

# Relatório descritivo de dataset

## Trabalho 1

### VISÃO GERAL

Existem diversos datasets em diversos formatos e sobre diversos assuntos. O objetivo deste trabalho é exercitar seus conhecimentos de Python, Numpy, Pandas, Matplotlib e Jupyter Notebook.

### OBJETIVOS

1. Trabalhar com um dataset.
2. Exercitar os conhecimentos das bibliotecas de processamento de dados.
3. Exercitar o Jupyter Notebook, tanto para programação quanto para geração de relatórios

### ESPECIFICAÇÕES

Neste trabalho, você deverá escolher um dataset (ou, se quiser, pode combinar vários), processá-lo, realizar uma limpeza (se necessário) e redigir um **relatório** mostrando todo o processamento do dado e fazendo uma análise descritiva de suas informações. Como exemplo, tomem como base a análise dos dados apresentada na aula do dia 08/06, em que foi feito um projeto de ponta a ponta (mas claro que **basta fazer a parte de análise dos dados**).

**OBS:** Não poderão ser reproduzidas análises feitas durante a disciplina, apresentadas pelos professores ou convidados. Caso os mesmos datasets sejam usados, é necessário que o relatório seja bem diferente, tanto em código quanto escrita.

### AVALIAÇÃO

Como não sabemos o dataset que será escolhido, decidimos não definir um número mínimo de análises a serem feitas. Porém, utilizaremos os seguintes critérios:

- 
- **Complexidade do dataset** - se o dataset for muito pequeno e/ou tiver poucas informações, haverá penalidade;
  - **Efetividade das análises** - garanta que você fez as análises mais importantes. Combinar sem motivo algum dois atributos e gerar um gráfico não será contabilizado para nota.
  - **Organização do notebook** - seu relatório será todo feito usando o Jupyter Notebook (o qual será o artefato de envio do trabalho). Logo, ele precisa está bem organizado, combinando código python com markdown;
  - **Reprodutividade do relatório** - seu notebook será executado no momento da avaliação. Garanta que as células estejam na ordem correta e que qualquer instrução adicional para a reprodutibilidade esteja devidamente apresentada no notebook;
  - **Criatividade das análises** - experimente combinar informações que podem parecer não estar relacionadas para obter *insights* sobre os dados;
  - **Tamanho da equipe** - trabalhos feitos por equipes possuem avaliações mais rígidas, sendo a dificuldade proporcional (não necessariamente de forma linear) ao número de membros.