

CC2 : Sujet Ours*Durée examen : 1H**Toute aide ou documents extérieurs à R sont interdits.**L'examen est à rendre sous forme d'un script R commenté et structuré. Une partie de la note tiendra compte du soin apporté au script (/2.5).*

Ce jeu de données correspond à un suivi d'ours brun dans plusieurs vallées alpines en Slovénie. Pour chaque ours capturé on indique sa vallée d'origine (identifiée par la première lettre de son nom), le nombre de prédation qui a été attribué à cet individu avant sa capture, son poids (en kg), son âge (en an), son sexe et si l'ours est en gestation ou non.

- 1) Estimez graphiquement si le poids des ours suit une distribution normale ou non. (/1)
- 2) Étudiez graphiquement et statistiquement la relation entre le poids des ours et le nombre de prédateurs qui leur est attribué. (/1.5)
- 3) Modélisez le poids des ours en fonction du nombre de prédateurs qui leur est attribué. Représentez graphiquement cette régression et analysez-en les sorties. (/2.5)
- 4) Codez par vous-mêmes la fonction permettant de calculer le coefficient de détermination d'une régression linéaire entre deux variables. Testez cette fonction sur votre précédent modèle. (/2)

Rappel:

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

e = résidus de la régression linéaire

y = valeurs de la variable réponse (vecteur numérique)

- 5) A l'aide d'une boucle, créer une variable représentant la classe d'âge des individus, avec quatre classes d'âge possibles: juvéniles (1 ou 2 ans), subadultes (3 ans), adultes (> 3 ans, < 20 ans), âgés (> 19 ans). (/2.5)
- 6) Étudiez statistiquement et graphiquement la relation entre la classe d'âge des individus (à l'aide de la variable nouvellement créée) et sexe des individus (pour le graphique: ordonnez temporellement les classes d'âges en abscisse à l'aide de la fonction "factor()" et de l'argument "levels ="). Que pouvez-vous en conclure (décrire et interprétez les résultats) ? (/3)
- 7) On cherche à identifier les vallées dans lesquelles la présence de l'ours est viable. Plusieurs critères entrent en compte pour identifier de telles vallées: il faut que le sex ratio (nombre de mâles / nombre d'individus total) dans la vallée soit équilibré (compris entre

0.4 et 0.6), qu'au moins un quart de la population soit en âge de se reproduire (classe d'âge "adultes"), qu'au moins une femelle gestante y ait été détecté, et qu'une condition physique moyenne supérieure à la moyenne de l'ensemble de la population soit observé. A l'aide d'une fonction de la famille "apply", testez chacune de ces conditions une à une, puis testez quelles populations sont viables. **(/5)**

Toutes les données de ce sujet ont été simulées aléatoirement.