



TD2

ING1 EC551 Statistiques descriptives

Galharret Jean-Michel
département MSC

https://galharret.github.io/WEBSITE/cours_ONIRIS.html

Exercice 1 : Lien entre deux variables quantitatives

1. Calculer la matrice de corrélation entre les différents types d'acide gras contenus dans les huiles d'olives (fonction ***cor***).
2. A l'aide du package ***corrplot*** que vous chargerez au préalable, représenter la matrice de corrélation sous la forme ci-dessous (fonction ***corrplot***). Quelles corrélations vous semblent intéressantes ?

`corrplot 0.92 loaded`

3. A l'aide du package ***psych*** que vous chargerez au préalable, reproduire le graphique ci-dessous (fonction ***panels.pairs***). Ecrire le lien qui existe entre les variables considérées.

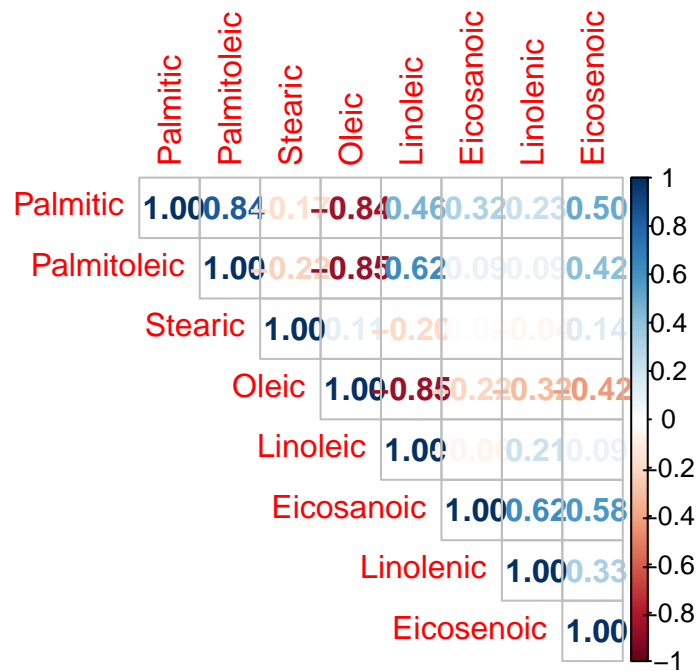


Figure 1: Matrice de corrélation

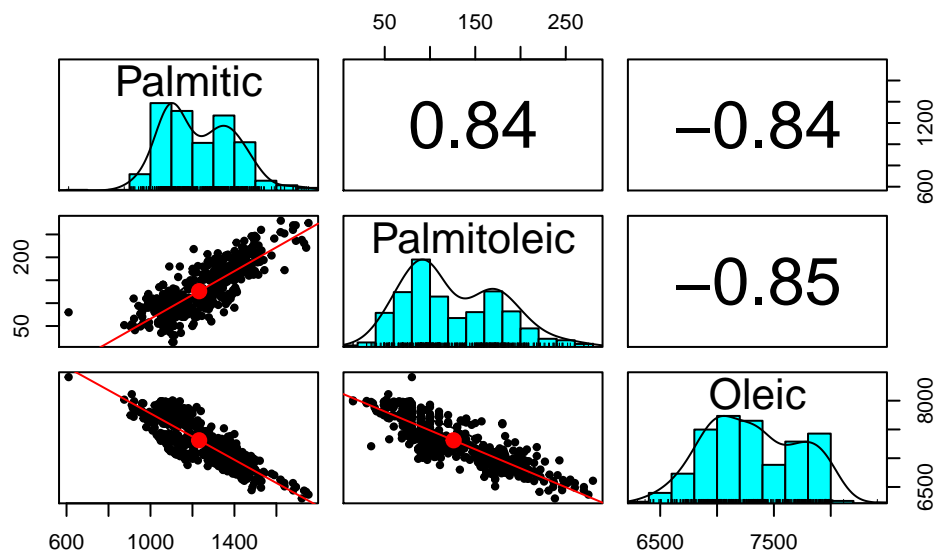


Figure 2: Graphe des corrélations

Exercice 2 : Lien entre 1 variable qualitative et une variable quantitative

1. Définissez sur R la fonction VAR telle que $Var(x) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$.
2. Calculer les moyennes et les variances de l'acide palmitique contenue dans les huiles en fonction de la région considérée (fonction **by**). Les résultats seront stockées dans deux vecteurs *Moy_Region* et *Var_Region* (pour transformer une liste en vecteur on utilisera la fonction **as.vector**).
3. En déduire quelle région a en moyenne le plus d'acide palmitique et dans quelle région les huiles sont le plus dispersées en termes d'acide palmitique.
4. En utilisant le cours (diapo 74) calculer les variances inter et intra de l'acide palmique selon la région considérée.
5. Quel est l'acide gras dont la présence est le plus lié à la région considérée ? On pourra utiliser la question précédente.

Exercice 3 : Lien entre deux variables qualitatives

Téléchargez et ouvrez sur Connect le fichier **mais.txt** :

1. Etablir la table de contingence des variables enracinements et couleur de l'épi. On ajoutera la somme marginale des colonnes

	Jaune	Jaune.rouge	Rouge	Somme
Faible	13	2	3	48
Fort	6	7	13	22
Moyen	17	3	8	29
Tres.fort	12	10	5	48

Etablir les profils de l'enracinement selon la couleur (fonction **proportions**).

2. Pour calculer la valeur du χ^2 on va utiliser la fonction **chisq.test**. A partir de l'aide déterminer la valeur correspondante.
3. En déduire V^2 (diapo 82).