

TD 1 - ANALYSE DE DONNÉES.

Exercice 1

1. Un enseignant a noté des copies sur 33 points. On lui demande de donner des notes sur 20, pour cela, il multiplie toutes les notes par un coefficient constant, de sorte qu'une copie ayant 33 sur 33 ait 20 sur 20. Que deviennent la moyenne et l'écart type des notes ?
2. Deux enseignants ont corrigé des copies. L'un mesure une moyenne de 7 et un écart type de 3.5, le second une moyenne de 12.5 et un écart type de 5. Quelle transformation affine doivent-ils appliquer aux notes pour obtenir une moyenne de 10 et un écart type de 2.5. Que devient pour chacun d'entre eux la note 20 ?

Exercice 2 Une entreprise est constituée de deux usines, appelées A et B . Le tableau suivant récapitule les salaires en euros par catégorie de personnel et par usine :

Usine A	Salaires	Effectifs	Usine B	Salaires	Effectifs
Ouvriers	700	200	Ouvriers	900	60
Employés	1400	20	Employés	1600	40
Cadres	5300	10	Cadres	7300	20

1. Calculer la moyenne des salaires dans chacune des usines, dans l'entreprise. Vérifier que la moyenne des salaires dans l'entreprise est la moyenne des salaires moyens de chaque usine.
2. Calculer la moyenne des salaires des ouvriers, puis des employés et enfin des cadres dans l'entreprise.
3. Calculer la variance des salaires dans chacune des usines et dans l'entreprise.
4. Vérifier que la variance des salaires dans l'entreprise est égale à la moyenne des variances des usines augmentée de la variance des moyennes calculées dans chaque usine. Quelle est la propriété du cours illustrée ici ?

Exercice 3 On étudie le tableau de notes de 6 élèves :

Français	Anglais	Maths	Sport
5.0	5.0	6.0	14.0
8.0	7.0	4.0	11.0
10.0	9.0	5.0	8.0
15.0	14.0	10.0	7.0
11.0	12.0	13.0	9.0
5.0	7.0	10.0	16.0

Dans un premier temps, la métrique utilisée est la métrique euclidienne standard.

1. Calculer le centre de gravité et le tableau des données centrées.
2. Calculer la matrice de variance/covariance des quatre variables ainsi que les coefficients de corrélation pour chaque paire de variables.
3. Calculer l'inertie du nuage des individus par rapport au centre de gravité.
4. Calculer les contributions des individus à l'inertie et les contributions des variables à la variance totale.
5. On utilise à présent la métrique donnée par l'inverse des variances de chaque variable. Que deviennent le tableau des données centrées, ainsi que la matrice de variance/covariance des quatre variables. Calculer le cosinus de l'angle formé entre les élèves 1 et 4 pour la métrique euclidienne.