









# TP 1 ☐ Présentation générale du logiciel , l'aide en ligne et les bases du langage

 est distribué gratuitement à partir du site du CRAN (Comprehensive R Archive Network) ☐ <http://www.r-project.org/>. Je vous suggère d'aller visiter le site officiel de .

 est un système d'analyse statistique créé par Ross Ihaka et Robert Gentleman distribué librement sous les termes de la GNU General Public Licence ☐ son développement et sa distribution sont assurés par plusieurs statisticiens rassemblés dans le R Development Core Team. Pour avoir une liste des chercheurs qui ont participé au développement de 

Tapez sur la fenêtre de commande `contributors()`


Toutes les fonctions de  sont stockées dans une grande bibliothèque, Cette bibliothèque contient des packages de fonctions. Allons donc voir les packages disponibles à l'installation du logiciel .


Tapez la commande `installed.packages()`

Notons les différences de «`priority`» entre les différents packages

Quel impact à la `priority` sur l'utilisation du package ☐

A quoi sert la fonction `library()` ☐ La fonction `library()` sert-elle à tous les packages ☐

Le package nommé «`base`» est le cœur de  et contient les fonctions de base du langage pour la lecture et la manipulation des données, la création de certains types de graphiques et certaines analyses statistiques.

Par ailleurs, de nombreux packages, développés par les utilisateurs de , sont disponibles sur le site internet du CRAN et téléchargeable gratuitement à l'adresse suivante ☐ <http://cran.us.r-project.org/src/contrib/PACKAGES.html>. Sur cette page est répertoriée la liste des packages disponibles. Allez donc la visiter....



### Création d'objets basiques

Affecter la valeur 27 à l'objet nommé  $x$ .  
Affecter la valeur 9 à l'objet nommé  $X$ .  
Visualiser les valeurs de  $x$  et  $X$   
Que constate t'on?

Affecter la valeur 5 à l'objet nommé  $x$ . Visualiser la valeur de  $x$   
Affecter la valeur 9 à l'objet nommé  $x$ . Visualiser la valeur de  $x$   
Que constate t'on?

On peut affecter, dans une même commande, des valeurs à plusieurs objets. Il suffit de séparer les affectations par des `;`.

Créer les variables `w = 8`, `name = "Arthur"` et `dicton = "Vive le logiciel R"` sur une ligne de commande.

Gestion de la mémoire les fonctions `ls()` et `rm()`

Consulter les objets en mémoire



Afin de connaître les objets chargés en mémoire on utilise la commande ls().

Stockez les 5 valeurs "aujourd'hui, nous sommes 22 étudiants, bienvenue!!!" "À vos claviers", "la séance risque d'être rude", "je suis là pour vous aider" et 22 dans, respectivement, les objets bienvenue, dicton, attention, aide et nombre.  
Affichez les objets en mémoire.

Affichez les objets dont le nom contient la lettre «i»

Affichez les objets dont le nom commence par la lettre «a»

Affichez la liste détaillée des objets en mémoire

### Effacer les objets en mémoire

Effacer tous les objets en mémoire

Créer 12 objets un = 1, deux = 2, trois = 3, quatre = 4, cinq = "cinq", six = 6, sept = "sept", huit = 8, neuf = "neuf", dix = 10, onze = 11 et douze = 12.

Effacer les objets dont le nom contient la lettre «i»

Effacer les objets dont le nom commence par la lettre «a»



## L'aide en ligne

fournit 2 outils d'aide en ligne : «`?>`» et «`help.search`». Précisons que «`?>`» recherche les fonctions uniquement dans les packages chargés en mémoire tandis que «`help.search`» recherche les fonctions dans tous les packages installés dans l'ordinateur.

Remarque 1 est un logiciel anglophone si bien que la requête d'une recherche est préférable en anglais.

Remarque 2 Une fois trouvées les fonctions associées aux différentes analyses, il sera intéressant d'exécuter les exemples joints au fiche détaillée (non pour comprendre les résultats mais simplement pour se familiariser avec les affichages, les sorties, ...)

Quelle est la fonction qui permet de calculer la moyenne d'un vecteur?

Quelle est la fonction qui permet de calculer la variance d'un vecteur?

Quelle est la fonction qui permet de calculer la médiane d'un vecteur?

Quelle est la fonction qui permet de générer un vecteur aléatoire de loi de poisson?

Quelle est la fonction qui permet de générer un vecteur aléatoire de loi normale?

Quelle est la fonction qui permet de calculer le déterminant d'une matrice?

Quelle est la fonction qui permet d'effectuer une analyse de la variance?

Quelle est la fonction qui permet d'effectuer une analyse en composante principale?

Quelle est la fonction qui permet d'effectuer une régression linéaire?

Quelle est la fonction qui permet d'effectuer une analyse discriminante?

Quelle est la fonction qui permet d'effectuer un clustering de type kmeans?

Quelle est la fonction qui permet d'effectuer une régression logistique?



## Création d'objets Les vecteurs, les matrices, les data.frames et les listes

### Les vecteurs vector()

Créer le vecteur x composé de 5000 zéros  
Créer le vecteur x composé de 5000 caractères.  
Créer le vecteur x suivant

```
x  
[1] 1.3 2.0 5.2 4.3 2.2
```

Créer le vecteur suivant

```
x  
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29  
[30] 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55  
[56] 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81  
[82] 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
```

Indication  Il n'est pas obligatoire de rentrer les éléments du vecteur à la main )

Créer un vecteur à 20 éléments, nommé xnorm, composées de nombre aléatoire extrait d'une loi normale de moyenne 3 et de variance 1.

Afficher la valeur du cinquième élément de xnorm.

Afficher la valeur des 5 derniers éléments de xnorm.

Afficher la valeur des éléments d'indices impairs

Afficher la valeur des éléments 1, 4, 8, 12 et 18

Affecter la valeur 0 aux éléments inférieurs à 3 et la valeur 1 aux éléments supérieurs à 3



## Les matrices `matrix()`

Regarder la fiche détaillée de `matrix`

Créer une la matrice identité de dimension  $5 \times 5$

Créer une matrice  $A$ , de dimension  $5 \times 5$  dont les éléments sont des nombres aléatoires extraits d'une loi normale de moyenne 3 et de variance 10

Afficher la dimension de la matrice  $A$

Afficher les éléments de  $A$  supérieurs à 3

Afficher le nombre d'éléments de  $A$  supérieurs à 3

Remplacer les éléments de  $A$  supérieurs à 3 par 1 et les éléments de  $A$  inférieurs à 3 par 0

Créer les matrices suivantes

```
B = matrix(1:25, 5, 5) et C = matrix(1:25, 5, 5, byrow = TRUE)
```

Que constatez-vous

Nommer les colonnes de  $B$  par  $C1, C2, C3, C4$  et  $C5$  et les lignes de  $B$  par  $L1, L2, L3, L4$  et  $L5$

Afficher les valeurs de la troisième colonne de  $B$  par le système d'indexation

Afficher les valeurs de la première et cinquième ligne de  $B$  par le système d'indexation



Afficher les valeurs de la première colonne de  $B$  par le nom. L'accès aux données par le nom fonctionne-t-il dans le cadre des matrices ?

Convertir la matrice  $B$  en une `data.frame` nommé  $C$

### Les variables qualitatives nominales `factor()`

La fonction `factor()` crée des variables qualitatives nominales.

Créer la variable nominale, `tignasse`, composée des 10 valeurs suivantes :  
blond, chatain, brun, brun, roux, blond, brun, chatain, brun, blond

### Les `data.frame` `data.frame()`

Regarder la fiche détaillée de `data.frame`

Accéder aux valeurs de la troisième colonne de la `data.frame`  $C$  par le nom et par l'indexation

Créer une `data.frame`  $D$  dont la première colonne est composée des chiffres de 1 à 6 et la deuxième colonne des 6 premières lettres de l'alphabet.  
Quel est le mode de la deuxième colonne de  $D$  ?

Créer une `data.frame`  $D$  dont la première colonne est composée de 6 chiffres extraits d'une loi uniforme et la deuxième colonne des 3 premières lettres de l'alphabet.  
Que constatez-vous ?



Créer une data.frame  $D$  dont la première colonne est composé de 5 chiffres extraits d'une loi de poisson de moyenne 8 et la deuxième colonne des 3 premières lettres de l'alphabet. Que constatez vous?

Créer une data.frame  $E$  à 5 lignes et 3 colonnes avec la contrainte que les colonnes soit de mode numeric, character et logical. Nommer les colonnes de  $E$ .

Accéder à la troisième colonne de  $E$  par l'indexation et le nom

Accéder aux valeurs de la première et cinquième ligne de  $E$  par l'indexation

Modifier la valeur du troisième élément de la deuxième colonne. Est-il possible de modifier les éléments sans tenir compte du mode?

Créer une data.frame, nommé  $F$ , composé des éléments  $a = c(1, 2, 3)$  et  $b = c("a", "b", "c")$ . Convertir  $F$  en une matrice. Que constatez-vous?

Convertir la data.frame  $E$  en une liste

### Les listes `list()`

Regarder la fiche détaillée de list

Accéder aux valeurs du troisième élément  $E$  par le nom et par l'indexation





Accéder au mode et à la longueur des différents éléments de E .

Considérons l'éléments de E de mode numeric. Remplacer les valeurs qui composent l'élément de mode numeric par sa moyenne.

Convertir la liste E en une data.frame

### Conversion de mode

Convertir le vecteur nommé logique en mode numeric puis en mode caractère  
`logique = c(TRUE , F, T, FALSE, TRUE, TRUE)`

Convertir le vecteur nommé caractere en mode numeric  
`caractere = c("1", "2", "3", "A", "/", "T", "%", "-")`  
Que remarquez-vous que signifie le symbole `NA`

Convertir le vecteur nommé numérique en mode logical  
`numerique = 0:5`

Convertir le vecteur nommé facteur en mode numeric (en conservant les valeurs de facteur).  
`facteur = factor(c(1, 1, 3, 10, 5, 7, 7, 7, 4), level = c(1, 3, 4, 5, 7, 10))`