

Análise inicial do banco de dados ‘Data Science Salaries 2024’

Apresentação

Essa foi a análise inicial do banco de dados geral (link: <https://www.kaggle.com/datasets/sazidthe1/data-science-salaries/data>).

Todo esse relatório foi feito no RMarkdown, fazendo com que os gráficos já fossem renderizados juntamente com as explicações das análises. Vou disponibilizar o script fonte juntamente com o resultado final (pdf) para que os códigos dos gráficos sejam extraídos e anexados no arquivo do relatório final.

Introdução

Aqui serão feitas análises conforme aquele roteiro de principais informações sobre os bancos de dados, principais variáveis, etc.

Ideias de possíveis análises com as variáveis disponíveis:

- Distribuição de salários por nível de experiência:
 - **Principal visualização:** Boxplot para cada nível de experiência.
- Média salarial por tipo de emprego (freelancer, meio período, etc.):
 - Calculamos algumas medidas resumo do salário para cada tipo de emprego (ex.: média, mediana, etc.).
 - **Principal visualização:** Gráficos de barras representando os salários médios por cada tipo de emprego.
- Variação salarial por tipo de cargo:
 - Aqui podem ser listados os cargos que pagam mais e que pagam menos.
 - Podemos analisar os cargos mais comuns e fazer uma análise de cada um, observando as suas particularidades.
 - **Possível visualização:** Gráfico de barras com as médias salariais de cada cargo (Ideia: podemos dividir em dois gráficos, um com os 5 cargos que mais pagam e outro com os 5 que menos pagam)
- Análise geográfica da distribuição dos salários:
 - Ideia geral: Observar como os salários variam a depender da localização da empresa e do país de residência do funcionário (muita atenção para considerar isso ao analisar os casos de trabalho remoto).
 - Podemos calcular os salários médios por país de residência e localização da empresa, por exemplo.
 - **Ideia de visualização:** Mapas de calor e gráficos de barra para representar as variações salariais geográficas (por exemplo: mapa mundi com os países que mais pagam em cores mais escuras, ou qualquer coisa similar)

- Avaliar as diferenças entre trabalho remoto e presencial:
 - podemos analisar a proporção de trabalhos remotos e o impacto salarial desse tipo de contratação.
 - Podemos comparar as variações salariais para cada categoria de ‘work_models’.
 - Podemos fazer gráficos de barras comparando as médias salariais para cada tipo de modelo.
- Analisar a relação entre o tamanho da empresa e as médias salariais:
 - Podemos calcular os salários médios para cada tamanho de empresa.
 - **Possível visualização:** Novamnete, gráficos de barras.
- Estudar o desempenho ao longo dos anos dos salários:
 - Buscar analisar como os salários evoluíram ao longo dos anos da pesquisa.
 - Podemos calcular a média salarial por ano.
 - **Melhor visualização:** Gráfico de linhas mostrando a tendência ao longo dos anos.
- Analisar a evolução do salário por tempo de experiência:
 - O objetivo aqui é analisar se com maiores experiências profissionais, os salários aumentam, e sobretudo, analisar como se essa progressão é proporcional.
 - **Visualização possível:** Boxplots com dois eixos salário X experiência.
- Analisar a distribuição do trabalho remoto por nível de experiência:
 - O objetivo aqui é analisar se profissionais mais ou menos experientes tendem a buscar vagas remotas.
 - Gráfico de barras da experiência X vagas remotas
- Estudar os tipos de contratos comparados com o nível de experiência:
 - O Objetivo é analisar qual o tipo de contratação mais comum para profissionais mais e menos experientes.
 - **Possível visualização:** Criar um gráfico de pizza avaliando a distribuição de cada profissional (junior, senior, pleno, etc.) em cada tipo de contrato.
 - * Fazemos um gráfico para cada tipo de contrato e avaliamos a distribuição dos níveis de experiência.
- **obs :** Uma visualização que pode ser interessante é um gráfico de pizza analisando a distribuição de cada tipo de modelo de trabalho (remoto, presencial, etc.) e também um gráfico de pizza representando a distribuição do tipo de emprego (Meio período, tempo integral, freelancer, etc.)