

TP2 – Echantillonnage, récolte et gestion des données

1. Stratégie d'échantillonnage

2. Récolte de données et encodage

A. Structurer la manière d'encoder des données

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

3. Importation et manipulation de tableaux dans R

A. Mise en forme des fichiers

B. Importation dans R

C. Manipulations de fichier

Fichiers nécessaires : **BD_Exemple_XLS.xlsx** et **BD_Exemple_CALC.ods**

Question scientifique :

Quel est l'effet du traitement forestier (régulier | jardiné ou irrégulier) sur la biodiversité (définie par la richesse en espèces végétales) ?

**Variable
dépendante = Variables
explicatives**

**Variabilité de S = F(Traitement) + variabilité
inexpliquée**

Formulation sous la forme d'un test d'hypothèse :

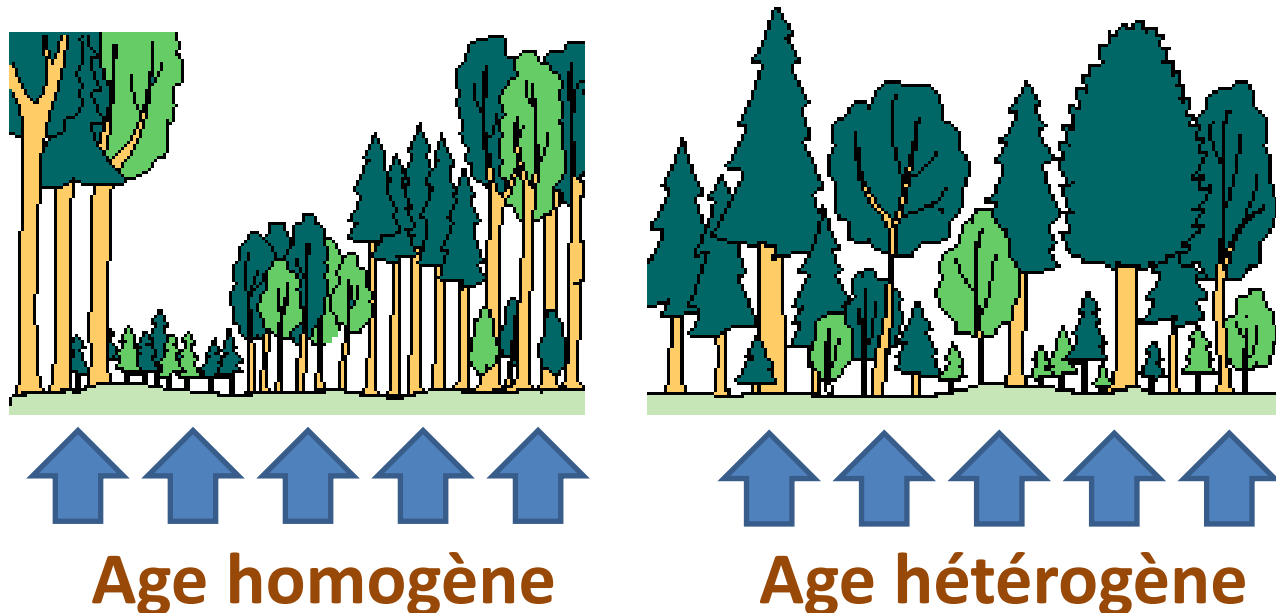
H_0 = pas d'effet

H_1 = jardiné > régulier

Objectif d'un protocole : réussir à rejeter H_0

= organiser l'échantillonnage pour contrôler les effets indésirables et maximiser les nombres de répétitions là où les questions sont pertinentes

1. Echantillonner des massifs forestiers en fonction du facteur contrôlé (Régulier/Jardiné-Irrégulier)



Objectif d'un protocole : réussir à rejeter H_0

2. Cibler des conditions environnementales et contrôler les facteurs « parasites » qui peuvent biaiser la réponse :

- Exemples ?
- Résineux, feuillus, mixte ?
 - Age des peuplements ?
 - Surface d'étude ?
 - Type de sol ?
 - Régions biogéographiques (altitude ?)

Indépendance des facteurs

| | | | |
|----------|-----|-----|-----|
| | Res | Feu | Mix |
| Jardiné | | | |
| Régulier | | | |

Facteurs hiérarchisés

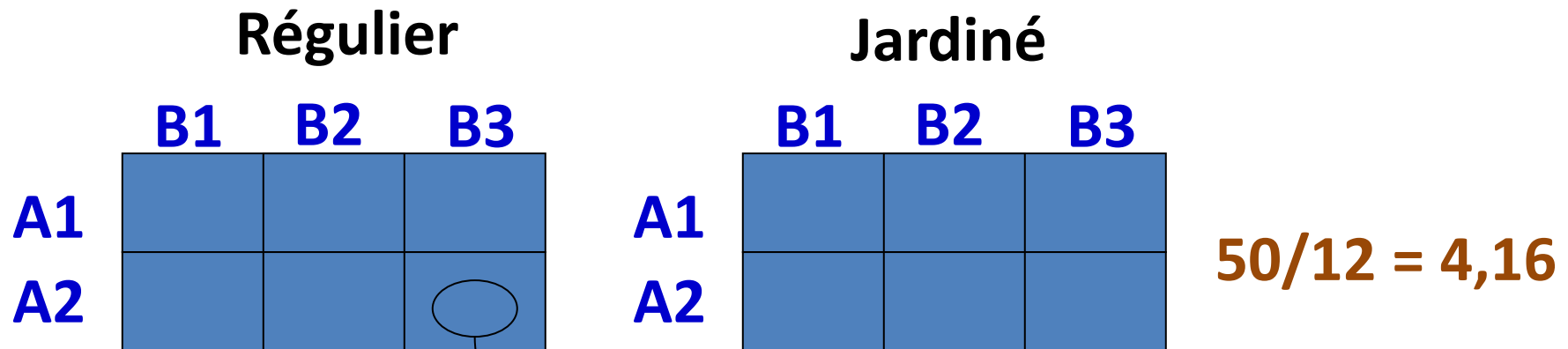
| | | |
|----------|---------|---------|
| | Brabant | Ardenne |
| Jardiné | Feu | Rési |
| Régulier | Feu | Rési |

Pour mesurer des interactions

Objectif d'un protocole : réussir à rejeter H_0

3. Définir le nombre d'unités d'échantillonnage et les répartir dans les différentes strates

Si j'ai des budgets pour faire au maximum 50 mesures



4 mesures par strate. Suffisant ?

Si non, diminuer le nombre de strates ...

Objectif d'un protocole : réussir à rejeter H_0

3. Définir le nombre d'unités d'échantillonnage et les répartir dans les différentes strates

- Régulier – Jardiné => 2 catégories
- Résineux – feuillus => 2 catégories
- Age des peuplements => 3 catégories (<20; 20-40; >40)
- Type de sol => 3 catégories (improductif, peu et fort productif)
- X Y (Z) => 5 régions biogéographiques

Nombres de combinaison :

$$2 * 2 * 3 * 3 * 5 = 180$$

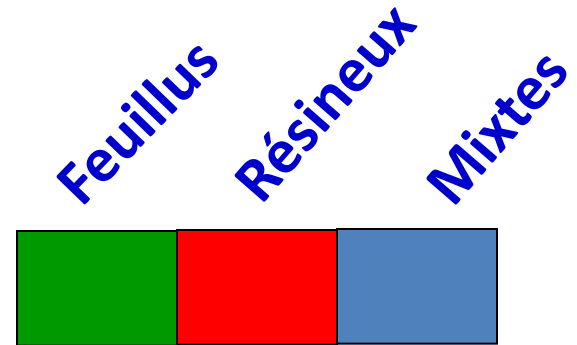
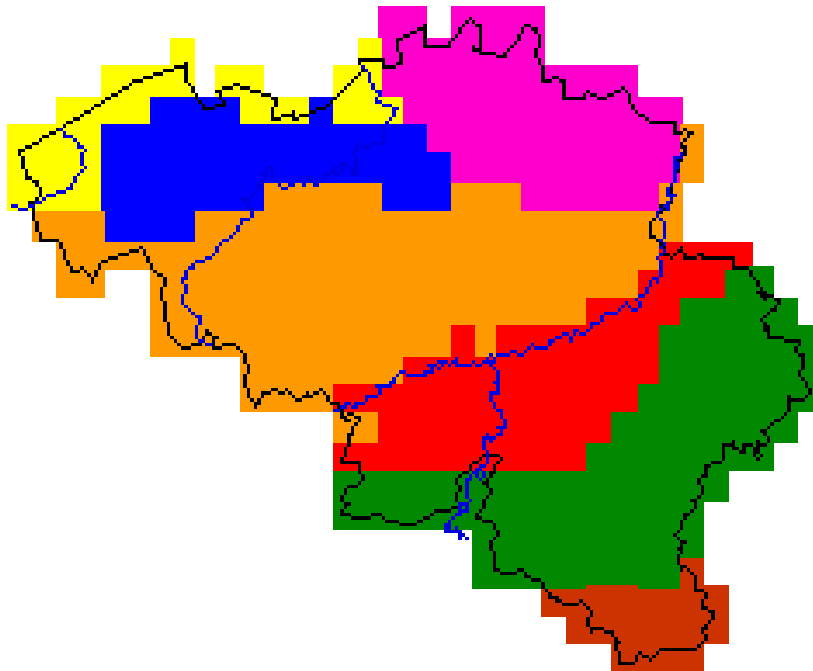
unités d'échantillonnage potentielles (sans répétition)

Si on dispose d'un budget pour 50 unités (avec 5 répétitions) on ne peut étudier que 10 combinaisons parmi les 180 potentielles ... => Il faut préciser la question de base ! (rien que les feuillus, un type de sol, ...)

Objectif d'un protocole : réussir à rejeter H_0

- Contrôler les facteurs « parasites » :

comme par exemple tenir compte de l'autocorrélation spatiale

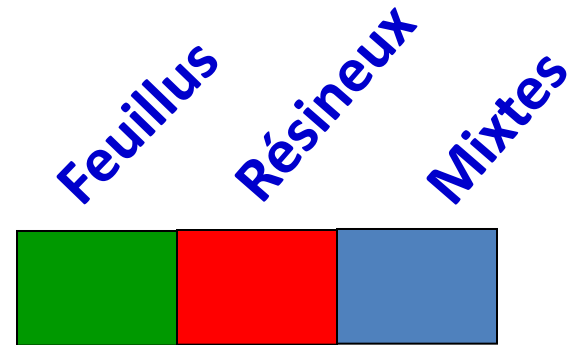
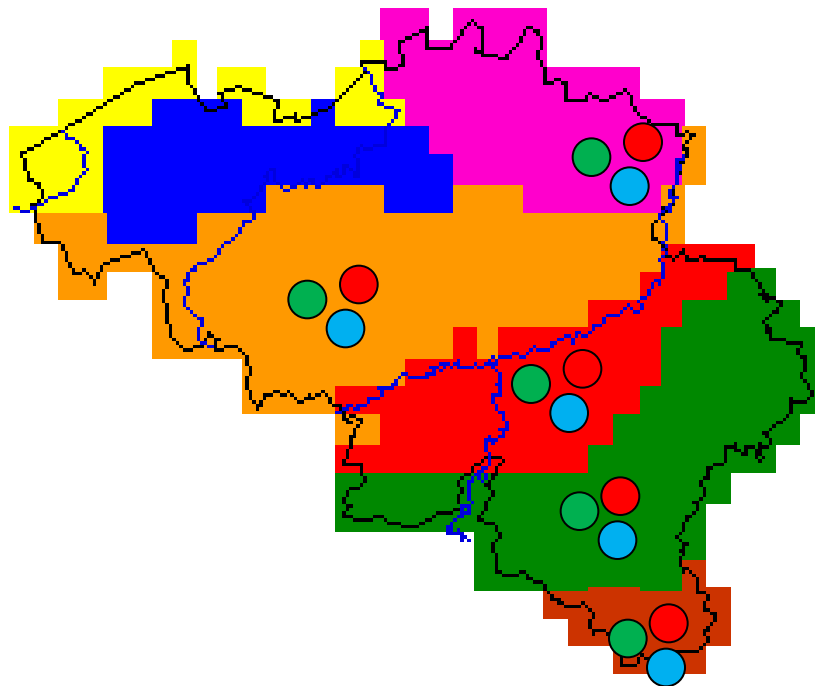


Comment les répartir
optimalement ?

Objectif d'un protocole : réussir à rejeter H_0

- Contrôler les facteurs « parasites » :

comme par exemple tenir compte de l'autocorrélation spatiale



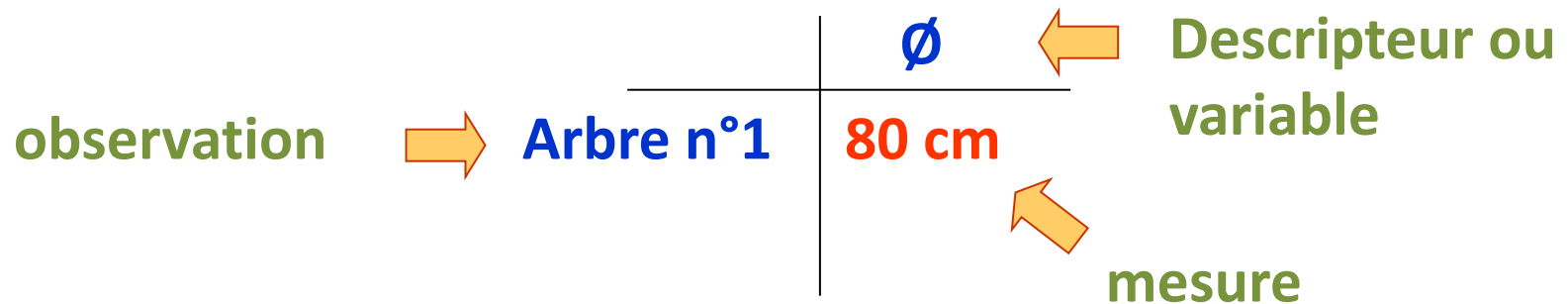
Au moins les 3 catégories
dans chacune des régions ET
les plus proches les unes des
autres

Si ce n'est pas possible, il
faudra contrôler à postériori
le rôle de X , Y , $X*Y$, X^2 , Y^2 , ..

A. Structurer la manière d'encoder des données

Qu'est ce qu'une donnée, une mesure ?

C'est la description d'une observation ! C'est un attribut d'un objet, une valeur que l'on mesure sur quelque chose et que l'on peut répéter sur autre chose ...



A. Structurer la manière d'encoder des données

On va généralement stocker les données sous forme matricielle

Descripteur ou variable

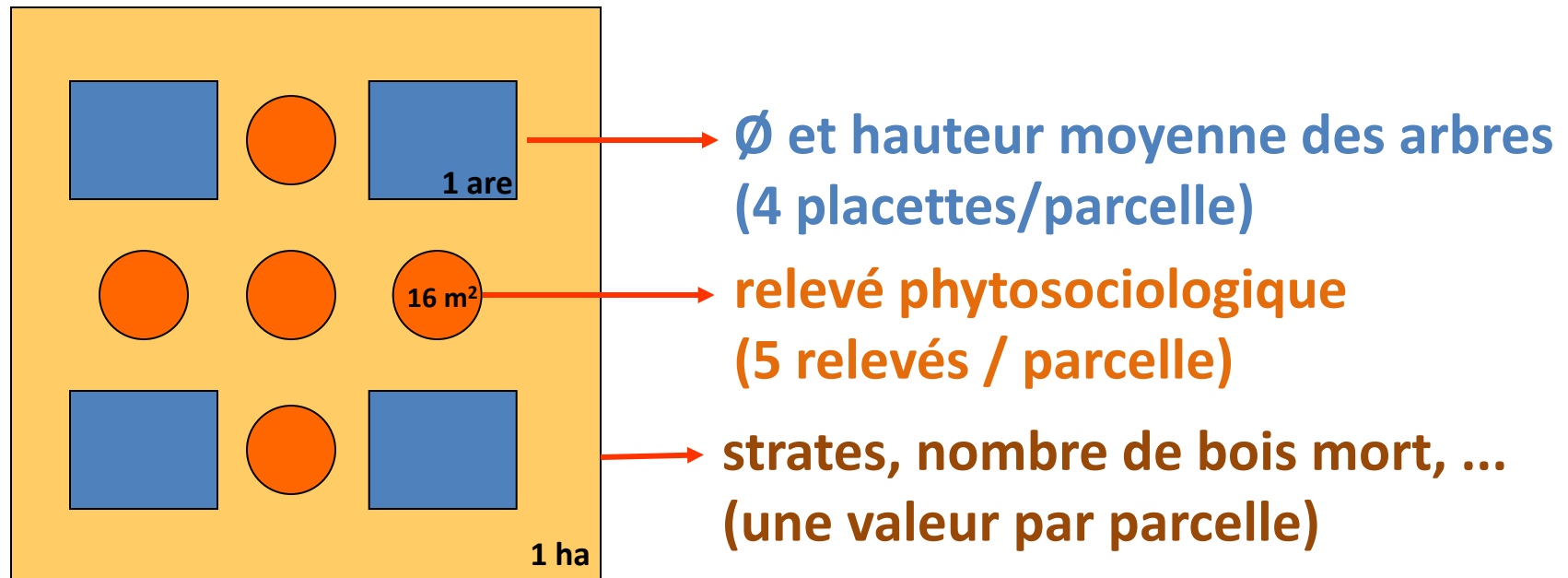
O
b
s
e
r
v
a
t
i
o
n
s

| Codesite | X | Y | Altitude | pHH2O | CsurN | V | Feuil600 | Feuil2000 | Conif600 | Conif2000 |
|----------|----------|----------|----------|----------|----------|----------|----------|-----------|----------|-----------|
| BL | 12,50614 | 11,90037 | 5,860786 | 1,667707 | 2,899772 | 1,931521 | 3,671225 | 5,321057 | 2,949688 | 5,72945 |
| BN | 12,21882 | 11,15698 | 6,042633 | 1,704748 | 2,701361 | 2,209373 | 3,990834 | 5,545958 | 3,374169 | 6,137511 |
| BR | 12,52373 | 11,85001 | 6,329721 | 1,642873 | 2,961658 | 2,045109 | 4,18662 | 5,31812 | 3,563883 | 6,769986 |
| CL | 12,41697 | 10,57378 | 5,888878 | 1,728109 | 2,619583 | 2,273156 | 4,136765 | 5,749393 | 0,530628 | 5,60433 |
| FB | 12,40967 | 10,75152 | 6,018593 | 1,89762 | 2,595255 | 3,845028 | 4,576771 | 6,053265 | 1,686399 | 5,931715 |
| FS | 11,94299 | 11,98331 | 4,75359 | 1,609438 | 2,667228 | 2,636196 | 4,685208 | 6,950464 | 1,823129 | 3,034794 |
| HL | 12,41403 | 10,56064 | 5,916202 | 1,691939 | 2,793004 | 2,180417 | 4,361824 | 5,721295 | 1,931521 | 5,551408 |
| LB | 12,23167 | 11,12985 | 5,942799 | 1,710188 | 2,62684 | 2,00283 | 2,533697 | 5,072044 | 4,222445 | 6,005367 |
| MH | 12,23206 | 11,29011 | 5,888878 | 1,680828 | 2,951258 | 2,19277 | 4,568506 | 6,743116 | 2,839078 | 5,460011 |
| MO | 12,56232 | 11,79811 | 6,381816 | 1,60342 | 2,888704 | 2,236445 | 2,85647 | 4,099332 | 4,25703 | 6,870884 |
| MU | 12,52693 | 11,94009 | 5,771441 | 1,667707 | 2,721295 | 2,230014 | 3,793239 | 5,263726 | 2,923162 | 5,249127 |
| PB | 12,52424 | 11,83411 | 6,338594 | 1,66203 | 2,866193 | 1,888584 | 3,671225 | 4,954418 | 3,459466 | 6,446196 |
| PO | 12,52016 | 11,86147 | 6,196444 | 1,623341 | 2,797281 | 1,832581 | 3,621671 | 5,297817 | 4,01998 | 6,798052 |
| PR | 12,41715 | 10,41304 | 5,70711 | 1,752672 | 2,391511 | 3,09603 | 4,546481 | 6,130357 | 2,208274 | 5,516649 |
| R1 | 12,50703 | 11,85254 | 6,133398 | 1,609438 | 2,906901 | 1,969906 | 3,688879 | 4,635699 | 4,028917 | 6,933618 |
| R2 | 12,5047 | 11,86338 | 6,066108 | 1,648659 | 2,894253 | 1,701105 | 3,98713 | 5,042134 | 3,960813 | 6,776165 |
| RA | 11,97162 | 12,01878 | 4,75359 | 1,61542 | 2,729812 | 2,227862 | 4,409897 | 6,675997 | 1,186776 | 3,918269 |
| RU | 12,53279 | 11,80672 | 6,398595 | 1,623341 | 2,869602 | 1,83098 | 3,198673 | 4,694096 | 3,983413 | 6,624065 |
| ST | 12,41425 | 10,72812 | 5,888878 | 1,780024 | 2,543961 | 3,579065 | 4,207673 | 5,885826 | 2,646175 | 4,481872 |
| TR | 12,40689 | 10,44249 | 5,70711 | 1,704748 | 2,727853 | 2,136531 | 3,788725 | 5,536152 | 3,295837 | 5,833055 |
| VR | 12,23038 | 11,26742 | 5,831882 | 1,704748 | 2,789323 | 2,396986 | 3,730501 | 6,42276 | 4,122284 | 5,499624 |
| VS | 12,51518 | 11,8704 | 6,018593 | 1,623341 | 3,431727 | 1,65058 | 4,333361 | 5,120983 | 3,520461 | 6,822197 |

entité homogène = une table

A. Structurer la manière d'encoder des données

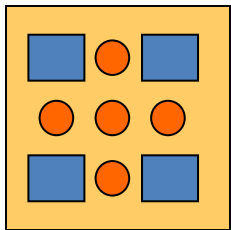
Pour tester l'impact du traitement forestier on va réaliser un certain nombre d'unités d'échantillonnage dans des forêts feuillues jardinées et équiennes (20 parcelles par type).



Comment structure-t-on ce jeu de données ?

A. Structurer la manière d'encoder des données

Concept d'entité homogène d'informations => tableau



On ne met dans une table qu'un seul ensemble homogène d'informations => les observations sont des répétitions décrites par les mêmes variables

Parcelles

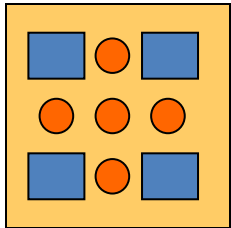
Placettes

Relevés

3 entités => 3 tables

A. Structurer la manière d'encoder des données

Création de relations entre les tables

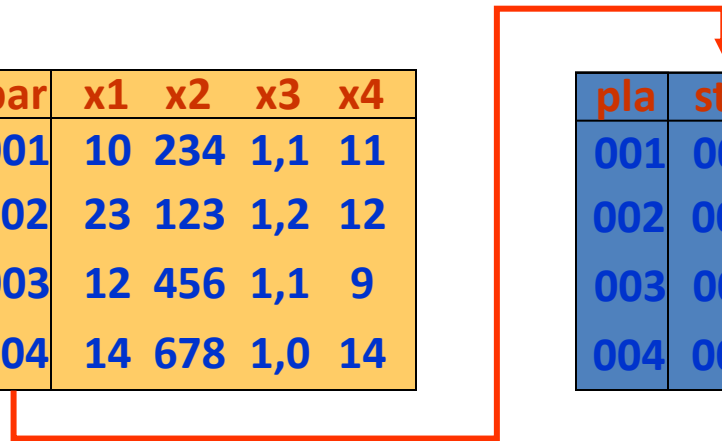


Etablir des relations entre les entités homogènes
=> création de clé relationnelle = tout simplement
des codes qui permettent d'associer des lignes
d'une table à des lignes d'une autre table.



| par | x1 | x2 | x3 | x4 |
|-----|----|-----|-----|----|
| 001 | 10 | 234 | 1,1 | 11 |
| 002 | 23 | 123 | 1,2 | 12 |
| 003 | 12 | 456 | 1,1 | 9 |
| 004 | 14 | 678 | 1,0 | 14 |

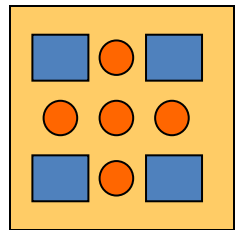
| pla | sta | y1 | y2 | y3 |
|-----|-----|-----|----|-----|
| 001 | 001 | 1,0 | 14 | 1,1 |
| 002 | 001 | 2,3 | 33 | 1,2 |
| 003 | 001 | 1,2 | 26 | 1,1 |
| 004 | 002 | 1,4 | 58 | 1,0 |



A. Structurer la manière d'encoder des données

Création de relations entre les tables

Clé relationnelle



| par | x1 | x2 | x3 | x4 |
|-----|----|-----|-----|----|
| 001 | 10 | 234 | 1,1 | 11 |
| 002 | 23 | 123 | 1,2 | 12 |
| 003 | 12 | 456 | 1,1 | 9 |
| 004 | 14 | 678 | 1,0 | 14 |

| pla | par | y1 | y2 | y3 |
|-----|-----|-----|----|-----|
| 001 | 001 | 1,0 | 14 | 1,1 |
| 002 | 001 | 2,3 | 33 | 1,2 |
| 003 | 001 | 1,2 | 26 | 1,1 |
| 004 | 002 | 1,4 | 58 | 1,0 |

Clé
primaire **univoque**

Permet de fusionner
les tables

| pla | par | y1 | y2 | y3 | x1 | x2 | x3 | x4 |
|-----|-----|-----|----|-----|----|-----|-----|----|
| 001 | 001 | 1,0 | 14 | 1,1 | 10 | 234 | 1,1 | 11 |
| 002 | 001 | 2,3 | 33 | 1,2 | 10 | 234 | 1,1 | 11 |
| 003 | 001 | 1,2 | 26 | 1,1 | 10 | 234 | 1,1 | 11 |
| 004 | 002 | 1,4 | 58 | 1,0 | 23 | 123 | 1,2 | 12 |

A. Structurer la manière d'encoder des données

Structure relationnelle simple entre deux fichiers

| Fichier ECHANTILLON | | | | | |
|---------------------|------------------|------------|-----------|--|--|
| Code | Localité | Date | Récolteur | | |
| 0001 | Rixensart | 02/03/1997 | Dupont J. | | |
| 0002 | Louvain-la-Neuve | 02/03/1997 | Durant P. | | |
| 0003 | Gembloux | 14/04/1997 | Dupont J. | | |
| ... | | | | | |

| Fichier OBSERVATION | | | | | |
|---------------------|------|------|-------|--------|-----------|
| Taxon | Code | Sexe | Stade | Nombre | Auteur |
| Libellula depressa | 0001 | M | A | 0003 | Dupont J. |
| Libellula depressa | 0001 | F | A | 0001 | Dupont J. |
| Buteo buteo | 0001 | ? | A | 0002 | Albert P. |
| Pieris napi | 0002 | M | A | 0001 | Henri A. |
| Papilio machaon | 0002 | ? | L | 0002 | Henri A. |
| ... | | | | | |

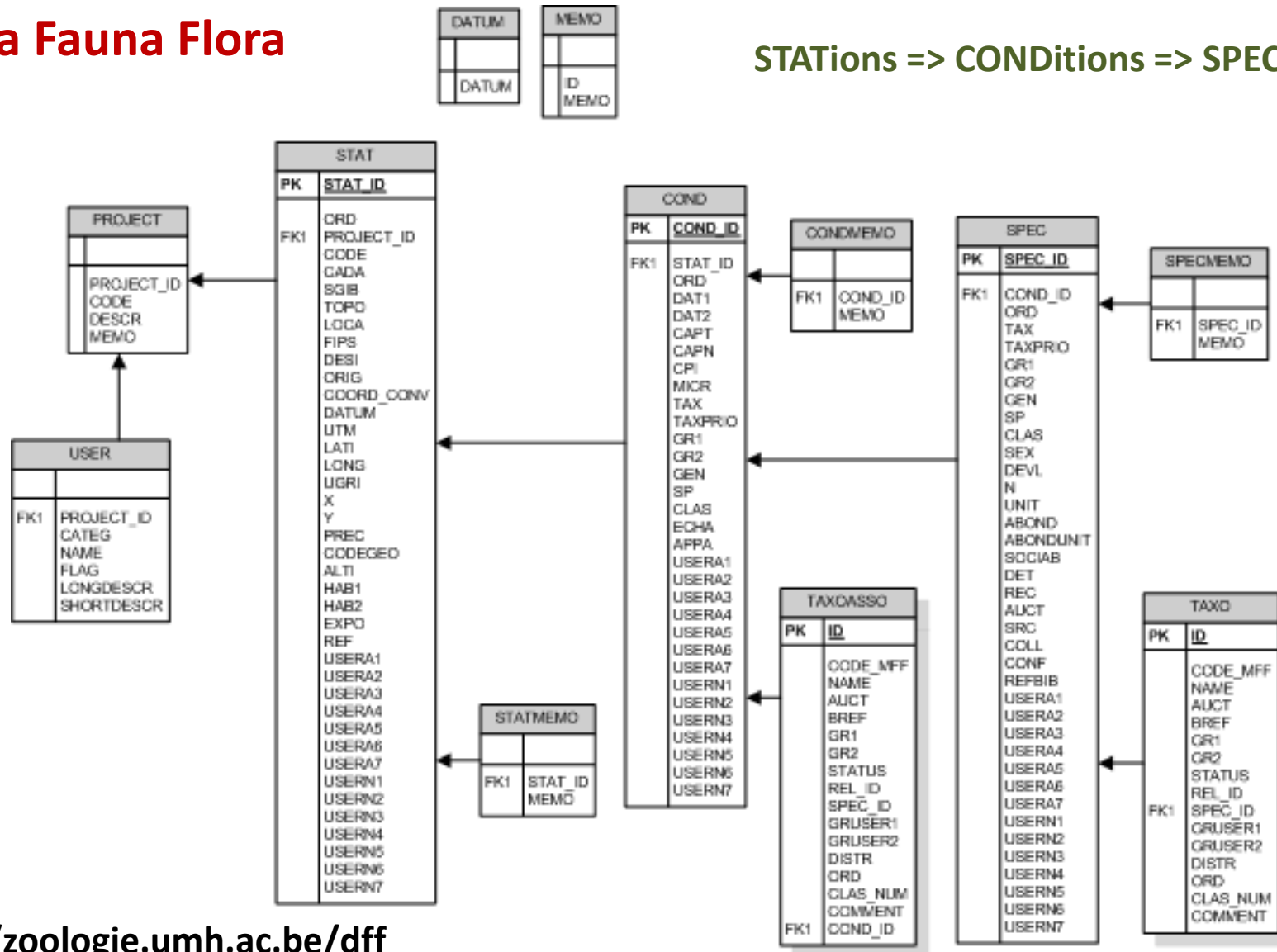
Chaque fois qu'une nouvelle station est échantillonnée, on ajoute une ligne dans le fichier "ECHANTILLON". On place le résultat de l'échantillonnage, soit la liste des taxons observés, dans le fichier "OBSERVATION". Ce faisant, on évite de répéter dans ce dernier fichier toute une série d'informations qui sont communes aux taxons récoltés dans un échantillon.

La variable "Code", commune aux deux fichiers, assure la relation entre le fichier ECHANTILLON et le fichier OBSERVATION. C'est elle qui permet de connaître – par exemple – la liste des espèces observées à Rixensart ou de connaître la liste des lieux d'observations de *Libellula depressa*.

A. Structurer la manière d'encoder des données

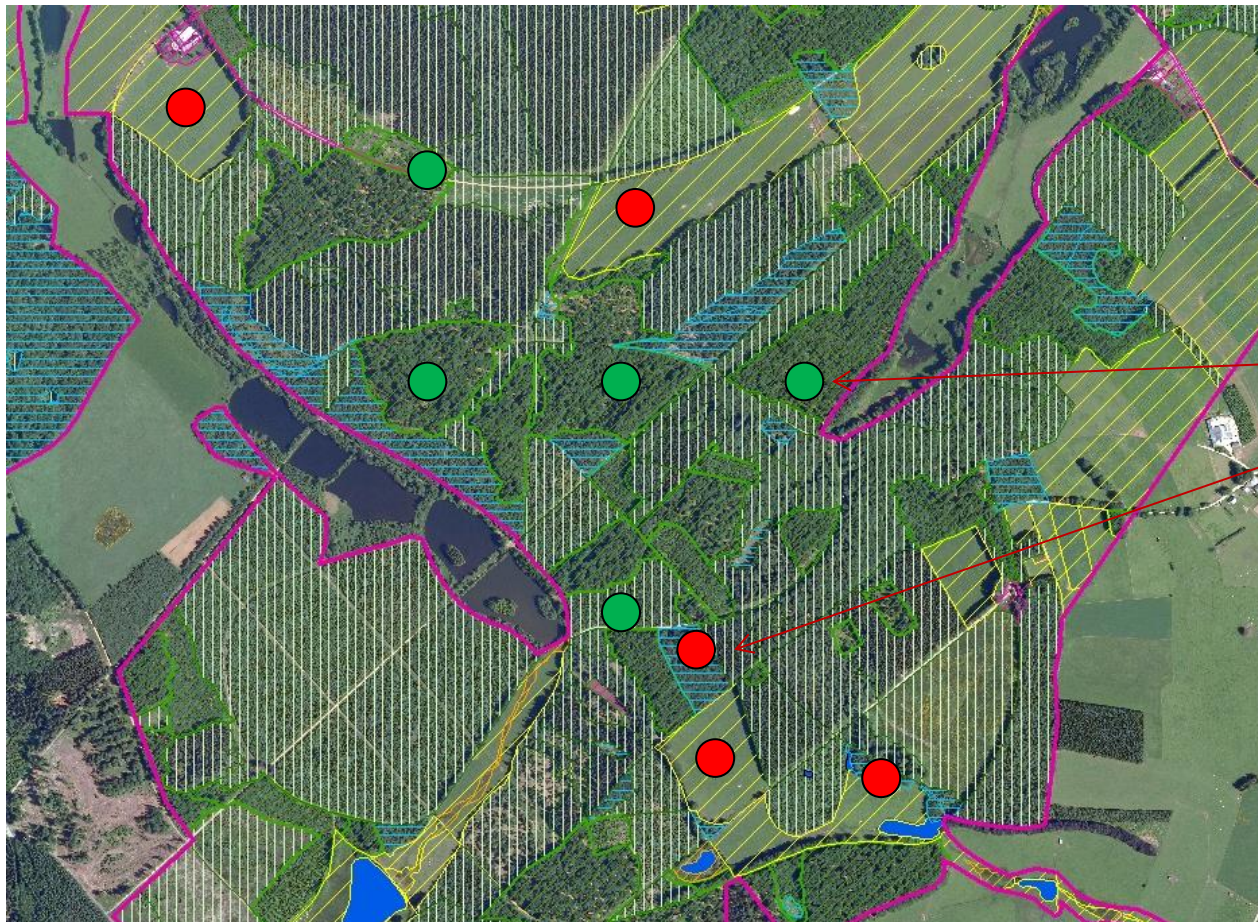
Data Fauna Flora

STATions => CONDitions => SPECimens



A. Structurer la manière d'encoder des données

Base de données descriptives des sites Natura 2000



Cartographie
d'Unités d'habitat
(UH)



Placettes
d'évaluation (PE)
de l'état de
conservation



Relevés
phytosociologiques

A. Structurer la manière d'encoder des données

Base de données descriptives des sites Natura 2000

Fiche décrivant une unité d'habitat (sauf C2 > 10 km²)

15.03.06

| | | | | | | |
|---|--------------------------|---|-------------------|-------------------|-----------------------------|--|
| Code UH labo | Code UH terrain | Type d'unité | | | Surface (si ≠ polyg) | |
| <i>Fourni par ArcView</i> | | POLYG | LIGNE | POINT | m ² | |
| Code site N2000 | Date cartographie | Opérateurs | | | | |
| EUNIS ACTUEL | | Si complexe, recouvrement des habitats (%) | | | | |
| potentiel 1 | potentiel 2 | potentiel 3 | objectif 1 | objectif 2 | objectif 3 | |
| Si haies (FA) : <input type="checkbox"/> Avec des arbres <input type="checkbox"/> Hauteur - de 4 m <input type="checkbox"/> Discontinuité <input type="checkbox"/> 0% <input type="checkbox"/> 0 à 40% <input type="checkbox"/> 40 à 75% | | | | | | |
| Commentaires | | | | | | |

Base de données descriptives des sites Natura 2000

Fiche décrivant une placette d'évaluation

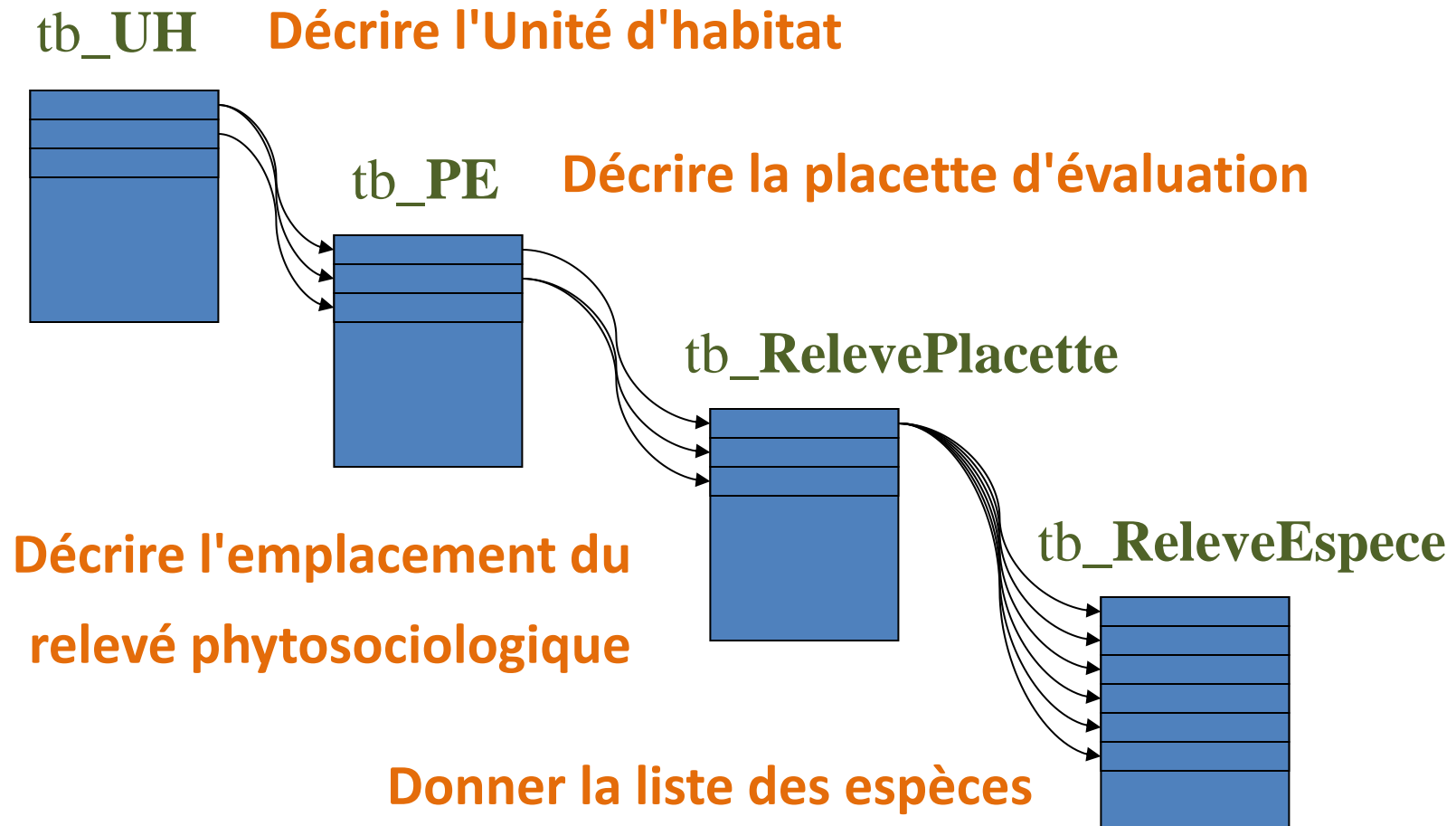
15.03.06

| Code UH terrain | Code placette | Type de placette | | | | Forme | | Surface |
|--|-----------------|-----------------------------------|------------|------------|---|------------|--------|----------------|
| | | UH | SYST | ALEA | REPR | CIRC | LIN | m ² |
| Code site N2000 | Date évaluation | EUNIS placette (<i>Si ≠ UH</i>) | | | | Opérateurs | | |
| X Lambert (m) | Y Lambert (m) | Précision N° GPS | | Pente | Exposition <input type="checkbox"/> N <input type="checkbox"/> NE <input type="checkbox"/> E <input type="checkbox"/> SE <input type="checkbox"/> S <input type="checkbox"/> SO <input type="checkbox"/> O <input type="checkbox"/> NO aucune | | | |
| <i>Si ≠ UH</i> | <i>Si ≠ UH</i> | | | degrés* | | | | |
| Topographie <div style="display: flex; justify-content: space-between;"> <div> <input type="checkbox"/> 1 = plateau ou plaine <input type="checkbox"/> 7 = tête de source </div> <div> <input type="checkbox"/> 2 = haut de versant <input type="checkbox"/> 5 = milieu de versant <input type="checkbox"/> 8 = bas de versant </div> <div> <input type="checkbox"/> 3 = terrasse alluviale haute <input type="checkbox"/> 6 = terrasse alluviale basse <input type="checkbox"/> 4 = dépression marécageuse </div> </div> | | | | | | | | |
| sol nu | musculaire | herbacée | ericoïde | arb. < 2 m | 2 à 8 m | 8-16 m | > 16 m | > 8 m |
| % | % | % | % | % | % | % | % | % |
| Perturbations | | | | | | | | |
| | A/N | % | Précisez ! | | | A/N | % | Précisez ! |
| Colonisation herbacée | | | | | Fauche excessive | | | |
| Colonisation ericoïde | | | | | Pâturage excessif | | | |
| Colonisation arbustive | | | | | Amendement important | | | |
| Colonisation arborée | | | | | Herbicides | | | |
| Nitrophytes | | | | | Eutrophisation (agricole) | | | |
| Xénophytes | | | | | Eutrophisation (eau) | | | |
| Plantation résineuse | | | | | Pollutions (routes) | | | |
| Plantation feuillue | | | | | Pollutions (autres) | | | |
| Dégâts de débardage | | | | | Dépôts déchets divers | | | |
| Coupe trop importante | | | | | Remblais | | | |
| Pression animale (vég) | | | | | Zone d'extraction | | | |
| Dégâts animaux (sol) | | | | | Passage véhicules | | | |
| Drainage superficiel | | | | | Fréquentation touristique | | | |
| Drainage profond | | | | | Escalade | | | |
| Atterrissement | | | | | | | | |

[illegible]

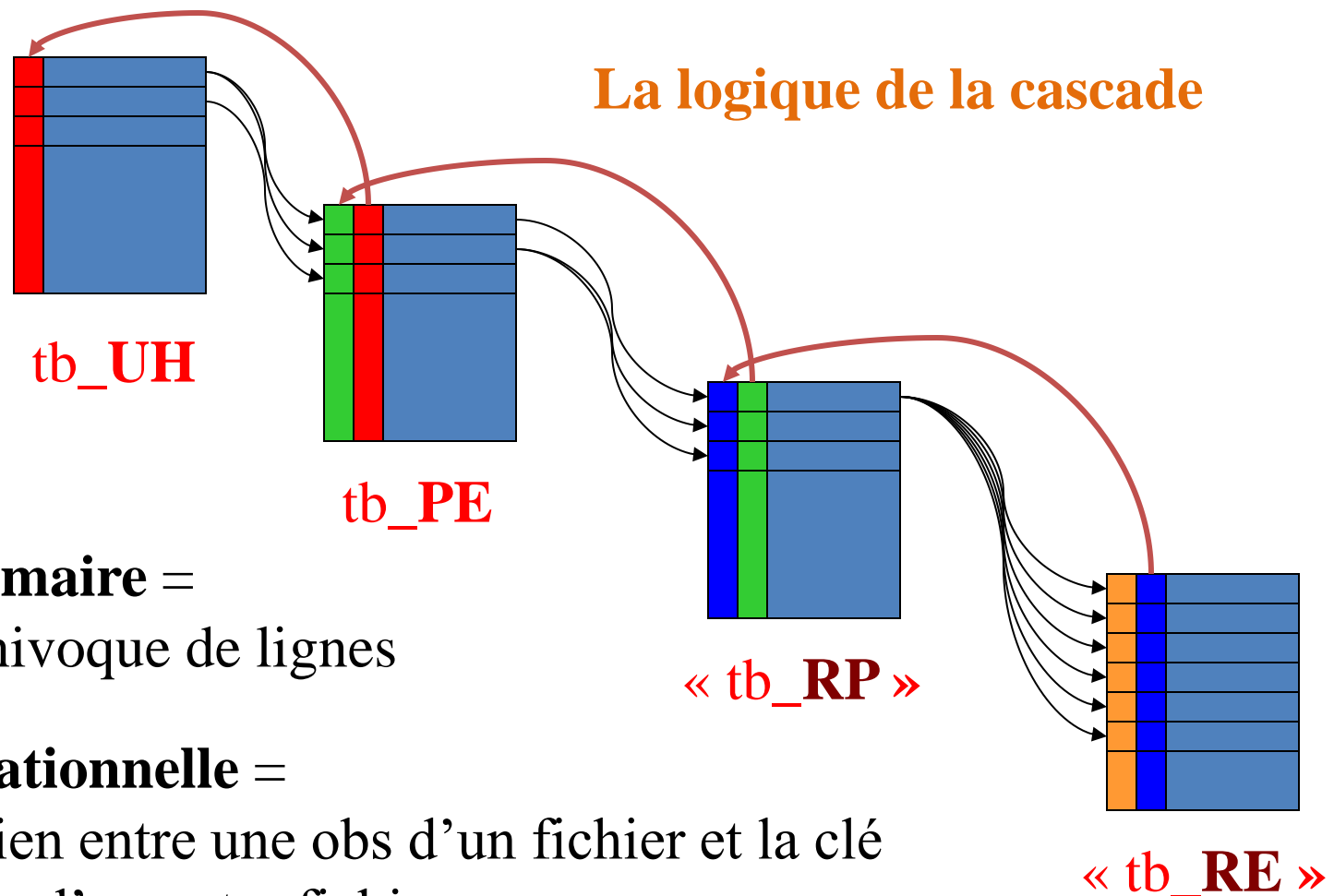
A. Structurer la manière d'encoder des données

Base de données descriptives des sites Natura 2000



A. Structurer la manière d'encoder des données

Base de données descriptives des sites Natura 2000



B. La préparation d'un tableau de données

Encodage des données => Tableur => une feuille par tableau

- **Très facile à structurer**
- **Très facile à modifier** (insérer et supprimer des colonnes ou des lignes)
- **Pré-encodage de certaines valeurs**
- **Permet de vérifier rapidement la cohérence des encodages** (filtre pour une liste univoque des valeurs d'une colonne)
- **Permet d'utiliser des formules pour standardiser des formats ou recréer de nouvelles variables**

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Savoir :

- distinguer les différents types d'informations que l'on peut déposer dans les cellules d'un tableur; les références relatives et les références absolues;
- comprendre ce qu'on peut faire dans un tableur

Savoir faire :

- se déplacer, copier/coller, importer/exporter, trier des jeux de données
- réaliser des calculs sur les cellules à l'aide du tableur ;
- utiliser les fonctions de base dans un tableau;
- tracer des graphiques et jouer avec les séries;
- produire des fichiers de synthèse

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

 Avant de commencer : attention à la gestion des décimales !

Séparateur : NE PAS UTILISER la virgule (,) MAIS BIEN le point (.)

- Seul le . est utilisable par les logiciels d'analyses de données
- La , est utilisée parfois comme séparateur dans certains formats d'export en CSV (il y a bien différents CSV !)

Sources d'erreurs lors de la lecture ou de l'importation de données.

 La configuration des options régionales de WINDOWS n'est pas intégrée par OPENOFFICE !

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)



Gestion des décimales (Windows- XLS)

Panneau de configuration

> Région et langues

> Paramètres supplémentaires

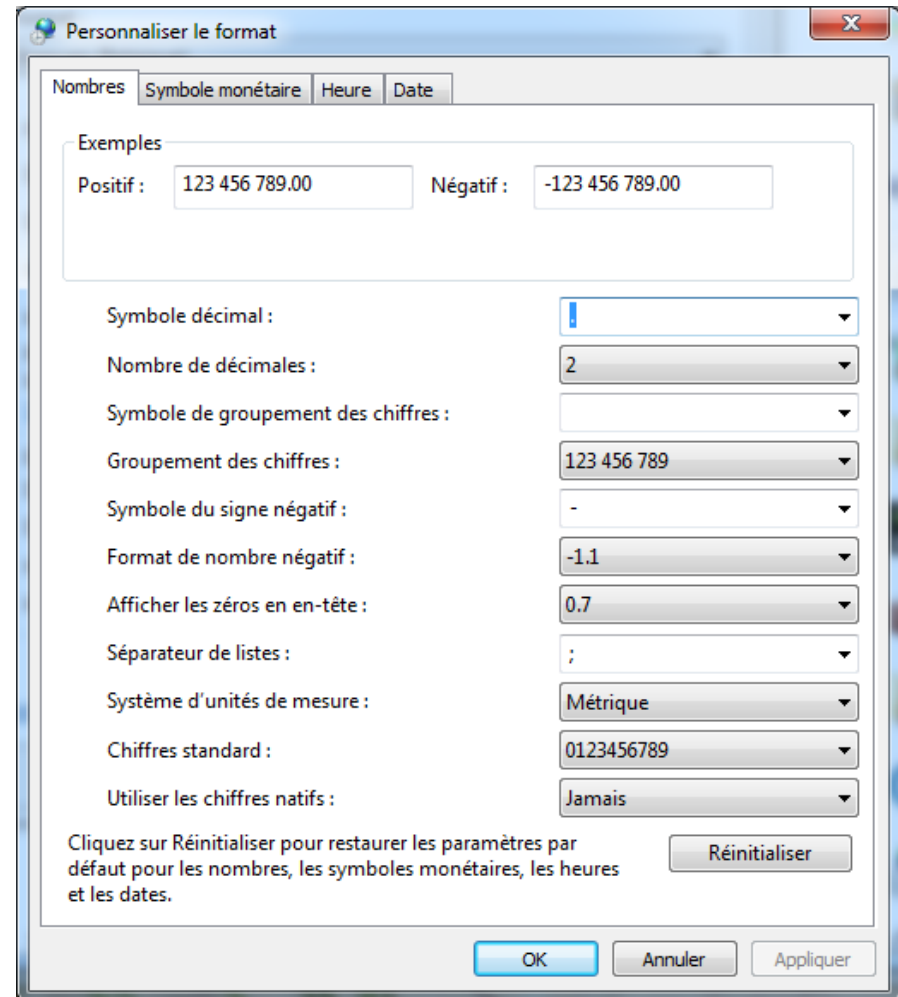
Mettre un .



Ne pas oublier de mettre un blanc
pour le séparateur de milliers ...



Bien vérifier que la gestion des
décimales soit bien réalisée
effectivement et pas simplement
sur l'affichage
=> copier-coller dans un fichier txt
ou exporter en format txt.



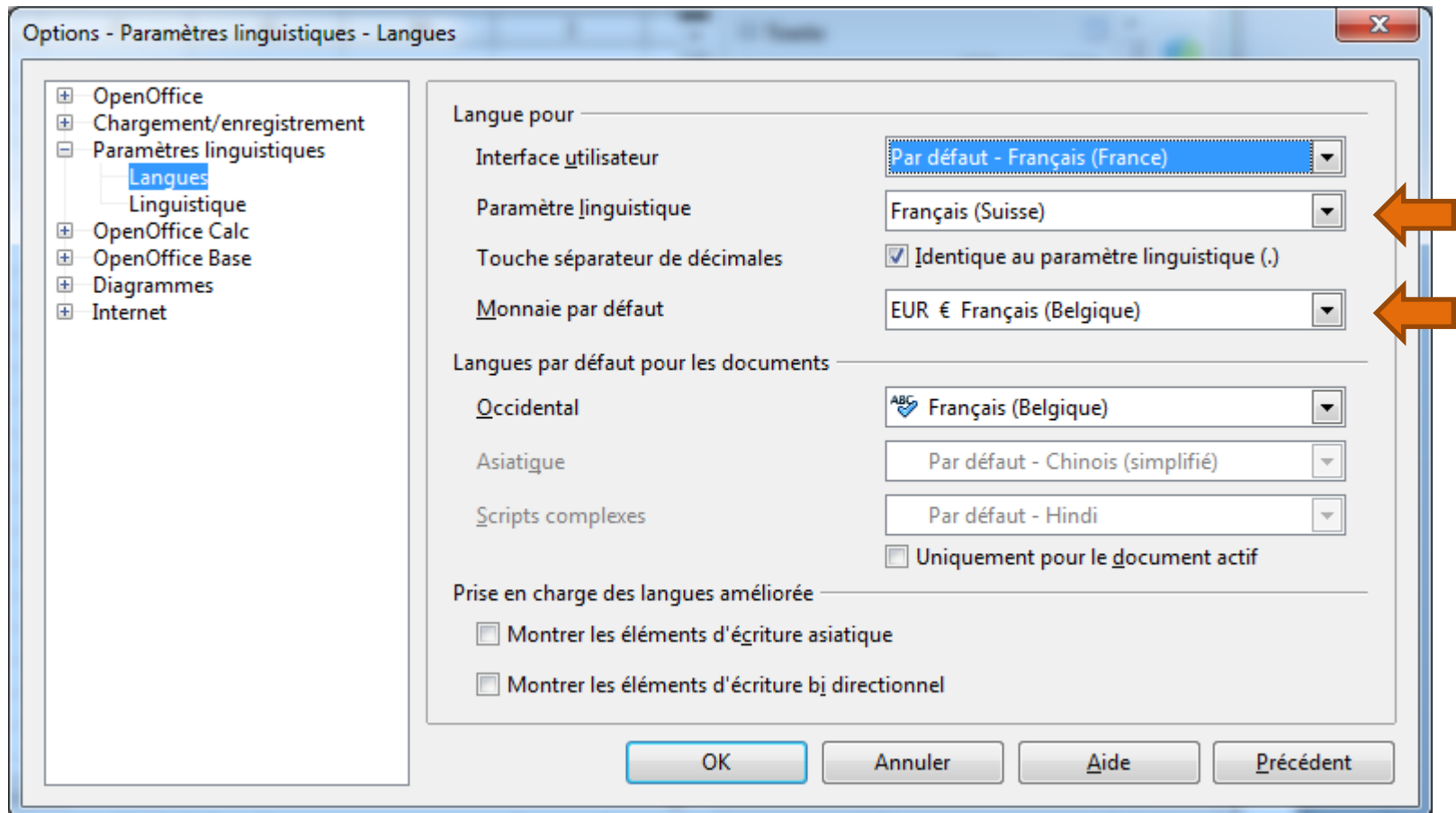
B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)



Gestion des décimales (OpenOffice - CALC)

Menu de CALC : Outils > Options > Paramètres Linguistiques > Langues >



B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Permet :

- de construire des tableaux de données;
- de réaliser des calculs;
- de réaliser des graphiques

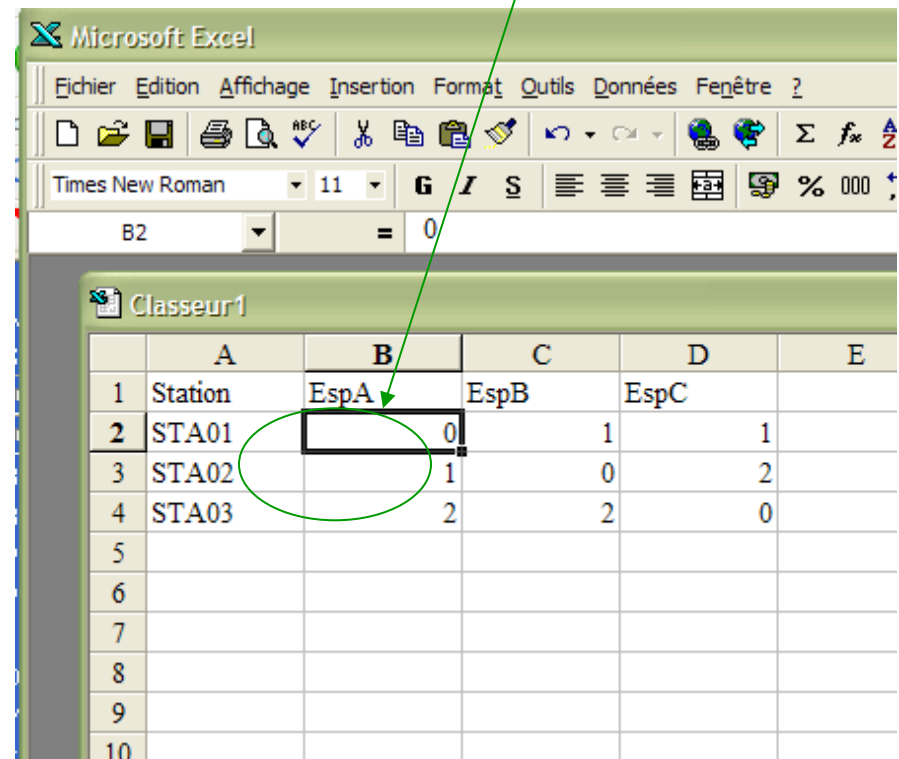
Se compose :

- colonnes (lettre = B);
- lignes (numéro = 2);
- cellule (croisement = B2)

Un tableau = 1 feuille

Déplacement rapide sur les
limites du tableau : CTRL+flèche
Sélection : Majuscule +CTRL+flèche

cellule active



Microsoft Excel

Echier Edition Affichage Insertion Format Outils Données Fenêtre ?

Times New Roman 11 G I S

B2 = 0

Classeur1

| | A | B | C | D | E |
|----|---------|------|------|------|---|
| 1 | Station | EspA | EspB | EspC | |
| 2 | STA01 | 0 | 1 | 1 | |
| 3 | STA02 | 1 | 0 | 2 | |
| 4 | STA03 | 2 | 2 | 0 | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Tous les types de données peuvent être encodés :

- **quantitatif** (discret ou continu) - aligné par défaut à droite
 - formaté (nombre de décimales, espacement milliers, %, ...)
- **qualitatif** (sous format texte) - aligné par défaut à gauche
- **spéciaux** : date, heure, monétaire, ... (**voir menu Format**)

Un format modifie la
donnée brute, pas
seulement son affichage



| | A | B | C |
|---|----------------------|-------------|---|
| 1 | <u>Donnees_brute</u> | Pourcentage | |
| 2 | 0.001 | 0.10% | |
| 3 | | | |
| 4 | | | |
| 5 | | | |

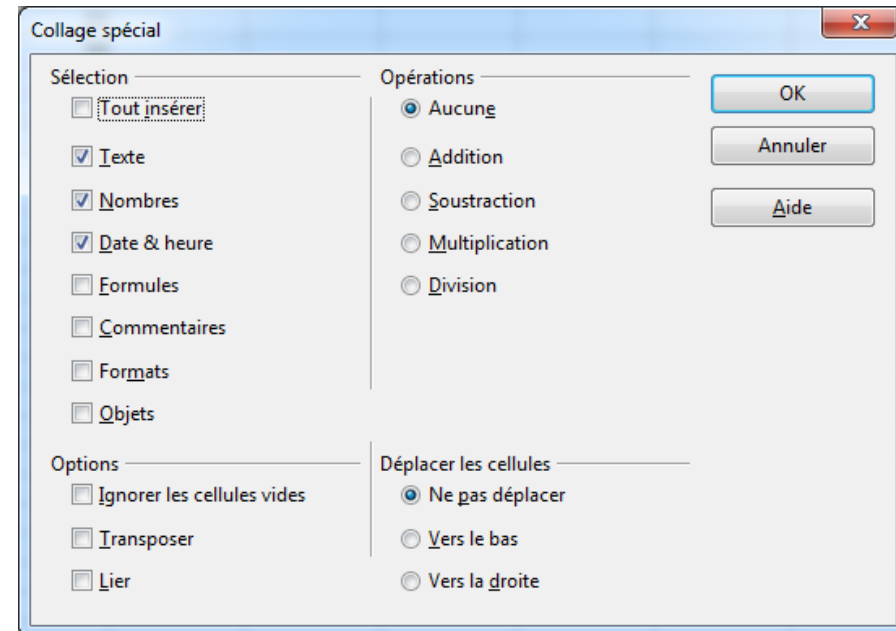
export : 0.10%

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Possibilité de recopier facilement des informations :

- par déplacement de la cellule active
- par les commandes **copier-coller** ou **recopier vers le bas / la droite**
- l'utilité de **collage spécial**
 - ⇒ pour éliminer les formules :
uniquement les textes
ou les nombres (CALC)
ou les valeurs (XLS)
uniquement les formats
 - ⇒ pour additionner, multiplier, ...
des cellules avec ce qui est copié
 - ⇒ **Transposition** pour inverser
les lignes et les colonnes



B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Utilisation de formules :

= opérations réalisées sur des cellules :

- opérateurs arithmétiques : **+, -, /, *, ^, ()**

- dans A1 écrire « 1 »;
- dans B1 écrire « = A1+1 »
- changer la valeur de A1

| | A | B | C |
|---|---|-------|---|
| 1 | 1 | =A1+1 | |
| 2 | 2 | 2 | |
| 3 | 3 | 4 | |
| 4 | 4 | 5 | |
| 5 | 5 | 6 | |

Possible de recopier vers le bas en tirant simplement la sélection vers le bas

- opérateurs de type texte :
& pour la concaténation de deux chaînes de caractères

| | A | B | C | D |
|---|-------|--------|-----------|---|
| 1 | Genre | Espece | Taxon | |
| 2 | Abax | ater | Abax ater | |
| 3 | | | | |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Utilisation de formules :

- statistiques descriptives (somme, moyenne, écartype, ...)
- mathématique (arrondi, trigonométrie, exposant, log, ...)
- texte (tronquer, extraire des chaînes de caractères, ...)
- calcul sur les dates (0=1/1/1900) et les heures
- logique : SI (A1>1;ce qu'on fait si c'est vrai; ce qu'on fait si c'est faux)

Formules s'écrivent de deux manières :

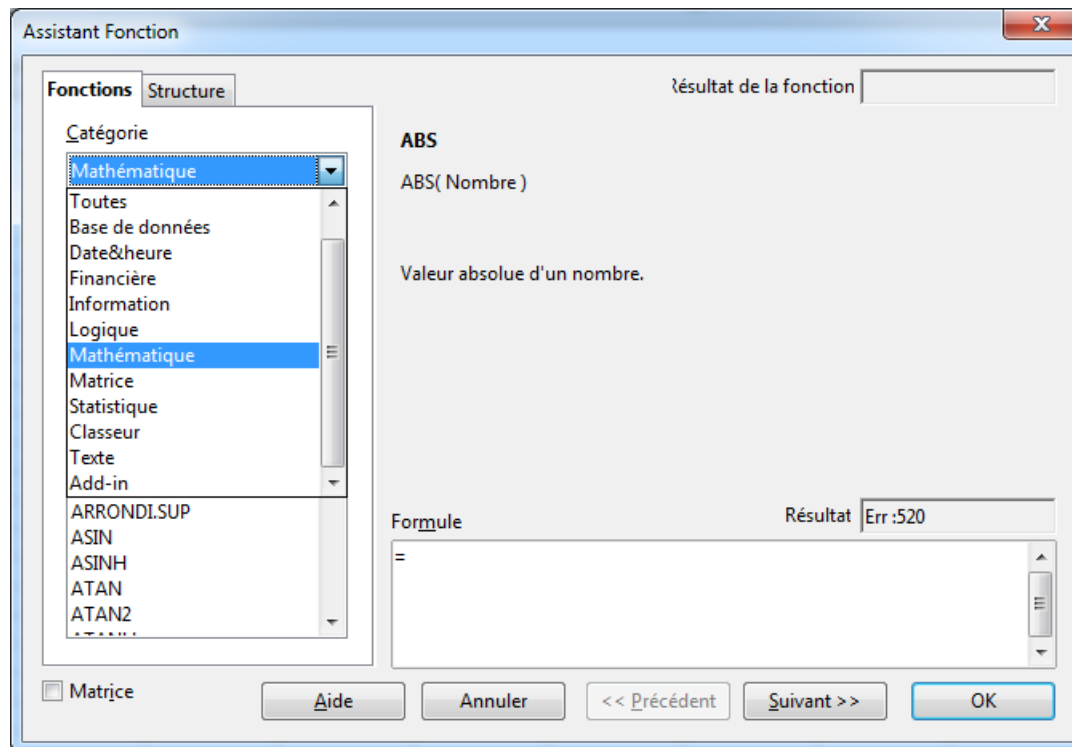
- soit directement : on écrit = somme(B1:B2)
- soit via l'appel « Insertion/Fonction » où on a un menu pour sélectionner les cellules concernées par la fonction et une aide contextuelle

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Utilisation de formules :

- utilisation de **fonctions** pré-programmées : choix très diversifié !



| | | | | |
|----|-----------------|--------|---|---------------|
| B5 | | \sum | = | =somme(B1:B4) |
| | A | B | C | |
| 1 | Abax ater | 10 | | |
| 2 | Abax ovalis | 12 | | |
| 3 | Abax parallelus | 2 | | |
| 4 | Amara anea | 15 | | |
| 5 | | 39 | | |
| 6 | | | | |

| | | | | |
|----|-----------------------|--------|---|---------------------|
| B6 | | \sum | = | =NB.SI(B1:B4;">10") |
| | A | B | C | |
| 1 | Abax ater | 10 | | |
| 2 | Abax ovalis | 12 | | |
| 3 | Abax parallelus | 2 | | |
| 4 | Amara anea | 15 | | |
| 5 | | | | |
| 6 | Nombre de taxons > 10 | 2 | | |
| 7 | | | | |

| | | | | |
|----|-----------------|--------|--------|---------|
| C2 | | \sum | = | =LN(B2) |
| | A | B | C | D |
| 1 | Taxons | NbrInd | Ln | |
| 2 | Abax ater | 10 | 2.3026 | |
| 3 | Abax ovalis | 12 | 2.4849 | |
| 4 | Abax parallelus | 2 | 0.6931 | |
| 5 | Amara anea | 15 | 2.7081 | |
| 6 | Somme | 39 | 8.1887 | |
| 7 | | | | |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Utilisation de formules :

- **référence relative – absolue** : Utilisation du \$ pour bloquer une référence (par exemple le calcul d'un pourcentage de chaque ligne dans un total)

simple recopiage
de la formule en C2



| | A | B | C |
|---|----------|---------------|---------|
| 1 | Taxon | Nbr individus | % |
| 2 | Espece 1 | 40 | 0.5 |
| 3 | Espece 2 | 30 | #DIV/0! |
| 4 | Espece 3 | 10 | #DIV/0! |
| 5 | Total | 80 | #DIV/0! |

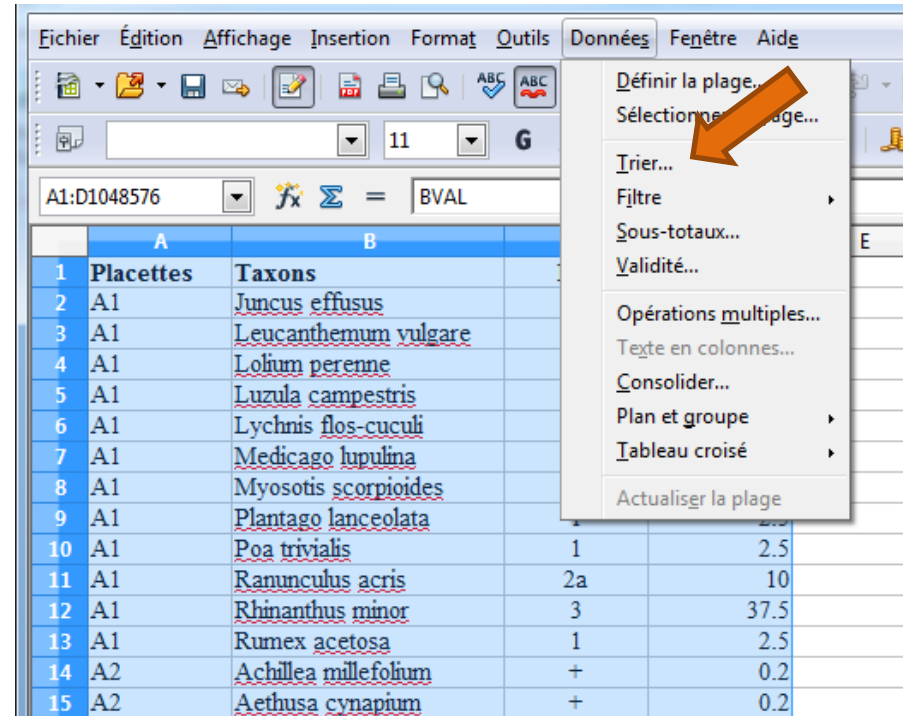
| | A | B | C | D |
|---|----------|---------------|--------|---------------|
| 1 | Taxon | Nbr individus | % | formule |
| 2 | Espece 1 | 40 | 50.0% | =B2/B\$5 |
| 3 | Espece 2 | 30 | 37.5% | =B3/B\$5 |
| 4 | Espece 3 | 10 | 12.5% | =B4/B\$5 |
| 5 | Total | 80 | 100.0% | =somme(C2:C4) |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Trier les tableaux :

- les lignes : simple et logique
- ne pas prendre en compte les lignes de titres
- nécessité de le faire pour l'ensemble des lignes d'un fichier sinon on perd les relations entre les données !



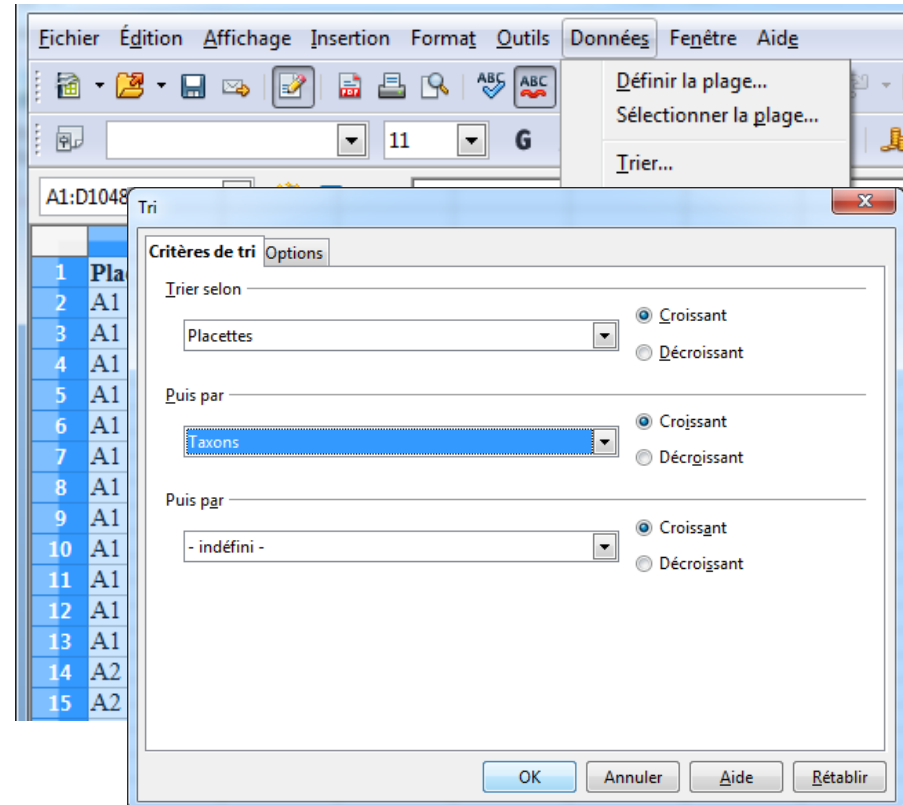
- possible de le faire aussi sur des colonnes.
- attention si il y a des formules utilisant des références relatives (d'abord copier - collage spécial rien que les valeurs)

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Trier les tableaux :

- les lignes : simple et logique
- ne pas prendre en compte les lignes de titres
- nécessité de le faire pour l'ensemble des lignes d'un fichier sinon on perd les relations entre les données !

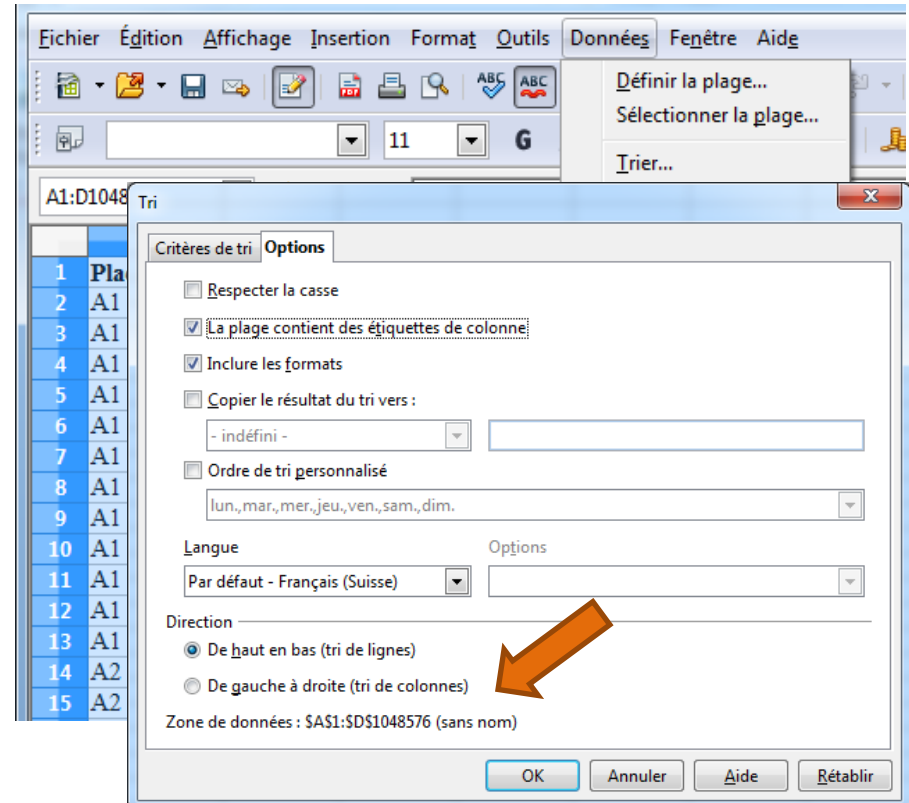


B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Trier les tableaux :

- les lignes : simple et logique
- ne pas prendre en compte les lignes de titres
- nécessité de le faire pour l'ensemble des lignes d'un fichier sinon on perd les relations entre les données !



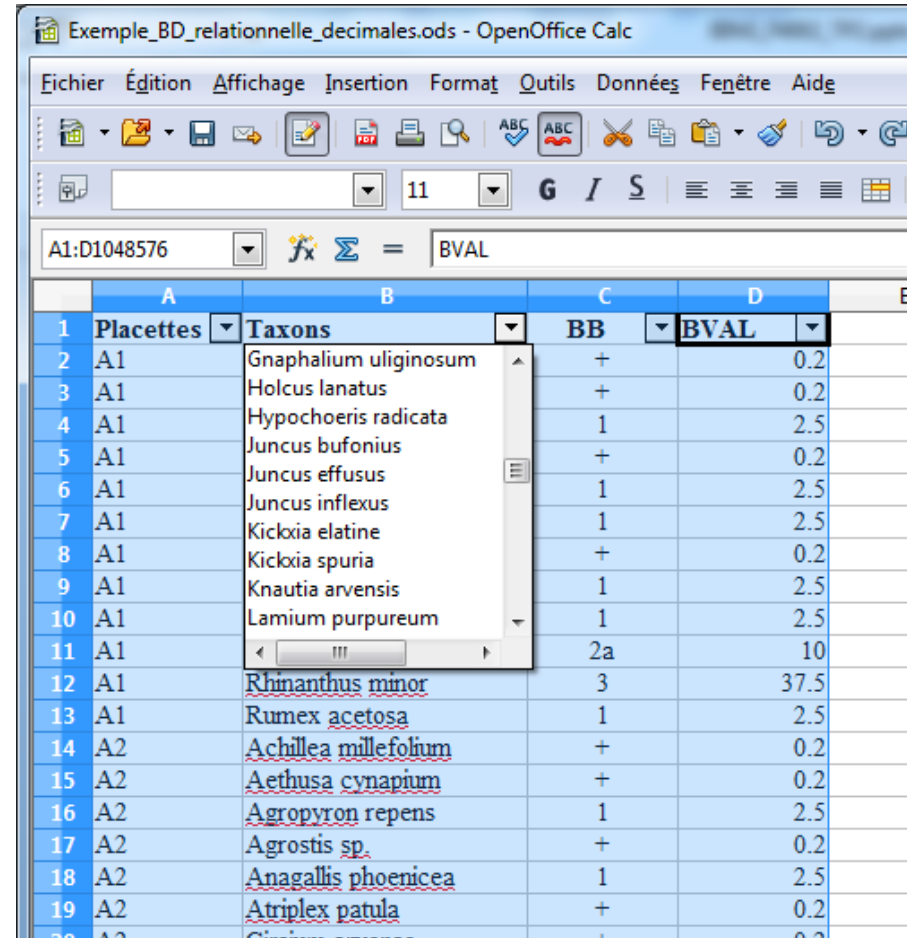
- possible de le faire aussi sur des colonnes.
- attention si il y a des formules utilisant des références relatives (d'abord copier - collage spécial rien que les valeurs)

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Filtrer les tableaux :

- Permet de sélectionner les lignes d'un tableau
- Chercher la liste des placettes où *Juncus effusus* a été observé
- Vérifier si la liste des codes BB (Braun-Blanquet) est correcte ...



The screenshot shows the OpenOffice Calc interface with a spreadsheet titled 'Exemple_BD_relationnelle_decimales.ods'. The spreadsheet has columns A, B, C, and D. Column A is labeled 'Placettes' and contains codes like A1, A2. Column B is labeled 'Taxons' and contains plant species names. Column C is labeled 'BB' and contains Braun-Blanquet codes. Column D is labeled 'BVAL' and contains numerical values. A dropdown menu is open for the 'Taxons' column, showing a list of species including *Gnaphalium uliginosum*, *Holcus lanatus*, *Hypochoeris radicata*, *Juncus bufonius*, *Juncus effusus*, *Juncus inflexus*, *Kickxia elatine*, *Kickxia spuria*, *Knautia arvensis*, *Lamium purpureum*, *Rhinanthus minor*, *Rumex acetosa*, *Achillea millefolium*, *Aethusa cynapium*, *Agropyron repens*, *Agrostis sp.*, *Anagallis phoenicea*, and *Atriplex patula*.

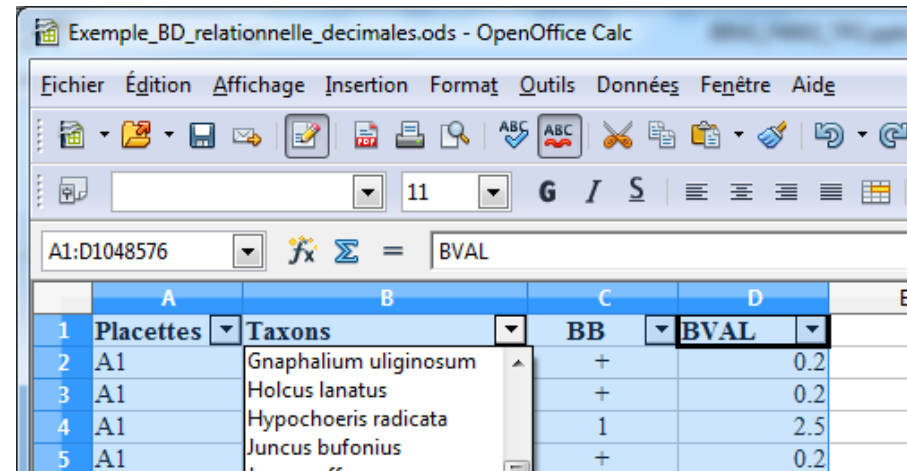
| Placettes | Taxons | BB | BVAL |
|-----------|------------------------------|----|------|
| A1 | <i>Gnaphalium uliginosum</i> | + | 0.2 |
| A1 | <i>Holcus lanatus</i> | + | 0.2 |
| A1 | <i>Hypochoeris radicata</i> | 1 | 2.5 |
| A1 | <i>Juncus bufonius</i> | + | 0.2 |
| A1 | <i>Juncus effusus</i> | 1 | 2.5 |
| A1 | <i>Juncus inflexus</i> | 1 | 2.5 |
| A1 | <i>Kickxia elatine</i> | + | 0.2 |
| A1 | <i>Kickxia spuria</i> | 1 | 2.5 |
| A1 | <i>Knautia arvensis</i> | 1 | 2.5 |
| A1 | <i>Lamium purpureum</i> | 2a | 10 |
| A1 | <i>Rhinanthus minor</i> | 3 | 37.5 |
| A1 | <i>Rumex acetosa</i> | 1 | 2.5 |
| A2 | <i>Achillea millefolium</i> | + | 0.2 |
| A2 | <i>Aethusa cynapium</i> | + | 0.2 |
| A2 | <i>Agropyron repens</i> | 1 | 2.5 |
| A2 | <i>Agrostis sp.</i> | + | 0.2 |
| A2 | <i>Anagallis phoenicea</i> | 1 | 2.5 |
| A2 | <i>Atriplex patula</i> | + | 0.2 |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Filtrer les tableaux :

- Permet de sélectionner les lignes d'un tableau
- Chercher la liste des placettes où *Achillea millefolium* a été observé



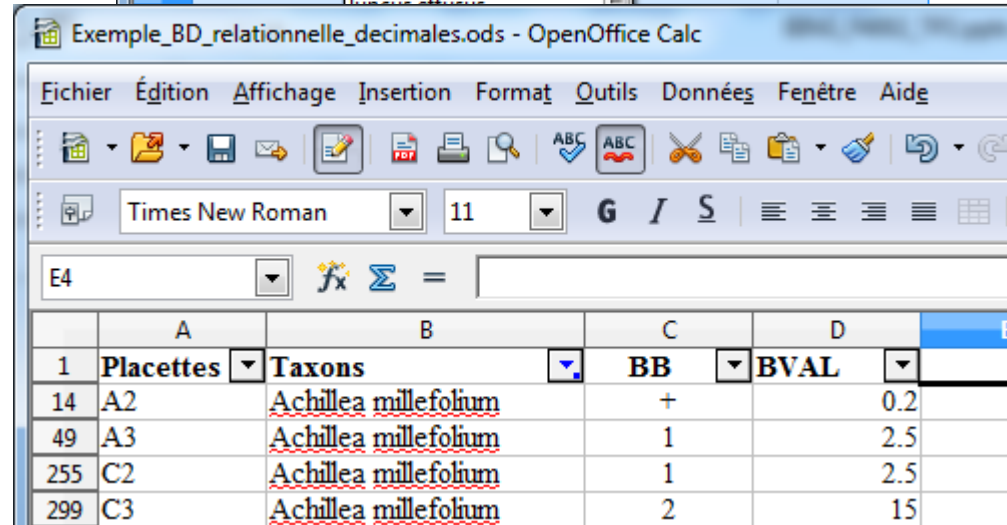
Exemple_BD_relationnelle_decimales.ods - OpenOffice Calc

Fichier Édition Affichage Insertion Format Outils Données Fenêtre Aide

11 G I S

A1:D1048576 = BVAL

| | A | B | C | D |
|---|-----------|-----------------------|----|------|
| 1 | Placettes | Taxons | BB | BVAL |
| 2 | A1 | Gnaphalium uliginosum | + | 0.2 |
| 3 | A1 | Holcus lanatus | + | 0.2 |
| 4 | A1 | Hypochoeris radicata | 1 | 2.5 |
| 5 | A1 | Juncus bufonius | + | 0.2 |



Exemple_BD_relationnelle_decimales.ods - OpenOffice Calc

Fichier Édition Affichage Insertion Format Outils Données Fenêtre Aide

Times New Roman 11 G I S

E4 =

| | A | B | C | D |
|-----|-----------|-----------------------------|----|------|
| 1 | Placettes | Taxons | BB | BVAL |
| 14 | A2 | <u>Achillea millefolium</u> | + | 0.2 |
| 49 | A3 | <u>Achillea millefolium</u> | 1 | 2.5 |
| 255 | C2 | <u>Achillea millefolium</u> | 1 | 2.5 |
| 299 | C3 | <u>Achillea millefolium</u> | 2 | 15 |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Filtrer les tableaux :

- Permet de sélectionner les lignes d'un tableau
- Voir si la liste des codes BB (Braun-Blanquet) est correcte

| Classes de recouvrement | Signification |
|-------------------------|----------------------|
| r | un individu |
| + | recouv. insignifiant |
| 1 | moins de 5 % |
| 2 | de 5 à 25 % |
| 3 | de 25 à 50 % |
| 4 | de 50 à 75 % |
| 5 | plus de 75 % |

2a (5 à 15%)

2b (15 à 25%)

Exemple_BD_relationnelle_decimales.ods - OpenOffice Calc

Fichier Édition Affichage Insertion Format Outils Données Fenêtre Aide

Times New Roman 11

E4

| | A | B | C | D |
|----|-----------|-----------------------------|--------------------|------|
| 1 | Placettes | Taxons | BB | BVAL |
| 2 | A1 | <u>Juncus effusus</u> | Tous | 0.2 |
| 3 | A1 | <u>Leucanthemum vulgare</u> | Top 10 | 0.2 |
| 4 | A1 | <u>Lolium perenne</u> | Filtre standard... | 2.5 |
| 5 | A1 | <u>Luzula campestris</u> | + | 0.2 |
| 6 | A1 | <u>Lychnis flos-cuculi</u> | 1 | 2.5 |
| 7 | A1 | <u>Medicago lupulina</u> | 2 | 2.5 |
| 8 | A1 | <u>Myosotis scorpioides</u> | 2a | 0.2 |
| 9 | A1 | <u>Plantago lanceolata</u> | 2b | 2.5 |
| 10 | A1 | <u>Poa trivialis</u> | 3 | 2.5 |
| 11 | A1 | <u>Ranunculus acris</u> | 5 | 10 |
| 12 | A1 | <u>Rhinanthus minor</u> | 3 | 37.5 |
| 13 | A1 | <u>Rumex acetosa</u> | 1 | 2.5 |
| 14 | A2 | <u>Achillea millefolium</u> | + | 0.2 |
| 15 | A2 | <u>Aethusa cynapium</u> | + | 0.2 |
| 16 | A2 | <u>Agropyron repens</u> | 1 | 2.5 |
| 17 | A2 | <u>Agrostis sp.</u> | + | 0.2 |
| 18 | A2 | <u>Anagallis phoenicea</u> | 1 | 2.5 |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Exercices :

- calculer la somme des abondances de la table "Relevés"
- générer une abréviation de noms d'espèces
(3 caractères pour le genre et 3 pour l'espèce)

| Si 2 colonnes | | | Si 1 colonne |
|---------------|-------------------|--------|------------------------|
| Genre | Especie | | Taxons |
| <i>Abax</i> | <i>ater</i> | Abaate | <i>Abax ater</i> |
| <i>Abax</i> | <i>ovalis</i> | Abaova | <i>Abax ovalis</i> |
| <i>Abax</i> | <i>parallelus</i> | Abapar | <i>Abax parallelus</i> |

- calculer le nombre de jours depuis que vous êtes nés
- quand sera votre 10000ème jouriversaire ?

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Exercices :

- générer une abréviation de noms d'espèces
(3 caractères pour le genre et 3 pour l'espèce)

| | A | B | C | |
|---|------|------------|--------|----------------------------|
| 1 | Abax | ater | Abaate | =STXT(A1;1;3)&STXT(B1;1;3) |
| 2 | Abax | parallelus | Abapar | =STXT(A2;1;3)&STXT(B2;1;3) |
| 3 | Abax | ovatus | Abaova | =STXT(A3;1;3)&STXT(B3;1;3) |

En majuscule ? =MAJUSCULE(STXT(A11;1;3)&STXT(B11;1;3))

Si le nom de taxon est en une seule colonne ?

| | A | B | C | D | |
|---|-----------------|-----------------|---------|---------|--|
| 1 | Taxon | position du " " | Genre | Especie | |
| 2 | Abax ater | 5 | Abax | ater | |
| 3 | Carabus auratus | 8 | Carabus | auratus | |
| 4 | Agonum gracile | 7 | Agonum | gracile | |

=CHERCHE(" ";A2;1)
 =STXT(A2;B2+1;NBCAR(A2))
 =STXT(A2;1;B2-1)

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Exercices :

- calculer le nombre de jours depuis que vous êtes nés

Les dates sont en fait exprimées en nombre de jours depuis le 1/1/1900 et cela est formaté en différents formats de date.

Donc, la différence entre deux dates est vite calculée ...

| Date naissance | Date du jour | Différence | formule utilisée |
|----------------|--------------|------------|------------------|
| 17/10/1988 | 17/10/2008 | 7305 | = B1-A1 |

- quand sera votre 10000 ème jouriversaire ?

| Date naissance | durée | 10000 ème jouriversaire | formule utilisée |
|----------------|-------|-------------------------|------------------|
| 17/10/1988 | 10000 | 04/03/16 | = A1+B1 |
| 32433 | 10000 | 42433 | |

=> 04/03/2016

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Conseils pour remplir son tableau de données :

- Une feuille différente pour chaque table
- Codage « intelligent » de sa structure d'échantillonnage
 - A, B, C, ... ou station 1, 2, 3a, 5b, ... ne sont pas « signifiants ».
 - Utiliser la structure d'échantillonnage pour les stations comme :
 - ESmArd1 pour la première répétition (1) d'une station équiennne (E), sur sols marginaux (Sm) en Ardenne (Ard). Ce code est signifiant dans les tableaux de synthèse, les résultats des analyses, les graphiques, ...

Règle = réserver toujours le même nbre de caractères pour une strate. Cela permet de décomposer facilement le code si nécessaire (fct STXT).

| D2 | | fx | | =STXT(B2;1;2) | | |
|----|----------|---------|--------|---------------|-------------|------|
| | A | B | C | D | E | F |
| 1 | OTU | Sample | NBRIND | Tillage | Crop_residu | Prof |
| 2 | Otu00017 | RTO107A | 1 | RT | O | A |
| 3 | Otu00019 | RTI094B | 1 | RT | I | B |
| 4 | Otu00022 | CTI097A | 1 | CT | I | A |
| 5 | Otu00022 | RTO091B | 1 | RT | O | B |
| 6 | Otu00025 | RTO100A | 1 | RT | O | A |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Exercices :

- Calculer la date de première observation d'une espèce ?

Feuille "Relevés_libellules" => BD_Exemple_CALC.ods ou BD_Exemple_XLS.xlsx

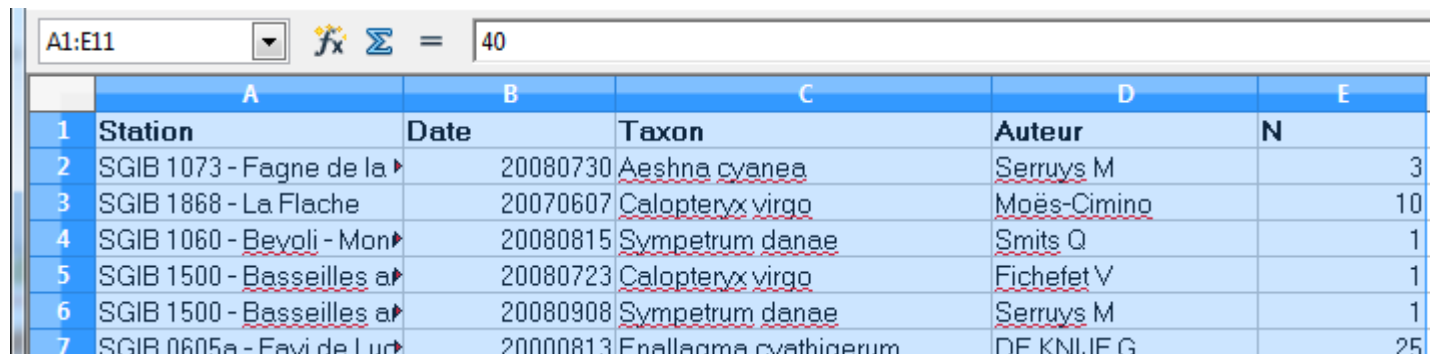
| A1:E11 | | | | | 40 |
|--------|--------------------------|----------|---------------------------------|--------------------|----|
| | A | B | C | D | E |
| 1 | Station | Date | Taxon | Auteur | N |
| 2 | SGIB 1073 - Fagne de la | 20080730 | <u>Aeshna cyanea</u> | <u>Serruys M</u> | 3 |
| 3 | SGIB 1868 - La Flache | 20070607 | <u>Calopteryx virgo</u> | <u>Moës-Cimino</u> | 10 |
| 4 | SGIB 1060 - Beyoli - Mon | 20080815 | <u>Sympetrum danae</u> | <u>Smits Q</u> | 1 |
| 5 | SGIB 1500 - Basseilles a | 20080723 | <u>Calopteryx virgo</u> | <u>Fichet V</u> | 1 |
| 6 | SGIB 1500 - Basseilles a | 20080908 | <u>Sympetrum danae</u> | <u>Serruys M</u> | 1 |
| 7 | SGIB 0605a - Fayi de Luc | 20000813 | <u>Enallagma cyathigerum</u> | <u>DE KNIJF G</u> | 25 |
| 8 | SGIB 1582 - Etang de Bil | 20080806 | <u>Cordulegaster boltonii</u> | <u>Heymans E</u> | 1 |
| 9 | SGIB 1582 - Etang de Bil | 20040709 | <u>Lestes sponsa</u> | <u>BAUGNEE JY</u> | 20 |
| 10 | SGIB 1059a - Tollihot | 20080702 | <u>Orthetrum coerulescens</u> | <u>Dierstein A</u> | 1 |
| 11 | SGIB 0605a - Fayi de Luc | 20080701 | <u>Libellula quadrimaculata</u> | <u>Dufrene M</u> | 40 |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Exercices :

- Calculer la date de première observation d'une espèce ?



| | A | B | C | D | E |
|---|--------------------------|----------|------------------------------|--------------------|----|
| 1 | Station | Date | Taxon | Auteur | N |
| 2 | SGIB 1073 - Fagne de la | 20080730 | <u>Aeshna cyanea</u> | <u>Serruys M</u> | 3 |
| 3 | SGIB 1868 - La Flache | 20070607 | <u>Calopteryx virgo</u> | <u>Moës-Cimino</u> | 10 |
| 4 | SGIB 1060 - Beyoli - Mon | 20080815 | <u>Sympetrum danae</u> | <u>Smits Q</u> | 1 |
| 5 | SGIB 1500 - Basseilles a | 20080723 | <u>Calopteryx virgo</u> | <u>Fichefet V</u> | 1 |
| 6 | SGIB 1500 - Basseilles a | 20080908 | <u>Sympetrum danae</u> | <u>Serruys M</u> | 1 |
| 7 | SGIB 0605a - Favi de Lur | 20000813 | <u>Enallagma cyathigerum</u> | <u>DF KNIJF G</u> | 25 |

- trier le fichier par *Taxon* et *Date*
- créer une variable qui identifie la première mention d'une espèce en comparant le nom de taxon d'une ligne et de la précédente [=SI(C1=C2;1;0)]
- copier - collage spécial de cette variable pour ne garder que les « valeurs »
- trier ou filtrer le fichier par cette variable ... Les lignes avec « 0 » = première observation.

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Conseils pour remplir son tableau de données :

- Format d'encodage doit être au plus proche de celui récolté sur le terrain ou au labo
- Nom des variables (colonnes) :
 - Commencer par une lettre, puis lettres et chiffres et éventuellement "_"; Pas de + ou – ou / ou () ou [] ou ... qui sont mal interprétés dans des formules et inacceptés lors de l'importation.
 - Majuscules tolérées mais il faudra les utiliser en fonction des logiciels d'analyses (R est "*case sensitive*" mais SAS ne l'est pas), ...
 - Etre intelligible, relativement court (pour les graphiques), **UNIVOQUE** (donc pour R : "PH" est différent de "Ph" ou de "pH" mais d'autres logiciels ne verront pas de différences !)

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Importer / exporter des données :

- copier-coller de word, de tableaux dans des pages web, ...
- sous la forme de fichiers en différents formats :
 - texte tabulé (txt ou tab)
 - dbf (access), csv, ...
- la logique => une feuille xls = un tableau de données

En cas d'exportation en format txt ou tab , on perd le formatage de ses données ...

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Données encodées en format "vectoriel" => "matriciel"

- type observation d'espèces :
- Placettes, taxons, nombre d'individus

| | A | B | C | D |
|---|-----------|-----------------------------|----|------|
| 1 | Placettes | Taxons | BB | BVAL |
| 2 | A1 | <u>Juncus effusus</u> | + | 0.2 |
| 3 | A1 | <u>Leucanthemum vulgare</u> | + | 0.2 |
| 4 | A1 | <u>Lolium perenne</u> | 1 | 2.5 |
| 5 | A1 | <u>Luzula campestris</u> | + | 0.2 |
| 6 | A1 | <u>Lychnis flos-cuculi</u> | 1 | 2.5 |
| 7 | A1 | Medicago lupulina | 1 | 2.5 |

| | A | B | C | D | E | F | G |
|-----|----------|----------------------|-------------------|------------------|------------------|--------------------|-------------|
| 361 | commune | lav | | | | | |
| 362 | Placette | Achillea millefolium | Achillea ptarmica | Aethusa cynapium | Agropyron repens | Agrostemma githago | Agrostis sp |
| 363 | A1 | | | | | | |
| 364 | A2 | 0.2 | | 0.2 | 2.5 | | 0.2 |
| 365 | A3 | 2.5 | | 2.5 | 2.5 | 2.5 | |
| 366 | B1 | | | | | | |
| 367 | B2 | | | | | | |
| 368 | B3 | | | | | | |
| 369 | C1 | | | | | | |
| 370 | C2 | 2.5 | 15 | | | | |
| 371 | C3 | 15 | | 15 | 0.2 | | |
| 372 | Total P | 20.2 | 15 | 17.7 | 5.2 | 2.5 | 0.2 |

- Construction d'une matrice
- Placettes x taxons avec le nombre d'individus total

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

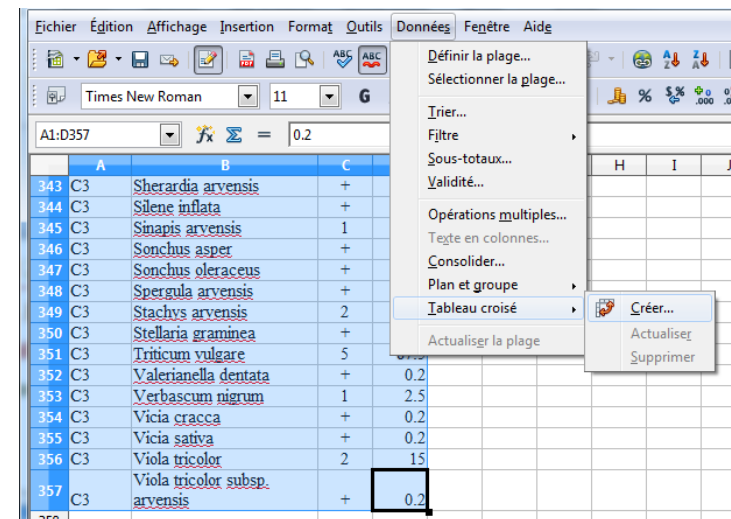
Données encodées en format "vectoriel" => "matriciel" (CALC)

- type observation d'espèces :
- Placettes, taxons, nombre d'individus

| | A | B | C | D |
|---|-----------|-----------------------------|----|------|
| 1 | Placettes | Taxons | BB | BVAL |
| 2 | A1 | <u>Juncus effusus</u> | + | 0.2 |
| 3 | A1 | <u>Leucanthemum vulgare</u> | + | 0.2 |
| 4 | A1 | <u>Lolium perenne</u> | 1 | 2.5 |
| 5 | A1 | <u>Luzula campestris</u> | + | 0.2 |
| 6 | A1 | <u>Lychnis flos-cuculi</u> | 1 | 2.5 |
| 7 | A1 | <u>Medicago lupulina</u> | 1 | 2.5 |

- Sélection du tableau puis

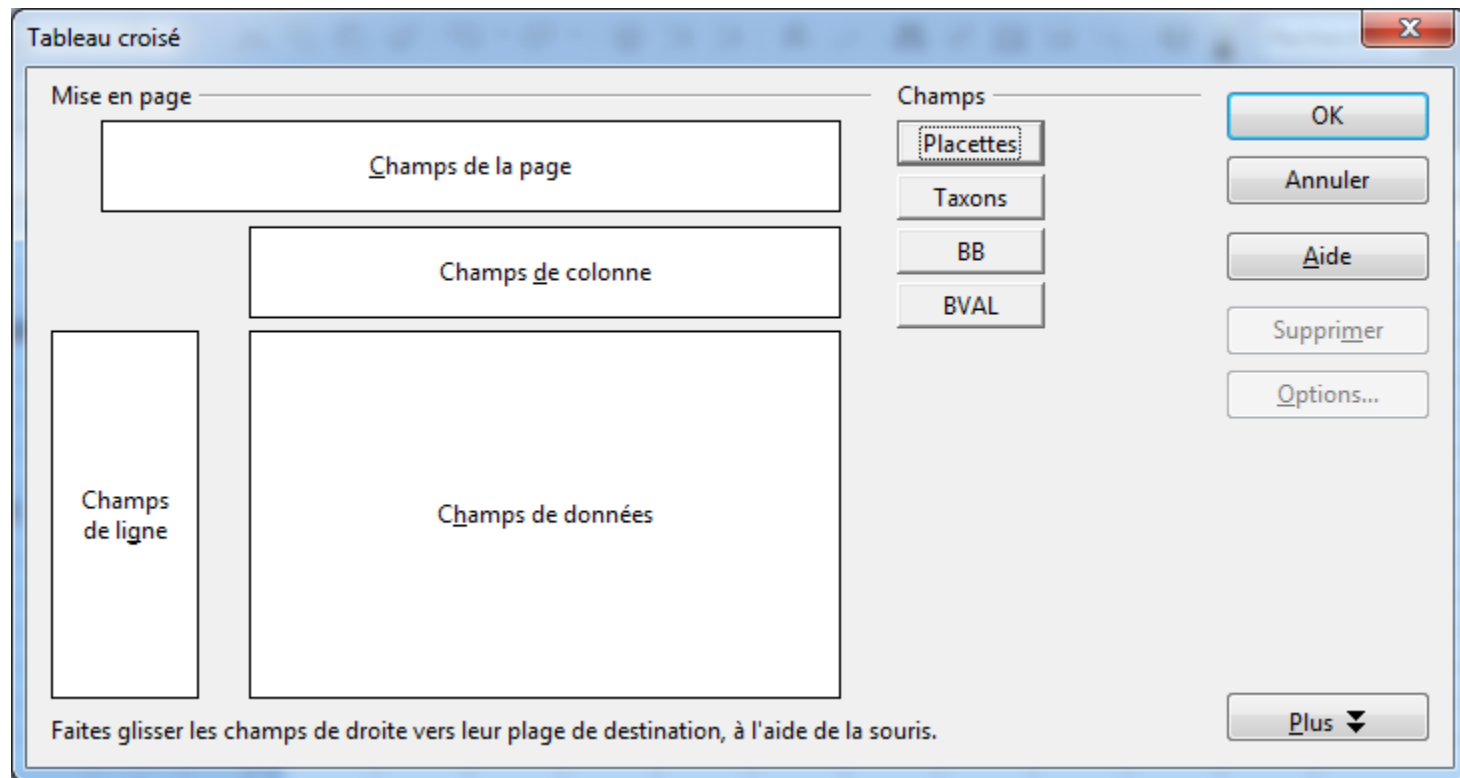
- CALC : Données > Tableau croisé
- XLS : Insertion > TablCroiséDynamique



B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

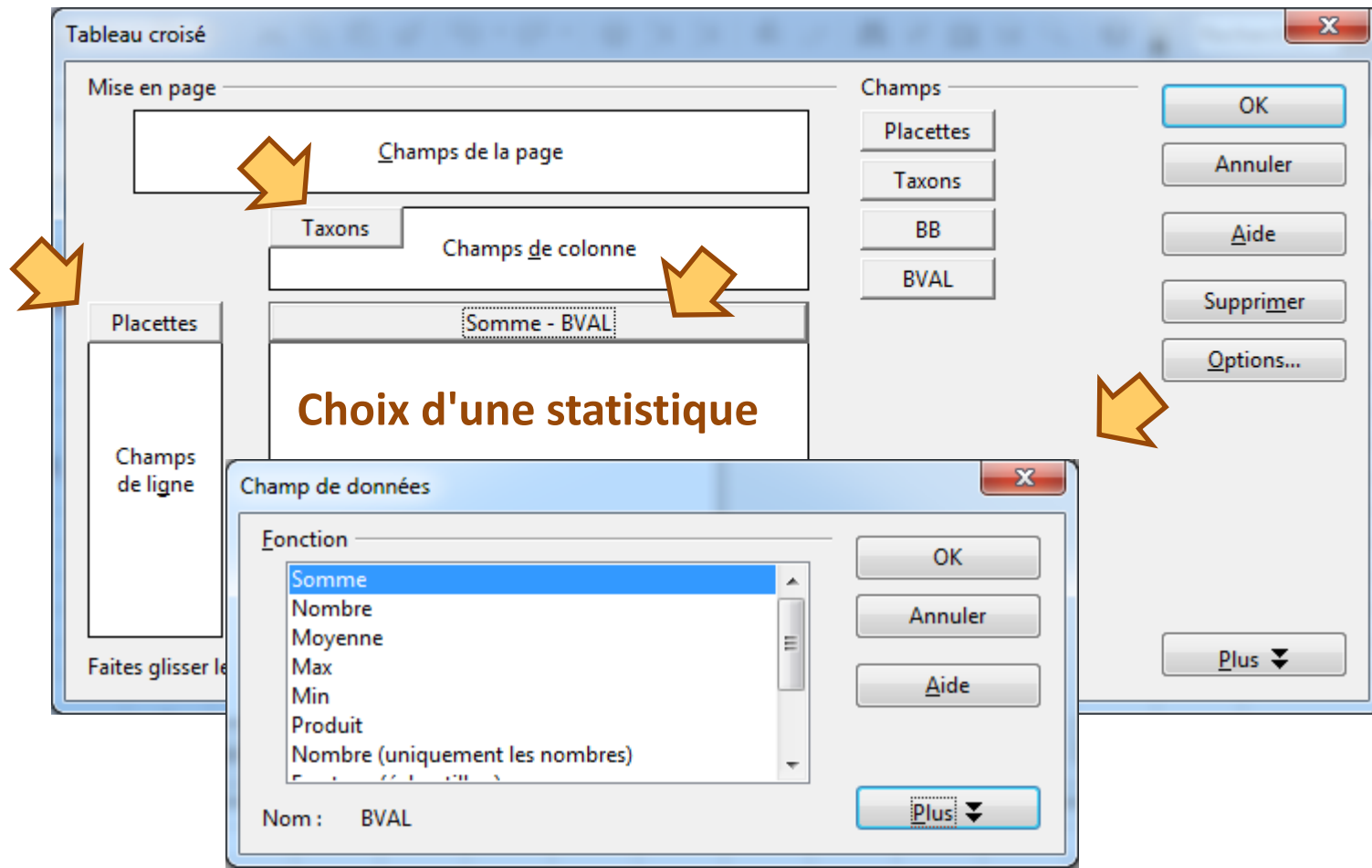
Données encodées en format "vectoriel" => "matriciel" (CALC)



B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Données encodées en format "vectoriel" => "matriciel" (CALC)



B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Données encodées en format "vectoriel" => "matriciel" (CALC)

Possibilité de
filtrer des
données

Localisation de
la matrice

Afficher totaux

Cellule vide = 0 => dans XLS

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Données encodées en format "vectoriel" => "matriciel" (CALC)

| Somm | Taxons | | | | | | | | | | | | | |
|----------|----------------------|-------------------|------------------|------------------|--------------------|--------------|----------------------|-------------------|---------------------|------------------------|----------------------|--------------------|---------------------|-----------------------|
| Placette | Achillea millefolium | Achillea ptarmica | Aethusa cynapium | Agropyron repens | Agrostemma githago | Agrostis sp. | Agrostis stolonifera | Agrostis vulgaris | Alchemilla arvensis | Alopecurus myosuroides | Alopecurus pratensis | Anagallis arvensis | Anagallis phoenicea | Anthriscus silvestris |
| A1 | | | | | | | | | | | | | | |
| A2 | 0.2 | 0.2 | 2.5 | | 0.2 | | | | | | | 2.5 | | |
| A3 | 2.5 | 2.5 | 2.5 | 2.5 | | | 0.2 | 15 | | | | 30 | 2.5 | |
| B1 | | | | | | | | | | 10 | | | | |
| B2 | | | | | 0.2 | | 15 | 0.2 | | | 15 | | 0.2 | |
| B3 | | | | | | | | | | | | | 2.5 | |
| C1 | | | | | | | | | | | | | | |
| C2 | 2.5 | 15 | | | 0.2 | | | 0.2 | | | | | 0.2 | |
| C3 | 15 | 15 | 0.2 | | | | | 15 | | | | 15 | 2.5 | |
| Total | 20.2 | 15 | 17.7 | 5.2 | 2.5 | 0.2 | 0.4 | 0.2 | 45 | 0.4 | 10 | 15 | 47.5 | 7.9 |

On vérifie si le total global est correct : 1758.4

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Données encodées en format "vectoriel" => "matriciel" (XLS)

The image shows a Microsoft Excel spreadsheet with two tables. The first table, on the left, is a data table with columns: OTU, Sample, NBRIND, Tillage, Crop_residue, and Prof. The second table, on the right, is a PivotTable created from the first table. The PivotTable has 'Tillage' as the row label and 'Crop_residue' as the column label. The values are the sum of NBRIND for each combination of Tillage and Crop_residue. A yellow arrow points from the 'NBRIND' column of the first table to the 'Somme de NBRIND' cell in the PivotTable.

| OTU | Sample | NBRIND | Tillage | Crop_residue | Prof |
|----------|---------|--------|---------|--------------|------|
| Otu00017 | RTO107A | 1 | RT | O | A |
| Otu00019 | RTI094B | 1 | RT | I | B |
| Otu00022 | CTI097A | 1 | CT | I | A |
| Otu00022 | RTO091B | 1 | RT | O | B |
| Otu00025 | RTO100A | 1 | RT | O | A |
| Otu00026 | RTI108A | 1 | RT | I | A |
| Otu00026 | RTI110A | 1 | RT | I | A |
| Otu00026 | RTO107A | 1 | RT | O | A |
| Otu00030 | CTO095B | 1 | CT | O | B |
| Otu00030 | RTI108A | 1 | RT | I | A |
| Otu00034 | RTO107A | 1 | RT | O | A |
| Otu00035 | RTO100A | 1 | RT | O | A |
| Otu00036 | RTI108A | 1 | RT | I | A |
| Otu00036 | RTI110A | 1 | RT | I | A |
| Otu00036 | RTI099A | 1 | RT | I | A |
| Otu00036 | RTO102A | 1 | RT | O | A |
| Otu00038 | RTO102A | 1 | RT | O | A |
| Otu00039 | CTO095B | 1 | CT | O | B |
| Otu00039 | RTI108A | 1 | RT | I | A |
| Otu00039 | RTO109A | 1 | RT | O | A |
| Otu00040 | RTI099A | 1 | RT | I | A |
| Otu00040 | RTO093B | 1 | RT | O | B |
| Otu00041 | RTO107A | 1 | RT | O | A |
| Otu00044 | RTO100A | 1 | RT | O | A |
| Otu00045 | RTI110A | 1 | RT | I | A |
| Otu00045 | RTI085B | 1 | RT | I | B |
| Otu00045 | RTO100A | 1 | RT | O | A |

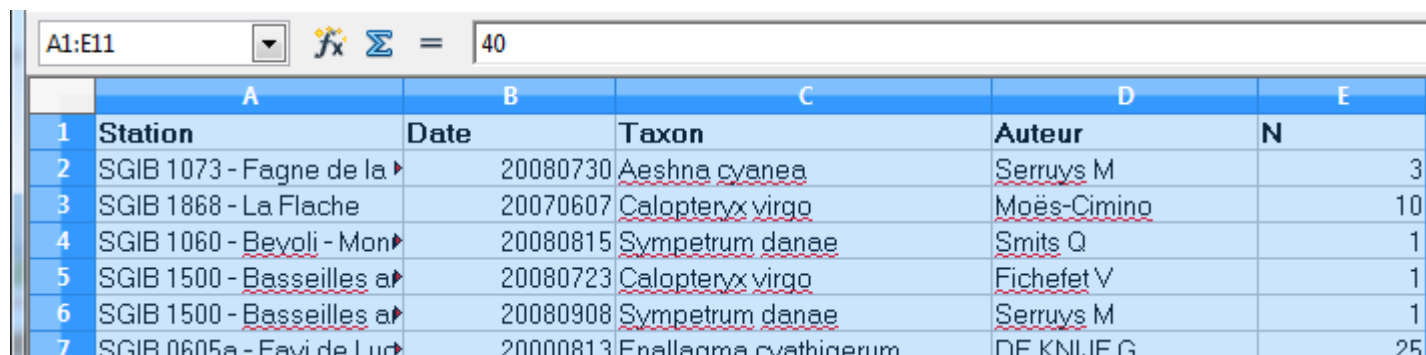
| | Tillage | CT | RT | | |
|----------------------|---------|---------|---------|---------|---------|
| Somme de NBRIND | 1 | 1 | 1 | | |
| Étiquettes de lignes | CTI081B | CTI087B | CTI090B | CTI096B | CTI097B |
| Otu00001 | | 157 | 133 | 218 | 142 |
| Otu00002 | | 92 | 70 | 62 | 49 |
| Otu00003 | | 77 | 41 | 40 | 53 |
| Otu00004 | | 35 | 25 | 26 | 33 |
| Otu00005 | | 71 | 33 | 39 | 51 |
| Otu00006 | | 29 | 24 | 47 | 36 |
| Otu00007 | | 41 | 30 | 32 | 34 |
| Otu00008 | | 30 | 19 | 14 | 32 |
| Otu00009 | | 33 | 20 | 20 | 47 |
| Otu00010 | | 11 | 14 | 11 | 26 |
| Otu00011 | | 13 | 31 | 18 | 17 |
| Otu00012 | | 11 | 9 | 4 | 16 |
| Otu00013 | | 19 | 12 | 10 | 18 |
| Otu00014 | | 14 | 19 | 13 | 32 |
| Otu00015 | | 5 | 9 | 7 | 9 |

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Exercices :

Feuille "Relevés_libellules" => BD_Exemple_CALC.ods ou BD_Exemple_XLS.xlsx



| | A | B | C | D | E |
|---|----------------------------------|----------|------------------------------|--------------------|----|
| 1 | Station | Date | Taxon | Auteur | N |
| 2 | SGIB 1073 - Fagne de la | 20080730 | <u>Aeshna cyanea</u> | <u>Serruys M</u> | 3 |
| 3 | SGIB 1868 - La Flache | 20070607 | <u>Calopteryx virgo</u> | <u>Moës-Cimino</u> | 10 |
| 4 | SGIB 1060 - <u>Beyoli - Mont</u> | 20080815 | <u>Sympetrum danae</u> | <u>Smits Q</u> | 1 |
| 5 | SGIB 1500 - <u>Basseilles a</u> | 20080723 | <u>Calopteryx virgo</u> | <u>Fichefet V</u> | 1 |
| 6 | SGIB 1500 - <u>Basseilles a</u> | 20080908 | <u>Sympetrum danae</u> | <u>Serruys M</u> | 1 |
| 7 | SGIB 0605a - Favi de Lur | 20000813 | <u>Enallagma cyathigerum</u> | <u>DE KNIJF G</u> | 25 |

- quelle est la somme des abondances pour chaque espèce ?
- quelle est la somme des abondances pour chaque observateur et chaque date ?
- quelle est la matrice de la somme des abondances par station et taxon ?

Cette fonction permet de gérer les données telles qu'elles sont récoltées (observation après observation) et de reconstruire les différents tableaux de synthèse pour les analyses très facilement

=> UNE SEULE SOURCE DE DONNEES !

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Gérer des relations entre les tables (XLS)

Feuille
« Parcelles »
triée par Parcelle

| A | B | C | D | E |
|------|----|-----|-----|----|
| parc | x1 | x2 | x3 | x4 |
| E01 | 10 | 234 | 1,1 | 11 |
| E02 | 23 | 123 | 1,2 | 12 |
| F01 | 12 | 456 | 1,1 | 9 |
| F02 | 14 | 678 | 1,0 | 14 |

Feuille
« Placettes »

| A | B | C | D | E |
|-----|------|-----|----|-----|
| pla | parc | y1 | y2 | y3 |
| A1 | E01 | 1,0 | 14 | 1,1 |
| A2 | E01 | 2,3 | 33 | 1,2 |
| A3 | E01 | 1,2 | 26 | 1,1 |
| A4 | E02 | 1,4 | 58 | 1,0 |

Feuille
« Placettes »

| A | B | C | D | E | F | G | | |
|-----|------|-----|----|-----|----|-----|-----|----|
| pla | parc | y1 | y2 | y3 | x1 | x2 | x3 | x4 |
| A1 | E01 | 1,0 | 14 | 1,1 | 10 | 234 | 1,1 | 11 |
| A2 | E01 | 2,3 | 33 | 1,2 | 10 | 234 | 1,1 | 11 |
| A3 | E01 | 1,2 | 26 | 1,1 | 10 | 234 | 1,1 | 11 |
| A4 | E02 | 1,4 | 58 | 1,0 | 23 | 123 | 1,2 | 12 |

XLS : RECHERCHE(valeur; vectcherche; vectreponse)

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Gérer des relations entre les tables (XLS)

Feuille
« Placettes »

Pour recopier vers le bas en gardant
les références absolues

| A | B | C | D | E | F | G | | |
|-----|------|-----|----|-----|----|-----|-----|----|
| pla | parc | y1 | y2 | y3 | x1 | x2 | x3 | x4 |
| A1 | E01 | 1,0 | 14 | 1,1 | 10 | 234 | 1,1 | 11 |
| A2 | E01 | 2,3 | 33 | 1,2 | 10 | 234 | 1,1 | 11 |
| A3 | E01 | 1,2 | 26 | 1,1 | 10 | 234 | 1,1 | 11 |
| A4 | E02 | 1,4 | 58 | 1,0 | 23 | 123 | 1,2 | 12 |

=RECHERCHE(B2;Parcelles!A\$2:A\$5;Parcelles!B\$2:B\$5)



Attention : si le code Par recherché n'existe pas dans la feuille « Parcelles » (ex: A22), XLS prend la valeur associée à ligne suivante (ici A3)

La solution = d'abord tester si on retrouve la valeur recherchée et si oui, on affiche la réponse de la colonne voulue

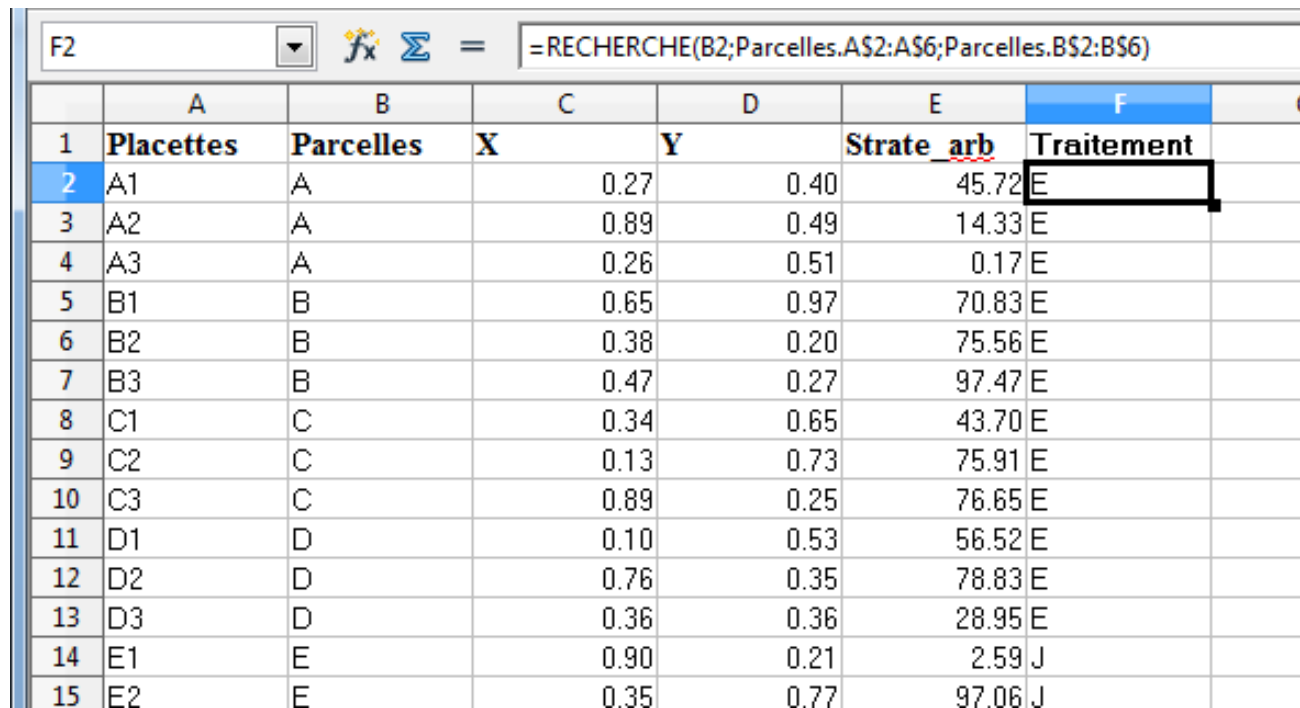
=SI(B2=RECHERCHE(B2;Parcelles!\$A\$2:\$A\$5;Parcelles!\$A\$2:\$A\$5);RECHERCHE(B2;Parcelles!\$A\$2:\$A\$5;Parcelles!B\$2:B\$5);"XXX")

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Gérer des relations entre les tables (CALC)

- Comme en XLS



| | A | B | C | D | E | F | G |
|----|-----------|-----------|------|------|------------|------------|---|
| 1 | Placettes | Parcelles | X | Y | Strate arb | Traitement | |
| 2 | A1 | A | 0.27 | 0.40 | 45.72 | E | |
| 3 | A2 | A | 0.89 | 0.49 | 14.33 | E | |
| 4 | A3 | A | 0.26 | 0.51 | 0.17 | E | |
| 5 | B1 | B | 0.65 | 0.97 | 70.83 | E | |
| 6 | B2 | B | 0.38 | 0.20 | 75.56 | E | |
| 7 | B3 | B | 0.47 | 0.27 | 97.47 | E | |
| 8 | C1 | C | 0.34 | 0.65 | 43.70 | E | |
| 9 | C2 | C | 0.13 | 0.73 | 75.91 | E | |
| 10 | C3 | C | 0.89 | 0.25 | 76.65 | E | |
| 11 | D1 | D | 0.10 | 0.53 | 56.52 | E | |
| 12 | D2 | D | 0.76 | 0.35 | 78.83 | E | |
| 13 | D3 | D | 0.36 | 0.36 | 28.95 | E | |
| 14 | E1 | E | 0.90 | 0.21 | 2.59 | J | |
| 15 | E2 | E | 0.35 | 0.77 | 97.06 | J | |

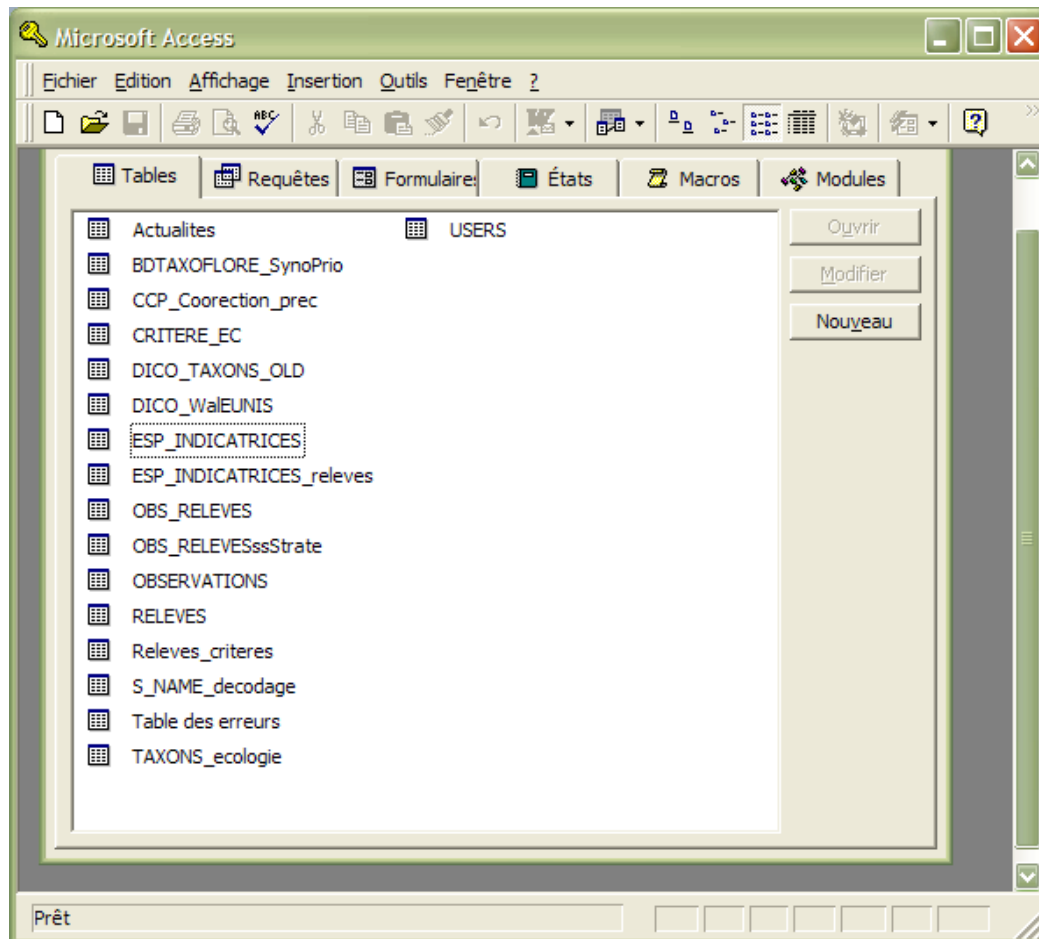
**VERIFIER
VOS
RESULTATS !**

- avec le même problème si une valeur n'existe pas ou n'est pas dans le bon ordre

B. La préparation d'un tableau de données

Utilisation d'un tableur : XLS ou CALC (kit de survie)

Pour gérer une base relationnelle => ACCESS (Win) ou BASE (OO)



Création ou
importation de tables

Création de requêtes
sur les tables (filtre,
tri, croisement,
tableau croisé, ...)

Pas très compliqué et
bon pour le CV ...

+ Création de
formulaires, ...

A. Mise en forme des fichiers

Exportation des données de XLS ou de CALC (ou autres BD relationnelles)

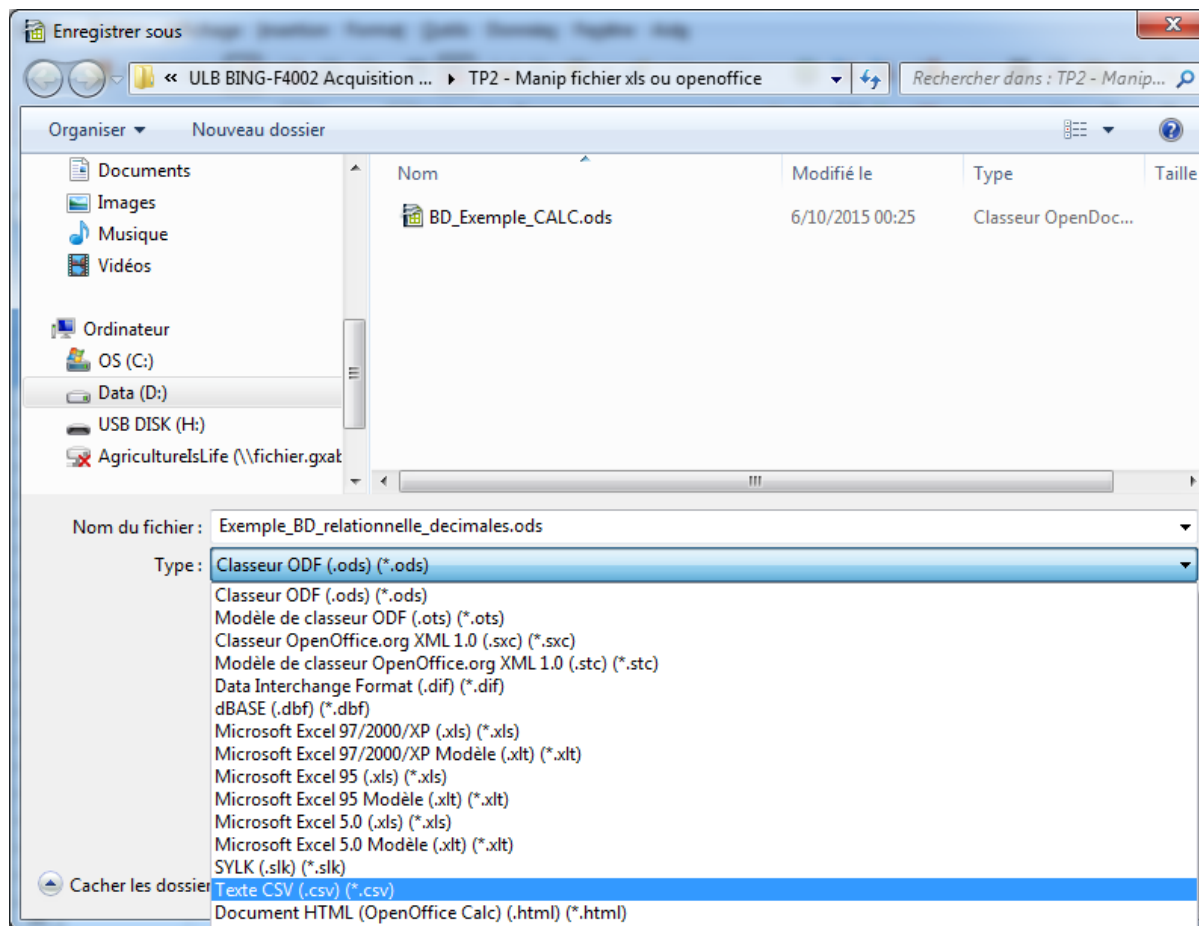
- Préférez le format "**text**" ou "**txt**" (sans format)
- Avec les colonnes séparées par des **tabulations** (pas des blancs si vous en avez dans vos colonnes. Ex : "Abax ater")
- **Eviter le CSV** (séparateur , ambigu ou ; avec des '"' autour des chaînes de caractères – différents CSV; pas tjrs reconnus)
- Exportation de tables avec les titres des colonnes et indices de lignes **clairs**, **UNIVOQUES**, respectant les contraintes de la plupart des logiciels (début = lettre, chiffre ou lettre ou "_" mais pas d'opérateurs (+, -, /, ...) ou de caractères bizarres)
- Abandonner les **caractères accentués** qui génèrent TOUJOURS des problèmes ...

=> Exporter les tables Parcelles, Placettes et Relevés

A. Mise en forme des fichiers

Exportation des données de XLS ou de CALC (ou autres BD relationnelles)

- **CALC : Fichier > Enregistrer sous > Texte CSV**

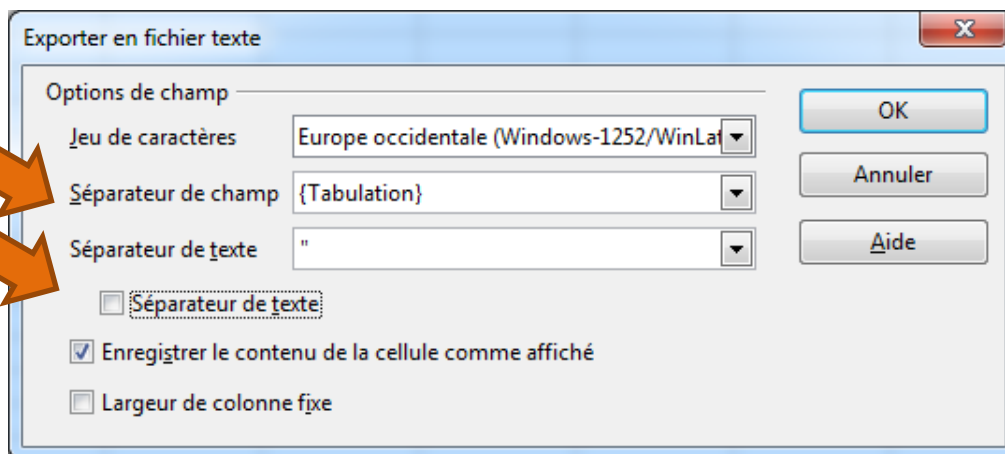
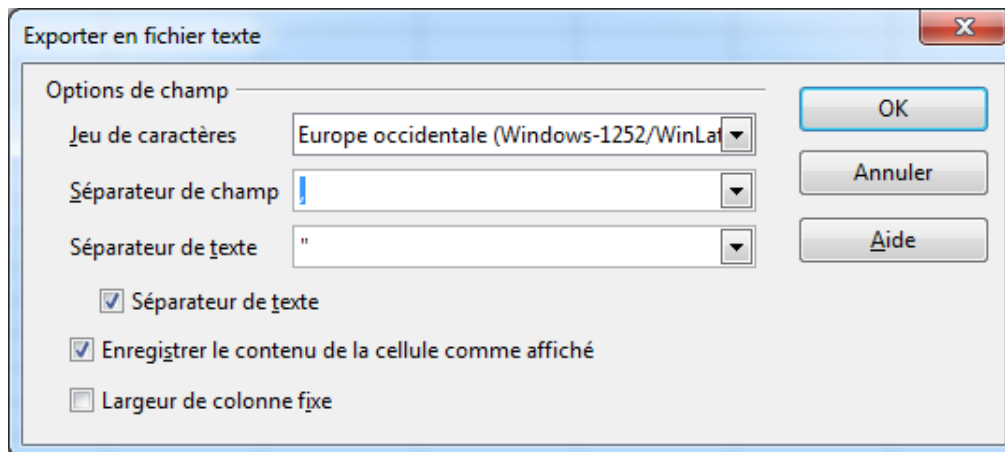


- **Format CSV ou autre format txt**
- **Conserver le format actuel**

A. Mise en forme des fichiers

Exportation des données de XLS ou de CALC (ou autres BD relationnelles)

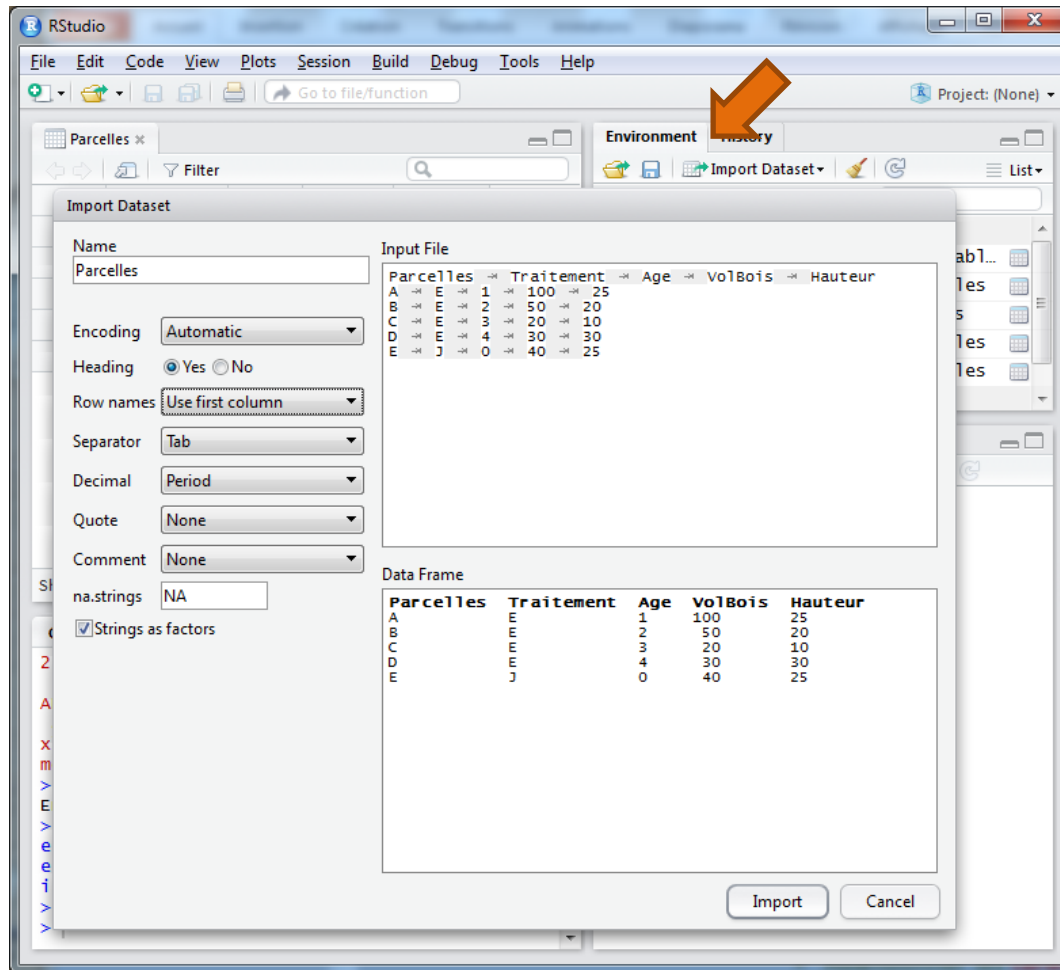
- **CALC : Fichier > Enregistrer sous > Texte CSV**



- **Séparateur => tabulation**
- **Pas de séparateur de texte**
- **Conserver le format actuel**
- **Vérifier dans un éditeur de texte ou dans R Studio**

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.



- Définit les formats du fichier txt
- Attention : pas de **caractères accentués** dans le chemin d'accès (WIN)

Cannot open file



'D:/Dufrene/GXABT

Enseignements/ULB BING-F4002

Acquisition et analyse de

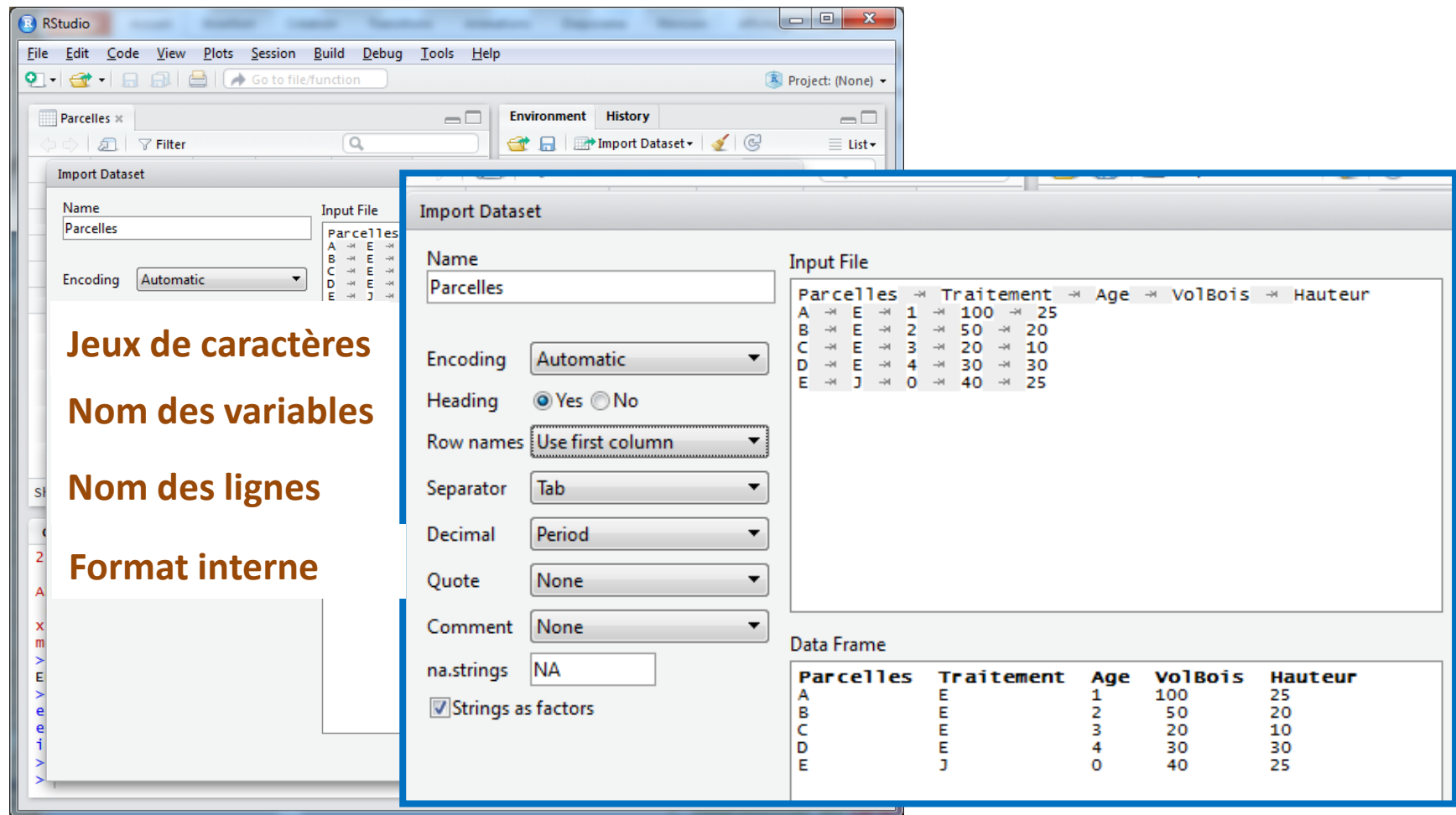
donnÃ©es/TP2 - Manip fichier xls

ou openoffice/Parcelles.txt': No

such file or directory

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction `read.table`.



Jeux de caractères

Nom des variables

Nom des lignes

Format interne

Import Dataset

Name: Parcelles

Input File: Parcelles

Encoding: Automatic

Heading: ☒ Yes ☐ No

Row names: Use first column

Separator: Tab

Decimal: Period

Quote: None

Comment: None

na.strings: NA

☒ Strings as factors

Data Frame

| Parcelles | Traitement | Age | VolBois | Hauteur |
|-----------|------------|-----|---------|---------|
| A | E | 1 | 100 | 25 |
| B | E | 2 | 50 | 20 |
| C | E | 3 | 20 | 10 |
| D | E | 4 | 30 | 30 |
| E | J | 0 | 40 | 25 |

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.

- Code R

> `file.show(file.choose())` pour voir le contenu d'un fichier

Définir le dossier de travail : **Menu Session > Set Working Directory** ou

> `setwd("D:/Dufrene/GXABT Enseignements/ULB BING-F4002
Acquisition et analyse de donnees/TP2 - Manip fichier xls ou
openoffice")`

> `read.table("Parcelles.txt")[1:3,]` pour voir les 3 premières lignes du fichier

```
> read.table("Parcelles.txt")[1:3, ]
      V1      V2  V3      V4      V5
1 Parcelles Traitement Age VolBois Hauteur
2      A      E    1     100      25
3      B      E    2      50      20
> |
```



← **Ligne de titre**

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.

- **Code R**

```
> read.table("BDExemple_parcelles.txt ", h=T)
```

| | Parcelles | Traitement | Age | VolBois |
|---|-----------|------------|-----|---------|
| 1 | A | E | 1 | 100 |
| 2 | B | E | 2 | 50 |
| 3 | C | E | 3 | 20 |
| 4 | D | E | 4 | 30 |
| 5 | E | J | 0 | 40 |

- **Option "header"**



La colonne avec les noms des Parcelles n'est pas comprise comme un identifiant unique des lignes mais bien comme une variable

Il faut préciser le séparateur => Option "row.names="

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.

- **Code R**

```
> read.table("Parcelles.txt", h=T, row.names="Parcelles")
```

| | Traitement | Age | VolBois | Hauteur |
|---|------------|-----|---------|---------|
| A | E | 1 | 100 | 25 |
| B | E | 2 | 50 | 20 |
| C | E | 3 | 20 | 10 |
| D | E | 4 | 30 | 30 |
| E | J | 0 | 40 | 25 |

- **Option "row.names"**



Si vous avez des blancs dans le contenu de vos variables, ils seront aussi interprétés comme des séparateurs par défaut

```
> read.table("Relevés.txt ", h=T) [1:10, ]  
Erreur dans scan(file, what, nmax, sep, dec, quote, skip,  
  la ligne 61 n'avait pas 5 éléments
```

Il faut préciser le séparateur => Option "sep ="

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.

- Code R

```
> read.table("Relevés.txt",h=T, sep="\t") [1:10, ]
  Placettes      Taxons BB BVAL
1      A1      Juncus effusus + 0.2
2      A1 Leucanthemum vulgare + 0.2
3      A1      Lolium perenne 1 2.5
4      A1      Luzula campestris + 0.2
5      A1 Lychnis flos-cuculi 1 2.5
6      A1      Medicago lupulina 1 2.5
7      A1 Myosotis scorpioides + 0.2
8      A1 Plantago lanceolata 1 2.5
9      A1      Poa trivialis 1 2.5
10     A1 Ranunculus acris 2a 10.0
```

- Option "sep"

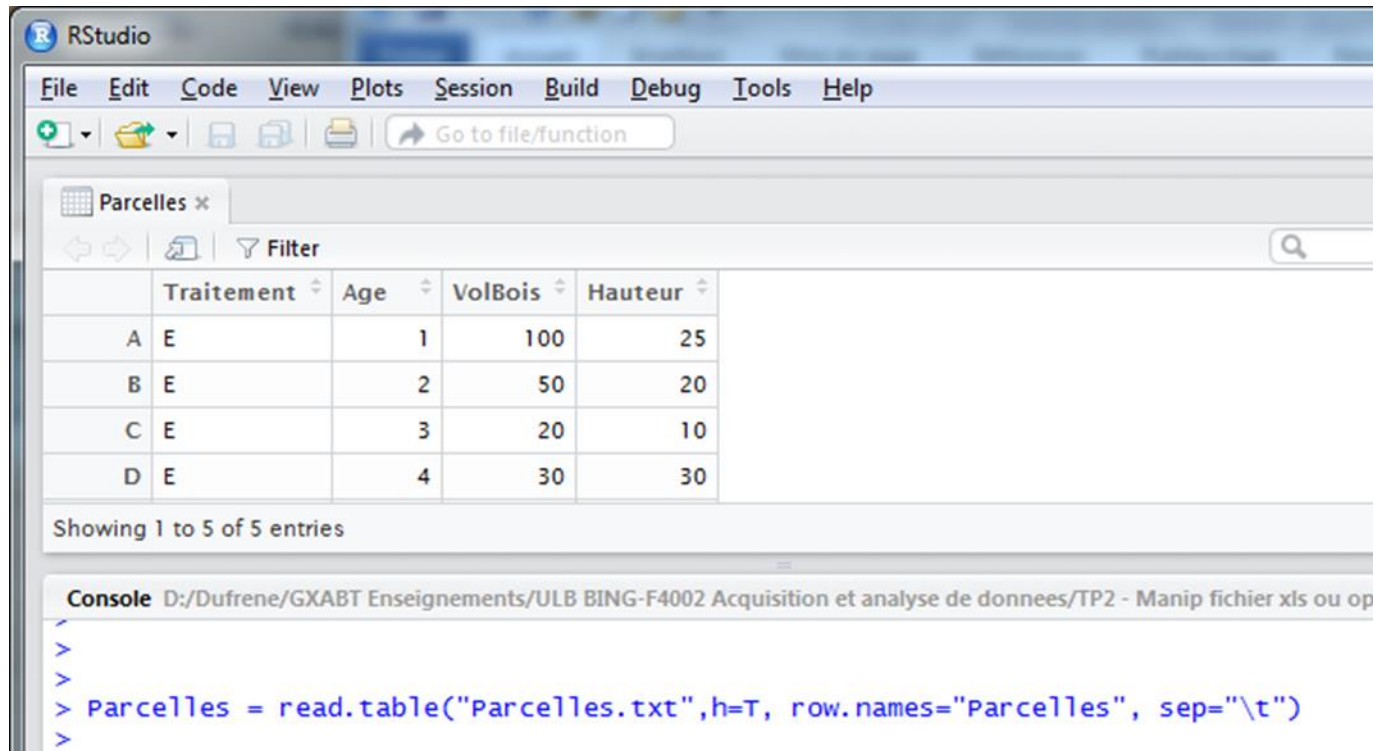
- Equivalent de l'interface d'importation :

```
Parcelles = read.table("Parcelles.txt",h=T, row.names="Parcelles", sep="\t")
```

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.

- Code R



RStudio interface showing the 'Parcelles' dataset imported from a text file. The dataset is displayed in the Environment pane with the following structure:

| | Traitement | Age | VolBois | Hauteur |
|---|------------|-----|---------|---------|
| A | E | 1 | 100 | 25 |
| B | E | 2 | 50 | 20 |
| C | E | 3 | 20 | 10 |
| D | E | 4 | 30 | 30 |

Showing 1 to 5 of 5 entries

Console output:

```
D:/Dufrene/GXABT Enseignements/ULB BING-F4002 Acquisition et analyse de donnees/TP2 - Manip fichier xls ou op
>
>
> Parcelles = read.table("Parcelles.txt", h=T, row.names="Parcelles", sep="\t")
>
```

=> Importer les 4 fichiers qui étaient dans CALC ou XLS

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.



Rappel important :

Quand les jeux de données sont longs, il est toujours difficile de vérifier si l'importation du contenu est correcte à l'œil. Utiliser de simples fonctions statistiques, comme la somme des nombres pour une ou plusieurs colonnes ou des calculs de fréquence de valeurs qualitatives dans le fichier de départ et après l'importation pour **TOUJOURS** vérifier si votre jeu de données est correct.

La somme totale par exemple des valeurs BBVAL du fichier « Releves.txt » est par exemple égale à 1758.4.

B. Importation dans R

L'importation de tables externes se fait simplement en utilisant **Import Dataset** de RStudio ou la fonction **read.table**.



```
> releves<-read.table("Relevés.txt", header=TRUE, sep="\t")
>
> dim(relevés)
[1] 356    4
> names(relevés)
[1] "Placettes" "Taxons"    "BB"        "BVAL"
>
> sum(relevés$BVAL)
[1] 1758.4
```

dim() = taille de la table

names() = nom des colonnes

sum() = somme d'une colonne

On notera qu'on peut aussi copier-coller des données (dans le clipborard ou le presse-papier) et les lire directement avec R.

```
monfichier<-read.table("clipboard",sep="\t",dec=",",header=T)
```

C. Manipulations de fichier

Fusion de tableaux => merge()

```
<
>
> Parcelles<-read.table("Parcelles.txt", header=TRUE, sep="\t")
> Placettes<-read.table("Placettes.txt", header=TRUE, sep="\t")
>
> Parcelles
  Parcelles Traitement Age VolBois Hauteur
1         A         E   1     100      25
2         B         E   2      50      20
3         C         E   3      20      10
4         D         E   4      30      30
5         E         J   0      40      25
> Placettes
  Placettes Parcelles      X      Y Strate_arb
1        A1         A 0.2659336 0.4024086 45.7170007
2        A2         A 0.8881164 0.4900394 14.3344909
3        A3         A 0.2613052 0.5059617  0.1679769
4        B1         B 0.6502693 0.9684797 70.8333599
5        B2         B 0.3836429 0.2010154 75.5601603
6        B3         B 0.4746960 0.2677185 97.4669143
7        C1         C 0.3351040 0.6458735 43.6976664
8        C2         C 0.1288499 0.7276161 75.9144194
9        C3         C 0.8883532 0.2469888 76.6454928
10       D1         D 0.1024372 0.5346595 56.5174319
11       D2         D 0.7642117 0.3491043 78.8318719
12       D3         D 0.3554768 0.3615868 28.9462846
13       E1         E 0.9018583 0.2057520  2.5940258
14       E2         E 0.3464000 0.7744766 97.0605503
15       E3         E 0.3116346 0.8119438 52.7975343
>
```


C. Manipulations de fichier

Fusion de tableaux => merge()

```
>
> Placettes_det = merge(Parcelles, Placettes, by.x = "Parcelles", by.y = "Parcelles")
> Placettes_det
```

| | Parcelles | Traitement | Age | volBois | Hauteur | Placettes | X | Y | Strate_arb |
|----|-----------|------------|-----|---------|---------|-----------|-----------|-----------|------------|
| 1 | A | E | 1 | 100 | 25 | A1 | 0.2659336 | 0.4024086 | 45.7170007 |
| 2 | A | E | 1 | 100 | 25 | A2 | 0.8881164 | 0.4900394 | 14.3344909 |
| 3 | A | E | 1 | 100 | 25 | A3 | 0.2613052 | 0.5059617 | 0.1679769 |
| 4 | B | E | 2 | 50 | 20 | B1 | 0.6502693 | 0.9684797 | 70.8333599 |
| 5 | B | E | 2 | 50 | 20 | B2 | 0.3836429 | 0.2010154 | 75.5601603 |
| 6 | B | E | 2 | 50 | 20 | B3 | 0.4746960 | 0.2677185 | 97.4669143 |
| 7 | C | E | 3 | 20 | 10 | C1 | 0.3351040 | 0.6458735 | 43.6976664 |
| 8 | C | E | 3 | 20 | 10 | C2 | 0.1288499 | 0.7276161 | 75.9144194 |
| 9 | C | E | 3 | 20 | 10 | C3 | 0.8883532 | 0.2469888 | 76.6454928 |
| 10 | D | E | 4 | 30 | 30 | D1 | 0.1024372 | 0.5346595 | 56.5174319 |
| 11 | D | E | 4 | 30 | 30 | D2 | 0.7642117 | 0.3491043 | 78.8318719 |
| 12 | D | E | 4 | 30 | 30 | D3 | 0.3554768 | 0.3615868 | 28.9462846 |
| 13 | E | J | 0 | 40 | 25 | E1 | 0.9018583 | 0.2057520 | 2.5940258 |
| 14 | E | J | 0 | 40 | 25 | E2 | 0.3464000 | 0.7744766 | 97.0605503 |
| 15 | E | J | 0 | 40 | 25 | E3 | 0.3116346 | 0.8119438 | 52.7975343 |

```
>
>
```

C. Manipulations de fichier

Tableau croisé => table()

```
Console D:/Dufrene/GXABT Enseignements/ULB BING-F4002 Acquisition et analyse de donnees/TP2 - Manip fichier
>
>
> # Lecture de la table décrivant les relevés phytosociologiques
> Releves = read.table("Releves.txt", header=TRUE, sep="\t")
> # Construction de la matrice « taxons x stations »
> tableau<-table(Releves$Taxons,Releves$Placettes)
> tableau
```

| | A1 | A2 | A3 | B1 | B2 | B3 | C1 | C2 | C3 |
|------------------------|----|----|----|----|----|----|----|----|----|
| Achillea millefolium | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| Achillea ptarmica | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Aethusa cynapium | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| Agropyron repens | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| Agrostemma githago | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Agrostis sp. | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Agrostis stolonifera | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| Agrostis vulgaris | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Alchemilla arvensis | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| Alopecurus myosuroides | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| Alopecurus pratensis | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Anagallis arvensis | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Anagallis phoenicea | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 1 |

Calcule la fréquence des combinaisons Taxons-Placettes

C. Manipulations de fichier

Tableau croisé => xtabs()

Calcule la somme des
abondances des combinaisons
Taxons-Placettes

Console D:/Dufrene/GXABT Enseignements/ULB BING-F4002 Acquisition et analyse de donnees/TP2 - Manip fichier xls ou openoffice/

```
>  
>  
> tableau<-xtabs(BVAL ~ Taxons + Placettes, releves)  
> tableau
```

| Taxons | Placettes | | | | | | | | |
|--|-----------|-----|------|------|------|-----|-----|------|------|
| | A1 | A2 | A3 | B1 | B2 | B3 | C1 | C2 | C3 |
| Achillea millefolium | 0.0 | 0.2 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 15.0 |
| Achillea ptarmica | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 | 0.0 |
| Aethusa cynapium | 0.0 | 0.2 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 |
| Agropyron repens | 0.0 | 2.5 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 |
| Agrostemma githago | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Agrostis sp. | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Agrostis stolonifera | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.2 | 0.0 |
| Agrostis vulgaris | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Alchemilla arvensis | 0.0 | 0.0 | 15.0 | 0.0 | 15.0 | 0.0 | 0.0 | 0.0 | 15.0 |
| Alopecurus myosuroides | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.2 | 0.0 |
| Alopecurus pratensis | 0.0 | 0.0 | 0.0 | 10.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Anagallis arvensis | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Anagallis phoenicea | 0.0 | 2.5 | 30.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 |
| Anthemis arvensis | 0.0 | 0.0 | 2.5 | 0.0 | 0.2 | 2.5 | 0.0 | 0.2 | 2.5 |
| Anthemis cotula | 0.0 | 0.0 | 15.0 | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 2.5 |
| Anthoxanthum odoratum | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Antirrhinum orontium | 0.0 | 0.0 | 2.7 | 0.0 | 0.2 | 0.2 | 0.0 | 0.0 | 0.2 |
| Apera spica-venti | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Arenaria serpyllifolia | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 |
| Arenaria serpyllifolia subsp.leptoclados | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Atriplex patula | 0.0 | 0.2 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 |
| Avena sativa | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Format liste à transformer en dataframe ...

C. Manipulations de fichier

Tableau croisé => xtabs()

Calcule la somme des
abondances des combinaisons
Taxons-Placettes

RStudio

```
tabrel=data.frame(tableau[1: nrow(tableau),1: ncol(tableau) ])
```

tabrel x

Filter

| | A1 | A2 | A3 | B1 | B2 | B3 | C1 | C2 | C3 |
|------------------------|-----|-----|------|------|------|-----|-----|------|------|
| Achillea millefolium | 0.0 | 0.2 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 2.5 | 15.0 |
| Achillea ptarmica | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 | 0.0 |
| Aethusa cynapium | 0.0 | 0.2 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 |
| Agropyron repens | 0.0 | 2.5 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 |
| Agrostemma githago | 0.0 | 0.0 | 2.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Agrostis sp. | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Agrostis stolonifera | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.2 | 0.0 |
| Agrostis vulgaris | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Alchemilla arvensis | 0.0 | 0.0 | 15.0 | 0.0 | 15.0 | 0.0 | 0.0 | 0.0 | 15.0 |
| Alopecurus myosuroides | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.2 | 0.0 |
| Alopecurus pratensis | 0.0 | 0.0 | 0.0 | 10.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Anagallis arvensis | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Anagallis phoenicea | 0.0 | 2.5 | 30.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 15.0 |

Showing 1 to 13 of 153 entries

Console D:/Dufrene/GXABT Enseignements/ULB BING-F4002 Acquisition et analyse de donnees/TP2 - Manip fichier xls ou openoffice/

```
>  
> tableau<-xtabs(BVAL ~ Taxons + Placettes, releves)  
> tabrel=data.frame(tableau[1: nrow(tableau),1: ncol(tableau) ])  
>
```

Tableau prêt pour tout usage ...

C. Manipulations de fichier

Exercices :

Feuille "Relevés_libellules" => BD_Exemple_CALC.ods ou BD_Exemple_XLS.xlsx

| | A | B | C | D | E |
|---|--------------------------|----------|------------------------------|--------------------|----|
| 1 | Station | Date | Taxon | Auteur | N |
| 2 | SGIB 1073 - Fagne de la | 20080730 | <u>Aeshna cyanea</u> | <u>Serruys M</u> | 3 |
| 3 | SGIB 1868 - La Flache | 20070607 | <u>Calopteryx virgo</u> | <u>Moës-Cimino</u> | 10 |
| 4 | SGIB 1060 - Beyoli - Mon | 20080815 | <u>Sympetrum danae</u> | <u>Smits Q</u> | 1 |
| 5 | SGIB 1500 - Basseilles a | 20080723 | <u>Calopteryx virgo</u> | <u>Fichet V</u> | 1 |
| 6 | SGIB 1500 - Basseilles a | 20080908 | <u>Sympetrum danae</u> | <u>Serruys M</u> | 1 |
| 7 | SGIB 0605a - Favi de Lur | 20000813 | <u>Enallagma cyathigerum</u> | <u>DE KNIJF G</u> | 25 |

- quelle est la somme des abondances pour chaque espèce ?
- quelle est la somme des abondances pour chaque observateur et chaque date ?
- quelle est la matrice de la somme des abondances par station et taxon ?