

Jeu de données

Un jeu de données (en anglais dataset ou data set) est un ensemble de valeurs (ou données) où chaque valeur est associée à une variable (ou attribut) et à une observation. Une variable décrit l'ensemble des valeurs décrivant le même attribut et une observation contient l'ensemble des valeurs décrivant les attributs d'une unité (ou individu statistique)¹.

Structure d'un jeu de données

Un jeu de données peut avoir une structure tabulaire, par exemple un fichier CSV, une structure d'arbre, comme dans un fichier JSON ou XML....