

1. Présentation du langage R

- Pourquoi êtes-vous
là?
- Bref historique
- Caractéristiques de R
- Démarrer une session

2. Bases du langage R

- Les fonctions
- Les objets R
- Commandes R
- Sommaire

3. Graphiques

- La fonction plot
- histogram et
boxplot
- Sommaire

4. Statistiques

- Moyenne, écart type
- Test t et analyse de
variance
- Modèles linéaires
- Sommaire

5. Rapports reproductibles

- Introduction

Atelier sur le logiciel R

Un introduction à la programmation en R

Sahir Rai Bhatnagar¹

29 juillet 2015

¹<https://github.com/sahirbhatnagar/atelier-R-GERAD>

Programmation de la journée

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- 09H00 : Début de l'atelier
- 10H00 - 10H20: pause
- 12H00 - 13H00: lunch
- 15H00 - 15H20: pause
- 17H00 : fin de l'atelier

Remerciements

1. Présentation du langage R

Pourquoi êtes-vous
là?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

- John Chambers
- Ross Ihaka et Robert Gentleman
- Greg Voisin
- Dr. Vahid Partovi Nia
- Toi



1. Présentation du langage R

Pourquoi êtes-vous
là ?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

- Ceci est une **introduction** au langage R

Avis #1

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

- Ceci est une **introduction** au langage R
- On va faire beaucoup d'exercices

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Ceci est une **introduction** au langage R
- On va faire beaucoup d'exercices
- N'hésitez pas à posez des questions

Avis #2

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction



R Markdown v2



Je n'ai aucune relation commerciale avec ces logiciels

Avis #3

1. Présentation du langage R

Pourquoi êtes-vous
là ?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

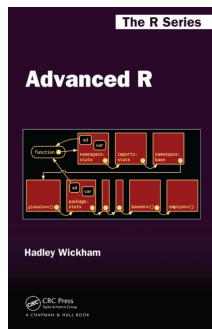
4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

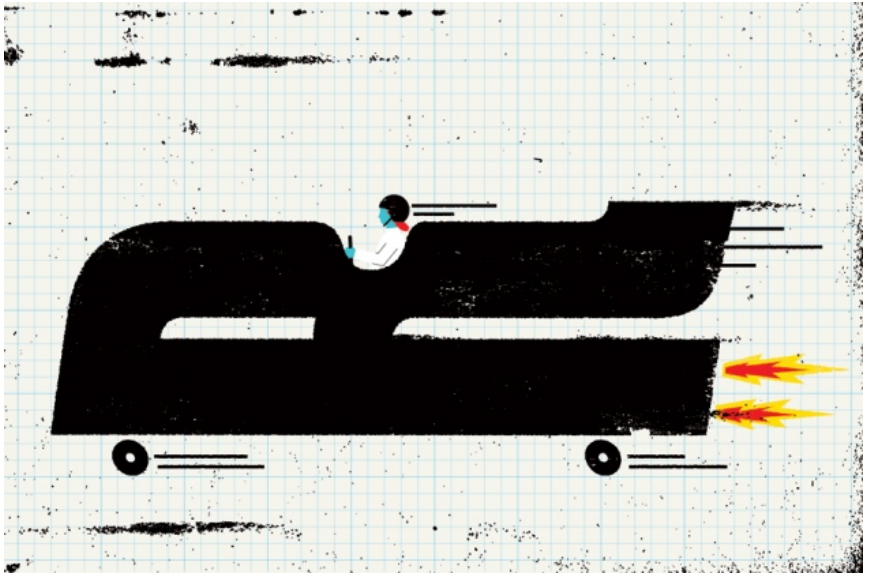
5. Rapports reproductibles

Introduction

- Le matériel pour cet atelier est basé sur plusieurs ressources
- Voir ce lien pour une liste complète de références:
<https://github.com/sahirbhatnagar/atelier-R-GERAD>
- Je vous suggère les livres de Vincent Goulet et Hadley Wickham



C'est parti



Après cet atelier vous devriez être capables de

- Comprendre, créer et modifier les 4 objets de bases en R (vector, data.frame, matrix, list)
- Utiliser des fonctions de bases
- Importer un jeux de données à partir d'un fichier externe
- Créer un graphique

1. Présentation du langage R

Objectifs du chapitre

- ① Comprendre les avantages d'apprendre R
- ② Connaître la provenance du R et ses caractéristiques
- ③ Démarrer une session R et exécuter des commandes simples
- ④ Créer, modifier et sauvegarder ses propres fichiers de script R

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Pourquoi vous êtes là?

Le langage R gagne en popularité

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type






















Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Language Rank	Types	Spectrum Ranking
1. Java	  	100.0
2. C	  	99.9
3. C++	  	99.4
4. Python	 	96.5
5. C#	  	91.3
6. R		84.8
7. PHP		84.5
8. JavaScript	 	83.0
9. Ruby	 	76.2
10. Matlab		72.4

Les meilleurs langages de programmation en 2015 selon
IEEE Spectrum

Plus de 100 000 questions posées dans les forums

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

[Questions](#)[Tags](#)[Users](#)[Badges](#)[Unanswered](#)[Ask Question](#)

Tagged Questions

[info](#)[newest](#)[featured](#)[frequent](#)[votes](#)[active](#)[unanswered](#)

R is a free, open-source programming language and software environment for statistical computing, bioinformatics, and graphics. Please supplement your question with a minimal reproducible example. For statistical questions please use <http://stats.stackexchange.com>.

[learn more...](#) [improve tag wiki](#) [top users](#) [synonyms \(2\)](#)

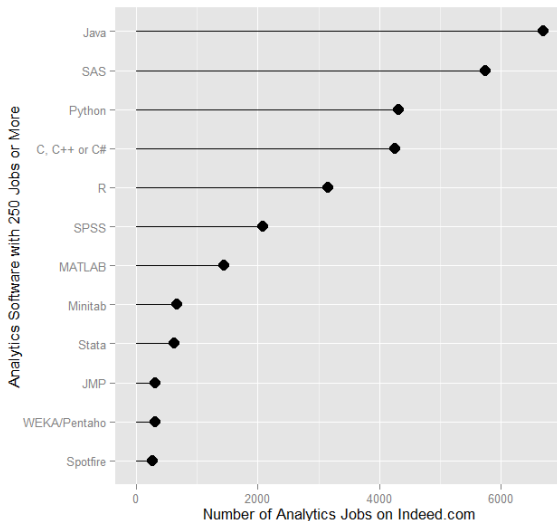
100,000
questions tagged

[r](#) [about »](#)

Hot Meta Posts

7 [What to do if I want to ask a](#)

Nombre d'emplois



référence: <http://r4stats.com/articles/popularity/>

Utilisé dans plusieurs domaines

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

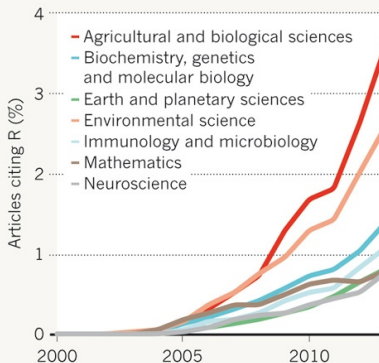
Sommaire

5. Rapports reproductibles

Introduction

A RISING TIDE OF R

An increasing proportion of research articles
explicitly reference R or an R package.



Publié dans *Nature*

Analyser vos données

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Plusieurs ressources d'outils statistiques
- Représenter graphiquement des jeux de données multivariées
- Intégrer votre code R dans des applications web
- Assurer la reproductivité de vos analyses

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Bref historique

À l'origine de R fut le S par John M. Chambers

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction



Figure 1 : S, un langage pour programmer avec des données, développé chez Bell Laboratories dans les années 1970 par une équipe de chercheurs menée par John M. Chambers

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction



Figure 2 : Inspirés par le S, Ross Ihaka (gauche) et Robert Gentleman (droite) de l'Université d'Auckland en Nouvelle-Zélande ont lancé la première version de R en 1996

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- 1990-2010: le S a principalement été popularisé par une mise en oeuvre commerciale nommée S-PLUS

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- 1990-2010: le S a principalement été popularisé par une mise en oeuvre commerciale nommée S-PLUS
- Fin des années 2000: L'utilisation de S-PLUS diminue en faveur du R, surtout dans les milieux académiques

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- 1990-2010: le S a principalement été popularisé par une mise en oeuvre commerciale nommée S-PLUS
- Fin des années 2000: L'utilisation de S-PLUS diminue en faveur du R, surtout dans les milieux académiques
- 2 raisons qui ont fortement contribué à la perte d'influence de S-PLUS
 - ① Disponible gratuitement

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- 1990-2010: le S a principalement été popularisé par une mise en oeuvre commerciale nommée S-PLUS
- Fin des années 2000: L'utilisation de S-PLUS diminue en faveur du R, surtout dans les milieux académiques
- 2 raisons qui ont fortement contribué à la perte d'influence de S-PLUS
 - ① Disponible gratuitement
 - ② Ouvert aux contributions de tous

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Caractéristiques de R

Langage de programmation orientée objet

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Cela permet de facilement trouver et ré-utiliser les résultats de tes analyses

Langage de programmation orientée objet

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Cela permet de facilement trouver et ré-utiliser les résultats de tes analyses
- Une fonction peut compléter plusieurs tâches

Langage de programmation interprété

- Langage de programmation interprété (versus C, C++, JAVA)
- Plus accessible qu'un langage compilé → ce qui permet aux économistes, écologistes, biologistes, statisticiens, épidémiologistes, etc. d'utiliser R

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Langage de programmation interprété

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Langage de programmation interprété (versus C, C++, JAVA)
- Plus accessible qu'un langage compilé → ce qui permet aux économistes, écologistes, biologistes, statisticiens, épidémiologistes, etc. d'utiliser R
- Le programme que l'on lance pour utiliser R est l'interprète

Langage de programmation interprété

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Langage de programmation interprété (versus C, C++, JAVA)
- Plus accessible qu'un langage compilé → ce qui permet aux économistes, écologistes, biologistes, statisticiens, épidémiologistes, etc. d'utiliser R
- Le programme que l'on lance pour utiliser R est l'interprète
- Celui-ci prend des commandes en R qu'il exécutera immédiatement

Langage de programmation interprété

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Langage de programmation interprété (versus C, C++, JAVA)
- Plus accessible qu'un langage compilé → ce qui permet aux économistes, écologistes, biologistes, statisticiens, épidémiologistes, etc. d'utiliser R
- Le programme que l'on lance pour utiliser R est l'interprète
- Celui-ci prend des commandes en R qu'il exécutera immédiatement
- Autre exemple: Python

Logiciel libre (*Open Source*)

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Développement actif pour la création de nouveaux outils dans plusieurs domaines
 - <https://cran.r-project.org/web/views/>
- Facilement voir le code des autres avec GitHub
 - <http://www.r-pkg.org/>
- Bien documenter avec beaucoup de ressources gratuites disponibles sur l'internet
 - [stackoverflow](https://stackoverflow.com/)
 - <http://www.rdocumentation.org/>
 - <http://www.r-bloggers.com/>
 - [twitter](https://twitter.com/rstats)
 - [R user groups](#)
 - [Google](#)

Outil statistique qui optimize l'approche matricielle

- Langage basé sur la notion de vecteur, ce qui simplifie les calculs mathématiques (non seulement la computation mais l'écriture aussi)

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Outil statistique qui optimize l'approche matricielle

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Langage basé sur la notion de vecteur, ce qui simplifie les calculs mathématiques (non seulement la computation mais l'écriture aussi)
- Réduit considérablement le recours aux structures itératives (boucles `for`, `while`, etc.)

Outil statistique qui optimize l'approche matricielle

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Langage basé sur la notion de vecteur, ce qui simplifie les calculs mathématiques (non seulement la computation mais l'écriture aussi)
- Réduit considérablement le recours aux structures itératives (boucles for, while, etc.)

R code 1.1

```
c(1,2,3) + c(4,5,6)
```

```
## [1] 5 7 9
```

Où trouver de l'aide pour une fonction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- `?nom_de_la_fonction`

Où trouver de l'aide pour une fonction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

- `?nom_de_la_fonction`

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

R code 1.2

```
# l'aide pour la fonction de régression linéaire 'lm'  
?lm
```

La rubrique d'aide

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- **Usage:** le nom de la fonction, tous ses arguments et les valeurs par défaut

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- **Usage:** le nom de la fonction, tous ses arguments et les valeurs par défaut
- **Value:** le type d'objet retourné par la fonction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Démarrer une session

R code 1.3

```
# Démarrer l'interface pour la documentation
# et naviguer les différentes ressources
help.start()

# trouver l'aide pour la fonction rnorm
?rnorm

# Connaitre le répertoire de travail
getwd()
```

R code 1.4

```
# On additionne
```

```
39 + 3
```

```
# On soustrait
```

```
58 - 16
```

```
# On multiplie
```

```
6 * 7
```

```
# Et on peut même diviser
```

```
8 / 3
```

R code 1.5

```
# Générer deux vecteurs de nombres pseudo-aléatoires  
# issus d'une loi normale centrée réduite.  
x <- rnorm(50)  
y <- rnorm(50)  
  
# Graphique des couples (x, y)  
plot(x, y)  
  
# Graphique d'un histogramme de x  
hist(x)
```

R code 1.6

```
# voir la matière de x  
x
```

```
# voir les objets de votre workspace  
ls()
```

```
# supprimer les deux vecteurs x et y  
rm(x,y)
```

```
# voir la matière de x  
x
```

```
# voir les objets de votre workspace  
ls()
```

R code 1.7

```
# Générer la suite 1, 2, ..., 20.  
x <- 1:20  
  
# créer un autre vecteur en fonction de x  
y <- 2*x+3  
  
# créer un data frame de deux colonnes et  
# voir sa matière  
dt <- data.frame(x, y)  
dt  
  
# estimer un modèle linéaire et voir les  
# résultats  
fit <- lm(y ~ x, data = dt)  
summary(fit)
```

2. Bases du langage R

Objectifs du chapitre

- ① Comprendre ce qui est une fonction et comment l'utiliser
- ② Identifier les principaux types d'objets dans R: `vector`, `matrix`, `data frame` et `list`
- ③ Créer et manipuler ces objets
- ④ Importer des fichiers `.txt` et `.csv`
- ⑤ Installer un *package*

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Les fonctions

Définition et appel de fonctions

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Plutôt que d'écrire plusieurs lignes de codes pour des choses assez simples et répétitives on fait appel à des **fonctions**

Définition et appel de fonctions

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Plutôt que d'écrire plusieurs lignes de codes pour des choses assez simples et répétitives on fait appel à des **fonctions**
- Une fonction est caractérisée par deux composantes:

Définition et appel de fonctions

1. Présentation du langage R

Pourquoi êtes-vous
là ?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

- Plutôt que d'écrire plusieurs lignes de codes pour des choses assez simples et répétitives on fait appel à des **fonctions**
- Une fonction est caractérisée par deux composantes:
 - ① son **nom**: c'est ce nom qui nous permettra d'appeler la fonction

Définition et appel de fonctions

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Plutôt que d'écrire plusieurs lignes de codes pour des choses assez simples et répétitives on fait appel à des **fonctions**
- Une fonction est caractérisée par deux composantes:
 - ① son **nom**: c'est ce nom qui nous permettra d'appeler la fonction
 - ② une **liste d'arguments**: ce sont les informations dont la fonction a besoin pour travailler et renvoyer son résultat

La syntaxe pour appeler une fonction

- Il faut écrire son nom suivi de parenthèses

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La syntaxe pour appeler une fonction

- Il faut écrire son nom suivi de parenthèses
- Entre ces parenthèses, nous indiquerons les arguments dont la fonction a besoin:

```
nom_de_la_fonction(arguments)
```

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La syntaxe pour appeler une fonction

- Il faut écrire son nom suivi de parenthèses
- Entre ces parenthèses, nous indiquerons les arguments dont la fonction a besoin:

`nom_de_la_fonction(arguments)`

R code 2.1

```
# chercher la rubrique d'aide pour racine carrée
?sqrt

# on voit que la fonction prend un argument
sqrt(49)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

Arguments facultatifs

- Certaines fonctions nécessitent plusieurs arguments et certains de ces arguments sont optionnels (aussi possible d'avoir aucun argument: `getwd()`)

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Arguments facultatifs

- Certaines fonctions nécessitent plusieurs arguments et certains de ces arguments sont optionnels (aussi possible d'avoir aucun argument: `getwd()`)
- Les arguments optionnels possèdent une valeur par défaut

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Arguments facultatifs

- Certaines fonctions nécessitent plusieurs arguments et certains de ces arguments sont optionnels (aussi possible d'avoir aucun argument: `getwd()`)
- Les arguments optionnels possèdent une valeur par défaut
- Si on ne spécifie pas de valeur, ces arguments prendront la valeur définie par défaut, sinon ils prendront la valeur spécifiée

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Arguments facultatifs

- Certaines fonctions nécessitent plusieurs arguments et certains de ces arguments sont optionnels (aussi possible d'avoir aucun argument: `getwd()`)
- Les arguments optionnels possèdent une valeur par défaut
- Si on ne spécifie pas de valeur, ces arguments prendront la valeur définie par défaut, sinon ils prendront la valeur spécifiée

R code 2.2

```
# chercher la rubrique d'aide pour logarithme
?log

# on ne précise pas de valeur pour le second argument
log(2)

# On peut spécifier le second argument en l'appelant par son nom
log(2, base = exp(1))
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Les objets R

Les objets R

1. Présentation du langage R

Pourquoi êtes-vous là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Table 1 : Les principaux objets en R et le nom de leurs fonctions

dimension	homogène ^a	hétérogène ^b
1d	Atomic vector <code>c()</code>	List <code>list()</code>
2d	Matrix <code>matrix()</code>	Data frame <code>data.frame()</code>

^a tous les éléments doivent être du même type

^b les éléments peuvent être de différents types

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- En R, à toutes fins pratiques, tout est un vecteur

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- En R, à toutes fins pratiques, tout est un vecteur
- La fonction de base pour créer un vecteur est `c()` (concaténation)

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- En R, à toutes fins pratiques, tout est un vecteur
- La fonction de base pour créer un vecteur est `c()` (concaténation)

R code 2.3

```
c(1, 2, 5)
```

```
## [1] 1 2 5
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les quatre types d'*atomic vectors* les plus utilisés:
 - ① *double* (également appelé *numeric*)
 - ② *integer*
 - ③ *character*
 - ④ *logical*

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les quatre types d'*atomic vectors* les plus utilisés:
 - ① *double* (également appelé *numeric*)
 - ② *integer*
 - ③ *character*
 - ④ *logical*

R code 2.4

```
c(1, 2.5, 4.5) # numeric
c(1L, 6L, 10L) # integer
c("ceci sont", "des characters") #character
c(TRUE, FALSE, T, F) # logical
```

- Tous les éléments d'un *atomic vector* doivent être du même type

1. Présentation du langage R

Pourquoi êtes-vous là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Tous les éléments d'un *atomic vector* doivent être du même type
- Lorsque vous essayez de combiner différents types, ils seront convertis au type le plus flexible

1. Présentation du langage R

Pourquoi êtes-vous là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Coercion

- Tous les éléments d'un *atomic vector* doivent être du même type
- Lorsque vous essayez de combiner différents types, ils seront convertis au type le plus flexible
- Types du moins au plus flexible sont:
 - ① *logical*
 - ② *integer*
 - ③ *double*
 - ④ *character*

R code 2.5

```
# combiner un character et integer donne un character
str(c("a", 1))
```

```
## chr [1:2] "a" "1"
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- La plupart des opérations mathématiques vont convertir un *atomic vector* à un *double* ou *integer*

R code 2.6

```
x <- c(FALSE, FALSE, TRUE)
as.numeric(x)
```

```
## [1] 0 0 1
```

```
# Nombre total de TRUE
sum(x)
```

```
## [1] 1
```

```
# La proportion de TRUE
mean(x)
```

```
## [1] 0.33
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les *lists* sont différent des *atomic vectors* parce que leurs éléments peuvent être de tout type
- La fonction pour créer un *list* est `list()`

1. Présentation du langage R

Pourquoi êtes-vous là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les *lists* sont différent des *atomic vectors* parce que leurs éléments peuvent être de tout type
- La fonction pour créer un *list* est `list()`

R code 2.7

```
(x <- list(1:3, "a", c(TRUE, FALSE, TRUE), c(2.3, 5.9)))

## [[1]]
## [1] 1 2 3
##
## [[2]]
## [1] "a"
##
## [[3]]
## [1] TRUE FALSE TRUE
##
## [[4]]
## [1] 2.3 5.9
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les matrices ne sont rien d'autre que des vecteurs en 2 dimensions

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les matrices ne sont rien d'autre que des vecteurs en 2 dimensions
- Surtout utilisées pour faire les calculs mathématiques (pensez à vos cours d'algèbre linéaire)

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les matrices ne sont rien d'autre que des vecteurs en 2 dimensions
- Surtout utilisées pour faire les calculs mathématiques (pensez à vos cours d'algèbre linéaire)
- La fonction pour créer un *matrix* est `matrix()`

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les matrices ne sont rien d'autre que des vecteurs en 2 dimensions
- Surtout utilisées pour faire les calculs mathématiques (pensez à vos cours d'algèbre linéaire)
- La fonction pour créer un *matrix* est `matrix()`

R code 2.8

```
# rempli par colonne par défaut  
matrix(c(1,2,3,4,5,6), nrow = 2, ncol = 3)
```

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

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

R code 2.9

```
# automatiquement converti en données homogènes  
matrix(c(1,2,3,"a","b","c"), nrow = 2, ncol = 3)
```

```
##           [,1] [,2] [,3]  
## [1,]  "1"   "3"   "b"  
## [2,]  "2"   "a"   "c"
```

- Le *data frame* est un tableau de données hétérogènes

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Data frame

- Le *data frame* est un tableau de données hétérogènes
- La fonction pour créer un *data frame* est `data.frame()`

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Data frame

- Le *data frame* est un tableau de données hétérogènes
- La fonction pour créer un *data frame* est `data.frame()`

R code 2.10

```
# une colonne 'id' pour identifier les sujets  
# une colonne 'age' pour leur age  
data.frame(id = c("didier", "patrice", "laurent"),  
           age = c(37, 34, 32))
```

```
##           id age  
## 1 didier   37  
## 2 patrice  34  
## 3 laurent 32
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La différence entre *data frame* et *matrix*?

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- `matrix`: calculs matriciels
- `data.frame`: toutes autres analyses de données hétérogènes

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Commandes R

Expression et affectation

- 1 Une **expression** est immédiatement évaluée et le résultat est affiché à l'écran:

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Expression et affectation

- 1 Une **expression** est immédiatement évaluée et le résultat est affiché à l'écran:

R code 2.11

```
2 + 3
```

```
## [1] 5
```

```
pi
```

```
## [1] 3.1
```

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

```
## [1] 0.71
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Expression et affectation

② 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 `<-`
- les deux caractères `<` et `-` placés obligatoirement l'un à la suite de l'autre:

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Expression et affectation

③ 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 `<-`
- les deux caractères `<` et `-` placés obligatoirement l'un à la suite de l'autre:

R code 2.12

```
a <- 5
```

```
a
```

```
## [1] 5
```

```
b <- a - 2
```

```
b
```

```
## [1] 3
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Indexation d'un *vector*

- L'indexation sert principalement à deux choses

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Indexation d'un *vector*

- L'indexation sert principalement à deux choses
 - ① extraire des éléments
 - ② remplacer des éléments

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Indexation d'un *vector*

- L'indexation sert principalement à deux choses

- ① extraire des éléments
- ② remplacer des éléments

R code 2.13

```
# création d'un vecteur  
x <- c(a = -1, b = 2, c = 8, d = 10)
```

```
# extraction par position  
x[1]
```

```
# extraction par étiquette  
x["c"]
```

```
# remplacer le deuxième élément par 5  
x[2] <- 5
```

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Indexation d'un *data frame* et *matrix*

- L'indexation des observations se fait à partir des rangées et colonnes du *data frame* et *matrix*:

`d[rangée, colonne]`

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Indexation d'un *data frame* et *matrix*

- L'indexation des observations se fait à partir des rangées et colonnes du *data frame* et *matrix*:

`d[rangée,colonne]`

R code 2.14

```
# création d'un data frame
d <- data.frame(Noms = c("Pierre", "Jean", "Jacques"),
  Age = c(42, 34, 19),
  Fumeur = c(TRUE, TRUE, FALSE))

d[1, ] # première rangée

d[ ,1] # première colonne

d[3,2] # troisième rangée, deuxième colonne
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Indexation d'un *list*

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

R code 2.15

```
# création d'un list
x <- list(joueur = c("V", "C"),
          score = c(10, 12))

# premier élément de la list
x[[1]]

# 1er élément du 2e élément de la list
x[[2]][1]
```

Importer et enregistrer des données et du code

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Table 2 : Les fonctions de bases pour importer et enregistrer des données et du code R

	fichiers	objets ^c	script R
importer	<code>read.table()</code> ^a <code>read.csv()</code> ^b	<code>load()</code>	<code>source()</code>
enregistrer	<code>write.table()</code> <code>write.csv()</code>	<code>save()</code>	File -> Save As..

^a chaque valeur est séparée par un espace

^b chaque valeur est séparée par une virgule

^c objets dans ton espace de travail (Environment)

Importer des données

- Il est nécessaire de préciser le chemin d'accès au fichier s'il n'est pas dans le répertoire de travail

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Importer des données

- Il est nécessaire de préciser le chemin d'accès au fichier s'il n'est pas dans le répertoire de travail
 - `getwd()`: connaître le répertoire de travail

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Importer des données

- Il est nécessaire de préciser le chemin d'accès au fichier s'il n'est pas dans le répertoire de travail
 - `getwd()`: connaître le répertoire de travail
 - `setwd()`: modifier le répertoire de travail

1. Présentation du langage R

Pourquoi êtes-vous là ?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

Importer des données

- Il est nécessaire de préciser le chemin d'accès au fichier s'il n'est pas dans le répertoire de travail
 - `getwd()`: connaître le répertoire de travail
 - `setwd()`: modifier le répertoire de travail

R code 2.16

```
# modifier le répertoire de travail
setwd("~/git_repositories/atelier-R-GERAD/data")

# importer les fichiers
# affectation aux objets 'lung' et 'admit'
lung <- read.csv("lung.csv")

admit <- read.table("admit.txt")
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Importer du code R à partir d'un script

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

R code 2.17

```
# modifier le répertoire de travail
setwd("~/git_repositories/atelier-R-GERAD/script")

# exécute les commandes R du script 'mtcars.R'
# et montre le output
source("mtcars.R", echo = TRUE)

# sauvegarde les objets du environnement
save(df, fit, file = "mtcars.RData")

# supprime les objets de ton environnement
rm(df, fit)

# importer les objets R de 'mtcars.RData'
load("mtcars.RData")
```

Installer des *packages* additionnels dans R

- Un **package** R est un ensemble cohérent de fonctions, de jeux de données et de documentation permettant de compléter les fonctionnalités du système ou d'en ajouter de nouvelles
- La liste complète de ces packages est disponible sur le site [METACRAN](#)

1. Présentation du langage R

Pourquoi êtes-vous là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Installer des *packages* additionnels dans R

- Un **package** R est un ensemble cohérent de fonctions, de jeux de données et de documentation permettant de compléter les fonctionnalités du système ou d'en ajouter de nouvelles
- La liste complète de ces packages est disponible sur le site [METACRAN](#)

R code 2.18

```
# installer des packages pour créer des rapports  
# reproductibles  
install.packages(c("knitr", "rmarkdown"))  
  
# pour avoir accès aux fonctions dans ces packages  
library(knitr)  
library(rmarkdown)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Sommaire

Sommaire des commandes de bases

1. Présentation du langage R

Pourquoi êtes-vous
là?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

Espace de travail

```
sessionInfo()  
install.packages()  
library()  
setwd()  
getwd()  
rm()  
ls()
```

Objets R les plus utilisés

```
c()  
data.frame()  
matrix()  
list()
```

Accéder les données et scripts R

```
read.table()  
write.table()  
load()  
save()  
source()
```

3. Graphiques

Objectifs du chapitre

- ① Comprendre la syntaxe de la fonction plot
- ② Créer des boxplot, histogram et density plot
- ③ Comment sauvegarder des graphiques

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction plot

Introduction aux graphiques en R

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- R offre une variété de graphiques remarquables
- Pour avoir une petite idée des possibilités offertes, il suffit de taper la commande `demo(graphics)`

La fonction plot

- La fonction plot peut prendre plusieurs arguments et types d'objets

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction plot

- La fonction plot peut prendre plusieurs arguments et types d'objets

R code 3.1

```
# Plot des valeurs d'un vecteur contre leurs indices
# équivalent à plot(mtcars[,"mpg"])
plot(mtcars$mpg, xlab = "Index", ylab = "mpg",
      main = "Titre")

# Graphique des couples (x, y)
plot(mtcars$mpg, mtcars$disp, xlab = "mpg",
      ylab = "disp", main = "mpg vs. disp")
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction plot

- La fonction plot peut aussi prendre un objet du modèle linéaire comme argument

1. Présentation du langage R

Pourquoi êtes-vous là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction plot

- La fonction plot peut aussi prendre un objet du modèle linéaire comme argument
- Ceci est un exemple qui démontre pourquoi R est un langage orientée objet

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction plot

- La fonction plot peut aussi prendre un objet du modèle linéaire comme argument
- Ceci est un exemple qui démontre pourquoi R est un langage orientée objet

R code 3.2

```
# importer 'mtcars.RData'  
load("mtcars.RData")  
  
# Graphiques des 4 diagnostics  
# du modèle linéaire  
# placer dans 2 rangés et 2 colonnes  
par(mfrow=c(2,2))  
plot(fit)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

**histogram et
boxplot**

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

histogram et boxplot

La fonction `histogram`

- Utile pour voir la distribution des données
- Pour les données continues et univariées

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction `plot`

**`histogram` et
`boxplot`**

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction histogram

- Utile pour voir la distribution des données
- Pour les données continues et univariées

R code 3.3

```
# la taille de 237 étudiants disponibles dans le jeu  
# de données 'survey' du library(MASS)  
library(MASS)  
  
# voir le nom des colonnes  
names(survey)  
  
# histogram de la taille et montrer la fréquence  
# de chaque barre  
hist(survey$Height, labels = TRUE)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

La fonction boxplot

- Utile pour voir la différence d'une variable continue parmi plusieurs groupes

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction boxplot

- Utile pour voir la différence d'une variable continue parmi plusieurs groupes

R code 3.4

```
# visualiser la différence de taille entre les  
# hommes et les femmes dans le jeux  
# de données 'survey' du library(MASS)
```

```
boxplot(survey$Height ~ survey$Sex,  
        ylab = "Taille (cm)",  
        col = c("lightpink","lightblue"))
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Sauvegarder les graphiques

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- 2 fonctions principales: `pdf()` et `png()`
- l'argument de ces fonctions est le nom du fichier désiré

Sauvegarder les graphiques

- 2 fonctions principales: `pdf()` et `png()`
- l'argument de ces fonctions est le nom du fichier désiré

R code 3.5

```
# enregistrer dans le répertoire de travail courant
pdf("boxplot_survey.pdf")
boxplot(survey$Height ~ survey$Sex,
        ylab = "Taille (cm)")
dev.off()
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Sommaire

Sommaire des commandes de graphiques

1. Présentation du langage R

Pourquoi êtes-vous
là?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

Fonctions pour créer les graphiques

```
plot()  
hist()  
boxplot()
```

Arguments communs pour ces fonctions

```
xlab  
ylab  
main  
type  
col
```

Fonctions pour enregistrer les graphiques

```
pdf()  
png()  
dev.off()
```

4. Statistiques

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Moyenne, écart type

Moyenne, variance, écart type, minimum, maximum

- fonctions principales: `mean()`, `var()`, `sd()`, `min()`, `max()`

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Moyenne, variance, écart type, minimum, maximum

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- fonctions principales: `mean()`, `var()`, `sd()`, `min()`, `max()`
- la fonction `summary()` est utile pour calculer quelques statistiques de bases pour un *data frame*

Moyenne, variance, écart type, minimum, maximum

1. Présentation du langage R

Pourquoi êtes-vous
là?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

- fonctions principales: `mean()`, `var()`, `sd()`, `min()`, `max()`
- la fonction `summary()` est utile pour calculer quelques statistiques de bases pour un *data frame*

R code 4.1

```
# enregistrer dans le répertoire de travail courant  
summary(mtcars)
```

La fonction apply

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Pour calculer des statistiques plus compliquées sur un *data frame* (ou *matrix*), on utilise la fonction `apply`

La fonction `apply`

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction `plot`

histogram et

`boxplot`

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Pour calculer des statistiques plus compliquées sur un *data frame* (ou *matrix*), on utilise la fonction `apply`
- La fonction `apply` sert à appliquer une fonction quelconque sur une partie d'un *matrix* ou *data frame*

La fonction apply

- La syntaxe de la fonction est la suivante:

`apply(X, MARGIN, FUN)`

- X: un matrix ou data frame
- MARGIN: 1 si l'on veut faire des calculs sur les rangées, 2 sur les colonnes
- FUN: est la fonction à appliquer

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction apply

- La syntaxe de la fonction est la suivante:

`apply(X, MARGIN, FUN)`

- X: un matrix ou data frame
- MARGIN: 1 si l'on veut faire des calculs sur les rangées, 2 sur les colonnes
- FUN: est la fonction à appliquer

R code 4.2

```
# variance par ligne  
apply(mtcars, 1, var)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

La fonction apply

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et
boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

R code 4.3

```
# variance par colonne
```

```
apply(mtcars, 2, var)
```

```
# écart type par colonne
```

```
apply(mtcars, 2, sd)
```

```
# minimum de chaque rangée
```

```
apply(mtcars, 1, min)
```

```
# maximum de chaque rangée
```

```
apply(mtcars, 1, max)
```

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

**Test t et analyse de
variance**

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Test t et analyse de variance

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

**Test t et analyse de
variance**

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- la fonction pour faire un test t est `t.test()`

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- la fonction pour faire un test t est `t.test()`

R code 4.4

```
# la taille de 237 étudiants disponible dans le jeux  
# de données 'survey' du library(MASS)  
# est-ce qu'il y a une différence de taille entre  
# les hommes et les femmes?
```

```
t.test(Height ~ Sex, data = survey)
```

Analyse de variance

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

**Test t et analyse de
variance**

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- La fonction pour l'analyse de variance est aov

Analyse de variance

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

**Test t et analyse de
variance**

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- La fonction pour l'analyse de variance est aov
- Prenons un jeux de données disponible dans R:
InsectSprays

Analyse de variance

1. Présentation du langage R

Pourquoi êtes-vous
là ?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- La fonction pour l'analyse de variance est aov
- Prenons un jeux de données disponible dans R:
InsectSprays
- 6 insecticides ont été testés 12 fois en culture, la réponse observée étant le nombre d'insectes

Analyse de variance

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- La fonction pour l'analyse de variance est `aov`
- Prenons un jeux de données disponible dans R:
`InsectSprays`
- 6 insecticides ont été testés 12 fois en culture, la réponse observée étant le nombre d'insectes
- Le but c'est de voir s'il y a une différence importante entre les 6 insecticides par rapport à le nombre d'insectes

Analyse de variance

1. Présentation du langage R

- Pourquoi êtes-vous
là?
- Bref historique
- Caractéristiques de R
- Démarrer une session

2. Bases du langage R

- Les fonctions
- Les objets R
- Commandes R
- Sommaire

3. Graphiques

- La fonction plot
- histogram et
boxplot
- Sommaire

4. Statistiques

- Moyenne, écart type
- Test t et analyse de
variance
- Modèles linéaires
- Sommaire

5. Rapports reproductibles

- Introduction

R code 4.5

```
# boxplot pour voir la différence entre  
# les groupes  
boxplot(InsectSprays$count ~ InsectSprays$spray)  
  
# Les résultats ne sont pas affichés, ceux-ci sont  
# copiés dans un objet nommé aov.spray  
aov.spray <- aov(count ~ spray, data = InsectSprays)  
  
# sommaire des résultats  
summary(aov.spray)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Modèles linéaires

Les formules

- Les formules représentent un élément-clé des analyses statistiques avec R

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Les formules

- Les formules représentent un élément-clé des analyses statistiques avec R
- La notation utilisée est la même pour (presque) toutes les fonctions

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Les formules

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les formules représentent un élément-clé des analyses statistiques avec R
- La notation utilisée est la même pour (presque) toutes les fonctions
- Une formule est typiquement de la forme

$$y \sim \text{model}$$

Les formules

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les formules représentent un élément-clé des analyses statistiques avec R
- La notation utilisée est la même pour (presque) toutes les fonctions
- Une formule est typiquement de la forme

$$y \sim \text{model}$$

- ① y : est la réponse analysée
- ② model : est un ensemble de termes pour lesquels les paramètres sont estimés

Les formules

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- Les formules représentent un élément-clé des analyses statistiques avec R
- La notation utilisée est la même pour (presque) toutes les fonctions
- Une formule est typiquement de la forme
$$y \sim \text{model}$$
 - ① y : est la réponse analysée
 - ② model : est un ensemble de termes pour lesquels les paramètres sont estimés
- Ces termes sont séparés par des symboles arithmétiques mais qui ont ici une signification particulière: $a+b$, $a*b$

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

R code 4.6

```
# importer 'admit.txt'
admit <- read.table("admit.txt", header = TRUE)

# est-ce que gpa et rank sont reliés à gre
fit <- lm(gre ~ gpa+rank, data = admit)

# voir les résultats
summary(fit)
```

1. Présentation du langage R

- Pourquoi êtes-vous
là?
- Bref historique
- Caractéristiques de R
- Démarrer une session

2. Bases du langage R

- Les fonctions
- Les objets R
- Commandes R
- Sommaire

3. Graphiques

- La fonction plot
- histogram et
boxplot
- Sommaire

4. Statistiques

- Moyenne, écart type
- Test t et analyse de
variance
- Modèles linéaires
- Sommaire

5. Rapports reproductibles

- Introduction

R code 4.7

```
# régression logistique
fit.glm <- glm(admit ~ gre+gpa+rank, data = admit,
               family = binomial(link = "logit"))

# voir les résultats
summary(fit.glm)
```

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Sommaire

Sommaire des commandes de statistiques

1. Présentation du langage R

Pourquoi êtes-vous
là ?
Bref historique
Caractéristiques de R
Démarrer une session

2. Bases du langage R

Les fonctions
Les objets R
Commandes R
Sommaire

3. Graphiques

La fonction plot
histogram et
boxplot
Sommaire

4. Statistiques

Moyenne, écart type
Test t et analyse de
variance
Modèles linéaires
Sommaire

5. Rapports reproductibles

Introduction

Fonctions pour des statistiques descriptives

```
mean()  
var()  
sd()  
min()  
max()  
median()  
apply()
```

Tests statistiques

```
aov()  
t.test()  
chisq.test()
```

Fonctions pour les modèles de régression

```
lm()  
glm()
```

5. Créer des rapports

1.Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2.Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

Introduction

Introduction

1. Présentation du langage R

Pourquoi êtes-vous
là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

- <http://rmarkdown.rstudio.com/>
- <http://shiny.rstudio.com/gallery/>

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de
variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

1. Présentation du langage R

Pourquoi êtes-vous là?

Bref historique

Caractéristiques de R

Démarrer une session

2. Bases du langage R

Les fonctions

Les objets R

Commandes R

Sommaire

3. Graphiques

La fonction plot

histogram et

boxplot

Sommaire

4. Statistiques

Moyenne, écart type

Test t et analyse de variance

Modèles linéaires

Sommaire

5. Rapports reproductibles

Introduction

- R version 3.2.1 (2015-06-18), x86_64-pc-linux-gnu
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: data.table 1.9.4, dplyr 0.4.1, ggplot2 1.0.1, knitr 1.10.5, xtable 1.7-4
- Loaded via a namespace (and not attached):
assertthat 0.1, chron 2.3-45, colorspace 1.2-6, DBI 0.3.1, digest 0.6.8, evaluate 0.7, formatR 1.2, grid 3.2.1, gtable 0.1.2, highr 0.5, magrittr 1.5, MASS 7.3-43, munsell 0.4.2, parallel 3.2.1, plyr 1.8.3, proto 0.3-10, Rcpp 0.11.6, reshape2 1.4.1, scales 0.2.5, stringi 0.5-5, stringr 1.0.0, tools 3.2.1