

Sobre o query “data engineer with experience in python and sql”

1- A Descoberta: "A análise da query revelou um 'cluster' (agrupamento) de vagas dominado pela empresa Turing, tanto no mercado de SP quanto no Remoto."

2- A Causa: "Isso é evidenciado pelos scores de similaridade e ângulos idênticos (ex: 0.3032 / 72.35°), provando que múltiplas vagas (Ranks 45, 60, 55) são, na verdade, cópias exatas do mesmo documento de texto."

3- A Implicação (Ruído nos Dados): "Isso também expõe um 'ruído' nos dados: o uso de texto 'boilerplate' (o parágrafo de introdução da Turing) faz com que essas vagas pareçam muito similares entre si e dominem o topo do ranking."

4- A Prova (Rank 132): "O algoritmo, no entanto, funciona como esperado. O Rank 132 (SP), com score 0.2430, mostra uma vaga real de 'Azure Data Engineer' que não é da Turing, provando que o sistema está corretamente ranqueando os documentos, mas que as vagas da Turing são matematicamente mais similares à query."