



# Semantix

## **Desafio Data Science**

BIG DATA & ANALYTICS



Olá!

Gostaríamos de realizar um teste que será usado para sabermos a sua proficiência nas habilidades para a vaga. O teste consiste na análise de um conjunto de dados e seus resultados deverão ser armazenados no *GitHub*. O link do seu repositório deverá ser compartilhado conosco ao final do teste.

Quando usar alguma referência ou biblioteca externa, informe no arquivo *README* do seu projeto. Se tiver alguma dúvida, use o bom senso e se precisar deixe isso registrado na documentação do seu projeto.

## *Dataset: Bank Marketing*

Fonte oficial do *dataset*: <https://archive.ics.uci.edu/ml/datasets/bank+marketing>

Dados: <https://archive.ics.uci.edu/ml/machine-learning-databases/00222/bank.zip>

Arquivos incluídos no link acima:

- **bank.csv**: uma versão reduzida do conjunto de dados;
- **bank-full.csv**: o conjunto completo;
- **bank-names.txt**: com a descrição dos campos do conjunto de dados.

Para a resolução deste teste utilize o arquivo **bank-full.csv** e para uma descrição sobre o *dataset* use **bank-names.txt** e/ou a fonte oficial.

## *Crítérios de Avaliação*

- Capacidade de entender e solucionar problemas
- Conhecimentos em *Machine Learning*
- Visualização de dados
- Conhecimentos em estatística
- Linguagens de programação
- Apresentação dos resultados



## Questões

1. Qual profissão tem mais tendência a fazer um empréstimo? De qual tipo?
2. Fazendo uma relação entre número de contatos e sucesso da campanha quais são os pontos relevantes a serem observados?
3. Baseando-se nos resultados de adesão desta campanha qual o número médio e o máximo de ligações que você indica para otimizar a adesão?
4. O resultado da campanha anterior tem relevância na campanha atual?
5. Qual o fator determinante para que o banco exija um seguro de crédito?
6. Quais são as características mais proeminentes de um cliente que possua empréstimo imobiliário?