|  |  |
| --- | --- |
|  |  |

# Quelques questions sur vos compétences actuelles et sur vos attentes

En premier lieu, nous tenons à vous remercier pour la confiance que vous nous accordez en vous inscrivant à la formation intitulée « Analyse NGS avec R ».

Cette formation nécessite une maîtrise suffisante du langage R. Afin de vous situer et de vous aider à évaluer votre niveau, nous vous invitons à effectuer le test ci-dessous et de le renvoyer à M. Aurélien Barré. Le cas échéant, si votre niveau n’est pas suffisant pour pouvoir suivre pleinement cette formation, nous vous inviterons, si ce n’est déjà fait, à vous inscrire à la formation « Langage R : introduction », qui précède ce stage.

Coordonnées : [aurelien.barre@u-bordeaux.fr](mailto:aurelien.barre@u-bordeaux.fr)

Dans le cas où vous votre niveau de maîtrise de R ne serait pas suffisant, nous vous invitons à suivre la formation « Langage R : initiation » programmée deux jours plus tôt.

1. *Quelle est la valeur retournée par les 4 séries d’instructions suivantes ?   
    (Sans les exécuter dans le terminal R)*

sum((1:5) \*2)

mean(c(4,5,6))

log(base=exp(1),x=3)

x<-2; 2x<-2\*x

1. *Quel est le but de cette fonction ?*

f <- function(x) {

odd.even <- length(x) %% 2

if (odd.even == 0)

(sort(x)[length(x)/2] + sort(x)[1+length(x)/2]) / 2

else

sort(x)[ceiling(length(x)/2)]

}

 La suite des nombres impairs jusqu'à une certaine valeur x.

 La médiane de la série d'observations contenues dans la variable passée en argument de la fonction f.

|  |
| --- |
|  Le nombre de valeurs ne dépassant pas la valeur médiane de la série d'observations contenues dans la variable passée en argument de la fonction f. |

1. *Quel est le rôle de ces instructions et quelle doit être la nature de x ?*

which(x<3)

na.omit(x)

1. *Quel est le nombre de ligne et de colonne du dataframe retourné par cette fonction si le fichier XXX est au format tabulé*

read.table(file="XXX,"header=T,sep=",",row.names=1,dec=".")

1. *Que va retourner cette suite d’instruction ?*

vecteur\_1 <- c(1,2,6,8,3,4,0,9,1)

vecteur\_2 <- c(6,8,3,6,8,4,0,9,1)

(sort(unique(vecteur\_1)) + vecteur\_2)[2]

1. *Pourquoi la première instruction fonctionne mais pas la seconde ?*

C(1:20)+C(0,1)

C(1:20)+C(0,1,2)

1. *Quel type d’objet est créé par cette fonction et à quel package est-elle associée ?*

qplot(data=mydata,x=length,y=width,geom="boxplot")

1. *Quelle commande doit-on utiliser pour transformer le vecteur grp, codé sous forme de facteurs, en une variable qualitative à 3 niveaux ordonnés où : A = "négatif", B et C = "neutre", D = "positif" (négatif < neutre < positif) ?*

> grp

[1] A A B B C C D D A A B B C C D D

> str(grp)

Factor w/ 4 levels "A","B","C","D": 1 1 2 2 3 3 4 4 1 1 ...

Levels: A B C D

 grp <- as.ordered(grp, levels=c("négatif", "neutre", "neutre", "positif"))

 levels(grp) <- c("négatif", "neutre", "neutre", "positif"); grp <- factor(grp, ordered=TRUE)

 levels(grp)[2:3] <- "neutre"; grp <- ordered(grp, levels=c(1,3), labels=c("négatif", "positif"))

1. *Que renvoi cette commande suivante ?*

mean(replicate(500,mean(rnorm(100, 69.5, 2.9))-mean(rnorm(100,63.9, 2.7))))

 La moyenne des différences de deux échantillons d'observations de taille 500 dont les valeurs ont été tirées aléatoirement dans des lois normales de moyennes 69.5 et 63.9, et d'écart-types 2.9 et 2.7, respectivement.

 La moyenne des différences de moyennes de deux échantillons d'observations de taille 500 dont les valeurs ont été tirées aléatoirement dans des lois normales de moyennes 69.5 et 63.9, et d'écart-types 2.9 et 2.7, respectivement.

 La moyenne des différences de moyennes de deux échantillons d'observations de taille 100 dont les valeurs ont été tirées aléatoirement dans des lois normales de moyennes 69.5 et 63.9, et d'écart-types 2.9 et 2.7, respectivement.

1. *En supposant que x1 soit codé sous la forme d'un facteur, que produit la commande suivante ?*

tapply(y, x1, function(x) c(sum(!is.na(x)), mean(x)))

  La moyenne des valeurs non manquantes de y par niveau de la variable x1.

 La moyenne et le nombre de données non manquantes pour la variable y par niveau de x1.

 La moyenne et le nombre de données manquantes pour la variable y par niveau de x1.

|  |
| --- |
|  |

*« Ce questionnaire est facultatif. Le destinataire de ces données est le responsable scientifique de la formation à laquelle vous êtes inscrit(e) ou préinscrit(e) ; elles lui permettront d’adapter au mieux le programme de la formation à vos attentes et à celles des autres stagiaires. Elles seront détruites à la fin du stage. Pour plus d’informations concernant l’utilisation de vos données personnelles, veuillez consulter la notice « Données personnelles » disponible sur notre site internet. »*