

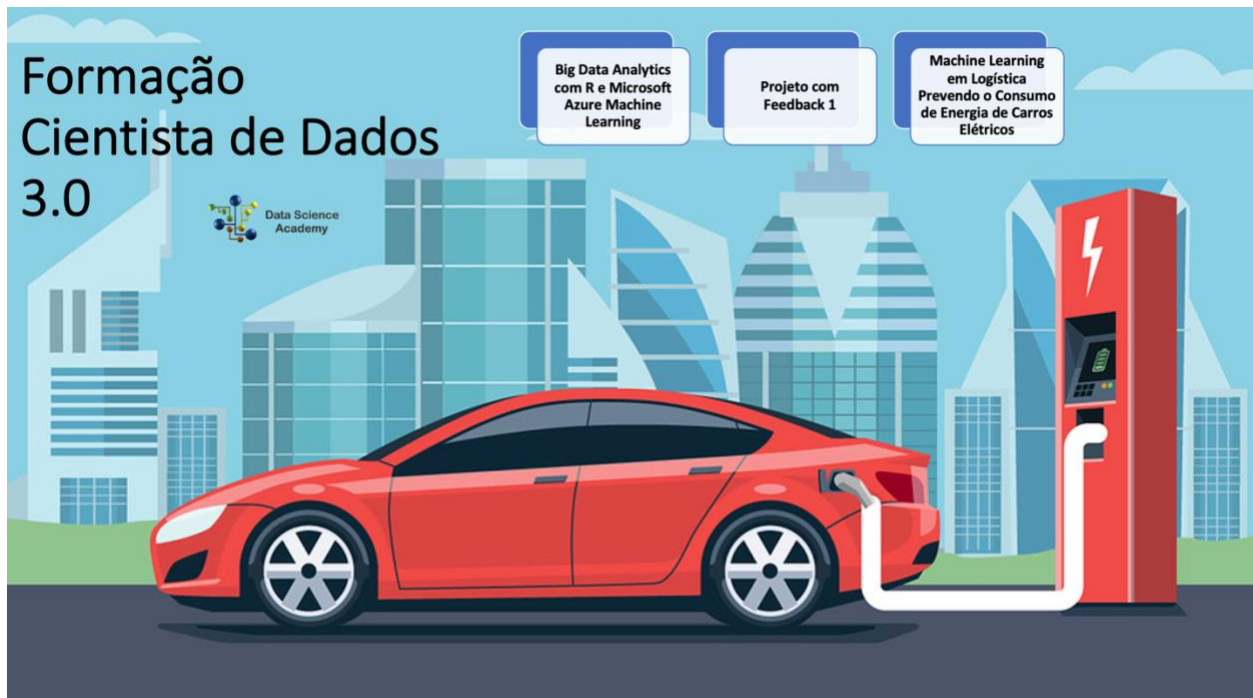


*Big Data Analytics com R e Microsoft Azure Machine Learning 3.0*

# Big Data Analytics com R e Microsoft Azure Machine Learning Versão 3.0

## Projeto com Feedback 1

### Machine Learning em Logística Prevendo o Consumo de Energia de Carros Elétricos



Uma empresa da área de transporte e logística deseja migrar sua frota para carros elétricos com o objetivo de reduzir os custos.

Antes de tomar a decisão, a empresa gostaria de prever o consumo de energia de carros elétricos com base em diversos fatores de utilização e características dos veículos.

Usando um incrível dataset com dados reais disponíveis publicamente, você deverá construir um modelo de Machine Learning capaz de prever o consumo de energia de carros elétricos com base em diversos fatores, tais como o tipo e número de motores elétricos do veículo, o peso do veículo, a capacidade de carga, entre outros atributos.

**Para a construção desse projeto, recomendamos a utilização da linguagem R e o dataset disponível para download no link abaixo:**

<https://data.mendeley.com/datasets/tb9yrptydn/2>

Este conjunto de dados lista todos os carros totalmente elétricos com seus atributos (propriedades) disponíveis atualmente no mercado. A coleção não contém dados sobre carros híbridos e carros elétricos dos chamados “extensores de alcance”. Os carros a hidrogênio também não foram incluídos no conjunto de dados devido ao número insuficiente de modelos produzidos em massa e à especificidade diferente (em comparação com veículo elétrico) do veículo, incluindo os diferentes métodos de carregamento.

O conjunto de dados inclui carros que, a partir de 2 de dezembro de 2020, poderiam ser adquiridos na Polônia como novos em um revendedor autorizado e aqueles disponíveis em pré-venda pública e geral, mas apenas se uma lista de preços publicamente disponível com versões de equipamentos e parâmetros técnicos completos estivesse disponível. A lista não inclui carros descontinuados que não podem ser adquiridos como novos de um revendedor autorizado (também quando não estão disponíveis em estoque).

O conjunto de dados de carros elétricos inclui todos os carros totalmente elétricos no mercado primário que foram obtidos de materiais oficiais (especificações técnicas e catálogos) fornecidos por fabricantes de automóveis com licença para vender carros na Polônia. Esses materiais foram baixados de seus sites oficiais. Caso os dados fornecidos pelo fabricante estivessem incompletos, as informações eram complementadas com dados do AutoCatálogo SAMAR (link disponível na seção Referências da fonte de dados).

Seu trabalho é construir um modelo de Machine Learning capaz de prever o consumo de energia de veículos elétricos.

Quando concluir o projeto, envie o(s) script(s) que você criar para o seguinte e-mail: [suporte@datascienceacademy.com.br](mailto:suporte@datascienceacademy.com.br).

Caso você tenha criado datasets auxiliares e esses sejam muito grandes, armazene em um diretório virtual (existem vários na internet, como Google Drive ou Dropbox) e envie o link para que nossa equipe possa baixar os datasets. Se os arquivos forem pequenos (uma amostra do dataset original), envie no anexo junto com o script. Documente seu script tanto quanto possível.

Caso prefira, disponibilize seu projeto no Github e envie o link do seu repositório para nossa equipe no e-mail [suporte@datascienceacademy.com.br](mailto:suporte@datascienceacademy.com.br). Nesse caso, o Readme do repositório deve constar que este trata-se de um projeto da Formação Cientista de Dados da Data Science Academy, caso contrário não será avaliado.

Em até 72 horas, daremos o feedback respondendo seu e-mail. Caso não receba a resposta em até 72 horas entre em contato com a nossa equipe para verificar se recebemos seu projeto.

Bom trabalho!