

# Tópicos Especiais em Sistemas de Informação I

## ATV A2 - Preparação, Análise Exploratória e Dashboard

Escolha um dataset e faça download. (guarde a URL do dataset usado)

1. Use o pandas para carregar o dataset e saber as informações básicas tais como números de instancias, média por atributo, etc. Por exemplo: `info()` e `describe()` do pandas.
2. Faça os tratamento de nulos. Ex. biblioteca `missino`
3. Faça o tratamento de anomalias. Sugestão: use `boxplot` e `histogramas` para certificar-se do problema. Use também medidas de `skewness` e `kurtosis`.
4. Faça transformações (reescala, codificação) necessárias. Tenha certeza que cada coluna numérica seja realmente um conteúdo numérico e não uma codificação com inteiros. Justifique os casos menos intuitivos.
5. Realize uma análise exploratória (univariada e bivariada) mostrando a distribuição de valores (`histograma`), valores mais frequentes (`countplot`) das colunas que você julgar mais relevantes e que podem oferecer padrões ou tendências mais destacadas. Mostrando `heatmap` de correlações entre as colunas numéricas, `scatter plots` para pares com maiores correlações, ou `boxplots` usando variáveis categóricas e variáveis numéricas.
6. Crie um dashboard seguindo os passos vistos em aula: (a) identificação de usuário; (b) questões de negócio; (c) métricas e indicadores; (d) gráficos; (e) layout; (f) demo funcionando em python, tableau ou power BI.

Observações:

- Toda as transformações devem preservar o dataset original.
- Entregue um único documento notebook (.ipnb) com link para o dataset original. A sessão inicial do documento deve conter a identificação do IFMA, Curso, Disciplina, Professor e sua matricula e nome.
- Organize o notebook com seções adicionais, sendo uma seção para cada passo (dataset, infos, anomalias, etc)
- Crie um dashboard em python, tableau ou power BI
- Submeta via google classroom até a data indicada um link para o google colab ou arquivo .ipnb + link do dataset
- Além das entregas, os trabalhos serão realizados em equipes de até 4 alunos e serão apresentados em aula para obtenção da nota.