

DADOS, PYTHON E ESTATÍSTICA: CAMINHOS PARA STEM

Ana Júlia Amaro Pereira Rocha

SEMANA 3

DIA 1

PERGUNTAS PARA UM DATASET

Toda análise de dados começa com boas perguntas, não com gráficos. Tipos de perguntas:

- **descritivas:** o que está acontecendo?
 - **Ex:** Quantos atendimentos ocorreram por mês em 2025?
- **comparativas:** quem é maior/menor?
 - **Ex:** Mulheres usam mais o serviço de saúde que homens?
- **relacionais:** uma coisa influencia a outra?
 - **Ex:** Escolaridade influencia vacinação?
- **preditivas:** O que pode acontecer?
 - **Ex:** Onde pode haver falta de médicos no futuro?

ÉTICA DOS DADOS

Ética significa tratar dados com respeito e responsabilidade. Algumas perguntas importantes são:

- Quem são os donos dos dados?
- É permitido usá-los?
- Estamos protegendo a privacidade das pessoas?
- Há exclusão de minorias?

Dados de saúde, renda, endereço e documentos, por exemplo, são sensíveis e devem ser anonimizados. Dados não devem reforçar preconceitos. Toda ação com dados deve ter transparência e justificativa.

BUSCANDO DATASETS

Ao procurar dados para trabalhar, precisamos saber onde encontrar **datasets confiáveis** para não espalhar informações e resultados falsos. Algumas plataformas que eu recomendo são:

- **Kaggle**: datasets públicos, competições de dados;
- **dados.gov.br**: dados oficiais do governo brasileiro;
- **IBGE**: saúde, educação, renda, população;
- **DATASUS**: dados públicos de saúde do SUS;
- **OMS (WHO)**: dados globais de saúde;

DOCUMENTAÇÃO DE UM PROJETO

Documentação é como contar a história da sua análise. Ela ajuda outras pessoas a entender o que foi feito e por que. Projetos com boa documentação têm mais valor no mundo profissional. As “etapas” são:

- **Problema:** O que queremos entender?
- **Justificativa:** Por que isso é importante para a sociedade?
- **Fonte dos dados:** De onde vieram os dados?
- **Variáveis:** Quais colunas foram usadas?
- **Metodologia:** O que foi feito com os dados?
- **Análise:** O que foi descoberto?
- **Visualizações:** Gráficos, tabelas, mapas
- **Conclusão:** O que isso significa para a comunidade?
- **Ética:** Como os dados foram tratados com responsabilidade?

PROFISSÕES QUE USAM DADOS

No mundo atual, governos usam dados para políticas públicas, hospitais usam dados para prevenção, escolas usam dados para reduzir evasão e empresas usam dados para decisões. Ou seja, dados são a base de decisões da sociedade moderna e o profissional que sabe trabalhar com dados consegue apoiar decisões baseadas em evidências. Trouxe alguns exemplos desses profissionais da atualidade e o salário médio deles de acordo com o site Glassdoor.

PROFISSÕES QUE USAM DADOS

- **Analista de Dados:** limpa dados, analisa, cria gráficos, produz relatórios e apoia decisões.
 - O salário médio no Brasil é R\$ 5000 - R\$ 8000 por mês.
 - Nos EUA é cerca de US\$ 92000 por ano.
- **Estatístico:** faz modelagem matemática, probabilidade, previsões, pesquisa científica e análise de risco.
 - O salário médio no Brasil é R\$ 7000 por mês.
 - No exterior, está entre cerca de US\$ 70000 e US\$ 110000 por ano.
- **Cientista de Dados:** Explora dados complexos, faz modelos de Machine Learning, previsões e interpreta padrões.
 - O salário médio no Brasil é R\$ 8000 - R\$ 15000 por mês.
 - No exterior, está entre cerca de US\$ 100000 e US\$ 140000 por ano.

PROJETO FINAL (PARTE 1)

Uma das partes mais importantes e que determinam todas as próximas etapas do processo de análise de dados é a coleta. No nosso caso, vocês não vão coletar propriamente os dados, mas quero que aprendam a pesquisar, identificar se um dataset tem o mínimo necessário para se trabalhar com ele e que escolham algum assunto que gostem para que seja mais interessante esse trabalho final.

Passei alguns sites confiáveis hoje nos quais vocês podem entrar e procurar um dataset. Quero que tenha pelo menos 1000 linhas e 7 colunas para terem bastante possibilidades e gere uma boa exploração. Me mostrem a base de dados que gostou até amanhã e eu vou dar uma conferida para apoiar vocês ao longo de cada etapa. Durante a semana vamos continuar avançando no projeto final, então começemos logo!