

Exercice d'auto-évaluation - Session 4

Un groupe industriel commercialisant du café souhaite comparer les cafés provenant de différents sites de production à partir de leur profil de composition physico-chimique : teneurs en acide chlorogénique (CGA), caféine (Cafeine), matières grasses (Fat), trigonelline (Trigonelline) et matières sèches (DM).

Les données sont fournies dans le fichier `coffee.txt`. Un script R important ces données est également à disposition. Dans un premier temps, on se concentre sur la question suivante : les variations de taux de matières grasses dépendent-elles de celles des teneurs en trigonelline ?

1. Dans cette problématique, quelle est la variable réponse ? Quelle est la variable explicative ? Vous donnerez aussi la nature (quantitative ou catégorielle) de ces variables.
2. Proposez un modèle statistique adapté à cette question. Vous donnerez le nom courant de ce modèle.
3. Estimez les paramètres du modèle de la question 2.
4. Décrivez par un graphique la relation entre le taux de matières grasses et la teneur en trigonelline.
5. Quelle est la valeur du coefficient de corrélation linéaire entre le taux de matières grasses et la teneur en trigonelline ?