



**Data Science  
Academy**

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

# Big Data Analytics com R e Microsoft Azure Machine Learning

## Definindo o Problema de Negócio

Avaliação de Risco de Crédito



## Sumário

Este experimento tem como objetivo criar um modelo preditivo para classificar o risco de crédito de clientes de uma instituição bancária.

## Descrição

Este experimento visa demonstrar o processo de construção de um modelo de classificação para prever o risco de concessão de crédito a clientes de um banco. Usaremos um conjunto de dados para construir e treinar nosso modelo.

## Dados

O conjunto de dados “German Credit Data” será usado para construir e treinar o modelo, neste experimento. Este dataset é baseado em dados reais gerados por um pesquisador da Universidade de Hamburgo, na Alemanha.

O dataset contém 1000 observações e 20 variáveis, representando os dados de clientes, tais como: status da conta corrente, histórico de crédito, quantidade de crédito atual, empregabilidade, residência, idade, etc..

Dataset: [https://archive.ics.uci.edu/ml/datasets/Statlog+\(German+Credit+Data\)](https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data))

## Objetivo

O objetivo será prever o risco que cada cliente oferece para o banco, na hora de conceder uma linha de crédito. O modelo preditivo deve ser bastante preciso, pois conceder crédito a um cliente com péssimo potencial de pagamento, pode trazer um grande prejuízo para o banco.