



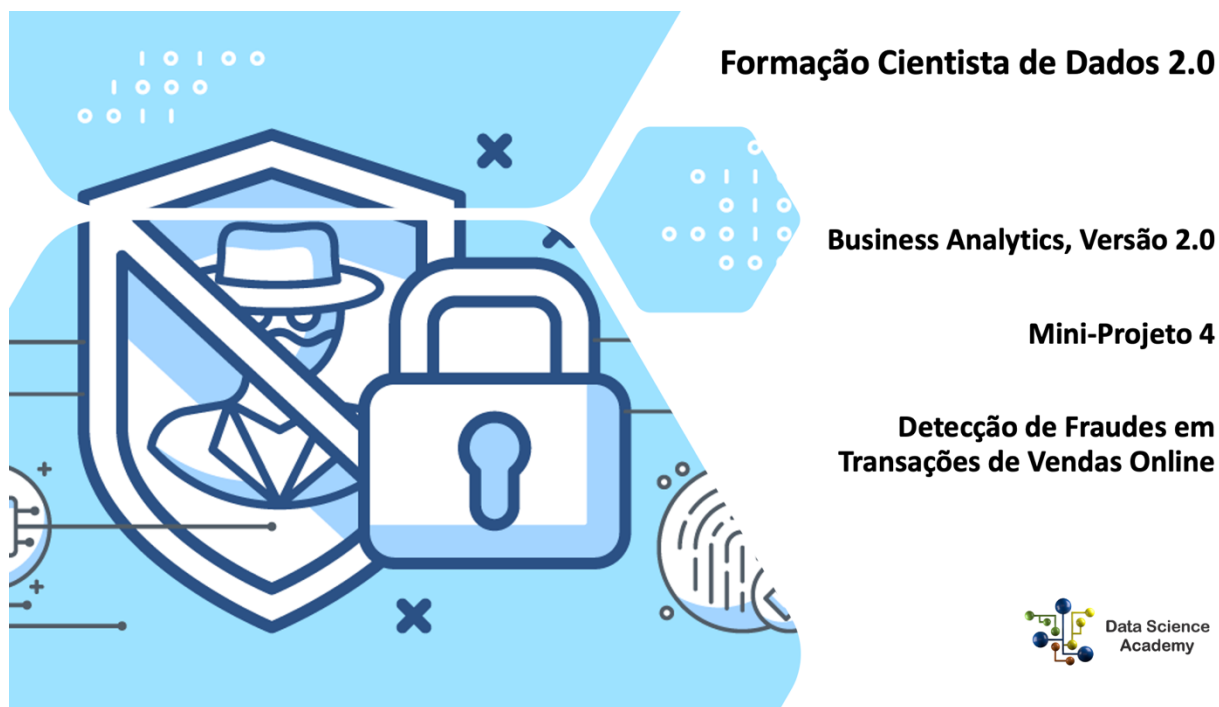
**Data Science
Academy**

www.datascienceacademy.com.br

Business Analytics

Mini-Projeto 4

Definição do Problema



Neste Mini-Projeto vamos construir um modelo capaz de prever a probabilidade de que uma transação online seja fraudulenta, conforme indicado pela variável target **isFraud**. Conforme você já sabe criaremos o modelo usando dados históricos e depois faremos previsões em novos dados.

Os dados históricos estão divididos em dois arquivos: de identidade (identifica o cliente) e de transação (identifica a transação), que são unidos pela coluna TransactionID. Nem todas as transações possuem informações de identidade correspondentes.

Como este é um problema de classificação (queremos prever sim/não para cada transação) usaremos um algoritmo de classificação em aprendizagem supervisionada.

As previsões devem ser salvas em arquivo csv e entregues ao cliente/área de negócio.

Na aula seguinte você encontra detalhes da fonte de dados.