Analyse en composantes principales sur le jeu de données Decathlon

Axel-Cleris Gailloty

2023-03-05

# Présentation du jeu de données

Le tableau de données contient 41 lignes et 13 colonnes. Les colonnes 1 à 12 sont des variables continues: les dix premières colonnes correspondent aux performances des athlètes pour les dix épreuves du décathlon et les colonnes 11 et 12 correspondent respectivement au rang et au nombre de points obtenus. La dernière colonne est une variable qualitative correspondant au nom de la compétition (Jeux Olympiques de 2004 ou Décastar 2004).

Le tableau de données vient avec le package FactoMineR, pas besoin donc d’importer le jeu de données à partir d’un fichier. Pour le charger en mémoire il suffit d’importer le package FactoMineR.

library(FactoMineR)  
data(decathlon)  
head(decathlon)

|  | 100m | Long.jump | Shot.put | High.jump | 400m | 110m.hurdle | Discus | Pole.vault | Javeline | 1500m | Rank | Points | Competition |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SEBRLE | 11.04 | 7.58 | 14.83 | 2.07 | 49.81 | 14.69 | 43.75 | 5.02 | 63.19 | 291.7 | 1 | 8217 | Decastar |
| CLAY | 10.76 | 7.40 | 14.26 | 1.86 | 49.37 | 14.05 | 50.72 | 4.92 | 60.15 | 301.5 | 2 | 8122 | Decastar |
| KARPOV | 11.02 | 7.30 | 14.77 | 2.04 | 48.37 | 14.09 | 48.95 | 4.92 | 50.31 | 300.2 | 3 | 8099 | Decastar |
| BERNARD | 11.02 | 7.23 | 14.25 | 1.92 | 48.93 | 14.99 | 40.87 | 5.32 | 62.77 | 280.1 | 4 | 8067 | Decastar |
| YURKOV | 11.34 | 7.09 | 15.19 | 2.10 | 50.42 | 15.31 | 46.26 | 4.72 | 63.44 | 276.4 | 5 | 8036 | Decastar |
| WARNERS | 11.11 | 7.60 | 14.31 | 1.98 | 48.68 | 14.23 | 41.10 | 4.92 | 51.77 | 278.1 | 6 | 8030 | Decastar |