

Le grand ménage de données : Excel

Carolyn Pecoskie, Bibliothécaire, métadonnées et ressources électroniques, l'Université McGill
carolyn.pecoskie@mcgill.ca

Cet atelier est offert le 1 août 2023 de 15 h 15 à 16 h 15 dans le cadre du deuxième [Programme Jumpstart BiblioTECH](#).

Table des matières

- [Attributions](#)
- [Objectifs éducatifs](#)
- [Installation et introduction à Excel](#)
- [Trucs et astuces](#)
 - [Formats de fichier](#)
 - [Types de données](#)
 - [Remplissage automatique \(ou « Autofill »\)](#)
 - [Figer des volets ou « Freeze Panes »](#)
 - [Insérer ou supprimer des lignes et des colonnes](#)
 - [Dépannage des erreurs](#)
- [Trier ou « Sorting » et Filtrer ou « Filtering »](#)
 - [Trier ou « Sorting »](#)
 - [Avertissement de tri](#)
 - [Filtrer ou « Filtering »](#)
- [Rechercher et Remplacer \(ou « Find » et « Replace »\)](#)
- [La mise en forme conditionnelle ou « Conditional Formatting »](#)
- [Formules et fonctions](#)
 - [Calculs mathématiques](#)
 - [Fonctions intégrées](#)
 - [Fonction SI.CONDITIONS \(ou « IFS Function »\)](#)
 - [Fonctions de fractionnement et de substitution](#)
 - [CONCATENER \(ou « CONCATENATE »\)](#)
 - [RECHERCHEV \(ou « VLOOKUP »\)](#)
- [Graphiques \(ou « Charts »\)](#)
- [Tableaux croisés dynamiques \(ou « PivotTables »\)](#)

Attributions

Certaines parties de ce contenu sont tirées de ces ateliers « Digital Scholarship » par Alisa Rod et Clara Turp: <https://github.com/Digital-Scholarship-Hub/WorkingWithDataInExcel> et <https://github.com/Digital-Scholarship-Hub/Excel-Intermediate>

Les thèmes de cet atelier sont mappés aux composants de cet atelier « Library Carpentry Open Refine » : <https://librarycarpentry.org/lc-open-refine/> (articles 1, 2, 4, 5, 7, 8)

Merci à la première cohorte du Programme Jumpstart BiblioTECH pour leurs contributions et suggestions.

Lectures complémentaires :

Pour plus de conseils sur l'utilisation de feuilles de calcul en tant que professionnel de l'information, veuillez consulter « Tidy Data for Librarians » par « The Carpentries » : <https://librarycarpentry.org/lc-spreadsheets/>

Objectifs éducatifs

L'objectif de l'atelier d'aujourd'hui est de vous présenter des trucs, astuces et outils utiles dans Excel, pour commencer à guider votre travail avec Excel en tant que professionnel de l'information. Cette session ne sera en aucun cas exhaustive, en termes de couverture de tout ce dont vous pourriez avoir besoin dans Excel, mais nous espérons qu'à la fin de la session, vous pourrez :

- Être conscient d'une gamme d'outils et de fonctions qui (d'après l'expérience de la présentatrice) sont très utiles pour le travail de la bibliothèque
- Soyez au courant de quelques trucs et astuces utiles pour gagner du temps et tirer le meilleur parti de ce qu'Excel peut faire pour vous
- Soyez plus confiant dans votre capacité à naviguer dans Excel et à consulter Google, le site d'aide de Microsoft Excel et d'autres sources chaque fois que vous avez besoin de trouver un nouvel outil ou une nouvelle fonction ou de résoudre une erreur

Cet atelier est proposé juste avant l'atelier Le grand ménage de données : OpenRefine. Nous proposons ces ateliers conjointement car Excel et OpenRefine peuvent tous deux être très utiles pour nettoyer et travailler avec des jeux de données, mais ils offrent des outils et des fonctions assez différents. Nous espérons que d'ici la fin de la journée, vous saurez quand utiliser l'un ou l'autre outil, en fonction de ce que vous aimeriez faire de vos données.

Installation et introduction à Excel

[Excel](#) est l'un des tableurs propriétaires les plus couramment utilisés et fait partie de la suite logicielle Microsoft Office. LibreOffice, OpenOffice.org, et Gnumeric sont d'autres exemples de tableurs « open source ». Ils ont des fonctionnalités similaires, bien que certains puissent être représentés légèrement différemment.

Vous devriez avoir accès à Excel à travers votre université:

- [McGill](#)
- [UdeM](#)

Si vous n'y avez pas accès, veuillez contacter un membre du comité organisateur de BiblioTECH.

Pour l'atelier, vous devez télécharger un document de données. [Cliquez ici pour les télécharger.](#)

Trucs et astuces

Formats de fichier

Page d'aide de Microsoft : [Formats de fichier pris en charge dans Excel](#)

Vous pouvez travailler avec une variété de types de fichiers dans Excel, notamment :

- Formats de fichier Excel (.xlsx ou l'ancien .xls sont les plus populaires)
- Formats de fichier texte (tels que .txt ou .csv)

Vous pouvez enregistrer un fichier Excel dans un autre format de fichier en cliquant sur **Fichier > Enregistrer sous** [« **File > Save As** »] et en choisissant l'un des formats de fichier disponibles dans le menu déroulant.

Excel vous encouragera à enregistrer les fichiers dans l'un de leurs propres formats, mais il est parfois préférable de le conserver dans le format d'origine (tel que .txt, .csv, etc.), surtout si vous souhaitez que certaines données restent formatées. d'une manière très particulière (en particulier les dates, car Excel peut avoir tendance à être un peu trop zélé en essayant de reformater vos dates pour vous).

Selon Microsoft: « Pour ouvrir un fichier créé dans un autre format de fichier, que ce soit dans une version antérieure d'Excel ou dans un autre programme, cliquez sur **Fichier > Ouvrir** [« **File > Open** »]. Si vous ouvrez un feuille de calcul Excel 97-2003, il s'ouvre automatiquement en **mode de compatibilité**. [...] Toutefois, vous avez également la possibilité de continuer à travailler en **mode de compatibilité**, qui conserve le format de fichier d'origine pour la compatibilité vers l'arrière. »

Types de données

Page d'aide de Microsoft : [Formats de nombres disponibles dans Excel](#)

Dans Excel, vous pouvez modifier le « format » d'une cellule, afin qu'elle reconnaisse son contenu comme des éléments tels que la devise, les pourcentages, les décimales, les dates, le texte, etc.

Pour ce faire, sélectionnez une cellule ou une plage de cellules que vous souhaitez modifier. Faites un clic droit sur la sélection et choisissez **Formater les cellules...** (« **Format Cells...** »). Dans le menu **Formater les cellules...** (« **Format Cells...** »), choisissez l'onglet **Nombre** (« **Number** »), qui est le premier onglet susceptible d'apparaître par défaut. Vous pouvez choisir le format de nombre que vous souhaitez utiliser et, pour certains formats de nombres, vous

pouvez encore plus personnaliser le format. À partir de la page d'aide Microsoft liée ci-dessus, voici les formats de nombres disponibles dans Excel :



Format	Description
Standard ou « General »	Format de nombre qu'Excel utilise par défaut lorsque vous tapez un nombre. Pour l'essentiel, les nombres mis en forme au format Standard s'affichent exactement tels que vous les avez tapés. Toutefois, si la cellule n'est pas assez large pour contenir l'ensemble du nombre, le format Standard arrondit les nombres à décimales. Le format de nombre Standard utilise aussi les notations scientifiques (exponentielles) pour les nombres importants (12 caractères ou plus).
Nombre ou « Number »	Format utilisé pour l'affichage général des nombres. Vous pouvez spécifier le nombre de décimales à utiliser, indiquer si vous voulez utiliser un séparateur de milliers et la façon dont les nombres négatifs doivent s'afficher.
Devise ou « Currency »	Format utilisé pour les valeurs monétaires générales et affiche le symbole de la devise par défaut en regard des nombres. Vous pouvez spécifier le nombre de décimales à utiliser, indiquer si vous voulez utiliser un séparateur de milliers et la façon dont les nombres négatifs doivent s'afficher.
Comptabilité ou « Accounting »	Format également utilisé pour les valeurs monétaires, mais il aligne les symboles des devises et des décimales pour les nombres d'une même colonne.
Date ou « Date »	Affiche les numéros de série date et heure comme valeurs date, selon le type et les paramètres régionaux (emplacement) que vous spécifiez. Les formats Date commençant par un astérisque (*) s'adaptent aux modifications des paramètres de date et d'heure spécifiés dans le Panneau de configuration. Les formats sans astérisque ne sont pas affectés par les paramètres du Panneau de configuration.

Format	Description
Heure ou « Time »	Affiche les numéros de série date et heure comme valeurs heure, selon le type et les paramètres régionaux (emplacement) que vous spécifiez. Les formats Heure commençant par un astérisque (*) s'adaptent aux modifications des paramètres de date et d'heure spécifiés dans le Panneau de configuration. Les formats sans astérisque ne sont pas affectés par les paramètres du Panneau de configuration.
Pourcentage ou « Percentage »	Multiplie la valeur de la cellule par 100 et affiche le résultat avec un symbole de pourcentage (%). Vous pouvez spécifier le nombre de décimales que vous souhaitez utiliser.
Fraction ou « Fraction »	Affiche un nombre sous forme de fraction, selon le type de fraction que vous spécifiez.
Scientifique ou « Scientific »	Affiche un nombre de façon exponentielle, en remplaçant une partie du nombre par E+n, où E (pour Exposant) multiplie le chiffre précédent par 10 à la puissance n. Par exemple, un format Scientifique à deux décimales affiche 12345678901 sous la forme 1,23E+10, soit 1,23 fois 10 à la puissance 10. Vous pouvez spécifier le nombre de décimales que vous souhaitez utiliser.
Texte ou « Text »	Traite le contenu d'une cellule comme du texte et l'affiche exactement comme vous le tapez, même si vous tapez des chiffres.
Spécial ou « Special »	Permet d'afficher des nombres tels que les codes postaux, des numéros de téléphone ou de sécurité sociale.
Personnalisé ou « Custom »	Permet de modifier la copie d'un code de format de nombre. Utilisez ce format pour créer un format de nombre personnalisé qui est ajouté à la liste des codes de formats de nombre. Vous pouvez ajouter entre 200 et 250 formats de nombre personnalisés, selon la langue de la version d'Excel installée sur l'ordinateur. Pour plus d'informations sur les formats personnalisés, voir Créer ou supprimer un format de nombre personnalisé .

Remplissage automatique (ou « Autofill »)

Page d'aide de Microsoft : [Recopier automatiquement des données dans les cellules d'une feuille de calcul](#)

La fonctionnalité de **Remplissage automatique** (ou « **Autofill** ») est très utile dans Excel, pour remplir les cellules avec des données qui suivent un modèle.

1. Tout d'abord, sélectionnez environ trois cellules côte à côte à utiliser comme base pour remplir les cellules supplémentaires qui seront à côté de ces trois cellules originales dans la feuille de calcul. Les modèles peuvent inclure des éléments tels que « 1, 2, 3 », « 15, 30, 45 », ou une série de formules.
2. Faire glisser la poignée de recopie : 
3. Le cas échéant, cliquez sur **Options de recopie incrémentée** (ou « **Auto Fill Options** ») , puis sélectionnez l'option de votre choix.

Notez que lorsque vous allez remplir automatiquement une formule, vous devrez ajouter la notation appropriée (« \$ ») dans le cas où vous souhaitez avoir une référence absolue, en termes de colonne ou de ligne ou les deux. Pour plus d'informations, veuillez consulter la page d'aide de Microsoft : [Basculement entre références relatives et absolues](#)

Figurer des volets ou « Freeze Panes »

Page d'aide de Microsoft : [Figurer des volets pour verrouiller des lignes et des colonnes](#)

Parfois, lorsque vous avez des données complexes, il est vraiment utile que les en-têtes restent en haut de la page, même si vous faites défiler les lignes de données. Excel a une fonctionnalité qui vous permet de le faire très rapidement. Cette fonctionnalité s'appelle **Figurer les volets** ou « **Freeze Panes** ».

Cliquez sur l'onglet **Affichage** (ou « **View** ») et accédez à la sous-section **Figurer les volets** (ou « **Freeze Panes** »). Vous pouvez ensuite choisir si vous souhaitez figurer la ligne de haut, la première colonne ou les deux volets de votre choix.

Insérer ou supprimer des lignes et des colonnes

Page d'aide de Microsoft : [Insérer ou supprimer des lignes et des colonnes](#) et [Insérer une ou plusieurs lignes, colonnes ou cellules dans Excel pour Mac](#)

Le moyen le plus simple d'insérer ou de supprimer une ligne ou une colonne entière est de cliquer avec le bouton droit sur le haut de la ligne ou de la colonne (c'est-à-dire de cliquer sur la lettre ou le numéro qui désigne cette ligne ou cette colonne) et de choisir **Insérer** (ou « **Insert** ») ou **Supprimer** (ou « **Delete** »). De nouvelles lignes seront insérées *au-dessus* de la ligne sélectionnée et de nouvelles colonnes seront insérées à *gauche* de la colonne sélectionnée.

Les références de cellule doivent s'ajuster automatiquement en fonction de l'endroit où les cellules sont décalées en raison d'une insertion ou d'une suppression.

Si vous souhaitez insérer ou supprimer plusieurs lignes ou colonnes à la fois, cela est également possible. Lors de l'insertion, il vous suffit de sélectionner le même nombre de lignes ou de colonnes que vous souhaitez insérer. Par exemple, pour insérer cinq colonnes (ou lignes) vides, vous devez sélectionner cinq colonnes (ou lignes). Ce n'est pas grave si les sélections contiennent des données, car les blancs seront insérés à gauche (pour les colonnes) ou au-dessus (pour les lignes) par rapport à la sélection.

Dépannage des erreurs

Excel vous donne toujours des informations quand vous faites une erreur. La première chose à remarquer est que les messages d'erreur commencent toujours par un hashtag (#) et se terminent par un signe de ponctuation (souvent ! ou ?). « **#DIV/0!** » est le message qui s'affiche lorsque vous essayez de diviser un nombre par zéro. En effet, les cellules vides sont traitées comme des zéros aux fins des calculs mathématiques. Le message d'erreur est court mais vous donne quelques informations. Les autres messages d'erreur courants sont : **#NAME?**, **#NULL!**, **#REF!**, **#VALUE!**. Si vous les voyez, ne paniquez pas, et essayez de copier l'erreur dans Google. Vous pouvez également passer votre curseur sur le petit triangle vert qui apparaît dans le coin supérieur gauche de la cellule avec un message d'erreur. Cela révélera un signe d'avertissement jaune avec un point d'explication, qui est un menu déroulant sur lequel vous pouvez cliquer pour obtenir plus d'informations sur l'erreur ou l'aide.

Trier ou « Sorting » et Filtrer ou « Filtering »

Trier (ou « **Sorting** ») et **Filtrer** (ou « **Filtering** ») sont des fonctions liées, mais ils ont des utilisations différentes. La principale différence entre les deux est que la fonction **Trier** (ou « **Sorting** ») réorganise vos données, modifiant finalement l'ordre dans lequel les données apparaissent dans votre feuille de calcul. La fonction **Filtrer** (ou « **Filtering** »), en revanche, modifie temporairement l'affichage des données sur la feuille, mais peut facilement être activé ou désactivé, ramenant ainsi très facilement la feuille de calcul à son état et à son ordre d'origine. Nous verrons plus en détail comment utiliser chacune de ces fonctions.

Trier ou « Sorting »

Page d'aide de Microsoft : [Trier les données d'une plage ou d'un tableau](#)

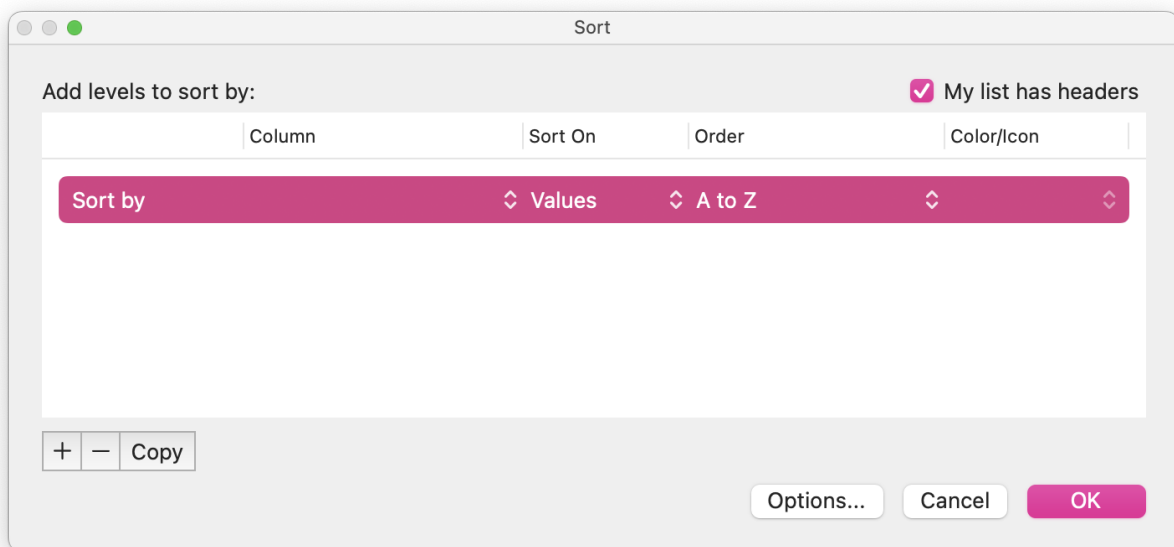
Il existe de nombreuses raisons pour lesquelles vous souhaitez peut-être trier vos données. Vous voudrez peut-être une liste de noms classés par ordre alphabétique, ou vous voudrez peut-être que les observations soient organisées par date afin de pouvoir entrer plus de variables.

Actuellement, le jeu de données est trié par « interview_date ». Nous allons plutôt trier ce jeu de données par « village ». Pour ce faire, nous devons sélectionner toutes les données, y compris les noms de variables (dans la ligne du haut). Toutes les cellules qui ne sont pas sélectionnées resteront dans l'ordre exact où elles se trouvent actuellement. Il existe de nombreuses façons de sélectionner les données. Nous allons cliquer sur la cellule A1. Ensuite, nous appuierons sur **CTRL (ou Command sur Mac) + Shift + (la touche Droite)**. Cela sélectionnera les cellules de A1 à BJ1.

Maintenant, nous voulons sélectionner toutes les lignes sous ces cellules. Appuyez maintenant sur **CTRL + Shift + (la touche Bas)**. Maintenant, toutes les données doivent être sélectionnées.

Maintenant, nous voulons trier les données. Dans l'onglet **Accueil** (ou « **Home** »), cliquez sur la flèche à côté du bouton **Trier et filtrer** (ou « **Sort & Filter** »). Choisissez ensuite l'option **Tri personnalisé** (ou « **Custom Sort** »).

Vous devriez voir la boîte de dialogue suivante :



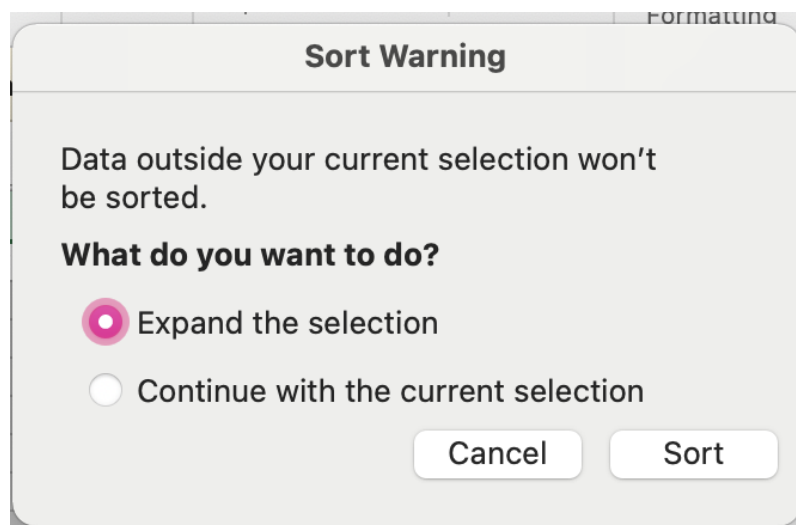
La première chose que vous voulez faire est de vérifier **Ma liste a des en-têtes** (ou « **My list has headers** ») afin que vos noms de variables ne soient pas triés comme une observation. Lorsque vous classez les noms de villages par ordre alphabétique, vous ne souhaitez pas que le nom de la variable « village » apparaisse dans la liste des noms.

Ensuite, à côté de **Trier par** (ou « **Sort by** »), choisissez « village » dans le menu déroulant **Colonne** (ou « **Column** »). Si vous souhaitez trier par plusieurs variables, vous pouvez cliquer sur le signe plus pour ajouter une deuxième variable à trier. Dans notre cas, nous nous en tiendrons à un, puis appuyez sur **OK**.

Vous devriez maintenant remarquer que le jeu de données a été trié.

Avertissement de tri

Une autre façon de trier par une seule variable consiste à sélectionner simplement la colonne que vous souhaitez trier, puis, dans l'onglet **Accueil** (ou « **Home** »), cliquez sur la flèche à côté de **Trier et filtrer** (ou « **Sort & Filter** »). Dans ce cas, un avertissement vous sera présenté :



Développer la sélection ou « Expand the Selection » : 9 fois sur 10 c'est ce que vous voulez faire, sinon la colonne que vous avez sélectionnée sera la seule à être réorganisée et les relations (potentiellement critiques) entre les colonnes seront perdues.

Filtrer ou « Filtering »

Pages d'aide de Microsoft : [Filtrer les données d'une plage ou d'un tableau](#) et [Filtrer des valeurs uniques ou supprimer des doublons](#)

La fonction **Filtrer** (ou « **Filtering** ») consiste à modifier temporairement votre vue de la feuille de calcul, mais ne modifie pas de manière permanente l'ordre d'un contenu. Vous pouvez activer ou désactiver la fonction **Filter** (ou « **Filtering** ») pour une ou plusieurs colonnes à la fois. Pour ce faire, sélectionnez la ou les colonnes auxquelles vous souhaitez appliquer le filtrage. Vous pouvez maintenir la touche **CTRL** (ou la touche **Command** sur Mac) enfoncée pour sélectionner plusieurs colonnes qui ne sont pas nécessairement côte à côte. Ensuite, dans l'onglet **Accueil** (ou « **Home** »), accédez à la flèche à côté du bouton **Trier et filtrer** (ou « **Sort & Filter** ») et sélectionnez **Filtrer** (ou « **Filter** »). Vous verrez de petites flèches apparaître à côté du ou des en-têtes de la ou des colonnes que vous avez sélectionnées. Ces flèches représentent votre capacité à filtrer le contenu dans la ou des colonnes.

Notez que si vous souhaitez modifier les colonnes qui ont l'option de filtrage, il est nécessaire de revenir à l'option **Trier et filtrer > Filtrer** (ou « **Sort & Filter > Filter** ») pour désactiver le filtrage pour la sélection actuelle, avant de pouvoir sélectionner de nouvelles colonnes à filtrer.

Pendant que vous regardez la vue filtrée de vos données, il est possible d'effectuer des opérations en masse (y compris [le remplissage automatique ou « autofill »](#)) qui ne devraient pas affecter les données qui sont à ce moment masquées car elles ont été filtrées.

Rechercher et Remplacer (ou « Find » et « Replace »)

Page d'aide de Microsoft : [Rechercher ou remplacer du texte et des nombres dans une feuille de calcul](#)

Rechercher (ou « **Find** ») et **Remplacer** (ou « **Replace** ») peuvent être utilisés pour rechercher quelque chose dans votre classeur, comme un nombre particulier ou une chaîne de texte. Ils peuvent être incroyablement utiles pour découvrir ce qui est contenu dans votre jeu de données (si vous voulez vérifier combien d'instances d'une valeur particulière apparaissent dans la feuille de calcul par exemple). Ils constituent également un moyen simple d'apporter des modifications en masse à une valeur particulière, par exemple pour corriger une série de coquilles ou de variations orthographiques.

CTRL+F (ou **Command+F** sur un Mac) vous permettra d'exécuter rapidement la fonction **Rechercher** (ou « **Find** »).

Les options complètes des fonctions **Rechercher** (ou « **Find** ») et **Remplacer** (ou « **Replace** ») sont accessibles à partir de l'onglet **Accueil** (ou « **Home** ») et en cliquant sur la flèche à côté du bouton **Rechercher et sélectionner** (ou « **Find & Select** »). Choisir **Remplacer...** (ou « **Replace...** ») vous amènera à la boîte de dialogue **Rechercher et remplacer** (ou « **Find & Replace** »).

Vous pouvez inclure des caractères de remplacement tels que des points d'interrogation (?), des tildes (~) et des astérisques (*), ou des chiffres dans vos termes de recherche. Plus d'informations peuvent être trouvées dans la page d'aide de Microsoft liée ci-dessus.

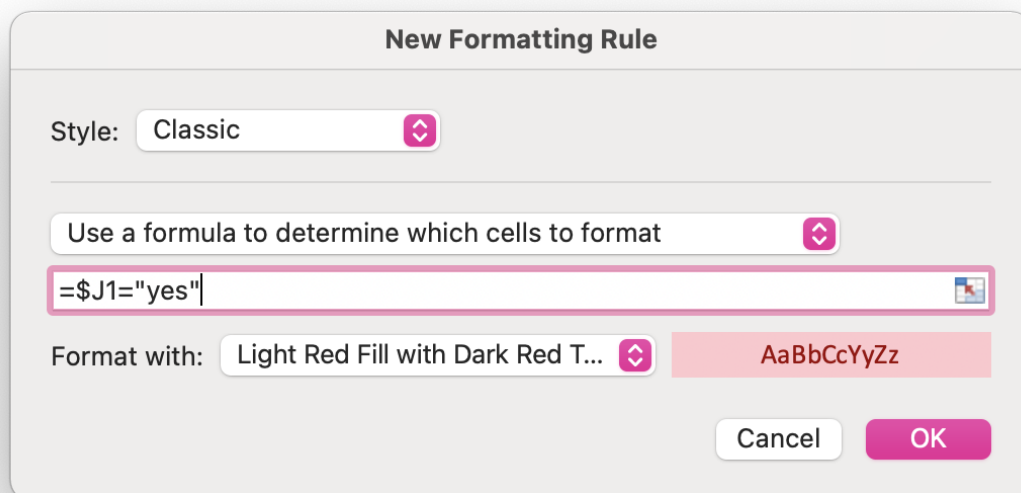
Vous pouvez effectuer une recherche par lignes et colonnes, rechercher dans les commentaires ou les valeurs et rechercher la totalité de la feuille de calcul ou du classeur (voir le bouton **Options** (même qu'en anglais) dans la boîte de dialogue **Rechercher et remplacer** (ou « **Find & Replace** »)).

La mise en forme conditionnelle ou « Conditional Formatting »

Page d'aide de Microsoft : [Appliquer la mise en forme conditionnelle pour faire ressortir des informations](#)

Selon Microsoft: « La mise en forme conditionnelle peut mettre en évidence les modèles et tendances dans vos données. Pour l'utiliser, vous devez créer des règles qui déterminent le format des cellules en fonction de leurs valeurs [...] Vous pouvez appliquer une mise en forme conditionnelle à une plage de cellules (une sélection ou une plage nommée), un tableau Excel, et dans Excel pour Windows à un rapport de tableau croisé dynamique. »

Dans notre exemple, disons que nous voulons sélectionner toutes les lignes où la valeur de « agr_assoc » = « yes ». Pour commencer, sélectionnez toutes les cellules de cette feuille Excel (vous pouvez utiliser **CTRL+A** ou **Command+A** on a Mac. Dans l'onglet **Accueil** (ou « **Home** »), accédez à la flèche à côté du bouton **Mise en forme conditionnelle** (ou « **Conditional Formatting** »). Choisissez l'option **Gérer les règles...** (ou « **Manage Rules...** »). Cliquez sur le signe **+** pour appliquer une nouvelle règle. Une boîte de dialogue **Nouvelle règle de mise en forme** (ou « **New Formatting Rule** ») s'ouvrira.



Sélectionnez **Classique** (ou « **Classic** ») dans le menu déroulant pour **Style**. Dans le menu déroulant suivant, choisissez l'option **Utiliser une formule pour déterminer pour quelles cellules le format sera appliqué** (ou « **Use a formula to determine which cells to format** »). Dans la fenêtre de texte, saisissez ce qui suit :

= \$J1="yes"

(Parce que « agr_assoc » apparaît dans la colonne J.)

Enfin, vous pouvez modifier le **Format** avec des options (pour un texte/couleur de remplissage différent) ou continuer avec les options par défaut. Cliquez sur **OK** pour appliquer ces modifications.

La mise en forme conditionnelle peut être supprimée en revenant à ce même menu **Gérer les règles...** (ou « **Manage Rules...** »).

Formules et fonctions

Calculs mathématiques

Excel est extrêmement puissant pour exécuter des calculs mathématiques. Les calculs sont effectués à l'aide de formules. Pour le moment, nous nous concentrerons sur des formules mathématiques simples. Les formules peuvent effectuer des tâches plus simples, telles que des additions ou des multiplications, ou des calculs plus complexes, tels que renvoyer le cosinus d'un angle.

1. Les formules commencent toujours par « = ». Cela indique à Excel que vous n'entrez pas seulement des nombres.
2. Vous pouvez écrire la fonction vous-même ou vous référer à une formule prédéfinie.
3. Opérateurs mathématiques dans Excel :
 - Pour ajouter, utilisez un signe plus : +
 - Pour soustraire, utilisez un signe moins (trait d'union) : -
 - Pour multiplier, utilisez l'astérisque : *
 - Pour diviser, utilisez la barre oblique inversée : /
 - Supérieur à est représenté par le signe suivant : >
 - Inférieur à est représenté par le signe suivant : <
 - L'opérateur mathématique pour « pas égal » (c'est-à-dire « n'est pas égal ») est représenté par les signes inférieur à et supérieur à ensemble, comme un diamant : <>

Fonctions intégrées

Page d'aide de Microsoft : [Formules et fonctions](#)

Plutôt que de taper chaque calcul à la main, nous pouvons utiliser les fonctions intégrées d'Excel. Les calculs courants tels que les moyennes, les médianes, les sommes et les maximums ont leurs propres fonctions Excel.

Si vous pensez qu'Excel peut avoir la fonction que vous souhaitez utiliser, vous pouvez accéder à l'onglet **Formules** (ou « **Formulas** ») et sélectionner **Insérer une fonction** (ou « **Insert**

Function »). Les fonctions sont organisées par catégorie dans la bibliothèque de fonctions (à droite). Essayez d'utiliser le bouton **Insérer une fonction** (ou « **Insert Function** ») pour rechercher comment calculer la moyenne. Cliquez sur une cellule vide, puis cliquez sur **Insérer une fonction** (ou « **Insert Function** »).

C'est la première des deux façons d'accéder aux formules intégrées d'Excel. **Moyenne** (ou « **Average** ») se trouve sous **Plus de fonctions > Statistiques** (ou « **More Functions > Statistical** »). Lorsque vous procédez ainsi, une boîte de dialogue apparaît et vous donne des informations sur l'utilisation de la formule et les données acceptées par la formule. Lorsque la formule demande un nombre, vous pouvez entrer un nombre, un identifiant de cellule, une plage, etc.

L'autre méthode consiste à taper le nom de la formule, dans ce cas **Moyenne** (ou « **Average** »). Une fois que vous commencez à taper **=M** (ou « **=A** »), une liste déroulante apparaîtra. Vous pouvez double-cliquer sur la formule souhaitée. Une explication apparaîtra, vous aidant à comprendre la formule.

Notez que « l'utilité » des fonctions variera considérablement en fonction de votre rôle et des types de tâches que vous essayez d'accomplir. Ceci étant dit, voici quelques fonctions utiles à connaître :

- Fonction **SI.CONDITIONS** (ou « **IFS Function** »)
- Splitting Functions
- **CONCATENER** (ou « **CONCATENATE** »)
- **RECHERCHEV** (ou « **VLOOKUP** »)

Fonction SI.CONDITIONS (ou « IFS Function »)

Pages d'aide de Microsoft : [Fonction SI](#) et [Fonction SI.CONDITIONS](#)

Selon Microsoft : « La fonction SI.CONDITIONS vérifie si une ou plusieurs conditions sont remplies et renvoie une valeur correspondant à la première condition vraie. L'utilisation de cette fonction revient à utiliser plusieurs instructions SI imbriquées, mais elle reste bien plus facile à lire quand plusieurs conditions se suivent. »

Syntaxe (simple)
=SI.CONDITIONS([test_logique1; valeur_si_vrai1; test_logique2; valeur_si_vrai2; test_logique3; valeur_si_vrai3])

Ou en anglais :

(Simple) Syntax
=IFS(Something is True1, Value if True1, Something is True2, Value if True2, Something is True3, Value if True3)

Notez que si aucune condition VRAI n'est trouvée, cette fonction renvoie l'erreur **#N/A**.

Nous allons tester l'utilisation d'une fonction **SI.CONDITIONS** (ou « **IFS** ») pour vérifier si les conditions suivantes sont remplies :

- parents_liv = yes
- sp_parents_live = yes
- grand_liv = yes
- sp_grand_liv = yes

Supposons que nous voulions voir si l'une de ces conditions est vraie. Insérons une nouvelle colonne vide dans la colonne S pour contenir les résultats de cette requête **SI.CONDITIONS** (ou « **IFS** »). Cliquez avec le bouton droit sur l'en-tête S et accédez à **Insérer** (ou « **Insert** »). Appelons cette colonne « SI.CONDITIONS » (ou « IFS »). À partir de S2, on tapera :

=SI.CONDITIONS(O2="yes", "yes", P2="yes", "yes", Q2="yes", "yes", R2="yes", "yes")

Ou en anglais :

=IFS(O2="yes", "yes", P2="yes", "yes", Q2="yes", "yes", R2="yes", "yes")

Et puis appuyez sur **Entrée** (ou « **Enter** »). Vous devriez voir le mot « yes » (sans les guillemets) en S2.

(Notez que lorsque nous saisissons du texte dans une formule, il est nécessaire de le mettre entre guillemets comme nous l'avons fait ci-dessus, afin qu'Excel le reconnaisse comme une chaîne de texte.)

Nous pouvons utiliser [le remplissage automatique \(ou « autofill »\)](#) pour appliquer cette formule dans le reste de la colonne.

Fonctions de fractionnement et de substitution

Pages d'aide de Microsoft : [Fractionner du texte en plusieurs colonnes en utilisant des fonctions](#) et [Fractionner du texte en plusieurs colonnes à l'aide de l'Assistant Conversion](#) et [Fonction SUBSTITUE](#)

Nous allons utiliser l'**Assistant Conversion** (l'assistant de conversion texte-colonne ou « **Convert Text to Columns Wizard** ») pour diviser le texte qui apparaît dans la colonne "liv_owned" (colonne AW). Pour commencer, nous aurons besoin de quelques colonnes vides supplémentaires pour contenir les données après la scission. Sélectionnez 10 colonnes pour être sûr, en commençant par la colonne à droite de "liv_owned" (colonne AX) et en allant vers la droite ("liv_owned_other") et en allant vers la droite jusqu'à "gps_Latitude" (colonne BG) inclusivement, puis cliquez avec le bouton droit et choisissez **Insérer** (ou « **Insert** »). Ensuite, cliquez sur la lettre au-dessus de l'en-tête "liv_owned" pour sélectionner cette colonne. Allez ensuite dans **Données > Convertir** (ou « **Data > Text to Columns...** »). Nous choisirons l'option **Délimité** (ou « **Delimited** »), puis cliquez sur **Suivant** (ou « **Next** »). Nous allons sélectionner notre délimiteur pour être **le point-virgule** dans ce cas, puis cliquez sur **Suivant**

(ou « **Next** »). Enfin, nous conserverons les données au format **Général** (ou « **General** »). Enfin, vous pouvez supprimer toutes les colonnes vides supplémentaires restantes.

Voyez-vous les limites de l'utilisation de cette fonction pour diviser les données en plusieurs colonnes ? Nous pourrions utiliser la fonction **SUBSTITUE** (ou « **SUBSTITUTE** ») pour supprimer les caractères indésirables qui restent. Par exemple, insérez une colonne à gauche de la colonne "liv_owned" d'origine et commencez avec cette formule dans la première cellule :

=SUBSTITUE(AW2,"[",",")

Ou en anglais :

=SUBSTITUTE(AW2,"[",",")

Alternativement, une combinaison de la fonction [Filtrer \(ou « Filtering »\)](#) et les fonctions [Rechercher et Remplacer \(ou « Find and Replace »\)](#) peut être utilisée pour nettoyer les données à ce stade. La fonction [Filtrer \(ou « Filtering »\)](#) peut également être utilisée pour identifier et supprimer les colonnes vides restantes du début de cette section.

CONCATENER (ou « CONCATENATE »)

Page d'aide de Microsoft : [Fonction CONCATENER](#)

Selon Microsoft : « La fonction CONCATENER, qui fait partie des [fonctions de texte](#) permet de joindre plusieurs chaînes au sein d'une seule chaîne. »

Syntaxe
=CONCATENER(texte1; [texte2]; ...)

Ou en anglais :

Syntax
=CONCATENATE(text1,[text2],...)

Nous utiliserons la fonction **CONCATENER** (ou « **CONCATENATE** ») pour créer une version complète de chaque adresse (y compris « village », « ward », « district », « province »). Commencez par la colonne « village ». Nous allons insérer une colonne à droite de « village ». Pour ce faire, sélectionnez la colonne « years_farm » (colonne I) et faites un clic droit et choisissez **Insérer** (ou « **Insert** »). Nous appellerons cette nouvelle colonne « adresse_complète » (vous pouvez la saisir dans la cellule I1). Ensuite dans la cellule I2 nous taperons la formule suivante :

=CONCATENER(H2," ",G2," ",F2," ",E2)

Ou en anglais :

=CONCATENATE(H2," ",G2," ",F2," ",E2)

Et puis appuyez sur **Entrée** (ou « **Enter** »). Vous devriez voir l'adresse donnée comme suit : **49, Bandula, Manica, Manica.**

Notez que « „ „ » (entre guillemets de style anglais, une virgule suivie d'un espace) est nécessaire pour avoir un séparateur entre chaque chaîne. Sinon, les mots seraient présentés les uns après les autres, sans espace entre eux.

Nous pouvons utiliser [le remplissage automatique \(ou « autofill »\)](#) pour appliquer cette formule dans le reste de la colonne.

RECHERCHEV (ou « VLOOKUP »)

Page d'aide de Microsoft : [Fonction RECHERCHEV](#)

Le tri et le filtrage de vos données peuvent être utiles pour l'organisation. Cependant, il vous suffit parfois de rechercher une valeur particulière, et si vous avez beaucoup de données, le tri et le filtrage prendront encore du temps. La fonction **RECHERCHEV** (ou « **VLOOKUP** ») vous permet de rechercher vos données par une variable et de renvoyer la valeur d'une autre variable. Une autre utilisation de **RECHERCHEV** (ou « **VLOOKUP** ») consiste à fusionner deux ensembles de données dont les informations se chevauchent.

Syntaxe
=RECHERCHEV(lookup_value, table_array, col_index_num, [range_lookup])

Ou en anglais :

Syntax
=VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

Notez que [l'Aide-mémoire : actualisateur RECHERCHEV](#) (ou [Quick Reference Card: VLOOKUP refresh](#)) est très pratique lorsque vous avez besoin d'un rappel de la signification de chaque élément de la syntaxe.

Notez également : **RECHERCHEX** (ou « **XLOOKUP** ») est une nouvelle version de **RECHERCHEV** (ou « **VLOOKUP** »), « qui fonctionne dans n'importe quelle direction et renvoie des correspondances exactes par défaut, ce qui rend plus pratique et facilite son utilisation par rapport à son prédécesseur ». Pour plus d'informations, veuillez consulter la page d'aide de Microsoft de cette fonction : [Fonction XLOOKUP](#)

Graphiques (ou « Charts »)

Excel offre la possibilité de faire des visualisations de données via la fonctionnalité **Graphiques** (ou « **Charts** »). Nous n'entrerons pas dans le détail des **Graphiques** (ou « **Charts** ») lors de

l'atelier d'aujourd'hui, mais si jamais vous avez besoin de créer un graphique, sachez que cette fonctionnalité existe. Microsoft propose une documentation utile qui peut vous guider dans la création d'un graphique : [Créer un graphique de bout en bout](#)

Tableaux croisés dynamiques (ou « PivotTables »)

Page d'aide de Microsoft : [Vue d'ensemble des tableaux croisés dynamiques et des graphiques croisés dynamiques](#)

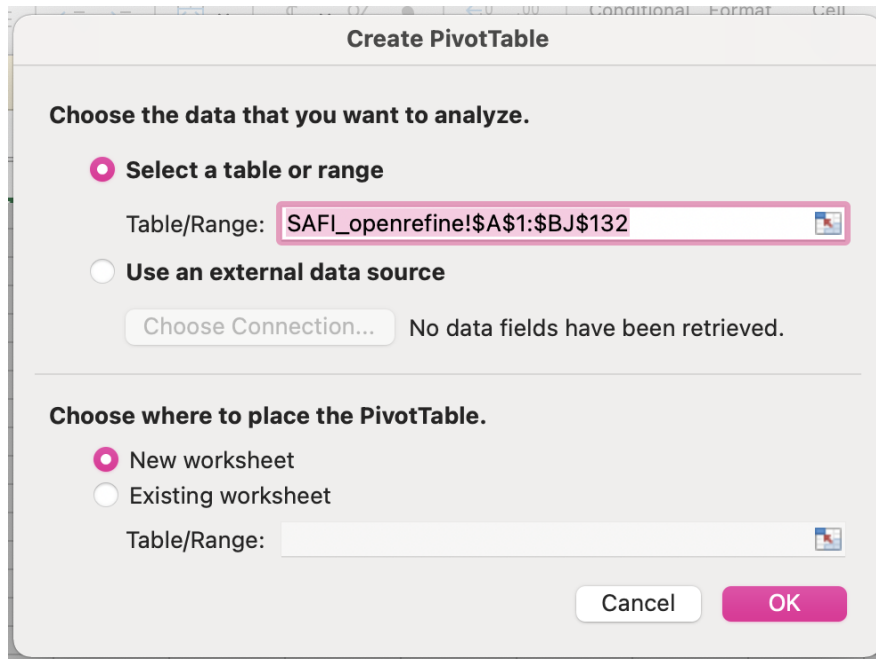
From Microsoft: « Vous pouvez utiliser un tableau croisé dynamique pour synthétiser, analyser, explorer et présenter des données de synthèse. Les graphiques croisés dynamiques complètent les tableaux croisés dynamiques en ajoutant des visualisations aux données de synthèse d'un tableau croisé dynamique. Ils vous permettent également d'effectuer des comparaisons et de dégager des tendances. »

[...]

« Un tableau croisé dynamique est un outil interactif qui permet de synthétiser rapidement de grandes quantités de données. Vous pouvez utiliser un tableau croisé dynamique pour analyser en détail des données numériques et répondre à des questions imprévues sur vos données. Un tableau croisé dynamique est particulièrement adapté aux opérations suivantes :

- interroger de grandes quantités de données de différentes manières conviviales ;
- obtenir des sous-totaux et agréger des données numériques, résumer les données par catégories et sous-catégories, et créer des calculs et des formules personnalisées ;
- développer et réduire les niveaux de données afin de cibler vos résultats et les analyser en détail à partir des données synthétisées par domaines ;
- déplacer des lignes dans des colonnes ou des colonnes dans des lignes (par « glissement ») afin d'afficher différents résumés des données sources ;
- filtrer, trier, regrouper et appliquer une mise en forme conditionnelle au sous-ensemble de données le plus utile et pertinent afin de vous concentrer sur les informations voulues ;
- présenter des rapports en ligne ou imprimés concis, annotés et bien organisés.

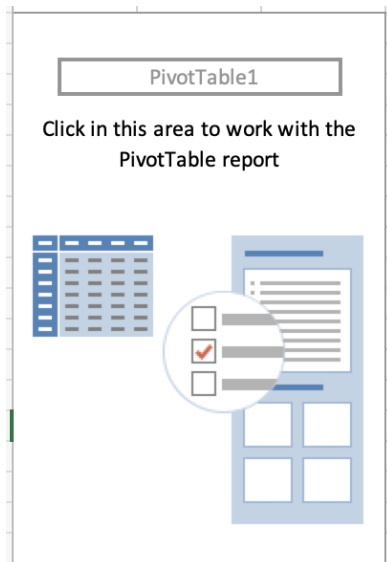
Pour commencer, sélectionnez toutes les données de votre feuille de calcul en utilisant **CTRL+A** (ou **Command+A** sur un Mac). Ensuite, allez dans l'onglet **Insertion** (ou « **Insert** ») et cliquez sur le bouton **Tableau croisé dynamique** (ou « **PivotTable** »). Cela ouvrira le menu **Créer un tableau croisé dynamique** (ou « **Create PivotTable** »).



Vérifiez que toutes les données de votre feuille de calcul ont été sélectionnées dans la première section. Pour la deuxième section, il est généralement plus simple de placer le tableau croisé dynamique dans une nouvelle feuille de calcul. Appuyez ensuite sur **OK**.

Votre interface de **tableau croisé dynamique** (ou « **PivotTable** ») doit se charger dans une nouvelle feuille de calcul (probablement appelée Sheet1). N'hésitez pas à renommer la feuille à ce stade, en cliquant avec le bouton droit sur l'onglet qui dit Sheet1 et en choisissant **Renommer** (ou « **Rename** »). Vous pouvez également faire glisser l'onglet de cette feuille sur le côté droit de l'onglet des données, si vous préférez (je le fais habituellement car cela me semble plus ergonomique).

Vous verrez une variété d'options dans le menu **Champs de tableau croisé dynamique** (ou « **PivotTable Fields** ») sur le côté droit de la page. Si jamais vous fermez accidentellement ce menu, faites simplement un clic droit sur l'espace réservé du tableau croisé dynamique (PivotTable) :



ou votre tableau croisé dynamique actif et choisissez **Afficher la liste des champs** (ou « **Show Field List** »). Notez que cliquer en dehors de l'espace réservé du tableau croisé dynamique ou du tableau croisé dynamique actif entraînera également le masquage du menu **Champs de tableau croisé dynamique** (ou « **PivotTable Fields** »), mais en cliquant à nouveau dessus, il s'affichera à nouveau.

Le menu **Champs de tableau croisé dynamique** (ou « **PivotTable Fields** ») comporte 5 composants importants :

- **Nom du champ** (ou « **Field Name** »)
 - Contient tous les noms de variables, c'est-à-dire tous les noms d'en-tête des colonnes de votre jeu de données
- **Filtres** (ou « **Filters** »)
- **Colonnes** (ou « **Columns** »)
- **Lignes** (ou « **Rows** »)
- **Valeurs** (ou « **Values** »)

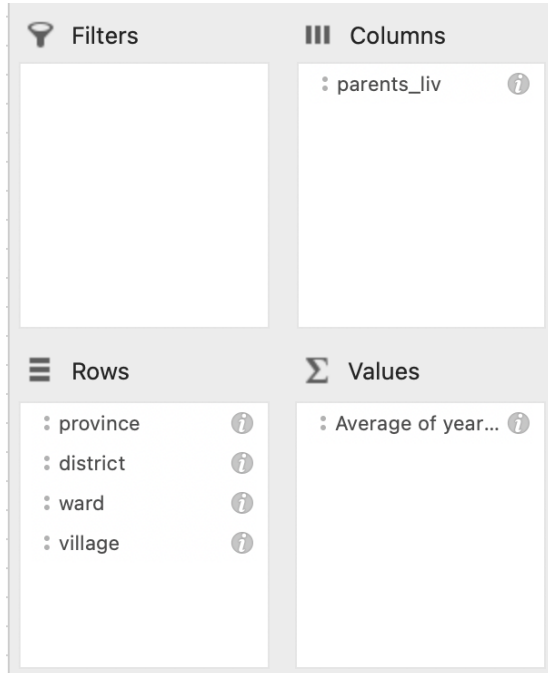
Nous allons expérimenter le déplacement des variables du **Nom du champ** (ou « **Field Name** ») entre chacune des quatre autres zones pour créer un tableau croisé dynamique (PivotTable). Cet objectif est de découvrir des tendances et des modèles dans nos données que nous ne verrions peut-être pas autrement.

Quelques points à noter :

- Pour inclure une variable, cochez la case à côté du nom de la variable dans la zone **Nom du champ** (ou « **Field Name** »).
- La variable apparaîtra dans l'une des quatre cases ci-dessous. Elle peut être glissée et déposée dans chaque catégorie.
- L'ordre dans chaque catégorie (**Filtres** ou « **Filters** », **Colonnes** ou « **Columns** », **Lignes** ou « **Rows** », **Valeurs** ou « **Values** ») est important. Vous pouvez également faire glisser et déposer l'ordre des noms de champs dans une catégorie donnée.

- Cliquer sur le petit *i* à droite de chaque nom de champ vous permet de personnaliser chaque champ, par exemple en passant du **nombre** (ou « **count** ») à la **moyenne** (ou « **average** »).

N'hésitez pas à expérimenter ! Pour aujourd'hui, je vais travailler à la création de ce qui suit :



Ce tableau croisé dynamique (PivotTable) utilise les variables suivantes dans la zone **Nom du champ** (ou « **Field Name** ») :

- province
- district
- ward
- village
- years_liv
- parents_liv

Pour les **Lignes** (ou « **Rows** »), vous aurez (dans cet ordre exact):

- province
- district
- ward
- village

Pour les **Colonnes** (ou « **Columns** »), vous aurez :

- parents_liv

Et enfin, pour les **Valeurs** (ou « **Values** ») vous aurez :

- year_liv

Et vous devrez aller au symbole *i* à la côté de year_liv pour changer l'option **Résumer par** (ou « **Summarize by** » par en **Moyenne** (ou « **Average** »).

Ce qui créera un tableau croisé dynamique (PivotTable) qui ressemble à :

Average of years_liv		Column Labels ▼	
Row Labels ▼	no	yes	Grand Total
[-] Manica	16.41176471	27.2875	23.05343511
[-] Bandula		15	15
[-] Bandula		15	15
Chirodzo		15	15
[-] Manica	16.41176471	27.44303797	23.11538462
[-] Bandula	16.41176471	26.93506494	22.7421875
49		26	26
Chirodzo	17.64705882	26.84210526	22.5
God	16.52941176	22.92	20.33333333
Ruaca	14.4375	31.48148148	25.13953488
Ruaca - Nhamuenda		16	16
Ruaca-Nhamuenda	25	21	22.33333333
Ruca		28.5	28.5
[-] Manica		47	47
Chiridozo		70	70
God		24	24
Grand Total	16.41176471	27.2875	23.05343511

Comme dernière étape de nettoyage, les cellules qui indiquent **les étiquettes de colonne** (ou « **Column Labels** ») et **les étiquettes de ligne** (« **Row Labels** ») peuvent être renommées simplement en cliquant et en tapant dans ces cellules. Par exemple, je pourrais changer **les étiquettes de colonne** (ou « **Column Labels** ») en **parent_liv** et **les étiquettes de ligne** (« **Row Labels** ») en **location**, comme ceci :

Average of years_liv	parent_liv		
location	no	yes	Grand Total
<input type="checkbox"/> Manica	16.41176471	27.2875	23.05343511
<input type="checkbox"/> Bandula		15	15
<input type="checkbox"/> Bandula		15	15
Chirodzo		15	15
<input type="checkbox"/> Manica	16.41176471	27.44303797	23.11538462
<input type="checkbox"/> Bandula	16.41176471	26.93506494	22.7421875
49		26	26
Chirodzo	17.64705882	26.84210526	22.5
God	16.52941176	22.92	20.33333333
Ruaca	14.4375	31.48148148	25.13953488
Ruaca - Nhamuenda		16	16
Ruaca-Nhamuenda	25	21	22.33333333
Ruca		28.5	28.5
<input type="checkbox"/> Manica		47	47
Chiridozo		70	70
God		24	24
Grand Total	16.41176471	27.2875	23.05343511