

Projeto Integrador – Análise Exploratória de Dados

Entregas: 1) Faça um relatório respondendo cada pergunta separadamente. 2) Link para a base utilizada. 3) Código completo em Python.

1) Escolha do Tema

2) Escolha da Base

3) Compreensão dos dados

- a) Qual é a estrutura do dataset? (Quantas linhas e colunas existem?)
- b) Quais são os tipos de variáveis presentes? (Numéricas, categóricas, texto, etc.)
- c) Existe algum problema de valores ausentes no dataset? Como lidar com eles?
- d) Existem valores duplicados ou inconsistências nos dados?

4) Estatísticas Descritivas

- a) Quais são as estatísticas básicas das variáveis numéricas? (Média, mediana, mínimo, máximo, desvio padrão).

5) Visualização e Padrões

- a) Existe uma relação entre variáveis numéricas? (Correlação entre idade e salário, por exemplo).
- b) Como visualizar melhor os dados? (Histograma, boxplot, scatterplot, heatmap)

6) Outliers e Anomalias

- a) Existem outliers nos dados? (Boxplots podem ajudar a identificar).
- b) Os dados apresentam alguma anomalia ou erro de coleta?

7) Comparações e Segmentações

- a) Os dados mostram diferenças entre grupos? (Exemplo: Salários médios por gênero)
- b) Como os dados se comportam ao longo do tempo? (Se houver uma variável temporal)
- c) Quais insights podem ser extraídos para tomada de decisão?

8) Contextualização e Aplicação

- a) Como os insights obtidos podem ser usados em ações concretas?
- b) Os dados analisados fazem sentido para o problema de negócio?