



Para este primeiro Hands On, iremos trabalhar com o conjunto de dados do Titanic. É um clássico para iniciar os seus projetos e portfólio na Ciência de Dados. Este constituirá de duas partes e será toda trabalhada em python.

Começaremos com a primeira parte, que será focada em entender os dados, ou seja, focar no EDA – *Exploratory Data Analysis* (análise exploratória de dados). Para entender melhor o conceito e como trabalha-lo, recomendo a leitura destes links: <https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15> (inglês), <https://minerandodados.com.br/exploratory-data-analysis-eda/>, <https://medium.com/@gabriel.stankevix/analise-explorat%C3%B3ria-de-dados-732007ddbfaf>. Para saber mais detalhes sobre o conjunto de dados (aba Data) e ver códigos de outras pessoas (aba Code), confira aqui: <https://www.kaggle.com/c/titanic>

Por fim, os dados encontram-se na pasta Titanic, no repositório chamado HandsOn\_Projetos no GitHub. Este foi criado para organizar o projeto, uma ótima ideia é salvá-lo (Star) ou realizar um fork para seu perfil, porque será usado novamente no futuro: [https://github.com/anamioto/HandsOn\\_Projetos](https://github.com/anamioto/HandsOn_Projetos).

#### Dicas e observações:

1. Utilize o conjunto train.csv!
2. Pesquise e pesquise bastante;
3. Veja exemplos, poderá trazer insights;
4. Entenda bem cada uma das variáveis e os tipos delas;
5. Tente utilizar as bibliotecas pandas, numpy, matplotlib (caso esteja começando), seaborn (caso tenha mais experiência ou queira um desafio);
6. Faça com calma e capriche!

Vocês terão 40 dias para entregar, e assim será marcada a reunião para discussão e apresentação dos resultados! **Bons estudos e bom trabalho!**