



escola
britânica de
artes criativas
& tecnologia

Profissão: Cientista de Dados

Análise de agrupamentos

Análise de agrupamentos

Brincando com a ideia







Redução de dimensionalidade

Em PCA fazemos redução de dimensionalidade nas colunas

Com Clusters fazemos redução de dimensionalidade nas linhas

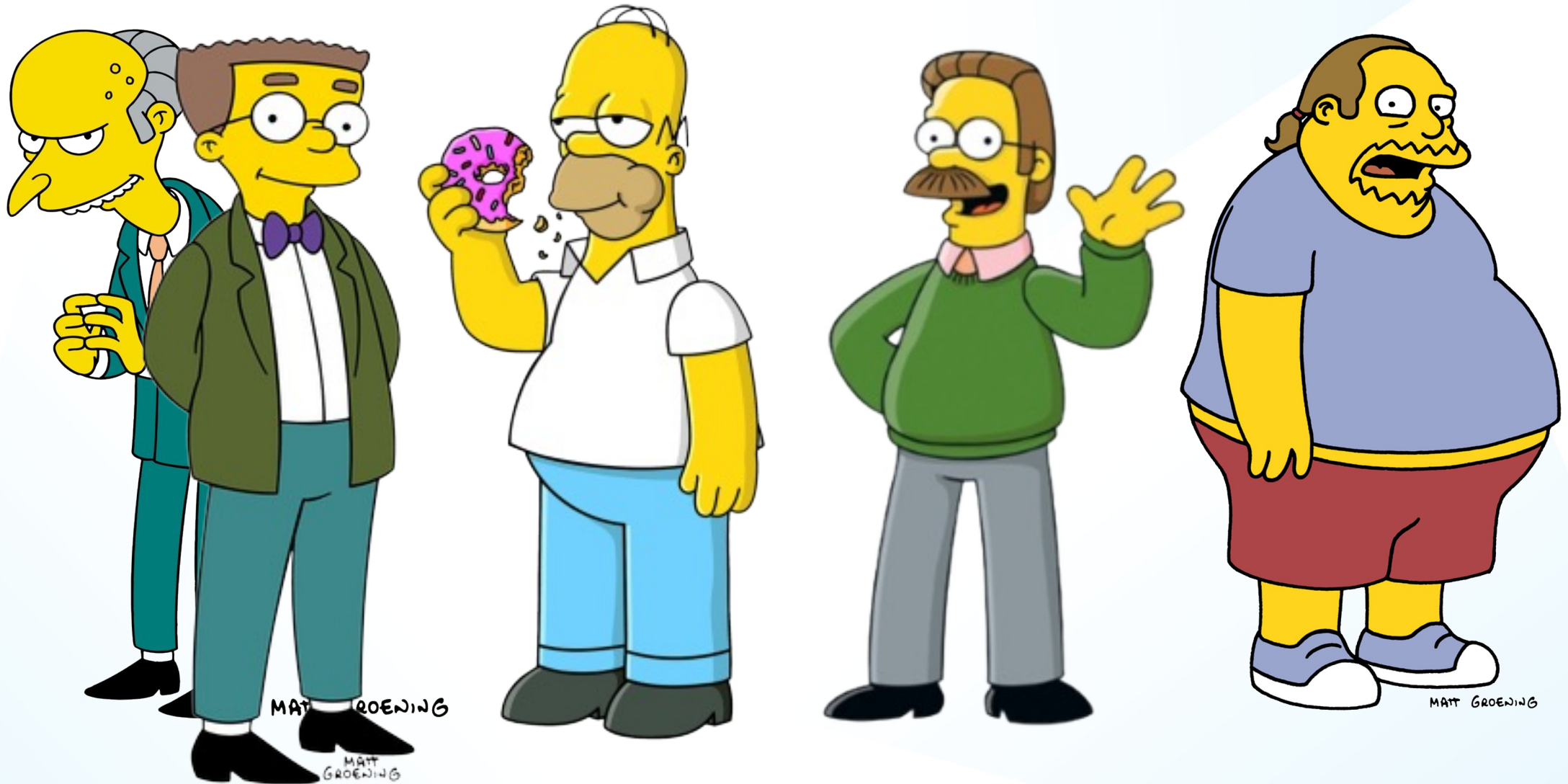
Ao invés de avaliar milhares de indivíduos, avaliamos grupos semelhantes

A photograph of a grocery store checkout area. In the foreground, a mesh bag filled with red apples sits on a metal scale. The bag has a green drawstring. In the background, there are cardboard boxes filled with more fruit, including yellow lemons or limes. A red sign is visible on the left side of the frame.

Supermercado

Ex1: Como são esses clients?

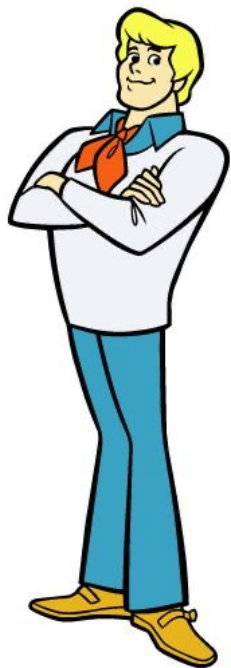
Redução de dimensionalidade



A top-down photograph of a red bowl filled with white popcorn, sitting on a dark wooden surface. A black TV remote control lies to the right of the bowl. Several popcorn kernels are scattered on the wood around the bowl and remote. The word "Streaming" is written in white, sans-serif font across the middle of the image, partially over the bowl and remote.

Streaming

Ex 2: Como são esses clientes?



Análise de agrupamentos

Classificação

Classificação dos algoritmos



Supervisionados

- Regressão
- GLM
- GLMM
- Support vector machines
- Naive Bayes
- K-nearest neighbors
- Redes Neurais
- Decision Trees



Não supervisionados

- K-Means
- Métodos hierárquicos
- Mistura Gaussiana
- DBScan
- Mini-Batch-K-Means

Classificação dos algoritmos



Métodos Machinelârnicos

- Árvores de decisão
- Bagging
- Boosting
- K-NN
- Redes Neurais
- Support Vector Machines

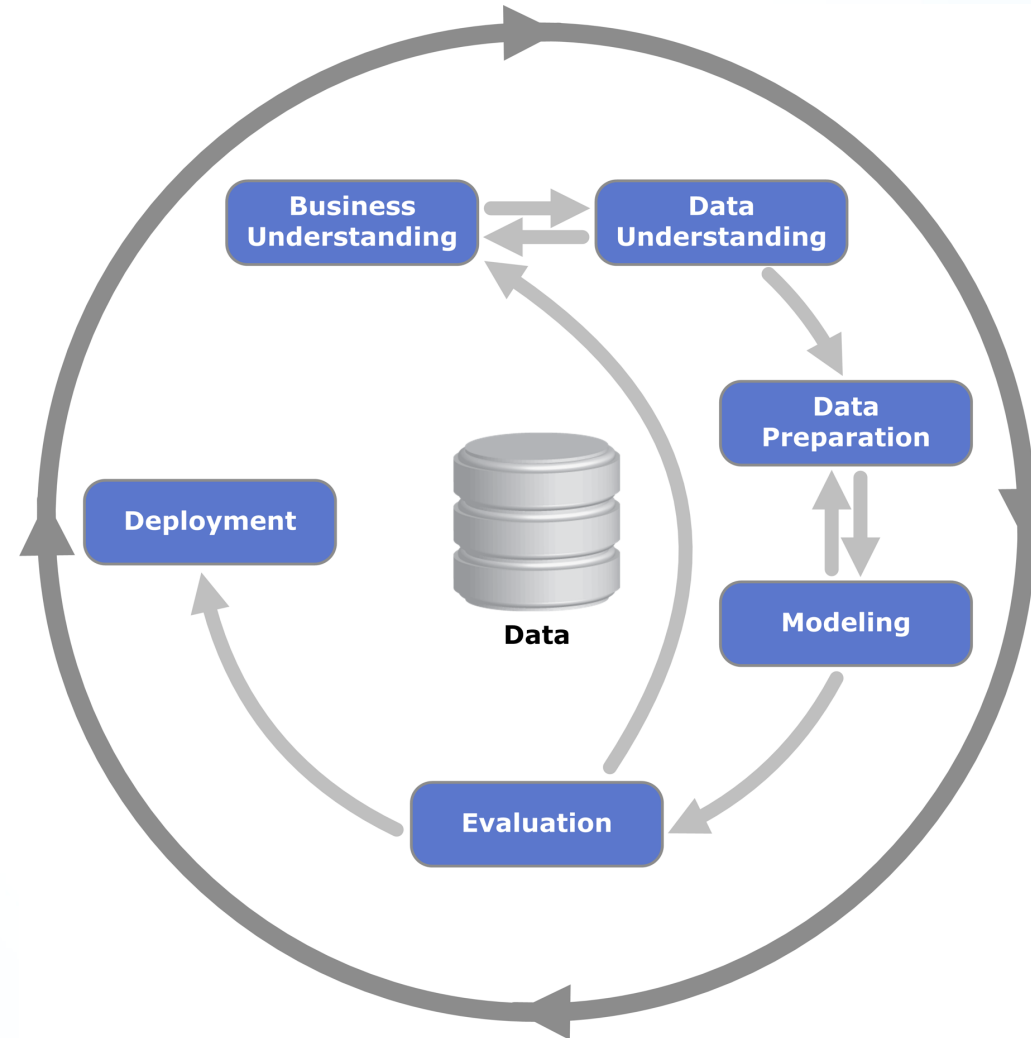


Métodos Estatísticos

- Regressão
- GLM
- GLMM
- ANOVA

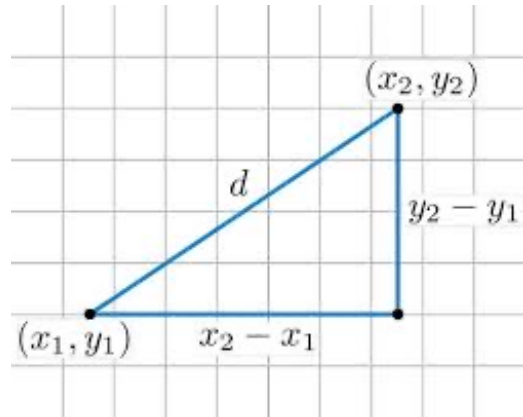
→ Estamos aqui!

CRISP-DM

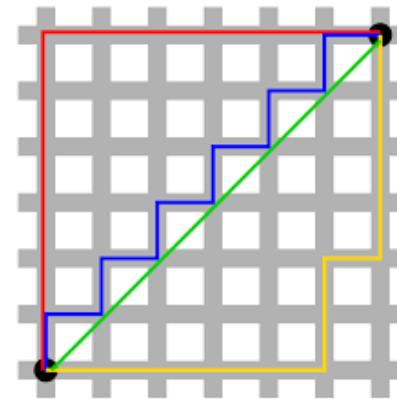


Fonte: <https://www.the-modeling-agency.com/crisp-dm.pdf>

Definições de distância

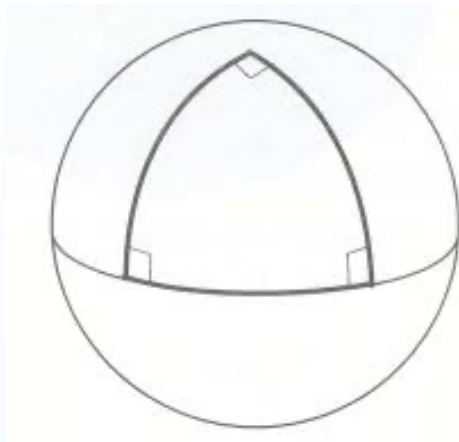


Distância Euclidiana



Distância de Manhattan

Distâncias Não Euclidiana



Distância de Mahalanobis

