

Big Data Analytics com R e Microsoft Azure Machine Learning



Data Science Academy



Azure Machine Learning



Data Science Academy



Seja Bem-Vindo



Data Science Academy



Introdução



Data Science Academy



Data Science Academy



E o que estudaremos agora?



Data Science Academy

O que estudaremos agora?

Capítulo 08 – Introdução ao Azure, criação de experimentos, importação/exportação de dados, criação e avaliação de modelos e execução de scripts R. Tudo isso no Azure. E ainda estudaremos o pacote Caret, para Machine Learning em R

Capítulo 09 – Manipulação e Transformação de Dados, Metadados, Tratamento de Dados Missing, Padronização, Erros e Outliers e SQL no Azure

Capítulo 10 – Regressão

Capítulo 11 – Classificação

Capítulo 12 – Deploy do Modelo como Web Service



Data Science Academy

O que estudaremos agora?

Capítulo 12 – Projetos 4 e 5 / Avaliação Final

Capítulo 13 – Bonus (Oracle R Enterprise e Microsoft R)



Data Science Academy

Mais Bonus?



Data Science Academy

Bonus

Capítulo	Algoritmos de Machine Learning em R
Capítulo 7	Redes Neurais Support Vector Machines
Capítulo 8	Clustering com K-Means Decision Trees
Capítulo 9	Regressão Logística Classificação com Naïve Bayes



Data Science Academy



O que é o Microsoft Azure Machine Learning?



Data Science Academy

O que é o Microsoft Azure Machine Learning?

Microsoft Azure Machine Learning é um framework de Machine Learning que pode ser usado para construir modelos de análise preditiva, através de conjuntos de dados de treinamento e uma variedade de fontes de dados



Data Science Academy



O que é o Microsoft Azure Machine Learning?

Machine Learning em nuvem (cloud)




Data Science Academy



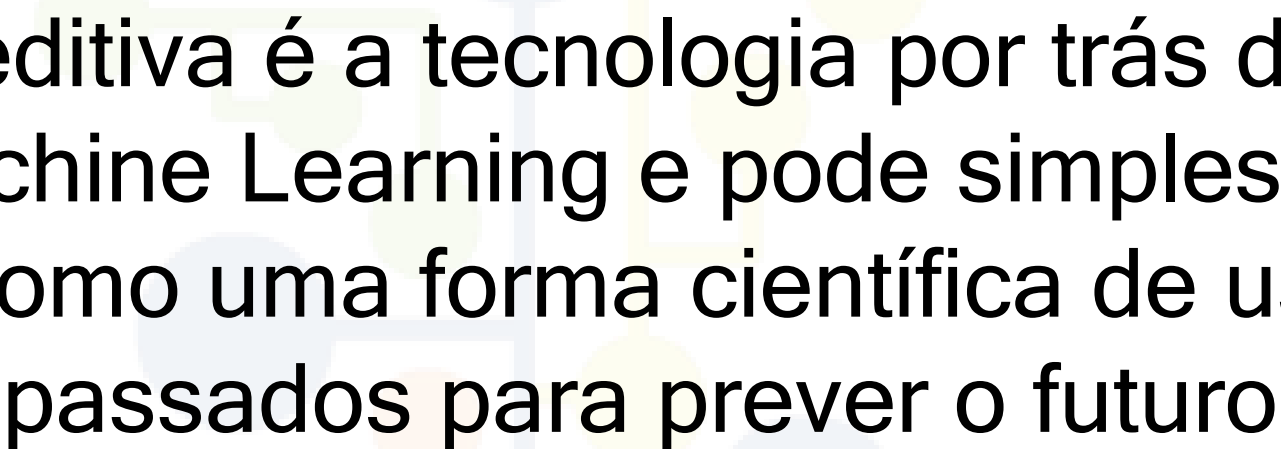
O Microsoft Azure Machine Learning oferece uma forma rápida de trabalhar com análise preditiva



Data Science Academy



Análise Preditiva é a tecnologia por trás do Microsoft Azure Machine Learning e pode simplesmente ser definida como uma forma científica de usar dados passados para prever o futuro



Data Science Academy

Análise Preditiva não é
advinhação. É Ciência sendo
usada para extrair previsões
a partir dos dados



Data Science Academy



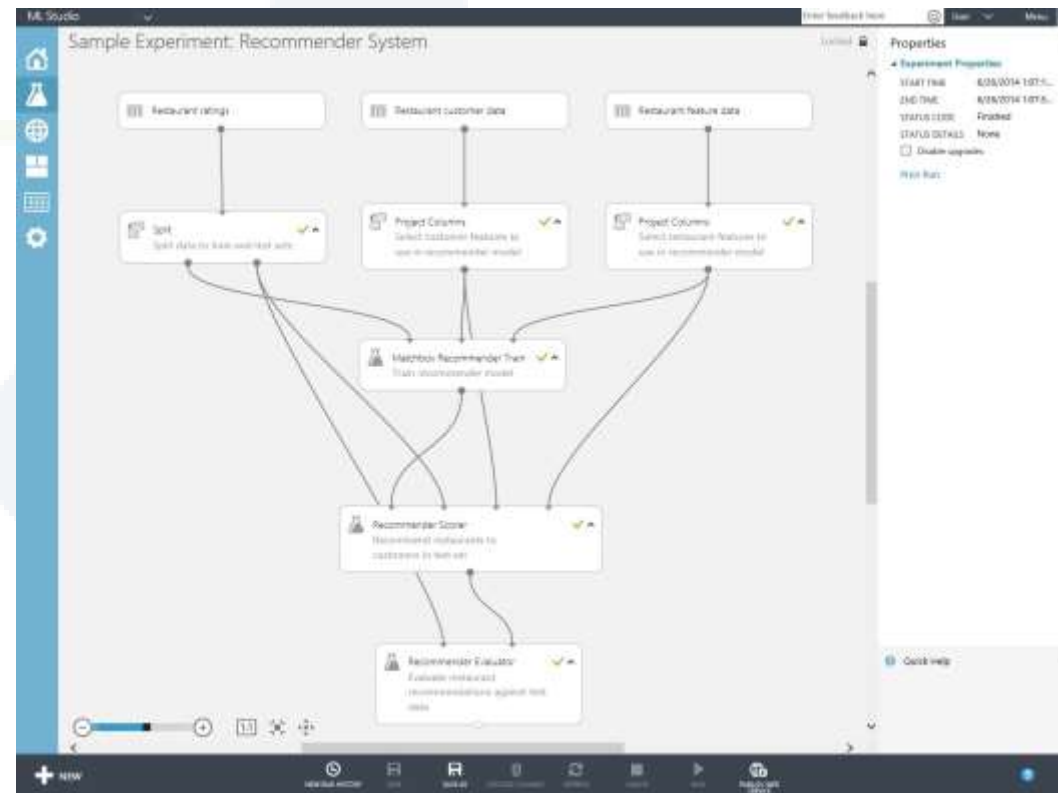
**FAIL FAST.
SUCCEED
FASTER**

lean**startup**machine



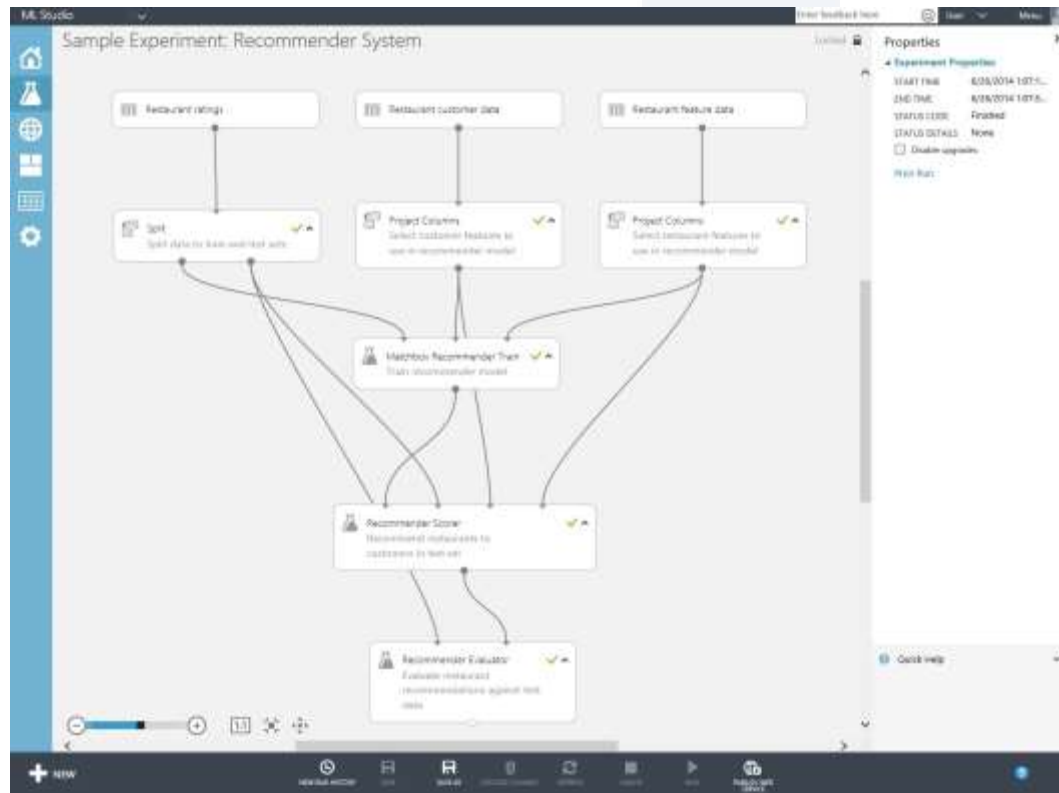
Data Science Academy

Azure Machine Learning Studio



Data Science Academy

Azure Machine Learning Studio



- R
- Python
- SQL



Data Science Academy



Não confunda o Azure Machine Learning
com o Azure



Data Science Academy

Mas não pense que as facilidades trazidas pelo Azure, significam que você não tenha que conhecer bem os algoritmos de [Machine Learning](#) e teorias necessárias para compreender claramente o que pode ser feito com o Azure ML.



Data Science Academy

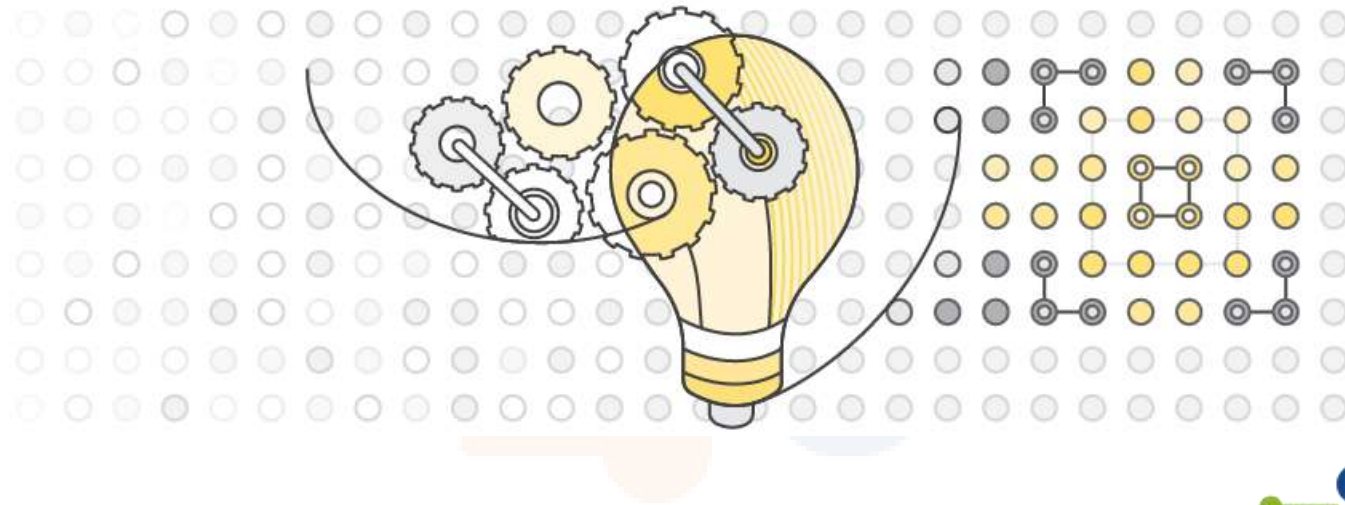


O Azure ML possui 3 grandes concorrentes



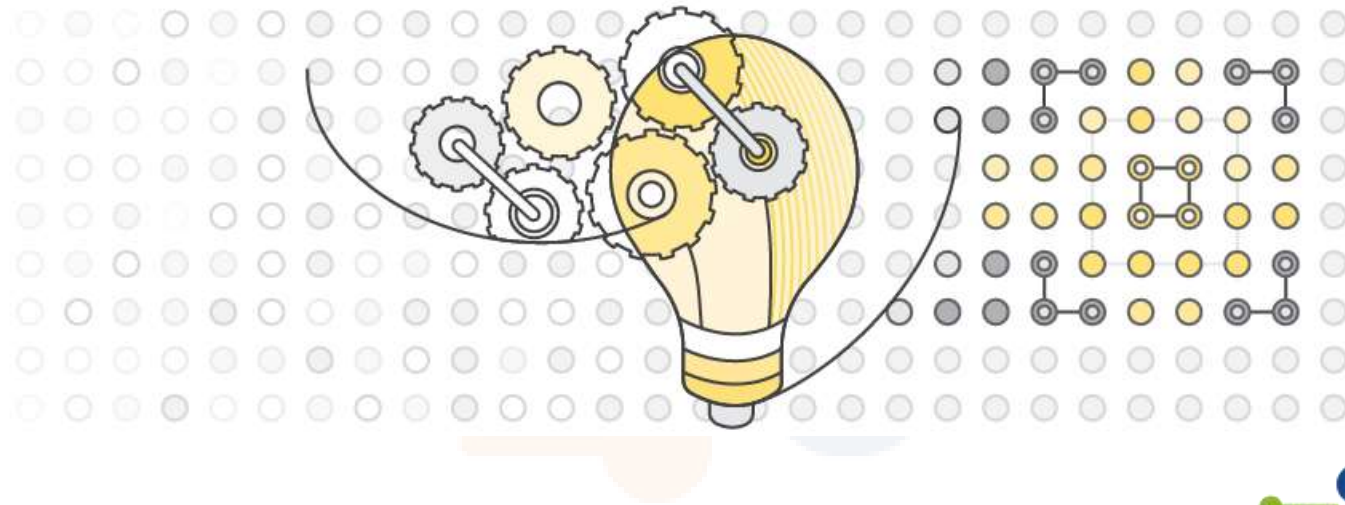
Data Science Academy

Amazon Machine Learning



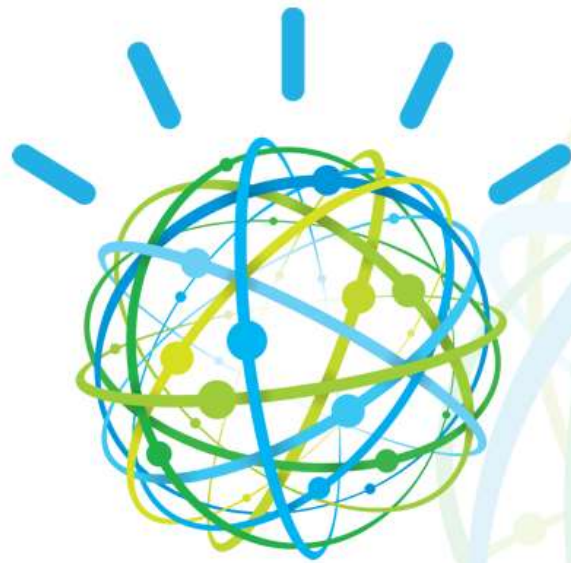
Data Science Academy

Amazon Machine Learning



Data Science Academy

IBM Watson Analytics

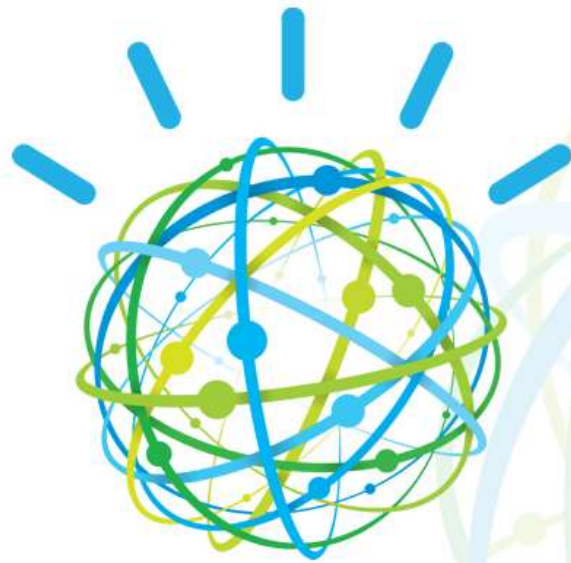


**IBM
Watson
Analytics**



Data Science Academy

IBM Watson Analytics



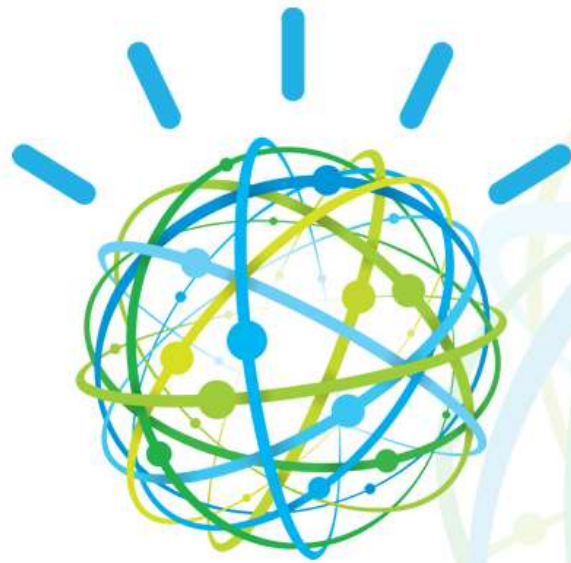
**IBM
Watson
Analytics**

AaaS



Data Science Academy

IBM Watson Analytics



**IBM
Watson
Analytics**

AaaS



Data Science Academy

Apache Spark



Data Science Academy

Apache Spark



Data Science Academy

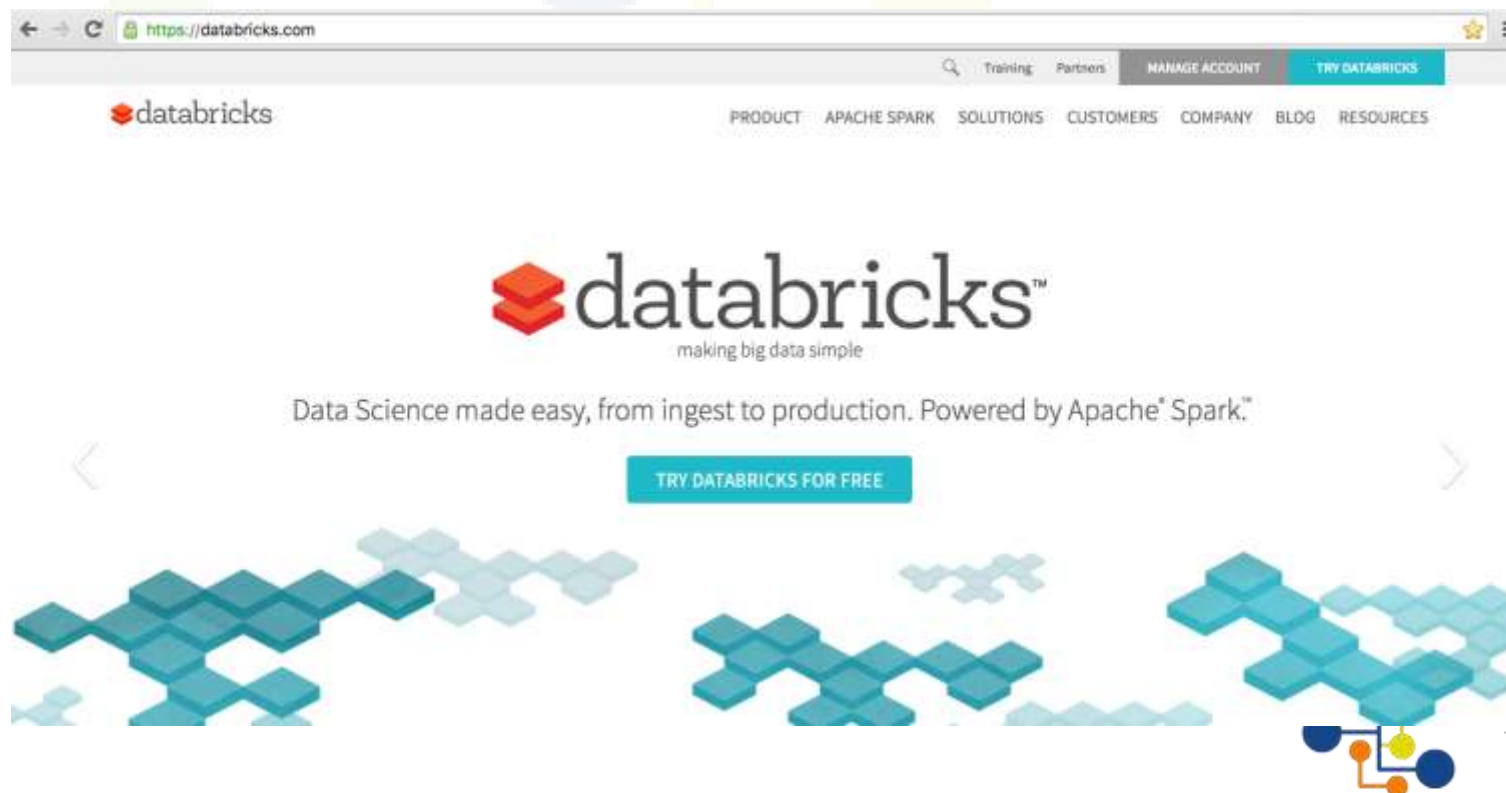
Apache Spark

trade-off



Data Science Academy

Apache Spark





Apache Spark

Spark suporta diferentes tarefas de [Data Science](#)



Data Science Academy



Resumindo



Data Science Academy

O resultado de um projeto de Data Science, pode ser a criação de um sistema de análise de dados em tempo real, que pode se tornar o componente de uma aplicação de negócio



Data Science Academy



Azure Machine Learning Fundamentos



Data Science Academy



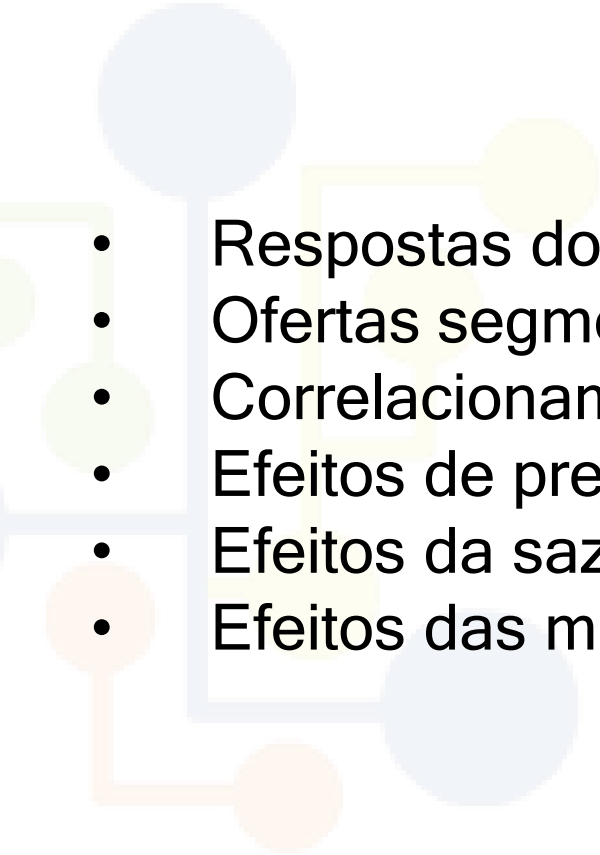
Azure Machine Learning Fundamentos



Data Science Academy



MARKETING & VENDAS

- 
- Respostas dos clientes a ofertas
 - Ofertas segmentadas por dados demográficos
 - Correlacionamento entre os clientes
 - Efeitos de preços e descontos
 - Efeitos da sazonalidade
 - Efeitos das mídias sociais



Data Science Academy



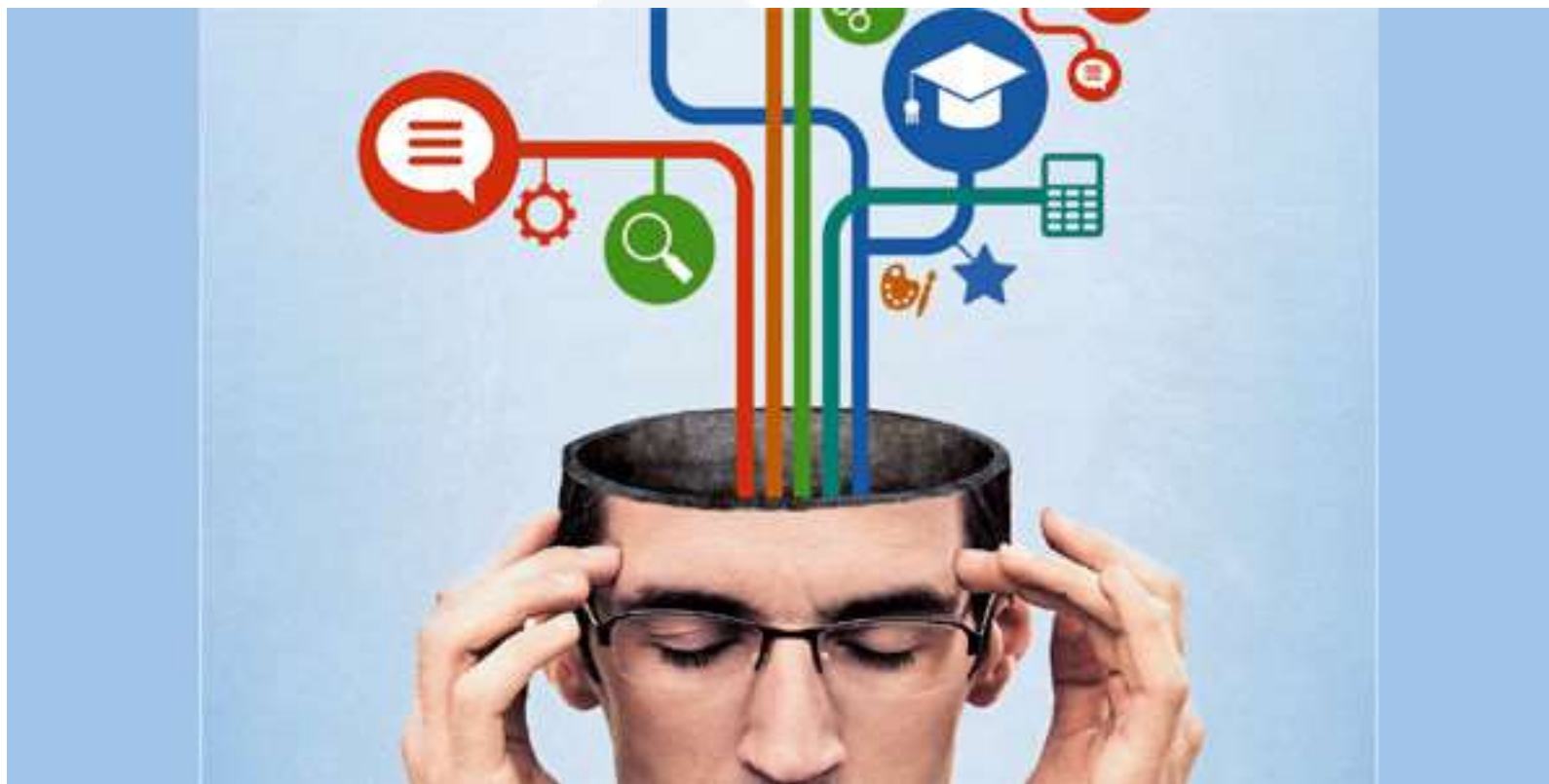
- Respostas dos clientes a ofertas
- Ofertas segmentadas por dados demográficos
- Correlacionamento entre os clientes
- Efeitos de preços e descontos
- Efeitos da sazonalidade
- Efeitos das mídias sociais



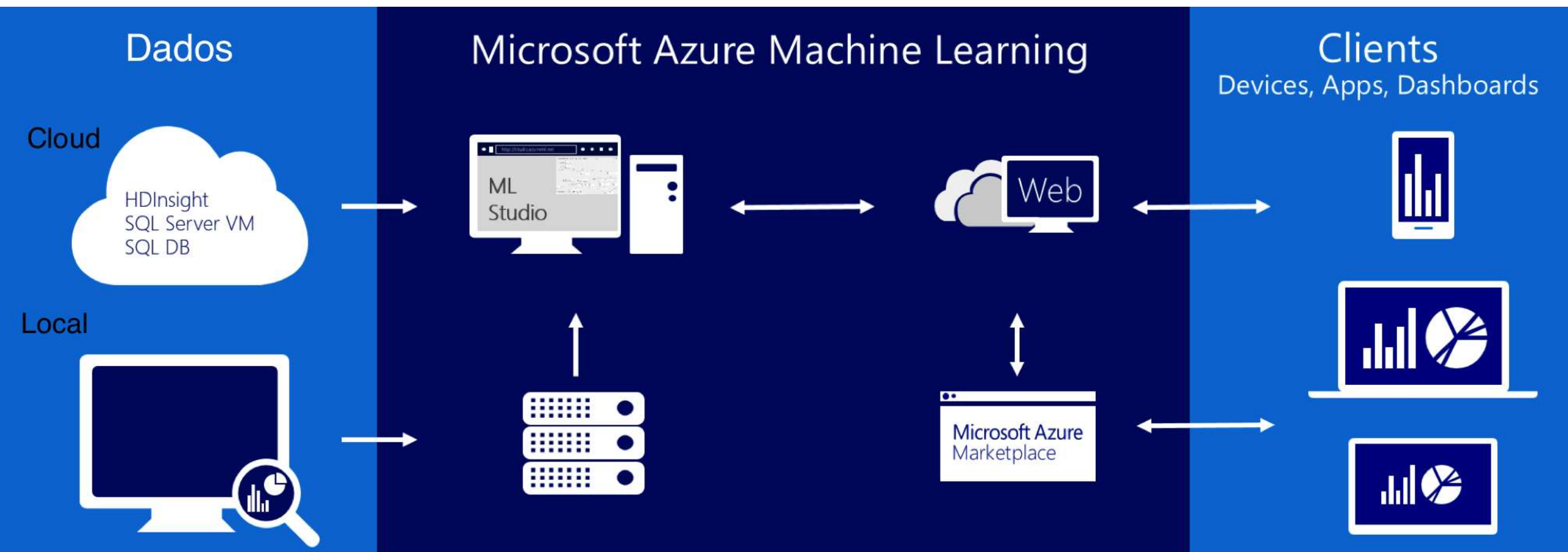
Data Science Academy



Data Science Academy



Data Science Academy



Data Science Academy

Azure Machine Learning Workflow



Data Science Academy

Azure Machine Learning Tool Kit

- **Classificação** - algoritmos usados para classificar dados em diferentes categorias e então fazer previsões sobre dados com variáveis qualitativas.
- **Regressão** - algoritmos usados para prever uma ou mais variáveis numéricas.
- **Clustering** - algoritmos usados para identificar grupos ou padrões em conjuntos de dados e então fazer previsões de classificação em grupos, para uma determinada variável.



Data Science Academy

Criar um Novo Modelo



Testar o Novo Modelo



Deploy do Novo Modelo



Data Science Academy



O Azure Machine Learning não é gratuito



Data Science Academy

Pacote Caret



Data Science Academy

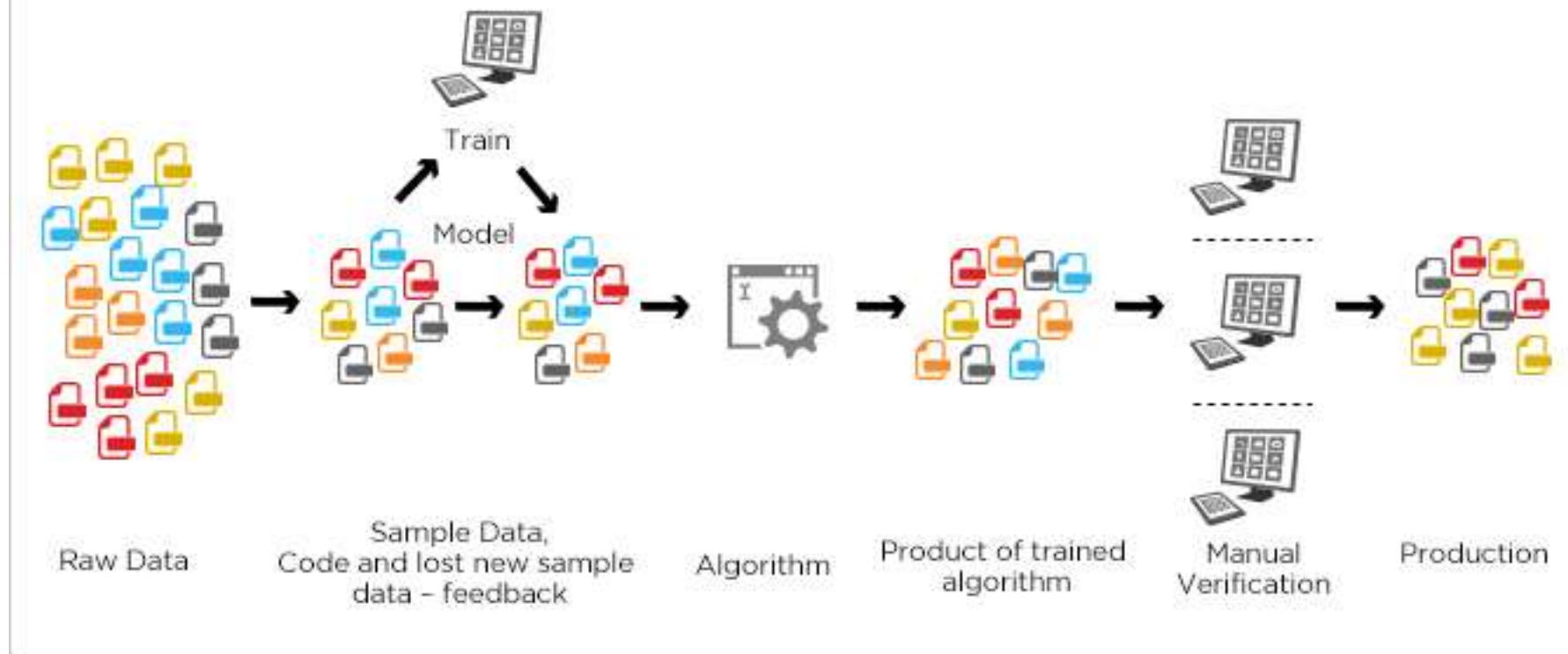


www.datascienceacademy.com.br



Data Science Academy

Supervised Learning



Data Science Academy

Pacote Caret



Data Science Academy

Pacote Caret

Supervised Learning



Unsupervised Learning



Data Science Academy

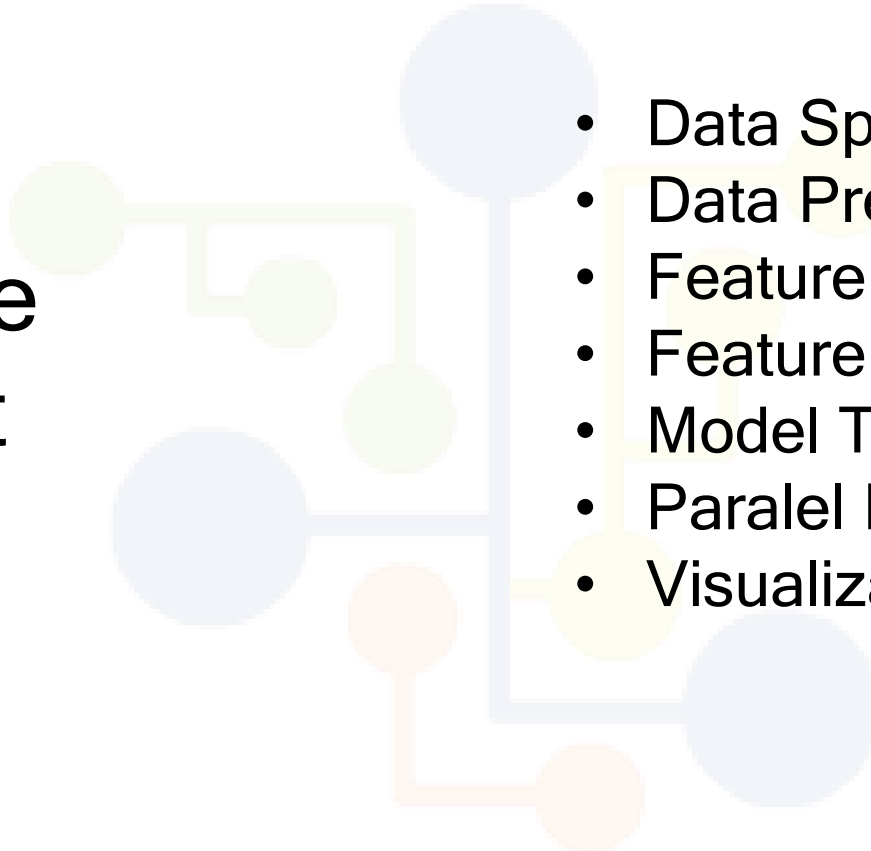
Pacote Caret



Data Science Academy



Pacote
Caret

- 
- Data Splitting
 - Data Pre-processing
 - Feature Selection
 - Feature Importance
 - Model Tuning
 - Parallel Processing
 - Visualization



Data Science Academy



Obrigado!



Data Science Academy