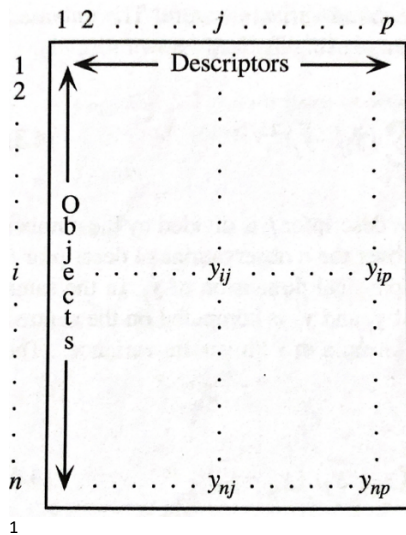


## Les bases de données écologiques

### CRÉER SA BASE DE DONNÉES

Ligne → : Objets ou unité d'échantillonnage

Colonne ↓ : Descripteurs ou variables



Objets (unités d'échantillonnage)

### Les types de descripteurs

- Les types mathématiques (Legendre et Legendre 2012)

The different mathematical types of descriptors.

Descriptor types	Ecological examples	
Binary (two states, presence-absence)	Species present or absent	
Multi-state (many states)		
Nonordered (qualitative, nominal, attributes)	Geological group	Facteur
Ordered		
Semiquantitative (rank-ordered, ordinal)	Importance or abundance scores	
Quantitative (metric, measurement)		Variable numérique
Discontinuous (meristic, discrete)	Equidistant abundance classes	
Continuous (metric)	Temperature, length	

<sup>1</sup> Legendre, Pierre, et Louis Legendre. 2012. *Numerical Ecology*. Third English edition. Amsterdam: Elsevier.

- Les types écologiques

Types de descripteurs	Description	Exemples écologiques
Métadonnées	Données qui fournissent de l'information sur une autre donnée (ou sur un groupe de données)	<ul style="list-style-type: none"> <li>- Observateur d'une mesure</li> <li>- Position GPS d'une parcelle</li> <li>- Parcelle à laquelle est associée un quadrat</li> <li>- Détails sur le protocole</li> </ul>
Données environnementales	Mesures, observations de l'environnement biotique ou abiotique de l'objet (de l'unité d'échantillonnage)	<ul style="list-style-type: none"> <li>- Variables physico-chimique</li> <li>- Ouverture de la canopée</li> <li>- Volume de bois mort au sol</li> </ul>
Données de communauté	Description de la communauté qui se trouve à l'unité d'échantillonnage	<ul style="list-style-type: none"> <li>- Présence-absence des espèces d'oiseaux</li> <li>- Abondance d'arthropodes à la famille</li> <li>- Recouvrement végétal</li> </ul>

### Organiser ses données

- Une ou plusieurs matrices de communautés
- Un ou plusieurs tableaux de données environnementales et de métadonnées
- Un fichier de métadonnées

Schéma de base de données

Schéma de votre base de données

**10 RÈGLES POUR LA SAISIE DE DONNÉES**

(Modifié de Broman et Woo 2018)<sup>2</sup>

1. Mettez un seul tableau par feuille Excel.
2. Indiquez une seule information par cellule.

Voici quelques règles à suivre :

- N'indiquez pas d'unité de mesure (indiquez les plutôt dans le fichier de métadonnées)
- Si vous avez des valeurs sous le seuil de détection d'un appareil, indiquez la valeur seuil et ajoutez une colonne pour renseigner si l'observation était sous le seuil ou pas.

3. Nommez correctement vos variables et vos fichiers.

Voici quelques règles à suivre :

- Utilisez des noms courts mais qui gardent du sens
- N'utilisez pas d'espace, si besoin utilisez un tiret (haut ou bas, mais toujours le même)
- N'employez pas de caractères spéciaux, et autres symboles ( \$ , @ , % , # , & , \* , ( , ) , ! , / , etc.)
- Gardez une cohérence à travers les variables, les feuilles et les fichiers.
- Gardez une cohérence pour coder vos variables (ex. pour le sexe des individus, utilisez toujours le même code et pas parfois « M », parfois « male » et parfois « m »).

4. Ne laissez pas de cellule vide et codez les données manquantes de manière unique.

- Faites bien la distinction entre 0 et NA

5. Utilisez un format texte pour les dates et la forme YYYY-MM-DD.

La gestion du format "date" sous Excel est problématique, vos données peuvent subir des modifications. Pour éviter ce problème, les dates doivent être entrées sous la forme YYYY-MM-DD (par exemple 2017-12-17), en format texte.

Pour forcer le format texte :

- Sélectionnez une colonne
- Dans la barre de menu, sélectionnez Format, puis Cellule
- Choisir "Texte"

Une autre alternative consiste à utiliser 3 colonnes : une pour l'année, l'autre pour le mois, et la troisième pour le jour.

---

<sup>2</sup> Broman, Karl W., et Kara H. Woo. 2018. « Data Organization in Spreadsheets ». *The American Statistician* 72(1): 2-10. Traduction française: <https://delladata.fr/12-recommandations-pour-organiser-efficacement-vos-donnees-dans-un-tableur/>

## 6. Renseignez chaque variable dans un fichier de métadonnées.

Chaque ligne correspond à une colonne de votre jeu de données, autrement dit à une variable (descripteur) et les colonnes peuvent contenir les informations suivantes :

- Le nom de la variable dans le jeu de données
- Le nom de la variable non-abrégée
- L'unité
- Une explication
- Les valeurs min et max
- La signification des codes utilisés pour les niveaux d'un facteur

Ces informations sont primordiales pour que vous ou quelqu'un d'autre puissiez comprendre votre base de données.

## 7. N'utilisez pas les options de surlignement ou de couleurs pour coder une variable.

L'utilisation de mise en forme pour ajouter une information est visuelle, mais se perdra lors de l'importation des données sur R. Ajoutez plutôt une variable pour renseigner l'information.

## 8. Ne faites pas de calculs sur Excel

Vous augmentez les risques d'erreur et perdez la trace de vos modifications.

Marche à suivre :

- Saisissez toutes les données récoltées sur le terrain
- Faites une ou plusieurs sauvegardes (sur des supports différents), sans jamais utiliser le terme "final" dans le nom que vous donnerez au fichier
- Assurez-vous que tous les membres de votre équipe travaillent avec la version la plus récente

## 9. Sauvegardez vos données au format csv

En plus de la sauvegarde sous le format Excel, faites une sauvegarde dans un format csv. Il s'agit d'un format texte dans lequel chaque ligne correspond à une ligne du tableau, et dans lequel les valeurs sont séparées par un point-virgule, une virgule ou une tabulation. L'avantage de ce format est qu'il ne dépend pas d'un logiciel, il peut être lu par n'importe lequel.

Pour faire une sauvegarde en csv depuis Excel :

- Aller dans le menu "Fichier"
- Choisissez "Enregistrez sous"
- Choisissez l'emplacement
- Puis dans l'onglet "Type", choisissez CSV (séparateur: point-virgule) (\*.csv)