

## Trabalho I

Prof. Tetsu Sakamoto

### Instru  es:

- Este trabalho consiste no desenvolvimento de uma an lise de dados completa utilizando dados reais para avaliar a capacidade do discente em utilizar as ferramentas computacionais abordadas durante a aula para importar, organizar, analisar e sumarizar os dados.
- Este trabalho ser  realizado em grupos de 2 a 3 pessoas. Os grupos ser o sorteados
- Realize todos os procedimentos utilizando a linguagem Python.
- **Ser  considerado pl gio a utiliza  o de an lises e vinhetas j  publicadas.**
- As apresenta  es referentes a este trabalho ser o realizadas nos dias 23/07/24 e 25/07/24.

### Dados:

Os conjuntos de dados selecionados para este trabalho ser o:

- Grupo 1 e 2:  
<https://www.kaggle.com/datasets/rabieelkharoua/students-performance-dataset>
- Grupo 3 e 4:  
<https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset>
- Grupo 5 e 6:  
<https://www.kaggle.com/datasets/nelgiriyeewithana/most-streamed-spotify-songs-2024>
- Grupo 7 e 8:  
<https://www.kaggle.com/datasets/rabieelkharoua/air-quality-and-health-impact-dataset>
- Grupo 9 e 10:  
<https://www.kaggle.com/datasets/teejmahal20/airline-passenger-satisfaction>

Acesse o link do conjunto de dados selecionado e baixe os arquivos

necessários. Caso exista, não deixe de verificar a documentação e artigo(s) associado(s) aos dados.

### **Análise dos dados**

Nesta etapa você irá gerar perguntas sobre seus dados. Procure respostas visualizando, transformando e modelando seus dados.

Você pode consultar este vídeo que fala sobre as etapas de uma análise exploratória dos dados: <https://www.youtube.com/watch?v=lgCNTuLBMK4>

- Analise a estrutura e consistência do seu conjunto de dados. Caso necessário, realize um pré-processamento dos dados.
- Execute uma análise exploratória, a partir de análises gráficas e numéricas.
- Analise os valores mais comuns, discrepantes e ausentes. Identifique possíveis problemas com os seus dados. Analise variações, associações e identifique padrões. Gere as visualizações e interprete os resultados.
- A análise exploratória gerou alguma hipótese para posterior investigação?
- Elabore uma apresentação de cerca de 10 minutos, descrevendo seu conjunto de dados, as principais informações obtidas e hipóteses a serem investigadas.
- Submeta no SIGAA o script que você utilizou para obter os resultados apresentados no trabalho.

### **Avaliação**

O grupo será avaliado pelos seguintes critérios (todos possuem o mesmo peso):

- Domínio das ferramentas computacionais utilizadas para a análise de dados;
- Domínio sobre o conjunto de dados;
- Apresentação do trabalho;
- Participação na discussão da apresentação dos outros grupos.