

Introduction à la programmation en S

Vincent Goulet

École d'actuariat
Université Laval

Première partie I

BASES DU LANGAGE S

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux
- 6 Listes
- 7 Data frames
- 8 Indiçage

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

1 Commandes S

2 Conventions pour les noms d'objets

3 Les objets S

- Modes et types de données
- Longueur
- Attributs
- L'objet spécial NA
- L'objet spécial NULL

4 Vecteurs

5 Matrices et tableaux

6 Listes

7 Data frames

8 Indiçage

Affectations et expressions

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Toute commande S est soit une *affectation*, soit une *expression*.

- Normalement, une expression est immédiatement évaluée et le résultat est affiché à l'écran :

```
> 2 + 3
```

```
[1] 5
```

```
> pi
```

```
[1] 3.141593
```

```
> cos(pi/4)
```

```
[1] 0.7071068
```

Affectations et expressions

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Lors d'une affectation, une expression est évaluée, mais le résultat est stocké dans un objet (variable) et rien n'est affiché à l'écran.
- Le symbole d'affectation est `<-` (ou `->`).

```
> a <- 5
```

```
> a
```

```
[1] 5
```

```
> b <- a
```

```
> b
```

```
[1] 5
```

Deux symboles d'affectation à éviter

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

■ L'opérateur =

- peut porter à confusion.

■ Le caractère _

- permis dans S-Plus, mais plus dans R depuis la version 1.8.0
- emploi fortement découragé
- rend le code difficile à lire
- dans le mode ESS de Emacs, taper ce caractère génère carrément `_<-_`.

Deux symboles d'affectation à éviter

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

■ L'opérateur =

- peut porter à confusion.

■ Le caractère _

- permis dans S-Plus, mais plus dans R depuis la version 1.8.0
- emploi fortement découragé
- rend le code difficile à lire
- dans le mode ESS de Emacs, taper ce caractère génère carrément `_<-_`.

Astuce

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Pour affecter le résultat d'un calcul dans un objet et en même temps voir ce résultat, placer l'affectation entre parenthèses.

L'opération d'affectation devient alors une nouvelle expression :

```
> (a <- 2 + 3)
```

```
[1] 5
```

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

1 Commandes S

2 Conventions pour les noms d'objets

3 Les objets S

- Modes et types de données

- Longueur

- Attributs

- L'objet spécial NA

- L'objet spécial NULL

4 Vecteurs

5 Matrices et tableaux

6 Listes

7 Data frames

8 Indiçage

Caractères permis dans les noms d'objets

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Les lettres a–z, A–Z
- Les chiffres 0–9
- Le point «.»
- «_» est maintenant permis dans R, mais son utilisation est découragée.

Caractères permis dans les noms d'objets

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Les lettres a–z, A–Z
- Les chiffres 0–9
- Le point «.»
- «_» est maintenant permis dans R, mais son utilisation est découragée.

Caractères permis dans les noms d'objets

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Les lettres a–z, A–Z
- Les chiffres 0–9
- Le point «.»
- «_» est maintenant permis dans R, mais son utilisation est découragée.

Caractères permis dans les noms d'objets

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Les lettres a–z, A–Z
- Les chiffres 0–9
- Le point «.»
- «_» est maintenant permis dans R, mais son utilisation est découragée.

Règles pour les noms d'objets

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Les noms d'objets ne peuvent commencer par un chiffre.
- Le S est sensible à la casse : `foo`, `Foo` et `FOO` sont trois objets distincts.
- Moyen simple d'éviter des erreurs liées à la casse : employer seulement des lettres minuscules.

Règles pour les noms d'objets

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Les noms d'objets ne peuvent commencer par un chiffre.
- Le S est sensible à la casse : `f00`, `F00` et `FOO` sont trois objets distincts.
- Moyen simple d'éviter des erreurs liées à la casse : employer seulement des lettres minuscules.

Noms déjà utilisés et réservés

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indice

- Certains noms sont utilisés par le système, aussi vaut-il mieux éviter de les utiliser. En particulier, éviter d'utiliser

`c, q, t, C, D, I, diff, length, mean, pi, range, var.`

- Certains mots sont réservés pour le système et il est interdit de les utiliser comme nom d'objet :

`Inf, NA, NaN, NULL`

`break, else, for, function, if, in, next, repeat,
return, while.`

- Dans S-Plus 6.1 et plus, `T` et `TRUE` (vrai), ainsi que `F` et `FALSE` (faux) sont également des noms réservés.

Noms déjà utilisés et réservés

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indice

- Certains noms sont utilisés par le système, aussi vaut-il mieux éviter de les utiliser. En particulier, éviter d'utiliser

`c, q, t, C, D, I, diff, length, mean, pi, range, var.`

- Certains mots sont réservés pour le système et il est interdit de les utiliser comme nom d'objet :

`Inf, NA, NaN, NULL`

`break, else, for, function, if, in, next, repeat,
return, while.`

- Dans S-Plus 6.1 et plus, `T` et `TRUE` (vrai), ainsi que `F` et `FALSE` (faux) sont également des noms réservés.

Noms déjà utilisés et réservés

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indice

- Certains noms sont utilisés par le système, aussi vaut-il mieux éviter de les utiliser. En particulier, éviter d'utiliser

`c, q, t, C, D, I, diff, length, mean, pi, range, var.`

- Certains mots sont réservés pour le système et il est interdit de les utiliser comme nom d'objet :

`Inf, NA, NaN, NULL`

`break, else, for, function, if, in, next, repeat,
return, while.`

- Dans S-Plus 6.1 et plus, `T` et `TRUE` (vrai), ainsi que `F` et `FALSE` (faux) sont également des noms réservés.

TRUE et FALSE dans R

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Dans R, les noms TRUE et FALSE sont également réservés.
- Les variables T et F prennent par défaut les valeurs TRUE et FALSE, respectivement, mais peuvent être réaffectées.

```
> T  
[1] TRUE  
  
> TRUE <- 3  
Erreur dans TRUE <- 3 : membre gauche de  
l'assignation (do_set) incorrect  
  
> (T <- 3)  
[1] 3
```

TRUE et FALSE dans R

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Dans R, les noms TRUE et FALSE sont également réservés.
- Les variables T et F prennent par défaut les valeurs TRUE et FALSE, respectivement, mais peuvent être réaffectées.

```
> T
```

```
[1] TRUE
```

```
> TRUE <- 3
```

```
Erreur dans TRUE <- 3 : membre gauche de  
l'assignation (do_set) incorrect
```

```
> (T <- 3)
```

```
[1] 3
```

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S**
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux
- 6 Listes
- 7 Data frames
- 8 Indiçage

Tout est un objet

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Tout dans le langage S est un objet, même les fonctions et les opérateurs.
- Les objets possèdent au minimum un **mode** et une **longueur**.

Mode et longueur

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Le mode d'un objet est obtenu avec la fonction `mode`.

```
> v <- c(1, 2, 5, 9)
```

```
> mode(v)
```

```
[1] "numeric"
```

- La longueur d'un objet est obtenue avec la fonction `length`.

```
> length(v)
```

```
[1] 4
```

- Certains objets sont également dotés d'un ou plusieurs **attributs**.

Modes et types de données

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Le mode prescrit ce qu'un objet peut contenir.
- Un objet ne peut donc avoir qu'un seul mode.
- Modes disponibles en S :

<code>numeric</code>	nombres réels
<code>complex</code>	nombres complexes
<code>logical</code>	valeurs booléennes (vrai/faux)
<code>character</code>	chaînes de caractères
<code>function</code>	fonction
<code>list</code>	données quelconques

Longueur

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indice

La longueur d'un objet est égale au nombre d'éléments qu'il contient.

- La longueur d'une chaîne de caractères est toujours 1. Un objet de mode `character` doit contenir plusieurs chaînes de caractères pour que sa longueur soit supérieure à 1.

```
> v <- "actuariat"
```

```
> length(v)
```

```
[1] 1
```

```
> v <- c("a", "c", "t", "u", "a", "r", "i",  
+       "a", "t")
```

```
> length(v)
```

```
[1] 9
```

Longueur

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indice

La longueur d'un objet est égale au nombre d'éléments qu'il contient.

- La longueur d'une chaîne de caractères est toujours 1. Un objet de mode `character` doit contenir plusieurs chaînes de caractères pour que sa longueur soit supérieure à 1.

```
> v <- "actuariat"
```

```
> length(v)
```

```
[1] 1
```

```
> v <- c("a", "c", "t", "u", "a", "r", "i",  
+       "a", "t")
```

```
> length(v)
```

```
[1] 9
```

Objet vide

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Un objet peut être de longueur 0.
- Doit alors être interprété comme un contenant vide.

```
> v <- numeric(0)
> length(v)

[1] 0
```

Objet vide

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Un objet peut être de longueur 0.
- Doit alors être interprété comme un contenant vide.

```
> v <- numeric(0)
```

```
> length(v)
```

```
[1] 0
```

Attributs

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Éléments d'information additionnels liés à cet objet.
- Attributs les plus fréquemment rencontrés :

<code>class</code>	affecte le comportement d'un objet
<code>dim</code>	dimensions des matrices et tableaux
<code>dimnames</code>	étiquettes des dimensions des matrices et tableaux
<code>names</code>	étiquettes des éléments d'un objet

L'objet spécial NA

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NA est fréquemment utilisé pour représenter les données manquantes.

- Son mode est `logical`.
- Toute opération impliquant une donnée NA a comme résultat NA.
- Certaines fonctions (`sum`, `mean`, par exemple), ont par conséquent un argument `na.rm` qui, lorsque `TRUE`, élimine les données manquantes avant de faire un calcul.
- La fonction `is.na` permet de tester si les éléments d'un objet sont NA ou non.

L'objet spécial NA

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NA est fréquemment utilisé pour représenter les données manquantes.

- Son mode est `logical`.
- Toute opération impliquant une donnée NA a comme résultat NA.
- Certaines fonctions (`sum`, `mean`, par exemple), ont par conséquent un argument `na.rm` qui, lorsque `TRUE`, élimine les données manquantes avant de faire un calcul.
- La fonction `is.na` permet de tester si les éléments d'un objet sont NA ou non.

L'objet spécial NA

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NA est fréquemment utilisé pour représenter les données manquantes.

- Son mode est `logical`.
- Toute opération impliquant une donnée NA a comme résultat NA.
- Certaines fonctions (`sum`, `mean`, par exemple), ont par conséquent un argument `na.rm` qui, lorsque `TRUE`, élimine les données manquantes avant de faire un calcul.
- La fonction `is.na` permet de tester si les éléments d'un objet sont NA ou non.

L'objet spécial NA

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NA est fréquemment utilisé pour représenter les données manquantes.

- Son mode est `logical`.
- Toute opération impliquant une donnée NA a comme résultat NA.
- Certaines fonctions (`sum`, `mean`, par exemple), ont par conséquent un argument `na.rm` qui, lorsque `TRUE`, élimine les données manquantes avant de faire un calcul.
- La fonction `is.na` permet de tester si les éléments d'un objet sont NA ou non.

L'objet spécial NA

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NA est fréquemment utilisé pour représenter les données manquantes.

- Son mode est `logical`.
- Toute opération impliquant une donnée NA a comme résultat NA.
- Certaines fonctions (`sum`, `mean`, par exemple), ont par conséquent un argument `na.rm` qui, lorsque `TRUE`, élimine les données manquantes avant de faire un calcul.
- La fonction `is.na` permet de tester si les éléments d'un objet sont NA ou non.

L'objet spécial NULL

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un constant vide,
 - NULL est une constante.
- La fonction `is.null` teste si un objet est NULL ou non.

L'objet spécial NULL

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un contenant vide ;
 - NULL est «pas de contenant».
- La fonction `is.null` teste si un objet est NULL ou non.

L'objet spécial NULL

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un contenant vide ;
 - NULL est «pas de contenant».
- La fonction `is.null` teste si un objet est NULL ou non.

L'objet spécial NULL

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un contenant vide ;
 - NULL est «pas de contenant».
- La fonction `is.null` teste si un objet est NULL ou non.

L'objet spécial NULL

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un contenant vide ;
 - NULL est «pas de contenant».
- La fonction `is.null` teste si un objet est NULL ou non.

L'objet spécial NULL

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un contenant vide ;
 - NULL est «pas de contenant».
- La fonction `is.null` teste si un objet est NULL ou non.

L'objet spécial NULL

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

NULL représente «rien», ou le vide.

- Son mode est NULL.
- Sa longueur est 0.
- Différent d'un objet vide :
 - un objet de longueur 0 est un contenant vide ;
 - NULL est «pas de contenant».
- La fonction `is.null` teste si un objet est NULL ou non.

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux
- 6 Listes
- 7 Data frames
- 8 Indiçage

En S, tout est un vecteur

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Dans un vecteur simple, tous les éléments doivent être du même mode.
- Il est possible (et souvent souhaitable) de donner une étiquette à chacun des éléments d'un vecteur.

```
> (v <- c(a = 1, b = 2, c = 5))
```

```
a b c
```

```
1 2 5
```

```
> v <- c(1, 2, 5)
```

```
> names(v) <- c("a", "b", "c")
```

```
> v
```

```
a b c
```

```
1 2 5
```

Et comment crée-t-on ces vecteurs ?

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Les fonctions de base pour créer des vecteurs sont

- `c` (concaténation)
- `numeric` (vecteur de mode `numeric`)
- `logical` (vecteur de mode `logical`)
- `character` (vecteur de mode `character`).

Indiçage d'un vecteur

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Se fait avec `[]`.

- On peut extraire un élément d'un vecteur par

- sa position ou

- son étiquette, si elle existe (auquel cas cette approche est beaucoup plus sûre).

```
> v[3]
```

```
c
```

```
5
```

```
> v["c"]
```

```
c
```

```
5
```

Indiçage d'un vecteur

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Se fait avec `[]`.
- On peut extraire un élément d'un vecteur par
 - sa position ou
 - son étiquette, si elle existe (auquel cas cette approche est beaucoup plus sûre).

```
> v[3]
```

```
c
```

```
5
```

```
> v["c"]
```

```
c
```

```
5
```

Indiçage d'un vecteur

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Se fait avec `[]`.
- On peut extraire un élément d'un vecteur par
 - sa position ou
 - son étiquette, si elle existe (auquel cas cette approche est beaucoup plus sûre).

```
> v[3]
```

```
c
```

```
5
```

```
> v["c"]
```

```
c
```

```
5
```


Indiçage d'un vecteur

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Se fait avec `[]`.
- On peut extraire un élément d'un vecteur par
 - sa position ou
 - son étiquette, si elle existe (auquel cas cette approche est beaucoup plus sûre).

```
> v[3]
```

```
c
```

```
5
```

```
> v["c"]
```

```
c
```

```
5
```

Indiçage d'un vecteur

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- Se fait avec `[]`.
- On peut extraire un élément d'un vecteur par
 - sa position ou
 - son étiquette, si elle existe (auquel cas cette approche est beaucoup plus sûre).

```
> v[3]
```

```
c
```

```
5
```

```
> v["c"]
```

```
c
```

```
5
```

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux**
- 6 Listes
- 7 Data frames
- 8 Indiçage

Une matrice est un vecteur ?

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une matrice ou, de façon plus générale, un tableau (*array*) n'est rien d'autre qu'un vecteur doté d'un attribut `dim`.

- À l'interne, une matrice est donc stockée sous forme de vecteur.
- La fonction de base pour créer des matrices est `matrix`.
- La fonction de base pour créer des tableaux est `array`.

Une matrice est un vecteur ?

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une matrice ou, de façon plus générale, un tableau (*array*) n'est rien d'autre qu'un vecteur doté d'un attribut `dim`.

- À l'interne, une matrice est donc stockée sous forme de vecteur.
- La fonction de base pour créer des matrices est `matrix`.
- La fonction de base pour créer des tableaux est `array`.

Une matrice est un vecteur ?

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une matrice ou, de façon plus générale, un tableau (*array*) n'est rien d'autre qu'un vecteur doté d'un attribut `dim`.

- À l'interne, une matrice est donc stockée sous forme de vecteur.
- La fonction de base pour créer des matrices est `matrix`.
- La fonction de base pour créer des tableaux est `array`.

Une matrice est un vecteur ?

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une matrice ou, de façon plus générale, un tableau (*array*) n'est rien d'autre qu'un vecteur doté d'un attribut `dim`.

- À l'interne, une matrice est donc stockée sous forme de vecteur.
- La fonction de base pour créer des matrices est `matrix`.
- La fonction de base pour créer des tableaux est `array`.

Remplissage d'une matrice

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- **Important** : les matrices et tableaux sont remplis en faisant d'abord varier la première dimension, puis la seconde, etc.

```
> matrix(1:6, nrow = 2, ncol = 3)
```

	[,1]	[,2]	[,3]
[1,]	1	3	5
[2,]	2	4	6

```
> matrix(1:6, nrow = 2, ncol = 3,  
+       byrow = TRUE)
```

	[,1]	[,2]	[,3]
[1,]	1	2	3
[2,]	4	5	6

Indiçage d'une matrice

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- On extrait les éléments d'une matrice en précisant leurs positions sous la forme (ligne, colonne) dans la matrice, ou encore leurs positions dans le vecteur sous-jacent.

```
> (m <- matrix(c(40, 80, 45, 21, 55, 32),  
+             nrow = 2, ncol = 3))
```

	[,1]	[,2]	[,3]
[1,]	40	45	55
[2,]	80	21	32

```
> m[1, 2]
```

```
[1] 45
```

```
> m[3]
```

```
[1] 45
```

Fusion verticale de matrices

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- La fonction `rbind` permet de fusionner verticalement deux matrices (ou plus) ayant le même nombre de colonnes.

```
> n <- matrix(1:9, nrow = 3)  
> rbind(m, n)
```

	[,1]	[,2]	[,3]
[1,]	40	45	55
[2,]	80	21	32
[3,]	1	4	7
[4,]	2	5	8
[5,]	3	6	9

Fusion horizontale de matrices

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- La fonction `cbind` permet de fusionner horizontalement deux matrices (ou plus) ayant le même nombre de lignes.

```
> n <- matrix(1:4, nrow = 2)
> cbind(m, n)
```

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	40	45	55	1	3
[2,]	80	21	32	2	4

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux
- 6 Listes**
- 7 Data frames
- 8 Indiçage

Un vecteur très général

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une liste est un type de vecteur spécial dont les éléments peuvent être de n'importe quel mode, y compris le mode `list` (ce qui permet d'emboîter des listes).

- La fonction de base pour créer des listes est `list`.
- Généralement préférable de nommer les éléments d'une liste : plus simple et sûr d'extraire les éléments par leur étiquette.

Un vecteur très général

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une liste est un type de vecteur spécial dont les éléments peuvent être de n'importe quel mode, y compris le mode `list` (ce qui permet d'emboîter des listes).

- La fonction de base pour créer des listes est `list`.
- Généralement préférable de nommer les éléments d'une liste : plus simple et sûr d'extraire les éléments par leur étiquette.

Un vecteur très général

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Une liste est un type de vecteur spécial dont les éléments peuvent être de n'importe quel mode, y compris le mode `list` (ce qui permet d'emboîter des listes).

- La fonction de base pour créer des listes est `list`.
- Généralement préférable de nommer les éléments d'une liste : plus simple et sûr d'extraire les éléments par leur étiquette.

Indiçage d'une liste

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'extraction des éléments d'une liste peut se faire de deux façons :

- 1 avec des doubles crochets `[[]]`
- 2 par leur étiquette avec
`nom.liste$etiquette.element.`

Indiçage d'une liste

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'extraction des éléments d'une liste peut se faire de deux façons :

1 avec des doubles crochets `[[]]`

2 par leur étiquette avec
`nom.liste$etiquette.element.`

Indiçage d'une liste

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'extraction des éléments d'une liste peut se faire de deux façons :

- 1 avec des doubles crochets `[[]]`
- 2 par leur étiquette avec
`nom.liste$etiquette.element.`

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux
- 6 Listes
- 7 Data frames**
- 8 Indiçage

Liste ou matrice ? Un peu des deux !

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'analyse de données — la régression linéaire, par exemple — repose sur les *data frames*.
- Liste de classe `data.frame` dont tous les éléments sont de la même longueur.
- Généralement représenté sous forme d'un tableau à deux dimensions (visuellement similaire à une matrice).
- Plus général qu'une matrice puisque les colonnes peuvent être de modes différents (`numeric`, `complex`, `character` ou `logical`).
- Créé avec la fonction `data.frame` ou `as.data.frame`.
- Moins important lors de l'apprentissage du langage de programmation.

Liste ou matrice ? Un peu des deux !

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'analyse de données — la régression linéaire, par exemple — repose sur les *data frames*.
- Liste de classe `data.frame` dont tous les éléments sont de la même longueur.
- Généralement représenté sous forme d'un tableau à deux dimensions (visuellement similaire à une matrice).
- Plus général qu'une matrice puisque les colonnes peuvent être de modes différents (`numeric`, `complex`, `character` ou `logical`).
- Créé avec la fonction `data.frame` ou `as.data.frame`.
- Moins important lors de l'apprentissage du langage de programmation.

Liste ou matrice ? Un peu des deux !

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'analyse de données — la régression linéaire, par exemple — repose sur les *data frames*.
- Liste de classe `data.frame` dont tous les éléments sont de la même longueur.
- Généralement représenté sous forme d'un tableau à deux dimensions (visuellement similaire à une matrice).
- Plus général qu'une matrice puisque les colonnes peuvent être de modes différents (`numeric`, `complex`, `character` ou `logical`).
- Créé avec la fonction `data.frame` ou `as.data.frame`.
- Moins important lors de l'apprentissage du langage de programmation.

Liste ou matrice ? Un peu des deux !

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'analyse de données — la régression linéaire, par exemple — repose sur les *data frames*.
- Liste de classe `data.frame` dont tous les éléments sont de la même longueur.
- Généralement représenté sous forme d'un tableau à deux dimensions (visuellement similaire à une matrice).
- Plus général qu'une matrice puisque les colonnes peuvent être de modes différents (`numeric`, `complex`, `character` ou `logical`).
- Créé avec la fonction `data.frame` ou `as.data.frame`.
- Moins important lors de l'apprentissage du langage de programmation.

Liste ou matrice ? Un peu des deux !

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'analyse de données — la régression linéaire, par exemple — repose sur les *data frames*.
- Liste de classe `data.frame` dont tous les éléments sont de la même longueur.
- Généralement représenté sous forme d'un tableau à deux dimensions (visuellement similaire à une matrice).
- Plus général qu'une matrice puisque les colonnes peuvent être de modes différents (`numeric`, `complex`, `character` ou `logical`).
- Créé avec la fonction `data.frame` ou `as.data.frame`.
- Moins important lors de l'apprentissage du langage de programmation.

Liste ou matrice ? Un peu des deux !

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- L'analyse de données — la régression linéaire, par exemple — repose sur les *data frames*.
- Liste de classe `data.frame` dont tous les éléments sont de la même longueur.
- Généralement représenté sous forme d'un tableau à deux dimensions (visuellement similaire à une matrice).
- Plus général qu'une matrice puisque les colonnes peuvent être de modes différents (`numeric`, `complex`, `character` ou `logical`).
- Créé avec la fonction `data.frame` ou `as.data.frame`.
- Moins important lors de l'apprentissage du langage de programmation.

Sommaire

Introduction à
la programmation
en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

- 1 Commandes S
- 2 Conventions pour les noms d'objets
- 3 Les objets S
 - Modes et types de données
 - Longueur
 - Attributs
 - L'objet spécial NA
 - L'objet spécial NULL
- 4 Vecteurs
- 5 Matrices et tableaux
- 6 Listes
- 7 Data frames
- 8 Indiçage**

Quatre façons d'indicer un vecteur

Introduction à
la programmation en
S

Vincent
Goulet

Commandes
S

Conventions
pour les noms
d'objets

Les objets S

Vecteurs

Matrices et
tableaux

Listes

Data frames

Indiçage

Dans tous les cas, l'indiçage se fait avec des crochets [].

- 1 Avec un vecteur d'entiers positifs. Les éléments se trouvant aux positions correspondant aux entiers sont extraits du vecteur, dans l'ordre. C'est la technique la plus courante.

```
> letters[c(1:3, 22, 5)]
```

```
[1] "a" "b" "c" "v" "e"
```

- 2 Avec un vecteur d'entiers négatifs. Les éléments se trouvant aux positions correspondant aux entiers négatifs sont alors **éliminés** du vecteur.

```
> letters[c(-(1:3), -5, -22)]
```

```
[1] "d" "f" "g" "h" "i" "j" "k" "l" "m"  
[10] "n" "o" "p" "q" "r" "s" "t" "u" "w"  
[19] "x" "y" "z"
```

- 3 Avec un vecteur booléen. Le vecteur d'indiçage doit alors être de la même longueur que le vecteur indicé. Les éléments correspondant à une valeur TRUE sont **extraits** du vecteur, alors que ceux correspondant à FALSE sont **éliminés**.

```
> letters > "f" & letters < "q"
```

```
[1] FALSE FALSE FALSE FALSE FALSE FALSE  
[7]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  
[13]  TRUE  TRUE  TRUE  TRUE FALSE FALSE  
[19] FALSE FALSE FALSE FALSE FALSE FALSE  
[25] FALSE FALSE
```

```
> letters[letters > "f" & letters < "q"]
```

```
[1] "g" "h" "i" "j" "k" "l" "m" "n" "o"  
[10] "p"
```

- 4 Avec une chaîne de caractères. Utile pour extraire les éléments d'un vecteur à condition que ceux-ci soient nommés.

```
> x <- c(Rouge = 2, Bleu = 4, Vert = 9,  
+ Jaune = -5)
```

```
> x[c("Bleu", "Jaune")]
```

```
Bleu Jaune
```

```
4      -5
```