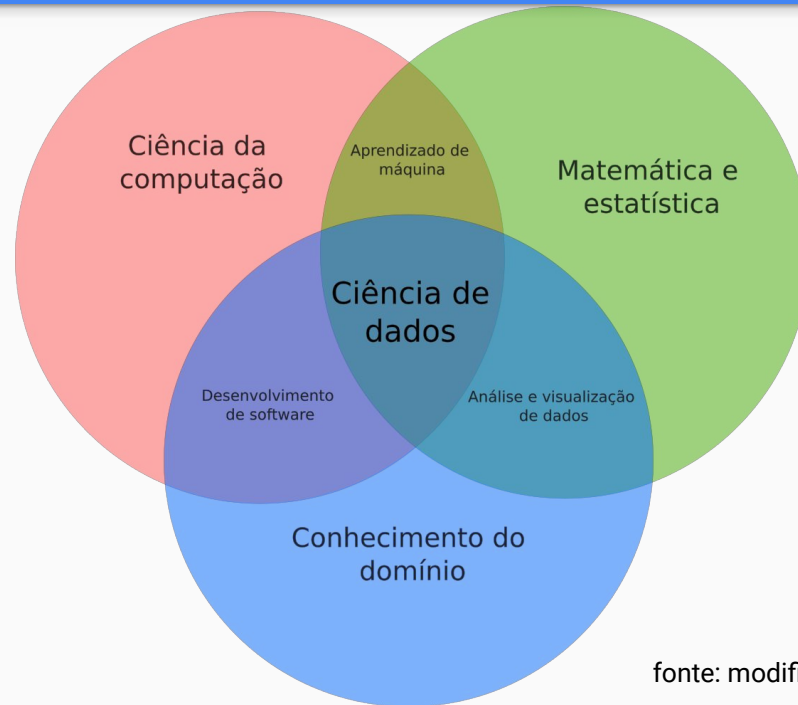


Dia 1 - Curso de IA e Mapeamento Preditivo

Introdução à ciência de dados e inteligência artificial

Victor S. dos Santos

Ciência de dados



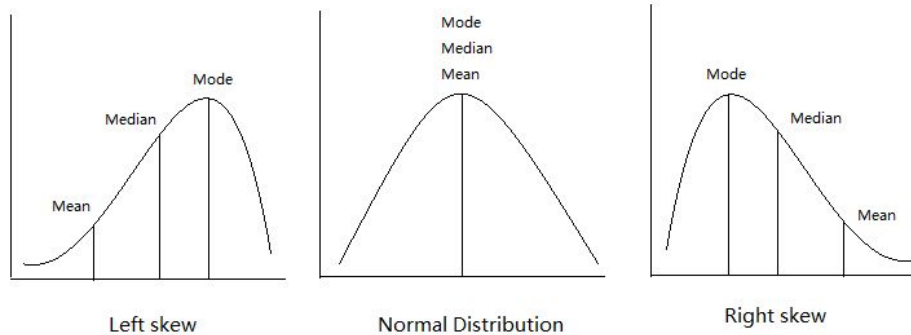
fonte: modificado de IA-IBM

Análise e visualização de dados: estatística descritiva

- **Análise univariada:** medidas centrais e de dispersão;
- **Análise bivariada:** gráficos de dispersão, medidas de dependência (correlação etc);
- **Análise multivariada:** teoria da informação e agrupamento.

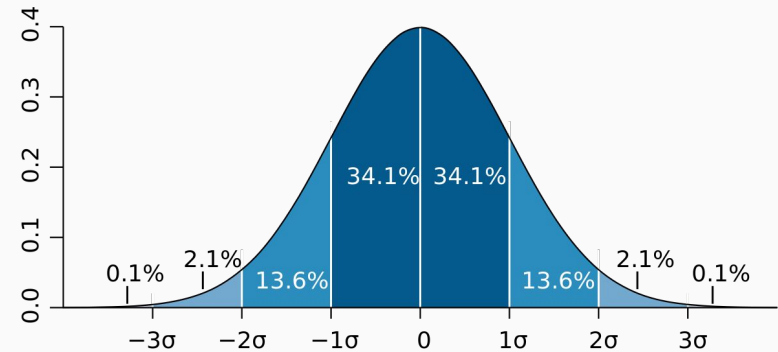
Tendência central e dispersão

Média, mediana e moda



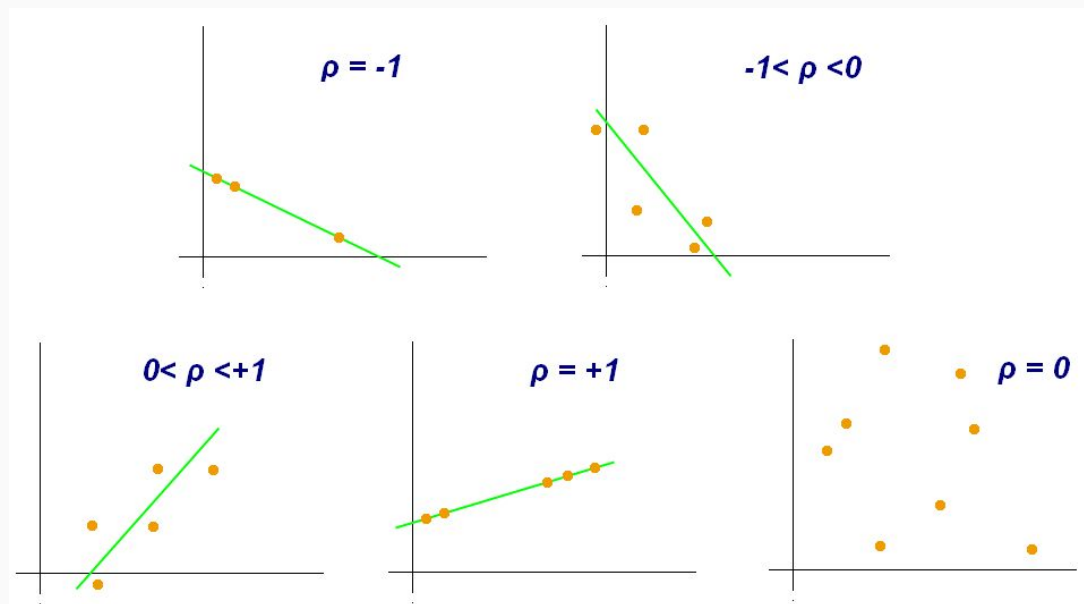
fonte: medium (Nhan Tran)

Desvio padrão



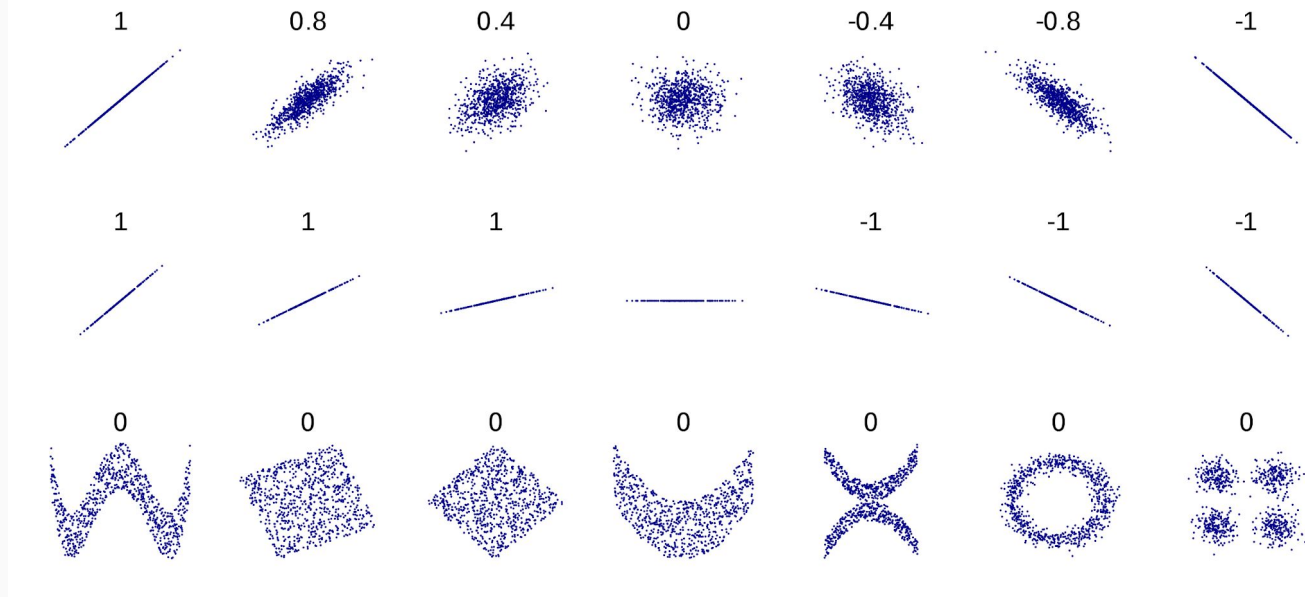
fonte: wikipedia

Gráficos de dispersão e medidas de dependência



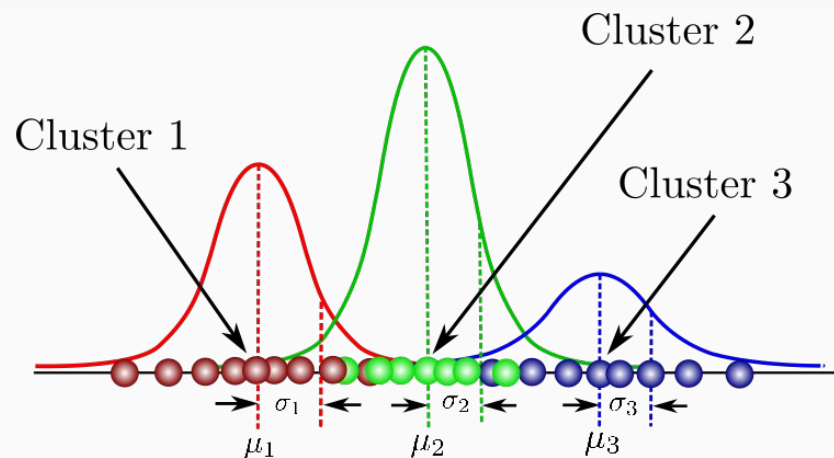
fonte: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient

Gráficos de dispersão e medidas de dependência



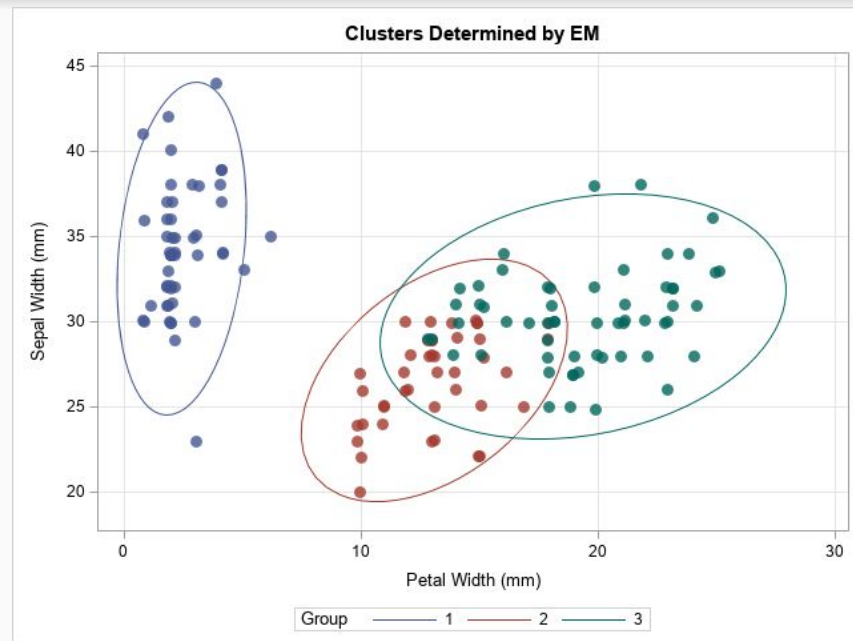
fonte: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient

Gaussian mixture



fonte: towards data science

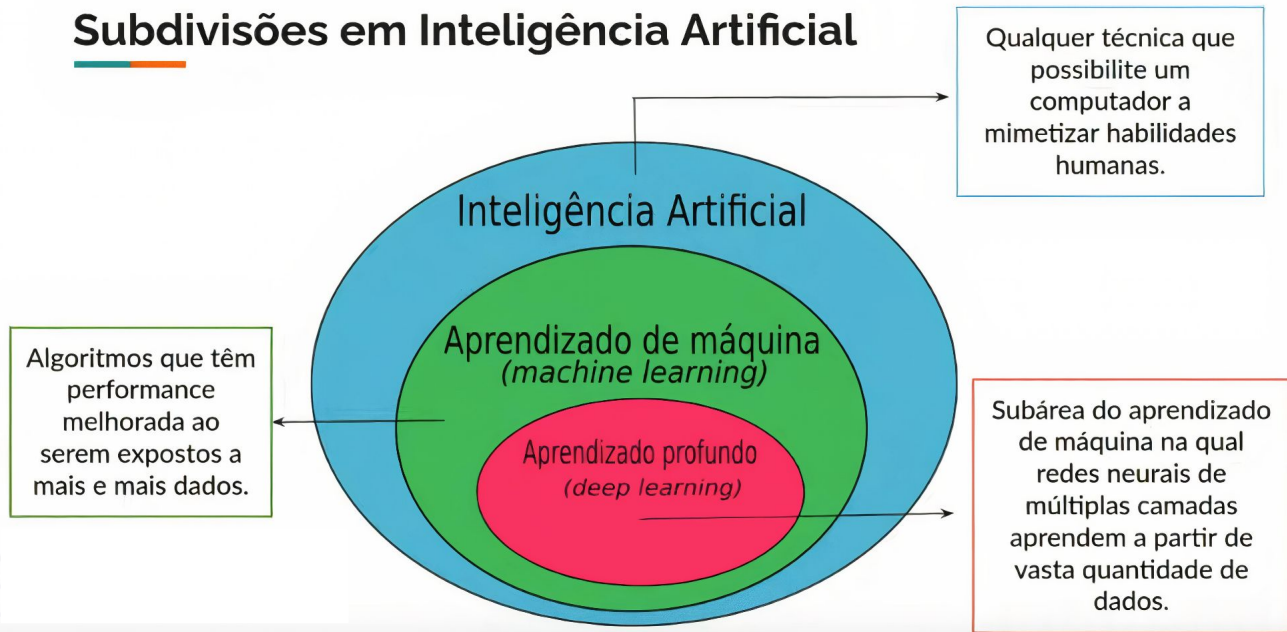
- lembrete: exemplo usando imagem de satélite.



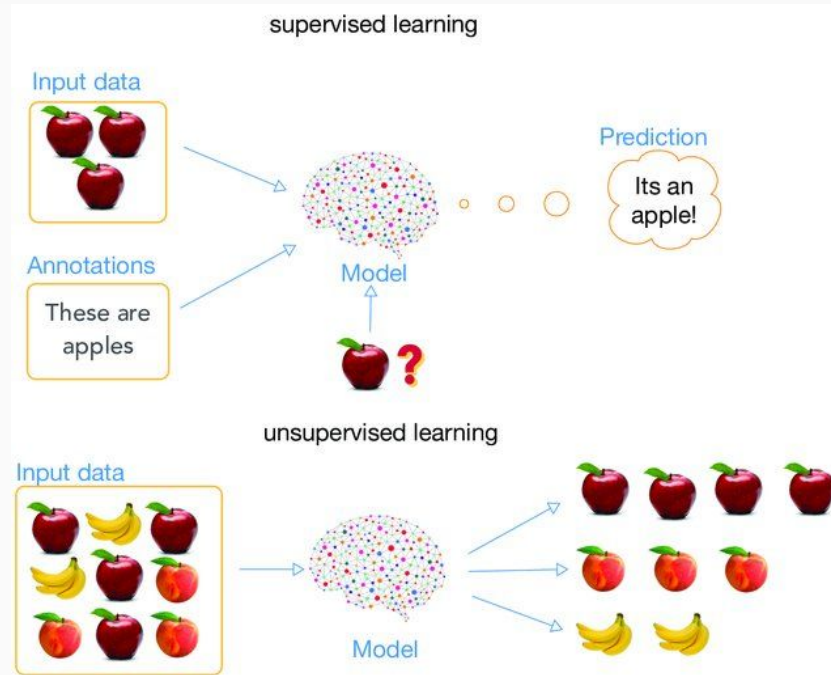
fonte: SAS blogs

Inteligência artificial

Subdivisões em Inteligência Artificial

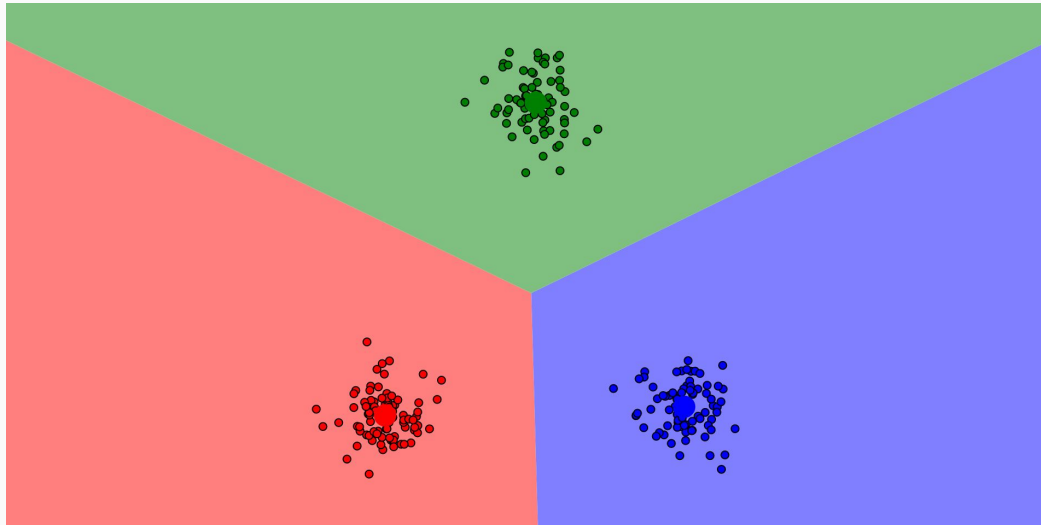


Métodos de machine learning



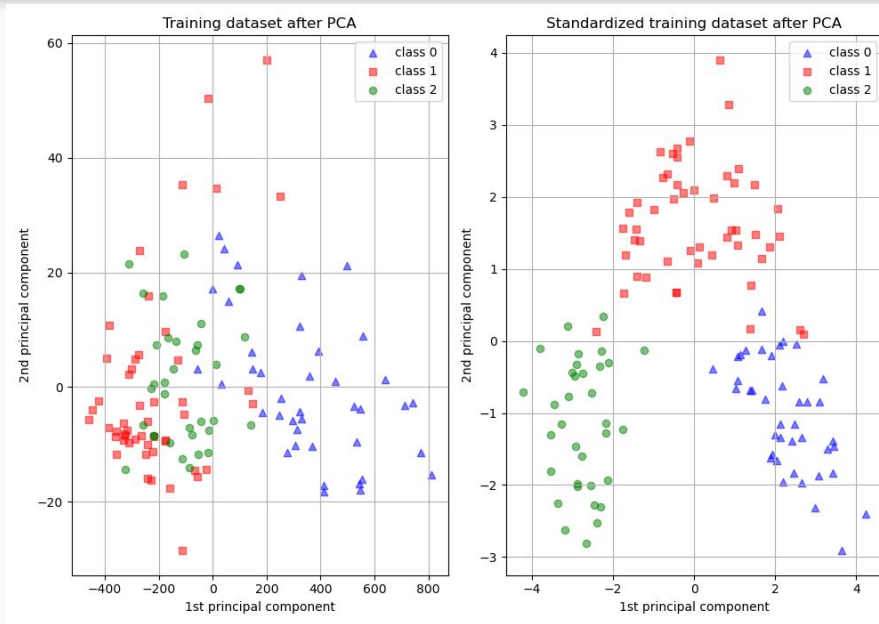
fonte: Ma, Y.; Liu, K.; Guan, Z.; Xu, X.; Qian, X.; Bao, H. Background Augmentation Generative Adversarial Networks (BAGANs): Effective Data Generation Based on GAN-Augmented 3D Synthesizing. *Symmetry* **2018**, *10*, 734.

K-means: buscando por agrupamentos



<https://www.naftaliharris.com/blog/visualizing-k-means-clustering/>

Rescaling



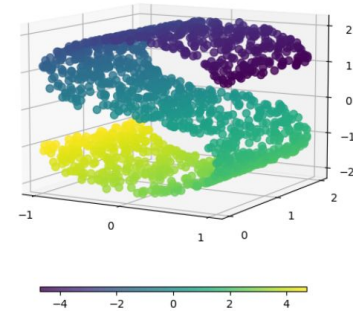
Manifold - redução de dimensionalidade



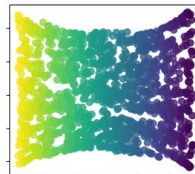
fonte: amazon

fonte:
sklearn

Original S-curve samples



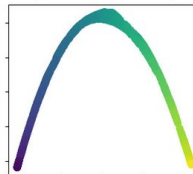
Isomap Embedding



Multidimensional scaling



Spectral Embedding



T-distributed Stochastic Neighbor Embedding

