

### **Projeto - Análise exploratória de dados - Turma de Segunda-feira**

#### Instruções

1. Grupos com, no máximo, 4 componentes.
2. Enviar apenas o link do repositório no GitHub. O repositório deverá conter o relatório em pdf e códigos em Python, pelo link do Moodle.
3. Enviar até **23/06/2025 às 23:59**
4. Trabalhos idênticos receberão nota 0.

#### O que fazer

1. O grupo deverá escolher, no mínimo, DOIS arquivos, sendo arquivos csv de duas pastas diferentes, relacionando-os. Por exemplo, o arquivo **excess\_mortality.csv** da pasta excess\_mortality e o **covid\_hospitalization.csv** da pasta hospitalizations.
  - a. Os dados estão disponíveis no formato csv no link  
<https://github.com/owid/covid-19-data/tree/master/public/data>
  - b. [https://catalog.data.gov/dataset/?res\\_format=CSV](https://catalog.data.gov/dataset/?res_format=CSV)

2. Preencher a planilha com o nome dos componentes de acordo com os arquivos csv escolhidos.

#### **Link:**

[https://docs.google.com/spreadsheets/d/1s-08Fkgwbqi87ttZO\\_YHzJ\\_Ti8wB2Nz\\_UBloB38-KogE/edit?gid=0#gid=0](https://docs.google.com/spreadsheets/d/1s-08Fkgwbqi87ttZO_YHzJ_Ti8wB2Nz_UBloB38-KogE/edit?gid=0#gid=0)

3. **Atenção; os mesmos arquivos csv só poderão ser utilizados por, no máximo, 2 grupos.**
4. O relatório deverá conter a descrição e explicação dos cálculos realizados e gráficos obtidos de acordo com o conteúdo da disciplina, inclusive testes de hipótese. Naturalmente, nem todos os itens devem ser abordados. O grupo deverá determinar as análises a serem realizadas, ou seja, a descrição do problema e a solução a partir do uso dos dados e do Python. Faça “perguntas” à base de dados e mostre o resultado.
5. O relatório deverá conter análises referentes aos conteúdos:
  - a. Estatística descritiva
  - b. Probabilidade
  - c. Inferência
7. Não serão aceitos links do Colab, Jupyter notebook ou equivalentes.

Bom trabalho!!!