

# Premiers pas avec R

---

## Introduction

---

Bienvenue dans ce premier module de la formation “Introduction au logiciel R”.

Nous y verrons :

1. Comment réaliser des opérations arithmétiques élémentaires avec R ?
2. Comment visualiser un tableau de données avec R ?
3. Comment réaliser des opérations arithmétiques élémentaires sur ce tableau de données ?

Vous apprendrez à utiliser \$, le premier symbole indispensable à connaître pour utiliser R.

---

---

## 1-Realiser des opérations arithmétiques élémentaires avec R

R peut à minima être utilisé comme une grosse calculatrice. Additions, soustractions, multiplication et divisions s'écrivent de la manière suivante (cliquez sur la flèche verte en haut à gauche de chaque cellule pour afficher le résultat) :

Addition

```
# Addition  
1+1
```

```
[1] 2
```

Soustraction

```
# Soustraction  
1-1
```

```
[1] 0
```

Multiplication

```
# Multiplication  
2*2
```

```
[1] 4
```

Division

```
# Division  
4/2
```

```
[1] 2
```

Pour la fonction puissance, on utilise le symbole ^

```
# Puissance  
2^3
```

```
[1] 8
```

Pour la racine carrée, on peut soit utiliser la fonction **sqrt**, soit la puissance 1/2

```
# Racine carrée avec la fonction sqrt  
sqrt(16)
```

```
[1] 4
```

```
# Racine carrée avec le symbole de puissance  
16^(1/2)
```

```
[1] 4
```

R propose également toutes les fonctions trigonométriques,

```
# Cosinus  
cos(20)
```

```
[1] 0.4080821
```

```
# Sinus  
sin(10)
```

```
[1] -0.5440211
```

```
# Tangente  
tan(10)
```

```
[1] 0.6483608
```

```
# Arc tangente  
atan(10)
```

```
[1] 1.471128
```

Et les fonctions logarithmiques :

```
# Logarithme népérien (base e)  
log(1)
```

```
[1] 0
```

```
# Logarithme en base 10 avec la fonction log10
log10(10)
```

```
[1] 1
```

```
# Logarithme en base 10 avec la fonction log
log(10, base = 10)
```

```
[1] 1
```

Voilà pour un premier aperçu des fonctions les plus classiques.

---

### ***Exercice d'application***

Calculez le logarithme en base 10 du sinus de 8 à la puissance 3

```
# Ecrivez votre formule en dessous de cette ligne :
```

Vous trouvez : **-1.099532 ? Bravo ! Sinon, cherchez l'erreur...**

---

## **2-Visualiser un tableau de données avec R**

R dispose de jeux de données préinstallées et prêtes à être utilisées dont les plus utilisés sont :

1. mtcars : sur les voitures
2. iris : sur les plantes
3. ToothGrowth : sur les dents
4. PlantGrowth : sur les plantes aussi
5. USArrest : sur les crimes

Pour en savoir plus sur ces jeux de données, vous pouvez utiliser la fonction **help** qui s’affiche dans un onglet séparé de Rstudio. La fonction **data()** vous donne accès à l’ensemble des jeux de données préinstallés.

```
help("ToothGrowth")
```

Pour accéder aux jeux de données il suffit de taper leur nom. Mais si vous faites ça (allez-y !), R affiche par défaut tout le jeu de données (les 150 lignes). Dur de s’y retrouver. Voici quelques fonctions utiles pour prendre connaissance d’un jeu de données. La fonction **dim** vous donne les dimensions du tableau. Ici 150 lignes et 5 colonnes.

```
# La fonction dim
dim(iris)
```

```
[1] 150    5
```

La fonction **names** vous donne le nom des colonnes du tableau :

```
# La fonction names
names(iris)
```

```
[1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

Enfin la fonction **head** vous donne un “aperçu” des 10 premières colonnes du tableau

```
head(iris)
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa

### Exercice d'application :

Choisissez l'un des jeux de données de R puis :

- Interrogez l'aide pour avoir des informations générales sur le jeu de données
- Trouvez le nombre de lignes et de colonnes du jeu de données
- Trouvez les noms des colonnes du jeu de données

```
# Ecrivez votre code ici :
```

---

---

### 3-Comment réaliser des opérations arithmétiques élémentaires sur ce tableau de données ?

Entrons maintenant dans le vif du sujet. Comment utiliser R pour manipuler et faire des opérations élémentaires sur ces données ?

---

**Le premier des quatre symboles indispensables à connaître pour utiliser R est le symbole “\$” (money, money, money !)**

---

Le symbole \$ sert à sélectionner une colonne d'un tableau. Par exemple si je souhaite récupérer la colonne “Sepal.Length” du dataset “iris”, j'écrirai `iris$Sepal.Length`.

```
# Récupérons de la colonne Sepal.Length du dataset iris
iris$Sepal.Length
```

```
[1] 5.1 4.9 4.7 4.6 5.0 5.4 4.6 5.0 4.4 4.9 5.4 4.8 4.8 4.3 5.8 5.7 5.4 5.1
[19] 5.7 5.1 5.4 5.1 4.6 5.1 4.8 5.0 5.0 5.2 5.2 4.7 4.8 5.4 5.2 5.5 4.9 5.0
[37] 5.5 4.9 4.4 5.1 5.0 4.5 4.4 5.0 5.1 4.8 5.1 4.6 5.3 5.0 7.0 6.4 6.9 5.5
[55] 6.5 5.7 6.3 4.9 6.6 5.2 5.0 5.9 6.0 6.1 5.6 6.7 5.6 5.8 6.2 5.6 5.9 6.1
[73] 6.3 6.1 6.4 6.6 6.8 6.7 6.0 5.7 5.5 5.5 5.8 6.0 5.4 6.0 6.7 6.3 5.6 5.5
[91] 5.5 6.1 5.8 5.0 5.6 5.7 5.7 6.2 5.1 5.7 6.3 5.8 7.1 6.3 6.5 7.6 4.9 7.3
[109] 6.7 7.2 6.5 6.4 6.8 5.7 5.8 6.4 6.5 7.7 7.7 6.0 6.9 5.6 7.7 6.3 6.7 7.2
[127] 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8
[145] 6.7 6.7 6.3 6.5 6.2 5.9
```

---

### Exercice d'application :

Calculez la longueur de `iris$Sepal.Length` et affichez seulement les 5 premières valeurs

```
#Ecrivez votre réponse ici :
```

---

Il est ensuite possible de faire directement **des opérations arithmétiques sur ces vecteurs**, comme par exemple :

La multiplication par un scalaire,

```
# Multiplier par quatre la colonne Petal.Length du dataset iris
4 * iris$Petal.Length
```

```
[1] 5.6 5.6 5.2 6.0 5.6 6.8 5.6 6.0 5.6 6.0 6.0 6.4 5.6 4.4 4.8
[16] 6.0 5.2 5.6 6.8 6.0 6.8 6.0 4.0 6.8 7.6 6.4 6.4 6.0 5.6 6.4
[31] 6.4 6.0 6.0 5.6 6.0 4.8 5.2 5.6 5.2 6.0 5.2 5.2 5.2 6.4 7.6
[46] 5.6 6.4 5.6 6.0 5.6 18.8 18.0 19.6 16.0 18.4 18.0 18.8 13.2 18.4 15.6
[61] 14.0 16.8 16.0 18.8 14.4 17.6 18.0 16.4 18.0 15.6 19.2 16.0 19.6 18.8 17.2
[76] 17.6 19.2 20.0 18.0 14.0 15.2 14.8 15.6 20.4 18.0 18.0 18.8 17.6 16.4 16.0
[91] 17.6 18.4 16.0 13.2 16.8 16.8 16.8 17.2 12.0 16.4 24.0 20.4 23.6 22.4 23.2
[106] 26.4 18.0 25.2 23.2 24.4 20.4 21.2 22.0 20.0 20.4 21.2 22.0 26.8 27.6 20.0
[121] 22.8 19.6 26.8 19.6 22.8 24.0 19.2 19.6 22.4 23.2 24.4 25.6 22.4 20.4 22.4
[136] 24.4 22.4 22.0 19.2 21.6 22.4 20.4 20.4 23.6 22.8 20.8 20.0 20.8 21.6 20.4
```

le calcul du sinus :

```
# Prendre le sinus de la colonne Petal.Length du dataset iris
sin(iris$Petal.Length)
```

```
[1] 0.9854497 0.9854497 0.9635582 0.9974950 0.9854497 0.9916648
[7] 0.9854497 0.9974950 0.9854497 0.9974950 0.9974950 0.9995736
[13] 0.9854497 0.8912074 0.9320391 0.9974950 0.9635582 0.9854497
[19] 0.9916648 0.9974950 0.9916648 0.9974950 0.8414710 0.9916648
[25] 0.9463001 0.9995736 0.9995736 0.9974950 0.9854497 0.9995736
[31] 0.9995736 0.9974950 0.9974950 0.9854497 0.9974950 0.9320391
[37] 0.9635582 0.9854497 0.9635582 0.9974950 0.9635582 0.9635582
[43] 0.9635582 0.9995736 0.9463001 0.9854497 0.9995736 0.9854497
```

```
[49]  0.9974950  0.9854497 -0.9999233 -0.9775301 -0.9824526 -0.7568025
[55] -0.9936910 -0.9775301 -0.9999233 -0.1577457 -0.9936910 -0.6877662
[61] -0.3507832 -0.8715758 -0.7568025 -0.9999233 -0.4425204 -0.9516021
[67] -0.9775301 -0.8182771 -0.9775301 -0.6877662 -0.9961646 -0.7568025
[73] -0.9824526 -0.9999233 -0.9161659 -0.9516021 -0.9961646 -0.9589243
[79] -0.9775301 -0.3507832 -0.6118579 -0.5298361 -0.6877662 -0.9258147
[85] -0.9775301 -0.9775301 -0.9999233 -0.9516021 -0.8182771 -0.7568025
[91] -0.9516021 -0.9936910 -0.7568025 -0.1577457 -0.8715758 -0.8715758
[97] -0.8715758 -0.9161659  0.1411200 -0.8182771 -0.2794155 -0.9258147
[103] -0.3738767 -0.6312666 -0.4646022  0.3115414 -0.9775301  0.0168139
[109] -0.4646022 -0.1821625 -0.9258147 -0.8322674 -0.7055403 -0.9589243
[115] -0.9258147 -0.8322674 -0.7055403  0.4048499  0.5784398 -0.9589243
[121] -0.5506855 -0.9824526  0.4048499 -0.9824526 -0.5506855 -0.2794155
[127] -0.9961646 -0.9824526 -0.6312666 -0.4646022 -0.1821625  0.1165492
[133] -0.6312666 -0.9258147 -0.6312666 -0.1821625 -0.6312666 -0.7055403
[139] -0.9961646 -0.7727645 -0.6312666 -0.9258147 -0.9258147 -0.3738767
[145] -0.5506855 -0.8834547 -0.9589243 -0.8834547 -0.7727645 -0.9258147
```

Il est également possible de faire de opérations entre les colonnes du tableau, par exemple :  
**sommer deux colonnes,**

```
# Somme de la colonne Petal.Length et Sepal.Length
iris$Petal.Length + iris$Sepal.Length
```

```
[1]  6.5  6.3  6.0  6.1  6.4  7.1  6.0  6.5  5.8  6.4  6.9  6.4  6.2  5.4  7.0
[16]  7.2  6.7  6.5  7.4  6.6  7.1  6.6  5.6  6.8  6.7  6.6  6.6  6.7  6.6  6.3
[31]  6.4  6.9  6.7  6.9  6.4  6.2  6.8  6.3  5.7  6.6  6.3  5.8  5.7  6.6  7.0
[46]  6.2  6.7  6.0  6.8  6.4 11.7 10.9 11.8  9.5 11.1 10.2 11.0  8.2 11.2  9.1
[61]  8.5 10.1 10.0 10.8  9.2 11.1 10.1  9.9 10.7  9.5 10.7 10.1 11.2 10.8 10.7
[76] 11.0 11.6 11.7 10.5  9.2  9.3  9.2  9.7 11.1  9.9 10.5 11.4 10.7  9.7  9.5
[91]  9.9 10.7  9.8  8.3  9.8  9.9  9.9 10.5  8.1  9.8 12.3 10.9 13.0 11.9 12.3
[106] 14.2  9.4 13.6 12.5 13.3 11.6 11.7 12.3 10.7 10.9 11.7 12.0 14.4 14.6 11.0
[121] 12.6 10.5 14.4 11.2 12.4 13.2 11.0 11.0 12.0 13.0 13.5 14.3 12.0 11.4 11.7
[136] 13.8 11.9 11.9 10.8 12.3 12.3 12.0 10.9 12.7 12.4 11.9 11.3 11.7 11.6 11.0
```

calculer le **produit scalaire** de deux colonnes :

```
# Produit de la colonne Petal.Length et Sepal.Length
iris$Petal.Length * iris$Sepal.Length
```

```
[1]  7.14  6.86  6.11  6.90  7.00  9.18  6.44  7.50  6.16  7.35  8.10  7.68
```



[13]	6.72	4.73	6.96	8.55	7.02	7.14	9.69	7.65	9.18	7.65	4.60	8.67
[25]	9.12	8.00	8.00	7.80	7.28	7.52	7.68	8.10	7.80	7.70	7.35	6.00
[37]	7.15	6.86	5.72	7.65	6.50	5.85	5.72	8.00	9.69	6.72	8.16	6.44
[49]	7.95	7.00	32.90	28.80	33.81	22.00	29.90	25.65	29.61	16.17	30.36	20.28
[61]	17.50	24.78	24.00	28.67	20.16	29.48	25.20	23.78	27.90	21.84	28.32	24.40
[73]	30.87	28.67	27.52	29.04	32.64	33.50	27.00	19.95	20.90	20.35	22.62	30.60
[85]	24.30	27.00	31.49	27.72	22.96	22.00	24.20	28.06	23.20	16.50	23.52	23.94
[97]	23.94	26.66	15.30	23.37	37.80	29.58	41.89	35.28	37.70	50.16	22.05	45.99
[109]	38.86	43.92	33.15	33.92	37.40	28.50	29.58	33.92	35.75	51.59	53.13	30.00
[121]	39.33	27.44	51.59	30.87	38.19	43.20	29.76	29.89	35.84	41.76	45.14	50.56
[133]	35.84	32.13	34.16	46.97	35.28	35.20	28.80	37.26	37.52	35.19	29.58	40.12
[145]	38.19	34.84	31.50	33.80	33.48	30.09						

---

***Exercice d'application :***

Calculez le sinus de la somme des colonnes Petal.Length et Sepal.Length du dataset iris

```
#Ecrivez votre réponse ici :
```