma cópia integral em todos os computadores dos programadores. Assim, o Git é um sistema distribuído de controle de versões porque está presente no computador de cada programador. (FREECODECAMP, 2024).

STREAMLIT

Streamlit é uma biblioteca de código aberto que permite criar aplicativos web interativos usando apenas código Python. Com ele podemos transformar scripts Python em aplicativos web interativos de forma rápida e fácil, sem a necessidade de conhecimentos avançados em desenvolvimento web.

Uma das principais vantagens é a sua simplicidade. Com apenas algumas linhas de código, você pode criar uma interface de usuário intuitiva e interativa para seus projetos. Além disso, oferece uma ampla gama de recursos e widgets que permitem personalizar e aprimorar a aparência e a funcionalidade dos seus aplicativos. Por exemplo, você pode adicionar temas personalizados, criar layouts complexos e até mesmo integrar seu aplicativo com outras bibliotecas populares, como o Pandas e o Matplotlib. E ainda, oferece suporte a recursos avançados, como autenticação de usuário e implantação em nuvem, para tornar seus aplicativos ainda mais poderosos e versáteis.

Ao criar sua primeira aplicação com o Streamlit, você pode aproveitar os recursos e widgets disponíveis para criar uma interface de usuário atraente e funcional. Você pode adicionar gráficos, tabelas, botões e muito mais para tornar seu aplicativo mais interativo e informativo. Além disso, o Streamlit permite que você atualize e visualize instantaneamente as alterações em seu aplicativo, facilitando o processo de desenvolvimento e depuração.

(WP RAIZ, 2024).

ANÁLISE DE DADOS E ESTATÍSTICA DESCRITIVA

Bancos de dados nada mais é que um conjunto de dados, geralmente organizados por coerência e finalidade específica. Quando criados, permite aos usuários: consultas e manipulações desses dados. O banco de dados armazena e gerencia os bens mais valiosos de uma empresa. Isso acontece porque o mercado está cada vez mais competitivo e acelerado, exigindo das empresas respostas rápidas e assertivas, além de estratégias bem planejadas e executadas. Como dissemos acima, o banco de dados armazena informações e, nessa batalha de competitividade, informação é poder. (IMPACTA, 2023).

Com a organização e análise de dados, pode-se gerar informações importantíssimas inclusive para tomadas de decisão pelos gestores das empresas. Para os negócios, a análise de dados tem como o objetivo geral estudar o desempenho em cada uma das etapas e assim conseguir informações relevantes. Com base nesse histórico, usando como base para projeções, os profissionais da área podem ter insights para o futuro da organização. (ESCOLA CONQUER, 2023).

A Estatística é uma ciência cujo campo de aplicação estende-se a muitas áreas do conhecimento humano. Entretanto, um equívoco comum que deparamos nos dias atuais é que, em função da facilidade que o advento dos computadores nos proporciona, permitindo desenvolver cálculos avançados e aplicações de processos sofisticados com razoável eficiência e rapidez, muitos pesquisadores consideram-se aptos a fazerem análises e inferências estatísticas sem um conhecimento mais aprofundado dos conceitos e teorias. Tal prática, em geral, culmina em interpretações equivocadas e muitas vezes errôneas... Em sua essência, a Estatística é a ciência que apresenta processos próprios para coletar, apresentar e interpretar adequadamente conjuntos de dados, sejam eles numéricos ou não. Pode se dizer que seu objetivo é o de apresentar informações sobre dados em análise para que se tenha maior compreensão dos fatos que os mesmos representam. A Estatística subdivide-se em três áreas: descritiva, probabilística e inferencial. A estatística descritiva, como o próprio nome já diz, se preocupa em descrever os dados. A estatística inferencial, fundamentada na teoria das

probabilidades, se preocupa com a análise destes dados e sua interpretação. A palavra estatística tem mais de um sentido. No singular se refere à teoria estatística e ao método pelo qual os dados são analisados enquanto que, no plural, se refere às estatísticas descritivas que são medidas obtidas de dados selecionados. A estatística descritiva, cujo objetivo básico é o de sintetizar uma série de valores de mesma natureza, permitindo dessa forma que se tenha uma visão global da variação desses valores, organiza e descreve os dados de três maneiras: por meio de tabelas, de gráficos e de medidas descritivas. A tabela é um quadro que resume um conjunto de observações, enquanto os gráficos são formas de apresentação dos dados, cujo objetivo é o de produzir uma impressão mais rápida e viva do fenômeno em estudo. Para ressaltar as tendências características observadas nas tabelas, isoladamente, ou em comparação com outras, é necessário expressar tais tendências através de números ou estatísticas. Estes números ou estatísticas são divididos em duas categorias: medidas de posição e medidas de dispersão. Para se obter bons resultados numa análise estatística, além dos métodos aplicados, também é necessário ter clareza nos conceitos utilizados.

A estatística trabalha com dados, os quais podem ser obtidos por meio de uma população ou de uma amostra, definida como: População: conjunto de elementos que têm pelo menos uma característica em comum. Esta característica deve delimitar corretamente quais são os elementos da população que podem ser animados ou inanimados. Amostra: subconjunto de elementos de uma população. Este subconjunto deve ter dimensão menor que o da população e seus elementos devem ser representativos da população. A seleção dos elementos que irão compor a amostra pode ser feita de várias maneiras e irá depender do conhecimento que se tem da população e da quantidade de recursos disponíveis. A estatística inferencial é a área que trata e apresenta a metodologia de amostragem. Em se tratando de conjuntos-subconjuntos, estes podem ser: Finitos: possuem um número limitado de elementos. Infinitos: possuem um número ilimitado de elementos. Segundo Medronho (2003), elemento significa cada uma das unidades observadas no estudo. Após a determinação dos elementos pergunta-se: o que fazer com estes? Pode-se medi-los, observá-los, contá-los, surgindo um conjunto de respostas que receberá a denominação de variável. Variável: é a característica que vai ser observada, medida ou contada nos elementos da população ou da amostra e que pode variar, ou seja, assumir um valor diferente de elemento para elemento. Não basta identificar a variável a ser trabalhada, é necessário fazer-se distinção entre os tipos de variáveis:

- Variável qualitativa: é uma variável que assume como possíveis valores, atributos ou qualidades. Também são denominadas variáveis categóricas.

- Variável quantitativa: é uma variável que assume como possíveis valores, números. Cada uma dessas variáveis pode ser subclassificada em:
- Variável qualitativa nominal: é uma variável que assume como possíveis valores, atributos ou qualidades e estes não apresentam uma ordem natural de ocorrência.
- Variável qualitativa ordinal: é uma variável que assume como possíveis valores atributos ou qualidades e estes apresentam uma ordem natural de ocorrência.
- Variável quantitativa discreta: é uma variável que assume como possíveis valores números, em geral inteiros, formando um conjunto finito ou enumerável.
- Variável quantitativa contínua: é uma variável que assume como possíveis valores números, em intervalos da reta real e, em geral, resultantes de mensurações.

(IME-USP, 2024)

2.4 METODOLOGIA

A metodologia adotada neste projeto foi estruturada em etapas, que incluem ouvir e interpretar o contexto e criar/prototipar, com foco na coleta e análise inicial dos dados.

Ouvir e Interpretar o Contexto

O projeto foi realizado com o objetivo de investigar a relação entre focos de queimadas na floresta amazônica e a qualidade do ar nas áreas afetadas. Utilizamos dados abertos fornecidos pelo Instituto Nacional de Pesquisas Espaciais (INPE) sobre queimadas, enquanto buscamos fontes compatíveis para dados de qualidade do ar, a fim de realizar uma análise cruzada no futuro.

Até o momento, as informações iniciais foram coletadas das seguintes fontes:

- Dados de queimadas: A base de dados do INPE oferece informações detalhadas sobre os focos de incêndio na Amazônia, incluindo localização, intensidade, e data.
- Dados de qualidade do ar: Estamos em processo de pesquisa e avaliação de bases de dados públicas e institucionais que forneçam informações sobre a concentração de poluentes atmosféricos, em diferentes regiões.

Criar / Prototipar

Na fase atual, focamos em uma análise quantitativa dos dados de queimadas. As seguintes etapas foram realizadas:

- Coleta e pré-processamento dos dados de queimadas: Os dados foram tratados para remover inconsistências e padronizar a estrutura temporal e geográfica, preparando-os para a etapa de cruzamento com os dados de qualidade do ar.
- Análise exploratória inicial: Utilizando Python e bibliotecas como Pandas, realizamos uma análise exploratória dos dados de queimadas, buscando identificar padrões temporais, geográficos e variações sazonais no comportamento dos incêndios. Visualizações gráficas, como histogramas e gráficos de séries temporais, foram geradas para facilitar a interpretação dos dados.

- Pesquisa de fontes de dados de qualidade do ar: Atualmente, estamos pesquisando bases de dados que forneçam medidas adequadas da qualidade do ar. Entre as opções avaliadas, estão dados de monitoramento ambiental fornecidos por instituições governamentais e institutos de pesquisa especializados em poluição atmosférica.

As principais ferramentas utilizadas até agora incluem:

- Python: Para manipulação dos dados.
- Bibliotecas Pandas, Matplotlib, Seaborn, Plotly: Para análise e visualização.
- Jupyter Notebooks: Para registro da análise e dos resultados.

Implementar / Testar (Futuro)

Ainda não entramos na etapa de implementação e testes, que será realizada assim que os dados de qualidade do ar forem coletados e adequados para o cruzamento com os dados de queimadas. A expectativa é testar correlações e analisar a influência das queimadas nos níveis de poluição atmosférica, com base nas informações disponíveis.

2.5 RESULTADOS PRELIMINARES: SOLUÇÃO INICIAL

Nesta seção, apresentamos os resultados preliminares, focando na criação de uma solução inicial. As etapas principais de coleta, análise e exploração dos dados de queimadas estão descritas a seguir, com destaque para os primeiros insights gerados.

Coleta de Dados

Os dados de queimadas foram extraídos da base pública do INPE, contendo informações sobre a localização, data e intensidade dos focos de incêndio na Amazônia ao longo dos anos. Até o momento, estamos no processo de pesquisa de fontes confiáveis e compatíveis para os dados de qualidade do ar, que serão cruzados com os dados de queimadas posteriormente.

Análise Exploratória

Os resultados preliminares foram gerados a partir de uma análise exploratória dos dados de queimadas. Utilizamos visualizações gráficas para identificar padrões temporais e geográficos, e os principais achados incluem:

- Picos sazonais: Observamos que a maior incidência de queimadas ocorre nos meses de seca, especialmente entre junho e setembro.

Visualizações

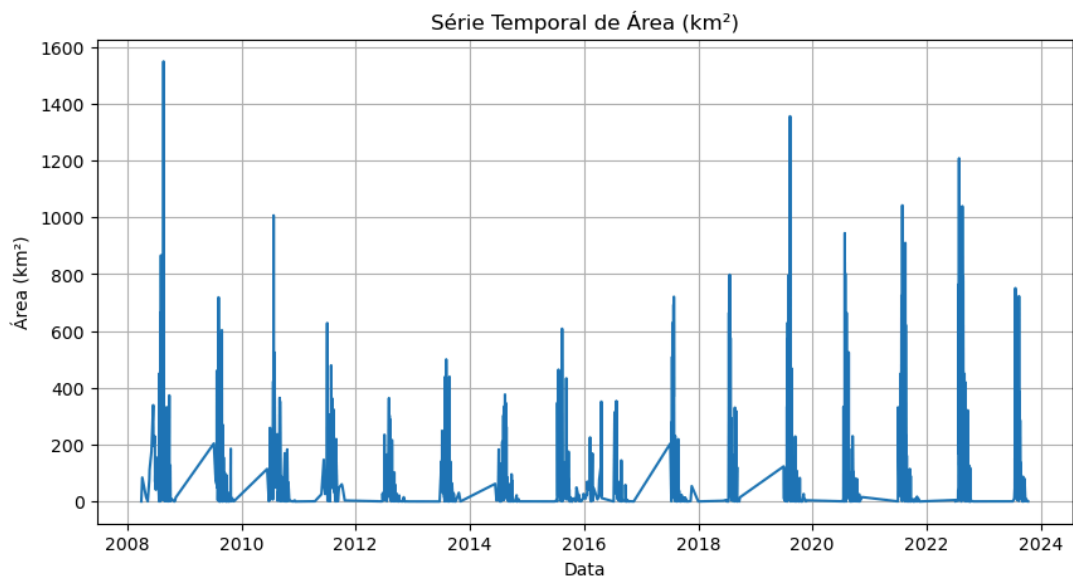


Figura 1: Gráfico mostrando a distribuição temporal dos focos de queimadas nos últimos anos.

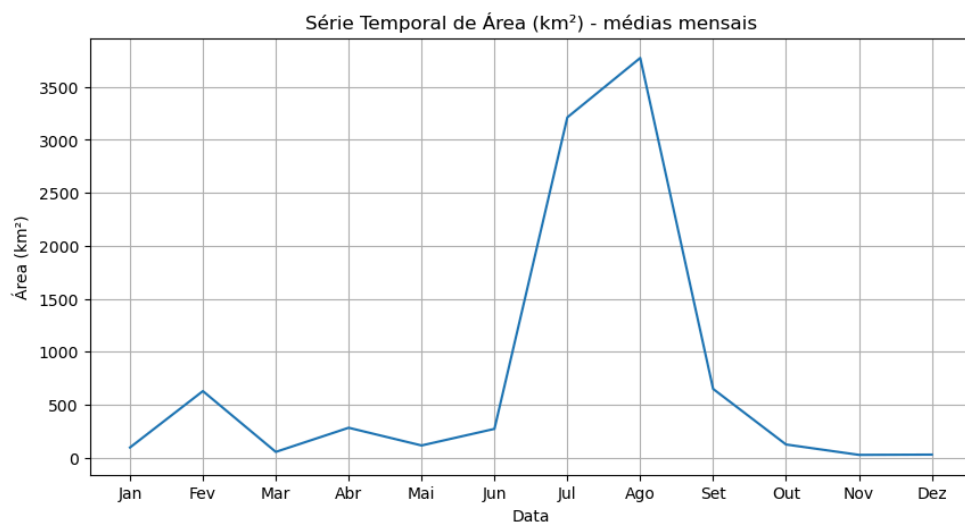


Figura 2: Gráfico mostrando a distribuição os meses com maior extensão de queimadas e desmatamento

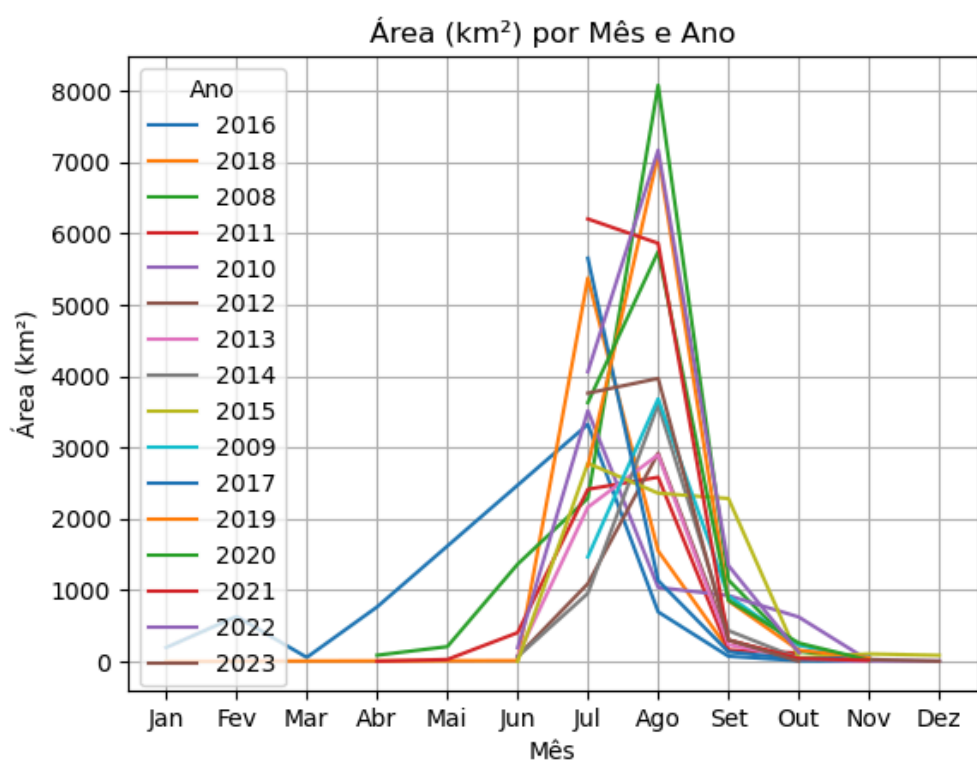


Figura 3: apresentação ano a ano dos meses com maior quantidade de desmatamento

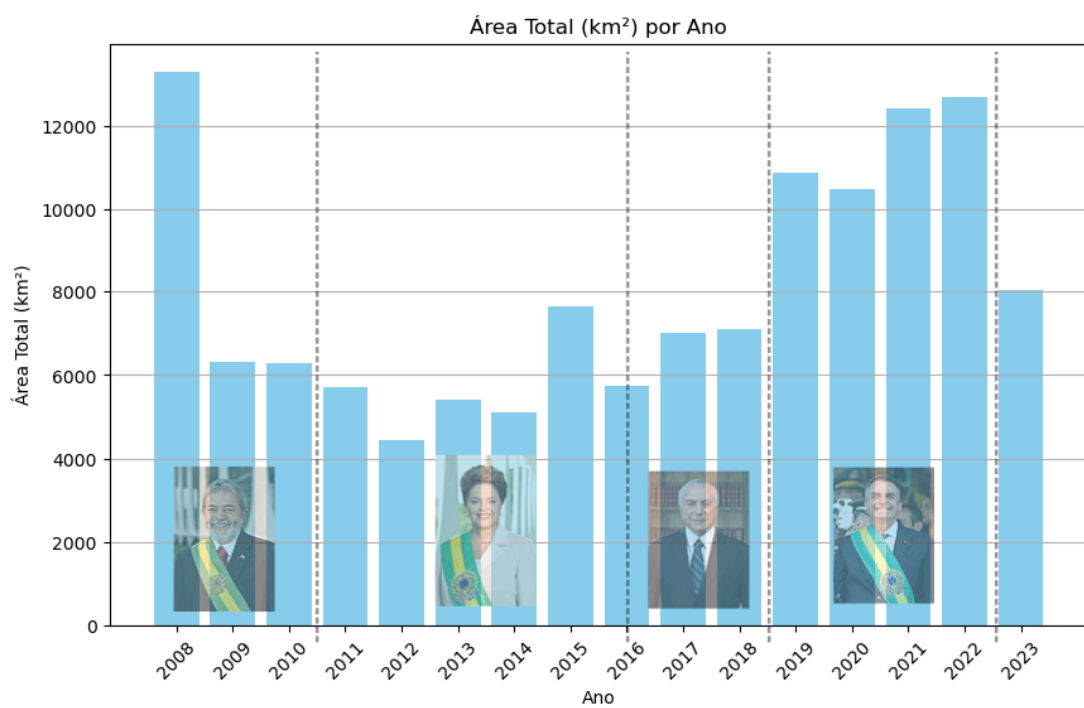


Figura 4: Total dos desmatamentos acumulados ano a ano, relacionado ao governante à época

Proposta de Solução Inicial

Com base nessa análise, propomos como solução inicial a criação de um modelo de correlação entre os focos de queimadas e os dados de qualidade do ar (ainda em fase de coleta). A hipótese inicial é que a ocorrência de queimadas em grandes quantidades durante os períodos de seca tem impacto direto na deterioração da qualidade do ar nas regiões afetadas.

3 Referências

ABNT – Associação Brasileira de Normas Técnicas. NBR 14724: Informação e documentação. Trabalhos Acadêmicos - Apresentação. Rio de Janeiro: ABNT, 2002.

ARRUDAS, Mariana. O que significa Design Thinking. Agência USP de Inovação, 2020. Disponível em: < <http://www.inovacao.usp.br/o-que-significa-design-thinking/>>. Acesso em: 04 ago. 2024.

BARI, A.; CHAOUCHI, M.; JUNG, T. Análise preditiva para leigos. 2a. ed. Rio de Janeiro: Alta Books, 2019.

BOYER, C. B.; UTA, C. M. História da Matemática [Trad. Helena Castro]. 3 ed. São Paulo: Blucher, 2012.

CARVALHO, André C. P. de L. F. et al. Inteligência artificial: uma abordagem de aprendizado de máquina. Rio de Janeiro: GEN, 2021.

Castro, Leandro Nunes de e Ferrari, Daniel Gomes. Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações. São Paulo: Editora Saraiva, 2016. ISBN 978-85-472-0100-5. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-5/>. Acesso em: 03set. 2024.

COSTA NETO, P. L. O.; CYMBALISTA, M. Probabilidades. 2. ed. rev. ampliada. São Paulo: Blücher, 2006.

LARSON, R.; FARBER, B. Estatística Aplicada. 6. ed. São Paulo: Pearson Education do Brasil, 2015.

DEITEL, H. M., DEITEL, P. J. C++: como programar. 5. ed. São Paulo: Pearson Prentice Hall, 2006.

DEVORE, Jay. L. Probabilidade e estatística para engenharia e ciências. Tradução da 9ª edição norte-americana. São Paulo: Cengage Learning Brasil, 2018. ISBN 9788522128044. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788522128044>>. Acesso em: 9 set. 2024.

EDELWEISS, N.; GALANTE, R. Estruturas de Dados. Porto Alegre: Bookman, 2011.

ELMASRI, R.; NAVATHE, S. B. Sistemas de banco de dados. 7.ed. São Paulo: Pearson Education do Brasil, 2018.

ESCOLA CONQUER. Análise de dados: o que é e qual a sua importância. Disponível em: <<https://escolaconquer.com.br/blog/analise-de-dados-o-que-e-e-qual-a-sua-importancia/#:~:text=Para%20os%20neg%C3%B3cios%2C%20a%20an%C3%A1lise%20de%20dados%20tem,podem%20ter%20insights%20para%20o%20futuro%20da%20organiza%C3%A7%C3%A3o>> Acesso em: 05 ago. 2024.

FACELI, K. et al. Inteligência artificial: uma abordagem de aprendizado de máquina. 2 ed. Rio de Janeiro: Grupo GEN/LTC, 2021. E-book. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521637509/>. Acesso em: 12 ago. 2024.

FERREIRA, R. G. C. et al. Preparação e análise exploratória de dados. Porto Alegre: SAGAH, 2021. E-book. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786556902890/>. Acesso em: 15 ago. 2024.

GERSTING, Judith L. Fundamentos Matemáticos para a Ciência da Computação: matemática discreta e suas aplicações. 7. ed. Rio de Janeiro: Grupo Gen, 2016. 908 p. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9788521633303/epubcfi/6/2/%5B%3Bvnd.vst.idref%3Dcover%5D!/4>. Acesso em: 19 ago. 2024.

GUPTA, C. B.; GUTTMAN, I. Estatística e Probabilidade com Aplicações para Engenheiros e Cientistas. Rio de Janeiro: Grupo GEN, 2016. ISBN 9788521632931. Disponível

em: <<https://integrada.minhabiblioteca.com.br/#/books/9788521632931/>>. Acesso em: 02 set. 2024.

HART-DAVIS, A. O Livro da Ciência. 2. ed. São Paulo: Globo, 2016.

IDEO. Design Thinking para Educadores. São Paulo: Instituto Educadigital, (2014). Disponível em: <https://educadigital.org.br/dteducadores/>. Acesso em: 28 ago. 2024 .

LARSON, R.; FARBER, B. Estatística Aplicada. 6. ed. São Paulo: Pearson Education do Brasil, 2015.

LOUZANE, F.; SAAB, P. Aprender na prática: projetos integradores 2018/2019. Recurso eletrônico. São Paulo: Univesp, 2019. Disponível em: <https://online.fliphtml5.com/wzky/hxmc/#p=1>. Acesso em: 05 ago. 2024.

MILANI, Alessandra Maciel Paz; SOARES, Juliane Adélia; ANDRADE, Gabriella Lopes; ROCHA, Elenise; COUTO, Júlia Mara Colleoni; LEDUR, Cleverson Lopes; MAITINO NETO, Roque. Visualização de dados. Porto Alegre: Sagah, 2020. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786556900278/>. Acesso em: 08 ago. 2024.

MORAIS, I.S. et al. Introdução a Big Data e Internet das Coisas (IoT). Porto Alegre: SAGAH, 2018.

OLIVEIRA, Israel Campos de. Aprimorando a elasticidade de aplicações de banco de dados utilizando virtualização em nível de sistema operacional. 2015. 96 f. Dissertação (Mestrado) - Ciência da Computação, Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2015. Disponível em: <<http://tede2.pucrs.br/tede2/bitstream/tede/6433/2/476714%20-%20Texto%20Completo.pdf>>. Acesso em: 12 ago. 2024..

PUGA, S.; FRANÇA, E.; GOYA, M. Banco de Dados: implementação em SQL, PL/SQL e Oracle 11g. São Paulo: Pearson Education do Brasil, 2013.

<https://www.usandopy.com/pt/manipulacao-e-visualizacao-de-dados-em-python/python-pandas-introducao-a-biblioteca-pandas/>. Acesso em: 30 ago. 2024

RIBEIRO, J. L. P. Áreas e Proporções nas Superquadras de Brasília Usando o Google Maps. Revista do Professor de Matemática. Rio de Janeiro, n. 92, p. 12-15, jan-abr. 2017.

SEVERINO, A. J. Metodologia do trabalho científico. 22. ed. rev. e ampl. São Paulo: Cortez, 2002.

RUSSELL, Stuart; NORVIG, Peter. Inteligência Artificial: uma abordagem moderna. Rio de Janeiro: GEN, 2022. E-book. ISBN 8595158878. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9788595159495/>, acesso em: 08 ago. 2024.

SOMERVILLE, I. Engenharia de software. 10. ed. São Paulo: Pearson Education do Brasil, 2018.

VanderPlas, J. Python Data Science Handbook. Sebastapol, CA: O'Reilly, 2016.

VICCI, C. (Org.). Banco de Dados. São Paulo: Pearson Education do Brasil, 2014.

ICODE. O que é programação e qual a sua importância para o futuro digital. Disponível em: <https://idocode.com.br/blog/programacao/o-que-e-programacao/>. Acesso em: 26 ago. 2024.

IMPACTA. Entenda a importância de um banco de dados em uma organização. Disponível em: <

<https://www.impacta.com.br/blog/entenda-a-importancia-de-um-banco-de-dados-em-uma-organizacao/#:~:text=Qual%20a%20import%C3%A2ncia%20de%20um%20banco%20de%20dados%3F,assertivas%2C%20al%C3%A9m%20de%20estrat%C3%A9gias%20bem%20planejadas%20e%20executadas>>. Acesso em: 05 ago. 2024.

SANTOS, Robson dos. 12 Melhores Frameworks Para Desenvolvimento. Disponível em:

<<https://www.brasilcode.com.br/12-melhores-frameworks-para-desenvolvimento-web/>>

Acesso em: 09 ago. 2024.

Por que a Amazônia é vital para o mundo? | Natureza | G1 (globo.com). Acesso em: 02 set. 2024.

<https://www.worldwildlife.org/stories/the-amazon-in-crisis-forest-loss-threatens-the-region-and-the-planet>. Acesso em: 06 set. 2024.

<https://g1.globo.com/natureza>. Acesso em: 08 set. 2024.

<https://www.ime.usp.br/>. Acesso em: 05 set. 2024.

<https://wpraiz.com.br/guia-completo-sobre-streamlit-crie-aplicacoes-interativas-com-python>.
Acesso em: 08 set. 2024.

Uma introdução ao Git: O que é, e como usá-lo. (freecodecamp.org) Acesso em: 02 set. 2024.

Olá, este é o Colaboratory - Colab (google.com). Acesso em: 05 set. 2024.

Como instalar o Jupyter Notebook? (Windows e Linux) | by Matheus Budkewicz | horaDeCodar | Medium. Acesso em: 02 set. 2024.

Plotly Python: Aprenda a criar gráficos interativos com essa poderosa biblioteca (awari.com.br). Acesso em: 02 set. 2024.

kit Design thinking para educadores, publicado em português pelo Instituto Educa digital.
Acesso em: 02 ago. 2024.

Mini Toolkit Design Thinking, da Escola Design Thinking. Acesso em: 02 ago. 2024.

capítulo 2, “Tipos de dados Python”, do livro Introdução à Computação usando Python: um foco no desenvolvimento de aplicações, de Ljubomir Perkovic. Este capítulo apresenta uma revisão dos principais tipos de dados da linguagem Python, essenciais para o desenvolvimento do PI (Projetos Integradores). Acesso em: 04 ago. 2024.

capítulo 3, “Programação imperativa”, do livro Introdução à Computação usando Python, de Ljubomir Perkovic. que apresenta uma revisão dos principais conceitos de programação imperativa da linguagem Python, essenciais para desenvolvimento do PI. Acesso em: 05 ago. 2024.

Python e Django, de Francisco M. B. Maciel. Acesso em: 05 ago. 2024.

ANEXOS (OPCIONAL)

Materiais coletados por meio de pesquisas em diversas fontes.

O grupo pode anexar qualquer tipo de material ilustrativo, tais como tabelas, lista de abreviações, documentos ou parte de documentos, resultados de pesquisas etc.

Podem ser incluídos separadamente e ordenados por letras, por exemplo, Anexo A, Anexo B etc.