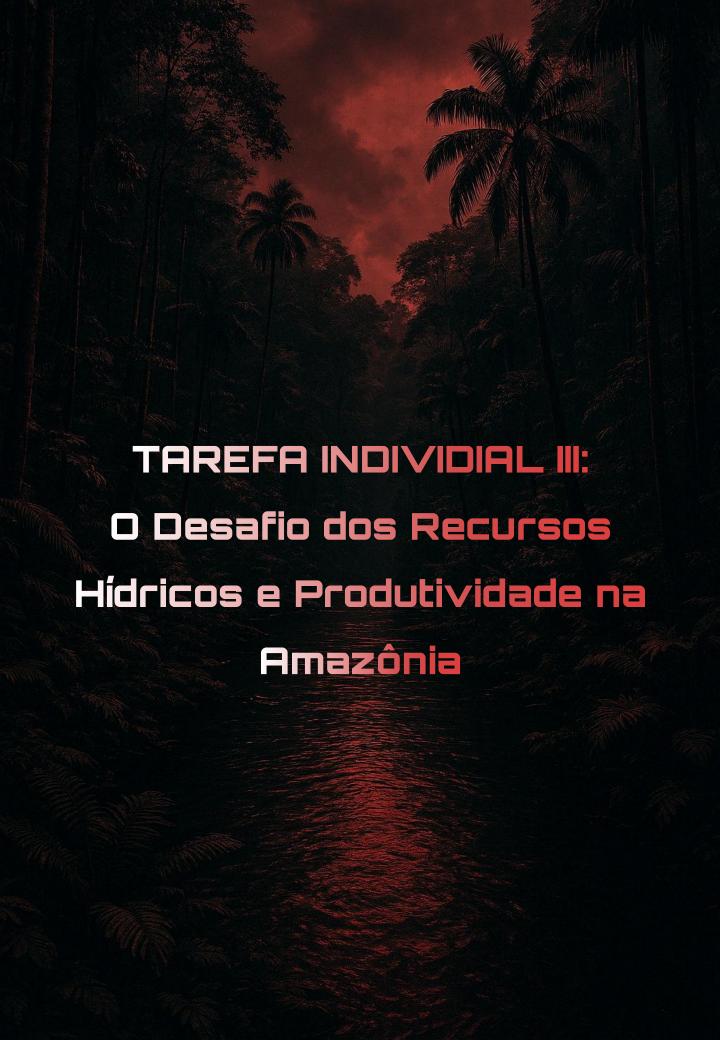


TAREFA INDIVIDUAL III



:: O CENÁRIO

A região amazônica enfrenta um problema crescente relacionado à gestão dos recursos hídricos e seus impactos diretos na segurança alimentar de comunidades ribeirinhas e agricultores familiares. Nos últimos anos, episódios de estiagens prolongadas e enchentes severas alteraram drasticamente os ciclos naturais, afetando tanto a disponibilidade de água quanto a produtividade agrícola local.

As comunidades relatam que, além das mudanças no regime de chuvas, há também uma percepção de redução da qualidade da água dos igarapés e rios, o que compromete o uso para irrigação, consumo humano e atividades produtivas. Junto disso, surge um aumento na ocorrência de doenças transmitidas por água contaminada, afetando diretamente o bem-estar da população.

Dois líderes comunitários solicitaram apoio para entender, com base em dados, como essas variações climáticas e os desafios hídricos estão afetando a produção agrícola e a qualidade de vida nas comunidades. Eles querem sair do campo da percepção e trabalhar com dados reais para fortalecer suas tomadas de decisão.

Para isso, foi realizada uma coleta de duas bases de dados. A primeira base contém dados climáticos diários da região: chuvas previstas, chuvas reais, temperatura média, presença de variações climáticas incomuns e índices de umidade do solo. A segunda base traz informações socioeconômicas e produtivas das comunidades: volume de produção agrícola, incidência de doenças relacionadas à água, acesso à água potável, e indicadores de segurança alimentar.

Durante a coleta, surgiram alguns desafios. As informações foram registradas por diferentes agentes locais, resultando em registros duplicados, erros de grafia, formatos de datas diferentes, e presença de valores ausentes em variáveis-chave, como volume de produção e chuvas reais. Também foram identificados possíveis outliers, como registros de chuvas superiores a 700 mm em um único dia, o que exige investigação.



:: O CENÁRIO

O seu desafio, como analista, será transformar esses dados "sujos" em informação de qualidade. A primeira missão é definir, com clareza, o problema central com base no entendimento das duas bases e suas conexões. Será necessário formular perguntas inteligentes e identificar as métricas que poderão guiar a análise.

Após a definição do problema, o próximo passo é estruturar uma estratégia de limpeza e preparação dos dados. Isso inclui remover duplicatas, padronizar categorias como "sim", "não" e "nao", tratar dados ausentes e converter formatos de datas para um padrão único. Também será preciso definir como tratar outliers, baseando-se na realidade do contexto amazônico.

A fase de Análise Exploratória (EDA) será fundamental. Você deverá explorar as distribuições das variáveis, entender a correlação entre chuvas, produtividade agrícola e incidência de doenças. Visualizações como histogramas, gráficos de dispersão e heatmaps ajudarão a revelar padrões ocultos, clusters e possíveis relações entre clima, acesso à água e segurança alimentar.

Por fim, você deverá construir um relatório com os achados da EDA, incluindo gráficos, tabelas e uma narrativa sobre o que os dados revelam. Esse relatório servirá como base para pensar, futuramente, em modelos preditivos e ações concretas que possam ser tomadas pelas comunidades para enfrentar os desafios de forma mais sustentável e inteligente.



:: Descrição das Variáveis do Dataset

BASE CLIMÁTICA

- data (datetime): data do registro.
- chuvas_previstas_mm (float): precipitação prevista em mm (0–200).
- chuvas_reais_mm (float): precipitação medida em mm.
- temperatura_media_C (float): temperatura média diária em °C (20–35).
- variacao_climatica (string): indicador de variação climática ("sim", "não")
- indice_umidade_solo (float): umidade do solo (%) (10–90).

BASE SOCIOECONÔMICA

- data (datetime): data do registro.
- volume_producao_tons (float): volume produzido em toneladas (0.5–20).
- incidencia_doencas (int/float): número de casos de doenças hídricas (Poisson λ=2)
- acesso_agua_potavel (string): acesso à água potável ("sim", "não").
- indicador_seguranca_alimentar (float): índice de segurança alimentar (0–100).



:: OBJETIVO PEDAGÓGICO DA ATIVIDADE

- Desenvolver a capacidade de transformar dados brutos em informações confiáveis, aplicando técnicas de limpeza, padronização e tratamento de dados em um contexto real da Amazônia.
- Fortalecer o pensamento analítico e crítico, por meio da definição clara de problemas, formulação de hipóteses e exploração de relações entre variáveis socioambientais e climáticas.
- Aprimorar habilidades práticas em Análise Exploratória de Dados (EDA), utilizando ferramentas digitais para gerar visualizações, descobrir padrões e criar narrativas baseadas em dados, com foco em soluções para desafios socioambientais.

:: CRITÉRIOS DE AVALIAÇÃO

- Clareza na definição do problema.
- Coerência na análise dos dados.
- Qualidade dos insights gerados.
- Organização e apresentação da solução.

:: COMO SERÁ ENTREGUE A ATIVIDADE

- Documento em PDF
- O nome do arquivo deverá ser tarefa3 i2a2.
- O arquivo deverá ser remetido para <u>challenges@i2a2.academy</u> até o dia 29 de junho de 2025 (domingo).
- O aluno que não entregar a atividade, até às 23:59:59 do dia 29 de junho, estará eliminado do curso.



