

CC2 : Sujet Merlu*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.*

Ce jeu de données contient des informations sur les techniques (types d'engins de pêche utilisés, longueurs en mètres) et mode de pêche (zone géographique) utilisés lors d'événements de pêche au merlu. On a également accès à la quantité de poisson pêchés (en kg) et la présence ou non de captures accidentelles de dauphins (« Bycatch ») avec une indication sur le nombre d'individu capturés (« Nombre »).

- 1) Créer une nouvelle variable regroupant les zones économiques (ZEE) du nord de l'Europe ("ZEE_Ir" et "ZEE_Gb") et celles du sud de l'Europe ("ZEE_Fr" et "ZEE_Sp").
- 2) Donner le nombre d'événements de pêche avec ou sans capture par classe d'engin de pêche, respectivement pour les zones au nord de l'Europe et celles au sud de l'Europe. Faire une représentation graphique de l'impact du type d'engin de pêche sur le nombre d'événements avec ou sans capture accidentelle, pour chaque zone géographique (nord de l'Europe / sud de l'Europe). Que pouvez-vous en conclure ?
- 3) En utilisant une boucle estimez si, pour chaque engin de pêche, la distribution des longueurs d'engins suit une loi normale.
- 4) A l'aide d'une fonction de la famille "apply" comparez les moyennes des tailles d'engin en fonction des engins de pêche. Faites un graphique permettant de représenter les différences observées. Les tailles d'engin diffèrent-elles statistiquement entre engins de pêche ?
- 5) Représentez graphiquement l'évolution du nombre de dauphins capturés en fonction du poids de pêche. Faire une description statistique de cette relation. Que pouvez-vous en dire ?
- 6) Modélisez l'évolution du nombre de dauphin capturés en fonction du poids de pêche. Illustrez graphiquement le modèle et interprétez les sorties de ce modèle.
- 7) Décrivez graphiquement et statistiquement l'évolution du poids de pêche en fonction de la longueur de l'engin de pêche.
- 8) Tentez de modéliser l'évolution du poids de pêche en fonction de la longueur de l'engin de pêche à l'aide d'une régression linéaire. Que pouvez-vous en dire ?
- 9) Faites de même après une transformation cubique des données. Que pouvez vous en dire ?

10) Faites de même (en conservant la transformation cubique), en retirant les événements de pêche associés à des engins de pêche dont la longueur dépasse les 3800 mètres.

11) Créer une fonction dans R permettant de centrer-réduire une variable.

12) Comparez les résultats du dernier modèle effectué avec celui modélisant l'évolution du nombre de dauphins capturés en fonction du poids de pêche (adaptez les modèles si nécessaire). Que pouvez-vous dire sur les possibles mécanismes de causalité existant ?

13) Les lignes de code suivantes comportent 6 erreurs, corrigez les et expliquez ce que réalisent ces lignes.

```
for (i in unique(Suivi_Merlu$Engin_peche)) {  
  for ("j" in unique(Suivi_Merlu$Zone)) {  
    if (i == "Chaluts") {}  
    print(  
      elseif(mean(Suivi_Merlu[Suivi_Merlu$Engin_peche == i & Suivi_Merlu$Zone == j]  
$Longueur_engin) < 100,  
        paste("Dans la zone", j, "l'engin", i, "est généralement de petite taille", sep = " "),  
        paste("Dans la zone", j, "l'engin", i, "est généralement de grande taille", sep = " "))  
    )  
  }  
  else if (i = "Palangres") {  
    print(  
      elseif(mean(Suivi_Merlu[Suivi_Merlu$Engin_peche == i & Suivi_Merlu$Zone == j]  
$Longueur_engin) < 1500,  
        paste("Dans la zone", j, "l'engin", i, "est généralement de petite taille", sep = " "),  
        paste("Dans la zone", j, "l'engin", i, "est généralement de grande taille", sep = " "))  
    )  
  }  
  else (i == "Filets maillants") {  
    print(  
      elseif(mean(Suivi_Merlu[Suivi_Merlu$Engin_peche == i & Suivi_Merlu$Zone == j]  
$Longueur_engin) < 2000,  
        paste("Dans la zone", j, "l'engin", i, "est généralement de petite taille", sep = " "),  
        paste("Dans la zone", j, "l'engin", i, "est généralement de grande taille", sep = " "))  
    )  
  }  
}  
}
```

14) Expliquez ce que réalisent les fonction "any()" et "is.na()".

15) En utilisant une fonction de la famille "apply", déterminez pour chaque colonne du jeu de données si des valeurs manquantes sont présentes ou non dans la colonne.

16) En utilisant une fonction de la famille "apply" sur la colonne présentant des valeurs manquantes, produire un vecteur logique indiquant la présence ou non de valeur manquante pour chaque individu.