

Relatório Preliminar: Análise dos Impactos Socioeconômicos do Desmatamento no Estado do Pará

Zetta Lab 2025 - Desafio I: Ciência e Governança de Dados

Autor(a): Breno Carvalho Pedroso



Repositório do Projeto (contém Notebooks, Dados e Dashboard Interativo):

https://github.com/Boakpe/ZettaLab2025_Dasafio1_Dados

1. Introdução

Este relatório preliminar apresenta os resultados iniciais da primeira etapa do Desafio I do Zetta Lab 2025, focado em "Ciência e Governança de Dados". O objetivo central é responder à pergunta: "Como poderíamos avaliar e prever os impactos socioeconômicos do desmatamento no estado do Pará?". Esta fase concentrou-se na aquisição, organização, exploração e visualização de dados relevantes.

2. Escolha dos Dados

Para abordar a complexa relação entre desmatamento e seus impactos socioeconômicos, foram selecionadas diversas bases de dados brasileiras. Estas abrangem dimensões ambientais, econômicas e sociais em nível municipal para o estado do Pará, cobrindo o período de 2008 a 2024. A escolha visou capturar tanto as manifestações do desmatamento quanto seus potenciais vetores e consequências.

A principal variável ambiental, o desmatamento medido em km², foi obtida do INPE, através do Projeto PRODES/TerraBrasilis, que realiza o monitoramento anual do desmatamento na Amazônia Legal. Esse é um indicador essencial da pressão sobre a cobertura florestal, fundamental para a análise. Complementarmente, os focos de queimada, também do INPE (Programa Queimadas), foram incluídos, pois queimadas estão frequentemente associadas ao processo de desmatamento e preparo de terra.

No âmbito dos indicadores econômicos municipais, foram utilizados dados do IBGE referentes ao PIB Municipal, PIB per capita (em R\$) e ao Valor Adicionado Bruto (VAB) por setor (Agropecuária, Indústria e Serviços). Estes indicadores medem a atividade econômica e o bem-estar material, permitindo analisar como o desmatamento se correlaciona com o desenvolvimento econômico local e a estrutura produtiva.

Os indicadores demográficos e sociais selecionados incluem a População Estimada, fornecida pelo IBGE, e o número de beneficiários de programas de transferência de renda, como o Bolsa Família, cujos dados agregados por município foram obtidos do Ministério do Desenvolvimento Social (MDS) via VIS DATA 3. A população indica a escala humana afetada e a pressão demográfica, enquanto os programas sociais podem indicar vulnerabilidade e a resposta do Estado.

Para investigar os vetores agropecuários, foram coletados dados da Produção Agrícola Municipal (PAM) do IBGE sobre a área plantada/colhida de culturas relevantes, como soja e milho, e da Pesquisa da Pecuária Municipal (PPM), também do IBGE, sobre o efetivo de rebanhos, como o bovino (em cabeças). A expansão agropecuária é um dos principais vetores de desmatamento, e essas variáveis ajudam a quantificar essa relação. A granularidade municipal e a série temporal adotadas permitem análises de evolução e comparações espaciais, cruciais para entender a heterogeneidade dos impactos no Pará.

3. Metodologia de Aquisição e Preparação dos Dados

A aquisição dos dados seguiu um processo estruturado. Primeiramente, realizou-se a identificação e o acesso às fontes, utilizando os portais oficiais do INPE (TerraBrasilis, BDQueimadas) e IBGE (SIDRA, Contas Regionais), além de dados do MDS. Em seguida, procedeu-se à extração dos arquivos, geralmente em formatos como CSV, XLS e XLSX, aplicando filtros geográficos (Estado do Pará) e temporais (2008 em diante) durante a extração ou logo após.

A etapa subsequente foi a de limpeza e tratamento, que envolveu a padronização de nomes de colunas, a conversão de tipos de dados (por exemplo, de strings para numéricos e o tratamento de separadores decimais), a seleção das colunas relevantes para a análise e a verificação com tratamento preliminar de valores ausentes (NaNs), onde os valores foram extrapolados usando uma regressão linear ou preenchidos de acordo com o contexto do problema. Por fim, os dados foram consolidados com a criação de um DataFrame único em Python, unindo as diversas fontes por município e ano. Para a execução dessas tarefas, foram utilizadas as ferramentas Python, juntamente com as bibliotecas Pandas, Scikit-learn, Plotly e Streamlit.

4. Análise Exploratória de Dados (EDA) e Principais Insights Preliminares

A Análise Exploratória de Dados foi conduzida utilizando o DataFrame consolidado. O foco esteve em visualizações e estatísticas descritivas para identificar padrões, tendências e correlações. Para facilitar essa exploração, foi desenvolvido um dashboard interativo em Streamlit, com gráficos gerados pela biblioteca Plotly.

A análise do panorama do desmatamento revelou, primeiramente, uma tendência temporal estadual com flutuações. A evolução do desmatamento total no Pará ao longo dos anos apresentou períodos de aumento e redução, indicando a complexidade dos fatores envolvidos e a eficácia variável das políticas de contenção. Observou-se, por exemplo, um grande aumento no desmatamento entre 2018 e 2021, possivelmente associado a mudanças na conjuntura política e ao enfraquecimento de políticas ambientais, o que impactou diretamente a fiscalização e o controle do uso da terra. Verificou-se também uma concentração geográfica do desmatamento, que não é uniforme, com alguns municípios consistentemente figurando entre os mais desmatados, sugerindo "hotspots" de pressão ambiental, conforme ilustrado pelo mapa do desmatamento anual. Além disso, a evolução municipal mostrou que municípios específicos demonstram trajetórias distintas, alguns com picos agudos e outros com aumento mais gradual, como visualizado nos gráficos de linha por município.

Ao investigar as relações socioeconômicas, percebeu-se que a relação entre desmatamento e PIB per capita não é linear. Em alguns municípios, o aumento do desmatamento pode coincidir com o crescimento do PIB per capita, potencialmente impulsionado pela expansão de atividades primárias. Em outros, essa relação pode ser menos clara ou até inversa a longo prazo, especialmente se os recursos naturais se esgotarem ou os impactos ambientais negativos se acumularem, uma dinâmica que os gráficos de dispersão anuais ajudam a visualizar pontualmente. Observou-se uma correlação positiva entre o aumento do VAB da agropecuária e as taxas de desmatamento em muitos municípios, reforçando o papel da expansão agrícola e pecuária como um dos principais vetores. Quanto à relação entre desmatamento e programas sociais como o Bolsa Família, embora o programa tenha ampla cobertura na região, não foi possível identificar, com base nos dados analisados, uma relação direta ou consistente entre o número de beneficiários e as variações no desmatamento; a presença do programa pode refletir mais as condições socioeconômicas locais do que influenciar diretamente as dinâmicas ambientais.

No que tange aos vetores agropecuários, o aumento da área plantada de commodities como soja e milho, e a expansão do rebanho bovino, mostraram correlação espacial e temporal com as áreas de maior

desmatamento em diversos municípios. Os gráficos de linha por município para essas variáveis, quando comparados com o desmatamento, evidenciam essas conexões. Adicionalmente, há uma forte correlação entre o número de focos de queimada e as áreas desmatadas, indicando o uso do fogo como ferramenta para limpeza de terra e expansão de atividades.

A análise da estrutura econômica municipal, através do VAB, indicou que municípios com maior participação do VAB Agropecuário em sua economia tendem a apresentar taxas de desmatamento mais elevadas. A evolução da composição do VAB ao longo do tempo pode indicar se o desmatamento está associado a uma primarização da economia local. Um exemplo específico dessa dinâmica foi observado no município de Altamira, onde ocorreu um pico de desmatamento em 2018, coincidindo com um aumento expressivo na área plantada de soja (750%) e milho (500%). Simultaneamente, o PIB per capita apresentou um crescimento de menos de 1%, sugerindo que os benefícios econômicos da expansão podem não estar sendo amplamente distribuídos ou que há um aumento da vulnerabilidade em outros setores.

5. Limitações do Estudo Preliminar

Este estudo preliminar possui algumas limitações que devem ser consideradas. A Análise Exploratória de Dados identifica correlações, mas não estabelece relações de causa e efeito, sendo necessárias análises estatísticas mais robustas para inferir causalidade. A disponibilidade e a defasagem de dados também representam um desafio, pois alguns indicadores socioeconômicos, como o IDH-M, não são calculados anualmente, limitando análises temporais de alta frequência. Além disso, dados mais recentes, como os de 2024, podem ser preliminares ou parciais para algumas fontes. A análise em nível municipal, embora informativa, mascara variações internas significativas dentro dos próprios municípios. Finalmente, fatores externos como políticas públicas (nacionais e estaduais), preços de commodities, eventos climáticos extremos e outros elementos exógenos não foram explicitamente modelados nesta fase, mas certamente influenciam as dinâmicas observadas.

6. Conclusão Preliminar

A aquisição e exploração inicial dos dados revelaram complexas interações entre o desmatamento e diversos indicadores socioeconômicos no estado do Pará. O dashboard interativo desenvolvido permite uma visualização dinâmica dessas relações, fornecendo uma base sólida para investigações futuras. As descobertas preliminares confirmam a forte ligação do desmatamento com a expansão agropecuária e evidenciam a necessidade de políticas que considerem as particularidades econômicas e sociais de cada município para promover um desenvolvimento mais sustentável e mitigar os impactos negativos da perda florestal. A continuidade deste trabalho, com foco em modelagem preditiva e análises estatísticas mais aprofundadas, será crucial para subsidiar a tomada de decisão e a governança de dados no enfrentamento deste desafio.