

Introduction aux logiciels pour les statistiques





Bilan examen IS2A3 / 2020-2021

En raison du contexte sanitaire exceptionnel, les élèves ont eu toute latitude pour suivre le cours, les TP et l'examen en présentiel ou distanciel. Le sujet d'examen était inédit. Ont été évaluées les notions vues en cours et TP. Tous les documents étaient autorisés. Les élèves devaient utiliser au choix les langages R ou Python. Des questions d'un niveau plus avancé ont été posées. Toute trace de recherche raisonnée - même infructueuse - a été récompensée.

La note attribuée à cet enseignement est la somme des notes d'examen et du contrôle continu. Le sujet d'examen était d'un très bon niveau pour un cours d'introduction. Le sujet d'examen a été noté selon le barème en annexe. La somme des scores du QCM, des parties 1 et 2 a été ramenée sur 20 par règle de 3. Les points obtenus à la partie 3 ont été ajoutés à la note sur 20. Cette partie a été considérée in fine comme du bonus. Eu égard aux prises d'initiatives qu'elle appelait. La note de l'examen a été arrondie au demi-point inférieur ou supérieur le plus proche. A ensuite été ajoutée la note du contrôle continu. Le contrôle continu avait pris la forme d'un QCM noté sur 3 points. La moyenne des notes attribuée à cet enseignement s'élève à 15.56/20.



Quelques remarques :

-  Il est fortement conseillé de bien lire l'intégralité du sujet avant de répondre aux questions. Les annexes précisait les tables à utiliser pour chaque partie du sujet.
-  Outre le QCM, le sujet était composé de 3 parties. La 1^{ère} était consacrée à la construction d'un data.frame à partir de la table *disney*. Il était proposé notamment de manipuler des chaînes de caractères et écrire des fonctions. La 2^{ème} partie portait sur l'analyse du data.frame construit dans la 1^{ère} partie. Il était possible de répondre aux questions de la partie 2 sans avoir nécessairement réussi la partie 1. Les élèves ont eu à réaliser des simulations pour retrouver une probabilité conditionnelle théorique. La 3^{ème} partie du sujet portait sur l'étude du chiffre d'affaires de l'attraction.
-  L'utilisation d'Internet n'est pas interdite. Gardons à l'esprit que chercher de la documentation en ligne - en R ou Python - s'avère souvent très pratique quand on écrit un script. Cette démarche évite notamment d'implémenter des fonctions déjà existantes.
-  Une fonction retourne un objet utilisable ultérieurement dans le code. À l'inverse, une procédure ne retourne pas de résultat. Elle sert entre autres à modifier l'état d'un objet ou afficher un message (cf Memento R du cours).



- ✎ Les fonctions *Vert*, *Jaune* et *Bleu* devaient prendre en paramètre un nom de personnage - pas nécessairement dans les tables *disney* et *data* - et retourner le nombre de points associé.
- ✎ La question 5) de la partie 1 est un exemple de vectorisation impossible.
- ✎ On attendait aux questions 2) et 3) de la partie 2 la probabilité conditionnelle théorique que Merlin perde et non pas qu'il gagne. Utiliser la fonction *unlist()* s'avère ici inutile. Cette fonction appliquée sur une liste renvoie un seul vecteur contenant tous les éléments de la liste.
- ✎ Pensons à l'usuelle fonction *sample()* et son paramètre *replace* dès lors qu'on souhaite effectuer des tirages avec ou sans remise. Son utilisation était attendue aux questions 4)b) de la partie 2 et 2)a) de la partie 3.
- ✎ Plusieurs enseignements pouvaient être tirés à la question 1) de la partie 3. Citons par exemple l'augmentation du chiffres d'affaires (CA) de la boutique quand l'attraction est en déficit. Citons de même l'augmentation du CA de l'attraction quand la boutique est en déficit. Nonobstant ce « paradoxe », le parc dégage sur chacune des périodes de 2018 un excédent supérieur à celui de 2017 quand l'attraction n'existait pas dans le parc. On pouvait aussi analyser les périodes de fortes hausses des CA qui coïncidaient - ou pas - avec les périodes de fortes affluences du parc (Noël, Halloween, vacances estivales...). D'autres pistes ont été explorées par les élèves. Rappelons alors que les graphiques et les éléments d'explications sont incontournables pour un statisticien.



Barème

QCM (1 point)

1. Numéros des questions notées sur 0.25 points : 1) - 2) - 3) - 4)

Partie 1 (7 points)

1. Numéros des questions notées sur 0.5 point : 1) - 5)
2. Numéros des questions notées sur 1 point : 2) - 3) - 4) - 7)
3. Numéros de la question notée sur 2 points : 6)

Partie 2 (7.5 points)

1. Numéro des questions notées sur 0.25 points : 1) - 4)c) - 5)b)
2. Numéro de la question notée sur 0.75 points : 5)a)
3. Numéro des questions notées sur 1 point : 2) - 4)a)
4. Numéro des questions notées sur 2 points : 3) - 4)b)

Partie 3 (4.5 points)

1. Numéro des questions notées sur 0.5 points : 2)c) - 2)d)
2. Numéro des questions notées sur 1 point : 2a) - 2)b)
3. Numéro de la question notée sur 1.5 points : 1)

Source des images [en ligne] : www.grand-mercredi.com