

TD- ANALYSE EN COMPOSANTES MULTIPLES

On utilise des données les caractéristiques de 27 races de chiens au moyen de sept variables catégorisées :

taille, poids, vélocité, intelligence, affection, agressivité et fonction.

Les quatre premières variables ont trois modalités chacune (petite, 1 ; moyenne, 2 ; grande, 3), les deux suivantes, deux modalités (faible, 1 ; forte, 2), et la dernière, trois modalités (compagnie, 1 ; chasse, 2 ; utilité, 3).

Les données sont à télécharger sur votre ENT : le fichier excel : Race de chiens.xlsx

Par la suite, avec Xlstat, réaliser une analyse en composantes multiples

1. Quelle différence y a-t-il entre une analyse factorielles des correspondances (AFC) et une analyse des correspondances multiples (ACM) ?
2. Quelles sont les objectifs suivis par une ACM ?
3. Donner une description du contenu d'un tableau disjonctif complet (TDC), d'un tableau de Burt.
4. Qu'est-ce qu'un patron de réponse ? Combien y a-t-il de patrons possibles ?
5. Y a-t-il des individus identiques ? Est-ce un problème ?
6. a) Combien y a-t-il des modalités actives ?
b) Retrouver alors le nombre de valeurs propres non nulles
c) Combien vaut l'inertie totale ?
7. Combien de valeurs propres doit on retenir ? Quelle proportion de l'inertie est alors expliquée ?
8. a) Quelles sont les catégories qui déterminent le plus le premier plan principal ?
b) indiquer, en quoi consiste, en résumé l'axe 1. De même pour l'axe 2.
9. Citer 4 catégories qui sont particulièrement mal représentées par le premier plan principal. Conclure
10. Quels sont les individus qui contribuent le plus au premier plan principal ?