

# TABLE DES MATIÈRES

<b>Avant-propos .....</b>	11
<b>Données utilisées comme exemples dans ce livre .....</b>	13
<b>PREMIÈRE PARTIE</b>	
<b>Découvrir R .....</b>	17
 <b>1 Introduction à R .....</b>	19
1.1 Origines de R .....	19
1.1.1 R et S-Plus .....	19
1.1.2 CRAN et projet R .....	19
1.1.3 Logiciels utilisant le langage R .....	20
1.2 Fichiers gérés par R .....	21
1.2.1 Scripts .....	21
1.2.2 Objets .....	21
1.2.3 Environnements .....	21
1.2.4 Historique de commandes .....	22
1.2.5 Gestion de la mémoire .....	22
1.3 Interfaces pour la programmation R .....	22
1.3.1 Interface R basique .....	22
1.3.2 Interface RStudio .....	24
 <b>2 Principes du langage .....</b>	33
2.1 Scripts .....	33
2.1.1 Ponctuation des scripts et règles d'écriture .....	33
2.1.2 Commentaires .....	34
2.1.3 Conventions de notation de syntaxe dans ce livre .....	34
2.2 Stockage d'information .....	34
2.2.1 Notion d'objet .....	34
2.2.2 Noms des objets .....	35
2.2.3 Supprimer un objet .....	35
2.3 Fonctions : dans R tout est fonction .....	35
2.3.1 Fonctions et opérateurs .....	35
2.3.2 Résultat d'une fonction, affectation .....	36
2.3.3 Appel de fonction, transfert d'arguments, ellipse .....	36
2.3.4 Fonctions génériques .....	37
2.3.5 Vectorisation .....	38
2.3.6 Création de fonctions .....	38
2.4 Packages .....	39
2.4.1 Principe d'un package .....	39
2.4.2 Récupération d'un package .....	39
2.4.3 Chargement d'un package .....	40
2.4.4 Packages utilisés dans ce livre .....	41



## Table des matières

<b>2.4.5</b>	Version des packages .....	44
<b>2.4.6</b>	Identification « du » bon package .....	44
<b>2.5</b>	Principaux types de données .....	46
<b>2.5.1</b>	Textes .....	46
<b>2.5.2</b>	Numériques et entiers .....	46
<b>2.5.3</b>	Dates .....	46
<b>2.5.4</b>	Booléens .....	47
<b>2.5.5</b>	Facteurs .....	47
<b>2.5.6</b>	Formules .....	48
<b>2.5.7</b>	Chemins .....	48
<b>2.5.8</b>	Valeurs spéciales .....	49
<b>2.5.9</b>	Conversions .....	49
<b>2.6</b>	Principales structures de données .....	50
<b>2.6.1</b>	Vecteurs et listes .....	50
<b>2.6.2</b>	Matrices et arrays .....	51
<b>2.6.3</b>	Data.frame, data.table, tibbles .....	51
<b>2.6.4</b>	Types spécifiques à des fonctions .....	52
<b>2.6.5</b>	Fonctions de conversion .....	52
<b>DEUXIÈME PARTIE</b>		
	Manipuler des données .....	53
<b>3</b>	<b>Récupération de données</b> .....	55
<b>3.1</b>	Import de fichiers externes .....	55
<b>3.1.1</b>	Fichiers plats .....	55
<b>3.1.2</b>	Classeurs Excel .....	60
<b>3.1.3</b>	Fichiers statistiques: SAS, SPSS et Stata .....	61
<b>3.2</b>	Connexion à une base de données .....	63
<b>3.3</b>	Récupération d'objets R sauvegardés .....	64
<b>3.3.1</b>	Fichiers Rdata .....	64
<b>3.3.2</b>	Fichiers RDS .....	65
<b>4</b>	<b>Filtres et requêtes</b> .....	67
<b>4.1</b>	Listes et résumés de données .....	67
<b>4.1.1</b>	Liste de données .....	67
<b>4.1.2</b>	Affichage du type des données .....	70
<b>4.1.3</b>	Statistiques sommaires .....	72
<b>4.1.4</b>	Taille et noms associés à un objet .....	73
<b>4.2</b>	Dimensions entre [ ] .....	74
<b>4.2.1</b>	Utilisation d'indices .....	74
<b>4.2.2</b>	Utilisation de booléens .....	74
<b>4.2.3</b>	Utilisation de noms d'éléments .....	75
<b>4.2.4</b>	Utilisation des crochets selon le type d'objet manipulé .....	75
<b>4.3</b>	Filtres et conditions .....	82
<b>4.3.1</b>	Utilisation d'indices à partir d'une condition .....	82
<b>4.3.2</b>	Filtres avec la fonction subset .....	85
<b>4.3.3</b>	Filtres avec le package dplyr .....	87

4.3.4 Filtres sur un data.table.....	90
4.3.5 Choix de la meilleure syntaxe.....	93
4.4 Tirages d'échantillons.....	95
  <b>Création de variables</b> .....	99
5.1 Vie des variables.....	99
5.1.1 Création de variables.....	99
5.1.2 Changement de nom de certaines variables.....	101
5.1.3 Suppression de variables.....	102
5.2 Formules systématiques.....	103
5.2.1 Fonctions pour les nombres .....	103
5.2.2 Fonctions pour les textes .....	107
5.2.3 Fonctions pour les Dates.....	113
5.2.4 Conversions.....	119
5.2.5 Gestion des facteurs .....	121
5.2.6 Gestion des valeurs manquantes.....	127
5.3 Formules conditionnelles.....	128
5.3.1 Condition unique.....	128
5.3.2 Conditions multiples.....	129
5.3.3 Affectation partielle .....	130
5.3.4 Appartenance à un quantile .....	131
5.4 Rangs, cumuls et blocs.....	131
5.4.1 Rangs.....	132
5.4.2 Cumuls.....	133
5.4.3 Repérage dans des blocs .....	134
5.4.4 Opérations par blocs .....	136
5.5 Vectorisation .....	137
5.6 Création d'une fonction sur mesure.....	140
  <b>Tri, jointure et transposition</b> .....	143
6.1 Tri et doublons .....	143
6.1.1 Tri simple .....	143
6.1.2 Tri d'un data.table.....	146
6.1.3 Repérage et suppression de doublons.....	147
6.2 Empilement et jointure: combiner des données.....	148
6.2.1 Empilement.....	149
6.2.2 Jointures .....	150
6.3 Transposition.....	154
6.3.1 Verticalisation.....	155
6.3.2 Pivot et statistiques .....	157

**TROISIÈME PARTIE****Produire des statistiques et des graphiques ..... 161****7 Statistiques descriptives ..... 163**

7.1 Statistiques simples ..... 163
7.1.1 Fonctions de base ..... 163
7.1.2 Gestion des données manquantes ..... 165
7.1.3 Production de plusieurs statistiques à la fois ..... 166
7.1.4 Production des mêmes statistiques sur toutes les variables ..... 169
7.2 Statistiques par groupes ..... 172
7.2.1 Cas général ..... 172
7.2.2 Agrégations statistiques avec le package dplyr ..... 174
7.2.3 Cas particulier des data.tables ..... 177
7.3 Tableaux de fréquence ..... 179
7.3.1 Fréquences et proportions ..... 179
7.3.2 Tableaux complets ..... 182
7.4 Tableaux statistiques sur mesure ..... 184
7.4.1 Organisation du tableau ..... 184
7.4.2 Statistiques du tableau ..... 186
7.4.3 Esthétique du tableau ..... 188

**8 Graphiques ..... 191**

8.1 Packages graphiques et paramètres ..... 191
8.1.1 Différents packages, différentes logiques ..... 191
8.1.2 Paramètres graphiques ..... 191
8.2 Nuages de points et courbes ..... 193
8.2.1 Nuages de points ..... 193
8.2.2 Courbes ..... 197
8.3 Diagrammes en bâtons et circulaires ..... 199
8.3.1 Diagrammes en bâtons ..... 199
8.3.2 Diagrammes circulaires ou camemberts ..... 205
8.3.3 Nuages de points qualitatifs ou <i>dot plots</i> ..... 207
8.4 Graphiques statistiques ..... 209
8.4.1 Histogrammes ..... 209
8.4.2 Boîtes à moustaches ou boxplots ..... 210
8.5 Ajout d'éléments personnalisés ..... 212
8.5.1 Titres ..... 212
8.5.2 Marqueurs isolés ..... 214
8.5.3 Lignes de référence ..... 216
8.5.4 Légende ..... 219
8.5.5 Axes ..... 222
8.6 Introduction au package ggplot2 ..... 224
8.6.1 Grammaire des graphiques ..... 224
8.6.2 Nuage de points ..... 226
8.6.3 Courbes ..... 227
8.6.4 Diagrammes en bâtons ..... 229
8.6.5 Boîtes à moustaches ..... 232
8.6.6 Éclatement ou <i>facetting</i> ..... 235

 <b>Export et reporting .....</b>	239
<b>9.1 Sauvegarde et export de données .....</b>	239
<b>9.1.1 Sauvegarde d'objets.....</b>	239
<b>9.1.2 Export vers Excel.....</b>	240
<b>9.1.3 Export en fichier plat.....</b>	243
<b>9.2 Reporting .....</b>	244
<b>9.2.1 Construction d'un document Rmd.....</b>	246
<b>9.2.2 Insertion de résultats dans un document Rmd.....</b>	253
<b>9.2.3 Création d'un document Word ou PowerPoint avec ReporteRs.....</b>	258
<b>9.2.4 Insertion de résultats dans un document ReporteRs.....</b>	265
<b>9.2.5 Reporting interactif avec shiny .....</b>	272
<b>Index.....</b>	277