Description

How to determine enough sample size? Is it possible to find similarity score between two distributions? If yes, please elaborate.

Distribution similarity example: Suppose we want to find a similarity between two documents using their words frequency, then, one way to do that is by using the frequencies of each word to build a histogram for each document. Finally, we want to find which distribution (histogram) is similar to what. What will be the best solution of tool?

Date Created
2019/01/08
Author
seonetwork