CHALLENGE

Gaussian Mixture Model

The case study is about determining how Wine characteristics go together and figuring out the drivers behind higher quality wine

The Gaussian Mixture Model is a probabilistic clustering algorithm that will help us create segments for wine.

01 OPTIMAL NUMBER OF CLUSTERS

We need to determine the optimal number of clusters. The KPIs to help us in the task are the AIC and BIC, which find a balance between the goodness of fit and model complexity.
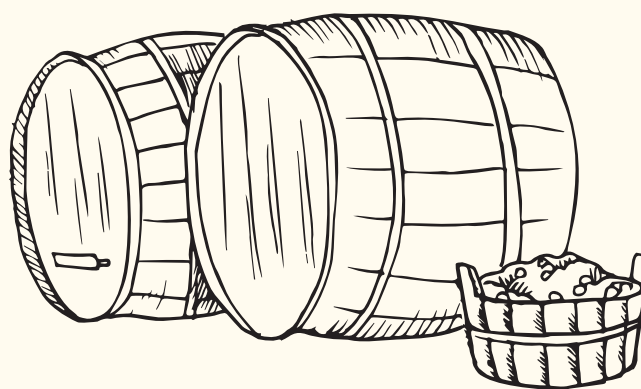
02 GAUSSIAN MIXTURE MODEL

Knowing the optimal number of clusters, we can easily create the Gaussian Mixture Model and segment the data into clusters. Remember that all inputs must be numberic.

03 INTERPRETATION

The most important part of any segmentation is the interpretation. If the segments do not make sense, then the application of the clusters is useless. If we cannot make a story, we need to go back to the drawing board and find new or even remove characteristics.

Diogo Resende