

**Exercice I**

On vous donne le vecteur S comme suit :  $S = 3,7,2,8,9,11,1,4$ .

1. Comparer les commandes suivantes :

```
matrix(S, nrow = 2)
```

```
matrix(S, byrow = TRUE)
```

2. Construire une matrice A comportant quatre lignes et trois colonnes remplies par lignes successives avec les éléments du vecteur 1 :12.
3. Construire une matrice B comportant quatre lignes et trois colonnes remplies par colonnes successives avec les éléments du vecteur 1 :12.
4. Extraire l'élément situé en deuxième ligne et troisième colonne de A.
5. Extraire la première colonne de A, puis la deuxième ligne de A.
6. Construire une matrice C constituée des lignes 1 et 4 de A.

**Exercice II**

1. Créer une matrice carrée M, qui contienne les chiffres de 1 à 9. Puis trouver la classe de l'objet M.

2. Essayer ces commandes :

```
M1=cbind("a","b","c","d","e")      M2=rbind("a","b","c","d","e")
```

```
M3=cbind(M,c(19,25,20))             M4=rbind(M, c(11,55,22))
```

3. Interpréter les résultats en comprenant le sens de chaque commande.

**Exercice III**

Créez une matrice X contenant les chiffres de 1 :12, avec 04 lignes et 03 colonnes.

1. Quelle est la longueur et le mode de la matrice ?
2. Extraire toutes les valeurs supérieures à 8 dans la matrice X.
3. Décaler les lieux des colonnes 1 et 3 en X.
4. Ajouter un vecteur avec trois zéros comme cinquième ligne de la matrice X.
5. Remplacez toutes les valeurs des deux premières colonnes de votre matrice par "NA".
6. Remplacez toutes les valeurs de la matrice X par 0 et convertissez-le en vecteur.

**Exercice IV**

1. Créez deux matrices  $2 \times 2$  différentes nommées K et L.

K doit contenir les valeurs 1 à 4 et L contient les valeurs de 5 à 8.

2. Essayez les commandes suivantes et, en regardant les résultats, voyez si vous pouvez comprendre ce qui se passe.

```
K * L    K / L    K %x% L    K + L    K - L    K == L.
```

3. Générez une matrice  $E_{10 \times 10}$  avec des nombres aléatoires.

(a) Ajoutez des alphabets comme des noms de lignes et de colonnes de la matrice E.

(b) Calculez la moyenne et la médiane des lignes de la matrice E.

(c) Enregistrez les résultats de la question précédente dans une nouvelle matrice W.

**Exercice V**

1. Utiliser les commandes suivantes pour créer un jeu de données (data frame) appelé D.

```
id <- 1 :11
```

```
age <- c(19, 54, 41, 40, 29, 39, 47, 38, 42, 38, 31)
```

```
Nom <- c("Farid", "Ahmed", "Mona", "Malek", "Raounek", "Rami", "Aymen", "Amina", "Billel", "Malik", "Aicha")
```

2. Trouver les nombres de ligne et de colonne de D, individuellement puis les deux au même temps.

3. Extraire :
  - (a) Le nom de deuxième colonne de D.
  - (b) La deuxième colonne en utilisant son numéro.
  - (c) La colonne age en utilisant son nom.
  - (d) L'élément de la sixième ligne et la première colonne.
  - (e) L'élément de la sixième colonne et la première ligne.
  - (f) Les éléments de 1 à 3 de la quatrième lignes.
4. Donner :
  - (a) Tous les éléments de la 3 ème ligne.
  - (b) les élément de la 2 ème colonne.
5. Essayer puis interpréter ces commandes :
  - (a) `class(D["age"])`
  - (b) `class(D[, "age"])`
  - (c) `head(D)`
  - (d) `tail(D)`

#### Exercice VI

---

1. On vous propose les vecteurs suivants :  
`V1=c(1,2,3,4,NA)`  
`V2=c(6,7,8,NA,10)`  
`V3=c(11,12,NA,14,15)`  
`V4=c(16,NA,18,19,20)`  
`V5=c(21,22,23,24,25)`
  - (a) Afficher chaque vecteur, puis trouver son classe.
  - (b) Combiner les vecteurs précédents pour avoir une data frame appelé df.
  - (c) Trouver : la classe, les nombres de colonnes et de lignes.
  - (d) Combiner trois data frame df par ligne. Appeler cette data ldf.
  - (e) Combiner trois data frame df par colonne. Appeler cette data cdf.
  - (f) Afficher l'entête et le pied des données ldf.
  - (g) Afficher les 4 première et les 2 dernière lignes.
  - (h) Trouver la classe de chaque vecteur de ldf.
2. Avec les données ldf :
  - (a) Trouver les noms : des variables, des colonne et des lignes.
  - (b) Remplacer les noms des colonnes par : a,b,c,d,e. Vérifier le changement.
3. Avec les données cdf :
  - (a) Calculer le transposé en l'appelant tcdf.
  - (b) Assurer vous que tcdf est de classe data frame.
  - (c) A partir de cdf créer une data frame (Newcdf) sans aucune valeur manquantes.

**Exercice VII**

1. Essayer puis interpréter les commandes suivantes :  
`data() attach(iris) View(iris) names(iris) help("iris")` or `?iris`.
2. Utilisation du jeu de données iris.
  - (a) Obtenez toutes les lignes d'espèce versicolor dans un nouveau bloc de données. Appelez cette data frame : iris.vers
  - (b) Obtenez un vecteur appelé sepal.dif avec la différence entre "Sepal.Length" et "Sepal.Width" des plantes versicolor.
  - (c) Mettre à jour (ajouter) iris.vers avec la nouvelle colonne sepal.dif.
  - (d) Utilisez dplyr pour filtrer toutes les données des espèces virginica ayant une largeur de sépale supérieure à 3,5.
  - (e) Trouver une data frame de toutes les données des espèces virginica avec un Sepal.Width supérieure à 3,5, mais sans la dernière colonne Espèces.
  - (f) Obtenez les ID des lignes correspondant aux deux critères de filtrage fournis ci-dessus.
  - (g) Manipulation de jeu de données au niveau de la colonne (dataset = "iris")
    - i. Répétez chaque fois la valeur de 'Sepal.Length' deux fois et répétez également la séquence complète deux fois.

**Exercice VIII**

Pour cet exercice, nous allons utiliser le jeu de données (intégré) VADeaths.

1. Assurez-vous que l'objet est un bloc de données, sinon changez-le en un bloc de données.
2. Créez une nouvelle variable, nommée Total, qui correspond à la somme de chaque ligne.
3. Modifiez l'ordre des colonnes pour que Total soit la première variable.

**Exercice IX**

1. Travailler avec le jeu de données mtcars
  - (a) Obtenez un histogramme des valeurs mpg de mtcars. Quel bac contient le plus d'observations ?
  - (b) Existe-t-il davantage de voitures à transmission automatique (0) ou manuelle (1) dans le jeu de données ? Indice : mtcars a 32 observations.
  - (c) Obtenez un diagramme de dispersion de hp vs weight (poids); de couleur rouge et de titres des axes : Car Weight in 1K lbs, Horsepower.
2. Voir l'entête de jeu de données (mtcars)
3. Vérifiez la classe de chaque variable dans mtcars.
4. Remplacez am, cyl et vs par un entier et stockez le nouvel ensemble de données sous le nom newmtc.
5. Arrondissez le bloc de données newmtc à un chiffre.

**Exercice X**

1. Déterminer les noms de variables de l'ensemble de données iris, ainsi que les dimensions (nombre de lignes et de colonnes).
2. Calculer la moyenne de toutes les variables numériques dans le jeu de données iris. Les variables numériques sont les quatre premières colonnes du jeu de données iris.
3. Calculer l'intervalle inter-quartile de toutes les variables numériques.
4. Il existe trois espèces d'iris
  - (a) Déterminer les statistiques descriptive pour l'espèce virginica.
  - (b) Déterminez également le résumé en cinq chiffres pour la largeur des pétales des virginicas. (La commande pour faire ceci est `fivenum`).
  - (c) Statistiques descriptive pour un sous-groupe de Petal.Width qui sont supérieur à 2.

**Exercice XI**

---

Utiliser votre fichier Excel nommé "filles" et le fichier csv "DataTabac"; pour faire respectivement ce qui suit :

1. Afficher les données "filles" (**res.** "DataTabac").
2. Interpréter le résultat de la commande : `str(filles)` (**res.** `str(DataTabac)`).
3. Afficher les noms de variables de ces données.
4. Afficher la statistique descriptive de chaque variable.
5. Afficher la statistique descriptive des données.
6. Tracer un histogramme convenable pour chaque variable de "filles" (**res.** "DataTabac"). Interpréter les résultats graphique.
7. Interpréter le résultat graphique de la commande suivante : `boxplot(filles)` (**res.** `boxplot(DataTabac)`).

**Page 4/4**

**Fin de TP 03.**

*Le 23 avril 1997 débute le Comprehensive R Archive Network (CRAN) puis le 5 décembre 1997, R est intégré au Projet GNU.*