

# Table des matières

<b>Introduction</b> . . . . .	9
<b>Chapitre 1. Observer et préparer un jeu de données</b> . . . . .	29
1.1. Constituer une base de données selon les objectifs posés . . . . .	29
1.2. Observer scrupuleusement la base de données . . . . .	31
1.2.1. Les données manquantes . . . . .	31
1.2.2. Hétérogénéité des variables . . . . .	32
1.2.3. Appréhender les répliques . . . . .	33
1.2.4. Nombre d'objets et de descripteurs . . . . .	34
1.2.5. La gestion des doubles zéros . . . . .	35
1.2.6. Adéquation objectifs/jeux de données . . . . .	35
1.3. Réduire le nombre de variables environnementales . . . . .	37
1.3.1. Quel intérêt ? . . . . .	37
1.3.2. La méthode d'Escoufier : réduire les variables en conservant le maximum d'informations . . . . .	37
1.3.3. La matrice de corrélation de Spearman : appréhender les redondances entre variables . . . . .	39
1.4. Éliminer les espèces rares . . . . .	42
1.4.1. Quel intérêt ? . . . . .	42
1.4.2. Méthode des valeurs médianes . . . . .	42
1.4.3. Méthode du tri par abondance (TPA) . . . . .	44
<b>Chapitre 2. Traitement préalable du jeu de données</b> . . . . .	47
2.1. Abondances, richesses et diversités spécifiques . . . . .	47
2.1.1. Abondances totales . . . . .	47

2.1.2. Richesses spécifiques . . . . .	49
2.1.3. Indices de diversité . . . . .	49
2.2. Transformation. . . . .	51
2.2.1. Les données quantitatives d'unités hétérogènes (par exemple, physico-chimie) . . . . .	51
2.2.1.1. Standardisation/centrage-réduction . . . . .	51
2.2.1.2. Conséquences en termes de déformation de données . . . . .	52
2.2.2. Les données quantitatives d'unités homogènes (par exemple, flore et faune) . . . . .	53
2.2.2.1. Réduire les gammes de variation . . . . .	53
2.2.2.2. Dégradation des données quantitatives en données qualitatives (binaires ou autres). . . . .	54
2.2.3. Les données qualitatives (par exemple, facteurs environnementaux) . . . . .	55
2.2.4. Tenter de normaliser des données . . . . .	56
2.3. Coefficients et matrices d'association . . . . .	56
2.3.1. Comment choisir son coefficient ? . . . . .	56
2.3.2. Mode direct. . . . .	59
2.3.2.1. Cas des coefficients symétriques : « double zéro » important . . . . .	59
2.3.2.2. Cas des coefficients asymétriques : « double zéro » non considéré. . . . .	59
2.3.3. Mode indirect. . . . .	60

**Chapitre 3. Structure sous forme de groupes d'objets/variables. . . . . 61**

3.1. Les analyses de groupement les plus utilisées . . . . .	61
3.1.1. Comment choisir l'algorithme de groupement ? . . . . .	62
3.1.2. Groupements hiérarchiques . . . . .	62
3.1.2.1. Les algorithmes les plus utilisés. . . . .	62
3.1.2.2. Réaliser les groupements et visualiser les classifications . . . . .	65
3.1.2.3. Choisir l'algorithme de groupement le plus approprié <i>via</i> les distances cophénétiqes . . . . .	68
3.1.2.4. Choisir le nombre de groupes à considérer : méthode des profils de similarités (test SIMPROF) . . . . .	69
3.1.3. Classification non hiérarchique k-means . . . . .	72
3.1.4. Classification floue c-means . . . . .	73
3.2. Informations sur les descripteurs générant les groupes obtenus . . . . .	74
3.2.1. Représentation graphique simple . . . . .	75
3.2.2. Les analyses de variance . . . . .	75

3.2.3. Déterminer des espèces indicatrices. . . . .	77
3.2.3.1. Espèces spécifiques à un groupe et fidèles aux stations de ce groupe (IndVal) . . . . .	77
3.2.3.2. Espèces contribuant à la dissimilarité entre groupes (SIMPER) . . . . .	78

## **Chapitre 4. Structure sous forme de gradients d'objets/variables . . . . .**

81

4.1. Alternative paramétrique : les analyses factorielles . . . . .	81
4.1.1. Principe mathématique . . . . .	83
4.1.2. L'analyse en composantes principales (ACP) . . . . .	85
4.1.2.1. Quand utiliser l'ACP ? . . . . .	85
4.1.2.2. Principe mathématique . . . . .	85
4.1.2.3. Les indicateurs de qualité . . . . .	87
4.1.2.4. Représentation et interprétation des résultats de l'ACP . . . . .	90
4.1.2.5. Les variables ou individus supplémentaires . . . . .	92
4.1.3. L'analyse factorielle des correspondances (AFC) . . . . .	92
4.1.3.1. Quand utiliser l'AFC ? . . . . .	92
4.1.3.2. Principe mathématique . . . . .	93
4.1.3.3. Les indicateurs de qualité . . . . .	95
4.1.3.4. Représentation et interprétation des résultats de l'AFC . . . . .	97
4.1.3.5. Cas particulier : l'effet fer à cheval ou effet Guttman . . . . .	99
4.1.3.6. Les variables et individus supplémentaires . . . . .	100
4.1.4. L'analyse factorielle des correspondances multiples (AFCM) . . . . .	100
4.1.4.1. Quand utiliser l'AFCM ? . . . . .	100
4.1.4.2. Les indicateurs de qualité . . . . .	101
4.1.4.3. Représentation et interprétation des résultats de l'AFCM . . . . .	102
4.1.5. L'analyse en coordonnées principales (ACoP) . . . . .	103
4.1.5.1. Quand utiliser l'ACoP ? . . . . .	103
4.1.5.2. Les indicateurs de qualité . . . . .	104
4.1.5.3. Représentation et interprétation des résultats de l'ACoP . . . . .	106
4.2. Alternative non paramétrique : le positionnement multidimensionnel non métrique (NMDS) . . . . .	107
4.2.1. Principe mathématique . . . . .	107
4.2.2. Mode direct . . . . .	109
4.2.2.1. À partir d'une matrice d'association (par exemple, communautés phytoplanctoniques) . . . . .	109
4.2.2.2. À partir d'une base de données (par exemple, paramètres biologiques) . . . . .	110
4.2.3. Mode indirect . . . . .	111

<b>Chapitre 5. Comprendre une structure</b> . . . . .	<b>113</b>
5.1. Corréler une structure à une ou plusieurs autres sans hypothèse de causalité . . . . .	114
5.1.1. Corréler des groupes. . . . .	114
5.1.2. Corréler des matrices d'association . . . . .	116
5.1.3. Corréler différents tableaux de données . . . . .	118
5.1.3.1. Alternative paramétrique : l'analyse factorielle multiple (AFM) . . . . .	118
5.1.3.2. Alternative non paramétrique : l'analyse procrustéenne généralisée. . . . .	122
5.2. Expliquer une structure par d'autres variables . . . . .	126
5.2.1. Facteurs structurants des groupes . . . . .	126
5.2.1.1. Arbres de décision . . . . .	126
5.2.1.2. Analyses de variance . . . . .	130
5.2.2. Facteurs quantitatifs structurants des gradients . . . . .	139
5.2.2.1. Corrélations passives ( <i>a posteriori</i> ). . . . .	139
5.2.2.2. Régression active aux axes d'une analyse factorielle. . . . .	145
5.2.3. Facteurs qualitatifs structurants des gradients . . . . .	157
5.2.3.1. Alternative paramétrique : MANOVA et analyse factorielle discriminante. . . . .	157
5.2.3.2. Alternative intermédiaire : MANOVA par permutations ou PERMANOVA . . . . .	163
 <b>Conclusion</b> . . . . .	 <b>167</b>
 <b>Bibliographie</b> . . . . .	 <b>169</b>
 <b>Index</b> . . . . .	 <b>173</b>