

# Lista 1 - Métodos Estatísticos Aplicados às Ciências Sociais

## Objetivo

A lista tem como objetivo realizar uma análise exploratória dos dados considerando os conceitos vistos nas primeiras 10 aulas do curso de Métodos Estatísticos Aplicados às Ciências Sociais. Dessa forma, a lista contemplará análises univariadas a partir de gráficos e de medidas descritivas.

## Descrição da lista

Para a realização dessa lista, cada pessoa trabalhará com uma base de dados para responder às mesmas questões. As bases disponíveis são:

1. Amazon Kindle Books Dataset
2. Anime Dataset 2023
3. Full TMDB Movies Dataset
4. Pokemon Dataset
5. Titanic Dataset

## Entrega

A entrega deve ser feita em **Colab Notebook** com a parte escrita como uma célula de texto (Markdown) dentro do próprio Colab.

Dica: Use e abuse das possibilidades do Markdown!

## Questões

1. Quem são os indivíduos/observações da sua base de dados?
2. Identifique os tipos das suas variáveis.
3. Em uma abordagem focada em problemas (**PPDAC**), que perguntas você pretende responder com a sua base de dados e que planejamento pretende seguir para respondê-las (etapas “Problema” e “Plano”) ?

OBS:

- Você deve realizar apenas análises univariadas a partir de gráficos e de medidas descritivas.
- Obrigatoriamente, use em seu planejamento, pelo menos, uma variável qualitativa e uma variável quantitativa.

Dicas:

- Aqui você pode incluir, por exemplo, detalhes de que variáveis planeja usar e que análises fará para responder às perguntas.
  - Em algumas bases, é possível realizar análises de séries temporais. E em outras, é possível realizar comparações das distribuições para diferentes categorias.
4. Realize suas análises e, para cada uma delas, descreva a sua interpretação dos resultados (etapa “Análise”).
    - a. Para as variáveis qualitativas, faça os gráficos da distribuição de frequência e descreva sua interpretação do resultado.
    - b. Para as variáveis quantitativas, além do gráfico da distribuição de frequência e a descrição do que percebe pela análise visual, use medidas descritivas de centro e de variabilidade e descreva o que essas medidas te dizem sobre as variáveis que está analisando.

OBS: Caso a sua base permita, realize análises comparativas das variáveis quantitativas para diferentes características das observações.

Dica: Talvez seja necessário fazer ajustes na base para responder aos problemas que você levantou. Lembre-se das funções do Pandas para realizar transformações em um DataFrame.

5. Escreva em texto corrido (máximo 3 parágrafos) as conclusões da sua análise considerando as perguntas que fez na questão 3 (etapa “Conclusão”).

OBS: Aqui também podem ser descritas novas ideias para dar continuidade à análise (como, por exemplo, outras perguntas que surgiram no caminho ou, até mesmo, relacionar a sua base com outra base específica para te possibilitar responder a novas perguntas).