

Introduction aux logiciels de statistiques

Examen GIS 2A3 / 2019-2020 / Tous documents autorisés

Le but de l'examen est de vous permettre de montrer vos acquis. Le sujet est volontairement long. Il est donc conseillé de traiter en premier lieu les questions sur lesquelles vous vous sentez le plus à l'aise. Le langage R est le seul autorisé pour répondre au sujet. Vous répondrez à chaque question en utilisant au moins une ligne de code R et en formulant au moins une phrase de réponse. Recopier une valeur lue visuellement dans la table de données ou l'environnement pour répondre à une question ne rapporte aucun point. Une valeur de la table ne pourra être utilisée qu'à condition de l'avoir trouvée par du code. Pour répondre à une question, vous pouvez admettre le résultat de la question précédente à condition de le mentionner. À la fin de l'examen, vous enregistrerez votre travail dans un fichier zip à votre nom. Vous l'enverrez ensuite à Sofiane.Maazi@polytech-lille.fr. La clarté des commentaires et la qualité du code constitueront un élément important dans l'appréciation de votre travail. Pour chaque question, vous êtes invité(e) à faire figurer toute trace de recherche même infructueuse. Le barème est donné à titre indicatif.



Descriptif des variables de la table personnages

- **Id** : numéro d'identification du personnage
- **Name** : nom du personnage
- **Gender** : genre du personnage
- **Job** : statut du personnage
- **House** : maison (Gryffindor, Slytherin, Ravenclaw, Hufflepuff) à laquelle appartient le personnage
- **Patronus** : caractéristique du sortilège patronus du personnage
- **Species** : espèce à laquelle appartient le personnage
- **Hair_colour** : couleur des cheveux du personnage
- **Eye_colour** : couleur des yeux du personnage
- **Sorts** : nombre de sorts lancés par le personnage durant l'année
- **Points** : nombre de points gagnés par le personnage durant l'année

Descriptif des variables de la table potions

- **Id** : numéro d'identification du personnage
- **Potions** : nom de la potion favorite du personnage

Partie 1

1. Importez les 2 tables. Vous utiliserez le point virgule comme délimiteur.
2. Quel est le nombre de lignes et colonnes de la table `personnages.csv` ?
3. Combien de personnages sont des hommes ?
4. Combien de personnages sont des femmes de la maison Gryffindor ?
5. (a) Quel est le numéro i de la ligne de la table consacrée à Gilderoy Lockhart ?
(b) À qui sont associés les $i^{\text{ème}}$ éléments des vecteurs `Hair_colour` et `Eye_color` ?
(c) En déduire, les couleurs des yeux et cheveux de Gilderoy Lockhart.
(d) En déduire les noms des personnages qui ont les yeux de même couleur que Gilderoy Lockhart.
6. Ernest Macmillan et Dolores Jane Umbridge ont-ils le même patronus ?
7. Représentez graphiquement la variable `Sorts`.
8. L'étudiante qui a lancé le moins de sorts est-elle de la maison Slytherin ?
9. Combien de sorts ont été lancés en moyenne par l'ensemble des personnages ? Précisez l'écart-type corrigé et commentez.
10. Écrivez une fonction qui renvoie la moyenne harmonique pondérée d'une série statistique lorsque tous les poids valent 1.5. Testez votre fonction sur la variable `Sorts`.
11. Le Directeur de Poudlard remet la coupe des 4 maisons à la maison qui a obtenu le plus de points. Seuls les personnages qui sont exclusivement des étudiants et étudiantes ont pu gagner des points. À quelle maison le Directeur remettra-t-il la coupe ?
12. Ron Weasley achète dans une confiserie un paquet de dragées surprises de Bertie Crochu. Il y en a 20 au goût chocolat, 30 au goût menthe 10 au goût orange, 10 au goût épinard et 30 au goût foie et tripe.
(a) Modélisez le paquet à l'aide d'un vecteur de R.
(b) Harry Potter prend une poignée de 7 dragées dans le paquet de Ron. Il en donne ensuite 2 au hasard à Hermione Granger. Elle en remet alors 1 au hasard dans le paquet. Simulez le partage des bonbons entre les 3 amis. En particulier, vous afficherez les bonbons de chacun à l'issue du partage.
13. Le Choixpeau est un artéfact posé sur la tête de chacun des nouveaux élèves de Poudlard. Le Choixpeau détermine alors dans laquelle des 4 maisons sera affecté l'élève qu'il coiffe. Pour cela, le Choixpeau simule 4 valeurs d'une variable aléatoire X suivant une loi uniforme sur $[0;3]$. Il simule ensuite 4 valeurs d'une variable aléatoire Y suivant une loi géométrique de paramètre 0.9. Il prend la valeur maximale de $Z=X+Y$ et l'arrondit à l'unité. Si le nombre obtenu est impair alors l'élève est affecté à Slytherin sauf si le nombre est divisible par 3. Dans ce dernier cas, il est alors affecté à Ravenclaw. Si le nombre obtenu est pair alors l'élève est affecté à Hufflepuff sauf si le nombre est divisible par 7. Dans ce dernier cas, il est alors affecté à Gryffindor. Implémentez le programme utilisé par le Choixpeau et testez le.



Partie 2

1. (a) Quelles sont les potions qui sont les plus appréciées par les étudiantes de Ravenclaw ?
(b) Quelles sont les potions qui sont appréciées par au moins 2 étudiantes de Ravenclaw (qui ont lancé au plus 72 sorts) et 3 étudiantes de Gryffindor ?
2. Problème ouvert : Minerva McGonagall est-elle le professeur qui a lancé le moins de sorts ?



Barème

Partie 1 (19.5 points)

1. Numéros des questions notées sur 0.25 points : 1 - 2 - 5) a) - 5) b) - 5) c) - 5) d)
2. Numéros des questions notées sur 1 point : 3 - 4 - 6 - 7 - 8 - 9
3. Numéros des questions notées sur 2 points : 10 - 11 - 12) a)
4. Numéros des questions notées sur 3 points : 12) b) - 13

Partie 2 (5 points)

1. Numéro de la question notée sur 2 points : 1) a)
2. Numéro de la question notée sur 3 points : 1) b)

Source des images [en ligne]

- pngriver.com
- smithschoolwriters.files.wordpress.com
- vignette.wikia.nocookie.net
- www.hp-lexicon.org