

Exercice 1. Chien ou loup ?

On a retrouvé un crâne fossile qui pourrait être celui d'un chien ou d'un loup. Différentes mesures ont été prises sur de nombreux crânes de loup et de chien. Les données se trouvent dans le fichier WOLF.csv, disponible sur la page web du cours. La dernière ligne concerne le crâne inconnu.

- (a) Charger les données en utilisant `wolf=read.csv("WOLF.csv")`. Que donne la fonction `plot(wolf)` ? Renseignez-vous sur la fonction `pairs`. Que donne la fonction `pairs(wolf[, -1])` ?
- (b) On ne voit pas grand chose dans le graphique précédent. On veut dessiner les données concernant le chien en bleu et celles concernant le loup en rouge. Pour cela définissez un vecteur `couleur` de même grandeur qu'une colonne du jeu de données qui prend la valeur "blue" (resp. "red") si la ligne correspond à un chien (resp. un loup).
- (c) Le crâne fossile appartenait-il à un chien ou à un loup ?

Pour votre information, les mesures correspondent à :

LCB	longueur condylo-basale
LSM	longuer de la mâchoire supérieure
LBM	largeur bi-maxillaire
LP	longueur de la carnassière supérieure
LM	longueur de la première molaire supérieure
LAM	largeur de la première molaire supérieure

Exercice 2. Triche.

Le rédacteur en chef d'un grand journal apprend que l'un de ses journalistes aurait demandé à un ami de rédiger un article à sa place durant le mois passé. Par chance il possède sur son ordinateur les 10 articles que le journaliste a transmis au journal avant qu'ils ne passent à la relecture et à la correction. Grâce à son logiciel, il peut immédiatement connaître le nombre de caractères pour chaque article, et de plus il sait combien de fautes de frappes ou d'orthographes ont été relevées par les correcteurs. Il résume tout ceci dans le tableau suivant :

sujet de l'article	nbre de caractères	nbre de fautes
sport / ski alpin	5459	42
people	2302	20
politique nationale	4012	43
reportage sur l'industrie agro-alimentaire	11202	101
économie	3192	22
interview	2989	21
sport / tennis	1509	14
critique cinéma	2498	65
politique régionale	3987	32
reportage au Mali	10782	91

Dessiner le nuage de points correspondant à ces données.

(Ecrire d'abord ces données dans un fichier `.txt`, les importer avec la fonction `read.table(...)` et utiliser la fonction `plot`).

Quel article n'a pas été écrit par le journaliste mais son ami ? (Motiver votre réponse.)

Exercice 3. On considère le jeu de données `survey` du package `MASS`, il contient les résultats d'une étude sur 237 étudiants de première années à l'université d'Adélaïde (Australie).

- (a) Plusieurs composantes de ce jeu de données sont des objets de la classe `factor`. Une de ces composantes, `Smoke`, contient les réponses à la question "Fumez-vous, si oui à quelle fréquence?". Que donne `table(survey$Smoke)` ?
- (b) Représenter graphique les réponses grâce à `barplot` et `pie`. Quel graphique vous paraît plus informatif ?
- (c) Faites la même chose avec les autres variables.

Exercice 4. Boîte à moustaches, Boxplot.

Durant la sance d'exercices, nous allons voir ce que sont les boîtes à moustaches et utiliser l'outil à l'exemple ci-dessous.

Les résultats ordonnés des examens de 10 élèves dans 7 cours différents sont résumés dans le tableau suivant.

Cours 1	Cours 2	Cours 3	Cours 4	Cours 5	Cours 6	Cours 7
4	42	19	33	21	14	18
12	44	23	47	24	75	19
23	46	25	59	27	76	21
35	47	27	67	29	77	23
46	49	31	69	77	78	24
52	51	43	73	79	79	25
67	54	48	75	83	80	27
75	56	51	77	85	81	29
83	57	63	83	86	83	30
92	58	73	85	87	84	93

- (a) Calculer (sans utiliser R) la médiane, le premier et le troisième quartile pour chacun des cours.
- (b) Utiliser R pour construire les boîtes à moustaches de chacun des cours (`boxplot(...)`) et comparer avec les valeurs obtenues en (a).
- (c) Peut-on dire que les résultats des différents cours sont similaires ?

À rendre jusqu'au vendredi 1er mars 2013, à 12h