

Sujet de DM

1 Données

Nous utiliserons trois jeux de données :

1. **Données hobbies**. Les données sont disponibles en R (on utilise `data(hobbies)` pour les charger). Ces données sont issues du questionnaire sur les loisirs. 8403 individus ont répondu à 18 questions concernant leurs hobbies. Les caractéristiques des individus ont également été enregistrées (sexe, classe d'âge, statut marital, profession). Le nombre total d'hobbies de chaque personne est disponible.
2. **Données German Credit**. Les données contiennent les informations sur 1000 clients d'une banque en ce qui concerne leurs caractéristiques et leur capacité de payer les crédits. Les données se trouvent dans le fichier `germanCredit.RData` qui contient les variables quantitatives (`liste_var_quant`) et qualitatives (`liste_var_qual`). Il n'est pas nécessaire d'effectuer l'analyse sur l'ensemble de données ! Pour plus de détail consulter la fiche de TP "Classification".

2 Objectifs

L'objectif est d'explorer les jeux de données en utilisant les analyses multidimensionnelles adaptées à chaque problématique.

L'interprétation et le résumé des analyses doit en général contenir : la justification de la méthode utilisée, la qualité globale de l'analyse, l'interprétation des résultats (les axes factorielles, les classes, *etc.*), les caractéristiques de la structure de données (individus, variables), *etc.* Consulter le cours et les TPs pour se rappeler de la structure du rapport d'analyse.

3 Consignes

Le travail est individuel. Un rapport et un script R doivent être fournis à la fin. Le rapport doit être synthétique et contenir l'évaluation de la qualité des analyses et l'interprétation des résultats

1. Le rapport doit contenir une page de garde et 2 sections correspondant à 2 jeux de données analysés dans l'ordre : 1. Hobbies ; 2. German credit.

Chaque section du rapport doit contenir 3 pages maximum (graphiques et tableaux compris). Synthétiser l'information !

Le rapport doit être fourni au format *.pdf. Le nom du rapport doit avoir la forme suivante : `rapport_nom.pdf`, avec `nom` votre nom.

2. Le script R doit contenir un programme propre et commenté correspondant aux analyses effectuées et au rapport fourni.

Le script doit être fourni au format *.R. Le nom du script doit avoir la forme suivante : `script_nom.R`.

L'ensemble des résultats (pdf + R) doivent être déposés sur Moodle.

ATTENTION : le non respect des consignes sera pénalisé !