Cluster Analysis

Table of Contents

* Required packages

knitr::opts\_chunk$set(echo = TRUE)  
#install.packages("dplyr","ade4","magrittr","cluster","factoextra","cluster.datasets")  
  
knitr::opts\_chunk$set(echo = TRUE)

# Definition of a distance

* A distance function or a metric on , is a function .
* A distance function must satisfy some required properties or axioms.
* There are three main axioms.
* A1. (identity of indiscernibles);
* A2. (symmetry);
* A3. (triangle inequality), where , and are all vectors of .
* We should use the term *dissimilarity* rather than *distance* when not all the three axioms A1-A3 are valid.
* Most of the time, we shall use, with some abuse of vocabulary, the term distance.