

Exercice d'auto-évaluation - Session 1

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 taux de matières sèches sont-ils différents selon le site de production ?

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. Quelle est la population concernée par cette problématique ?
3. Quelle est la taille de l'échantillon de données disponibles ?
4. Dans l'échantillon, quels sont les effectifs par site de production ?
5. Dans quel site de production le café a t'il le plus fort taux de matières sèches médian ?
6. Décrivez par un graphique les différences de répartition de taux de matières sèches par site de production.