

Exercício Avaliativo de Ciência de Dados

Prof. clenio.silva@uniube.br

Objetivo: Os alunos deverão aplicar conceitos de ciência de dados para analisar um conjunto de dados real de voos

1. Obtenção dos Dados

- a) Use o dataset público de voos dos EUA disponível em: Kaggle - Flights Dataset (<https://www.kaggle.com/datasets/usdot/flight-delays>).

2. Análise Exploratória dos Dados (EDA)

- a) Identifique as variáveis mais relevantes.
- b) Visualize tendências, correlações e padrões (atrasos por companhia, aeroporto, horário etc).
- c) Verifique dados faltantes e valores extremos.

3. Preparação dos Dados

- a) Faça o tratamento de valores ausentes.
- b) Crie variáveis relevantes (ex: atraso binário, horário categorizado, etc).
- c) Normalize ou codifique variáveis categóricas conforme necessário.

4. Visualização de dados

- a) Crie gráficos para visualização dos dados.