

Resumo do Desafio Técnico FADESP

Abordagem Metodológica

A solução foi estruturada para realizar a carga de dados oficiais em um banco relacional PostgreSQL usando ETL em Python. Foi realizada a limpeza, padronização e normalização das informações, seguidas de análises exploratórias e estatísticas para identificar padrões relevantes.

Principais Desafios

- Dados extensos exigiram divisão em lotes.
- Presença de valores ausentes e inconsistências nos tipos de colunas.
- Necessidade de mapeamento de códigos para nomes mais compreensíveis.

Principais Descobertas

- Identificou-se correlação positiva entre número de cursos e bolsistas.
- Visualizações geográficas mostraram concentração de cursos em estados com maior infraestrutura educacional.
- Diferenças significativas entre categorias administrativas e tipos de instituição.

Recomendações

- Melhorar a padronização de dados em bases públicas.
- Expandir variáveis para modelos de predição mais robustos.
- Investir em políticas que fortaleçam regiões com menor produção científica.