

## Exercício Avaliativo – Ciência de Dados

Prof. [clenio.silva@uniube.br](mailto:clenio.silva@uniube.br)

1. Você deverá realizar um estudo exploratório sobre um conjunto de dados escolhido, com foco em entender a estrutura dos dados, tratar inconsistências, identificar padrões e preparar os dados para uso em modelos de aprendizado de máquina. A entrega deve conter **um relatório técnico** (em PDF ou notebook comentado) com os seguintes itens:
  - Estatísticas descritivas (média, mediana, desvio padrão, etc.)
  - Visualizações (histogramas, boxplots, gráficos de dispersão, etc.)
  - Análise de correlação entre variáveis
  - Detecção e Remoção de Outliers
    - Métodos utilizados para identificar outliers (z-score, IQR, visualizações, etc.)
    - Comparação dos dados antes e depois da remoção
  - Tratamento de valores ausentes (remoção, imputação, etc.)
  - Codificação de variáveis categóricas (label encoding, one-hot encoding)

### Sugestões de Conjuntos de Dados (opcional):

- Titanic (sobrevivência de passageiros)
- Iris (classificação de flores)
- House Prices (valores de imóveis)
- Dados do IBGE ou Kaggle

Lembrando que os conjuntos de dados acima são apenas sugestões, existem uma infinidade de datasets na internet.