



Le fichier “**eaux.csv**” présente 57 marques d’eaux en bouteille décrites par les variables ci-après :

NOM : le nom complet de l’eau inscrit sur l’étiquette.

ACRO : Le nom en abrégé.

PAYS : le pays d’origine identifié par les lettres de l’immatriculation automobile officielle.

TYPE : à deux modalités : M pour eau minérale, S pour eau de source.

PG : à deux modalités : P pour eau plate, G pour eau gazeuse.

CA : ions calcium en mg/litre.

MG : ions magnésium en mg/litre.

NA : ions sodium en mg/litre.

K : ions potassium en mg/litre.

SUL : ions sulfates en mg/litre.

NO3 : ions nitrates en mg/litre.

HCO3 : ions carbonates en mg/litre.

CL : ions chlorures en mg/litre.

MOY : masse moyenne des ions en mg/litre.

- 1- Est-il pertinent, dans le cas présent, de recourir à une Analyse en Composantes Principales (ACP) pour réduire le nombre de variables ? Justifier.
- 2- Est-il nécessaire de centrer-réduire les données avant de procéder à une ACP ? Justifier.
- 3- Quelles sont les variables à exclure de l’ACP ? Pourquoi ?
- 4- Pour la suite du problème, la dernière variable (la masse moyenne des ions) sera considérée dans l’ACP comme “ variable supplémentaire”. Pouvez-vous argumenter ce choix ?
- 5- Faire une ACP. Quel pourcentage de variabilité est expliqué par les deux premières composantes ?
- 6- Combien de composantes suggériez-vous pour cette étude ? Justifiez.
- 7- Quelles sont les variables que la deuxième composante oppose ?
- 8- Quelle est la variable qui a la plus forte contribution dans la troisième composante.
- 9- Écrire un paragraphe (au plus 150 mots) pour résumer cette analyse.