

STRUCTURER ET GERER DES DONNEES AVEC EXCEL

Introduction

Excel dispose de nombreux outils pour vous permettre de stocker et d'exploiter facilement des données. Néanmoins, ces outils sont conçus pour fonctionner sur des données structurées de manière adéquate.

Ce cours est illustré sur Excel 2010, mais les concepts sont valables sur toutes les versions et la plupart des fonctionnalités présentées le sont également (mais peuvent être présentées différemment).

Structurer ses données

Une base de données conventionnelle, c'est :

- **Toutes les données dans un même et unique onglet.** Dupliquer et disséminer les onglets dans des fichiers voire des classeurs différents est contreproductif. Au mieux, ce sera complexe à exploiter (et source d'erreur), au pire carrément impossible...
- **Les colonnes correspondent à des types d'informations différents** : une colonne NOM, une autre PRENOM, une autre DATE... A éviter : plusieurs informations accolées à la suite dans une même cellule.

Nom	Prénom	Date
...	...	...

- **Les lignes correspondent à des enregistrements différents** : ligne 1 pour *DUPONT Hervé, le 01/01/2000*, ligne 2 pour *MARIE Marie, le 31/12/2012*, etc.

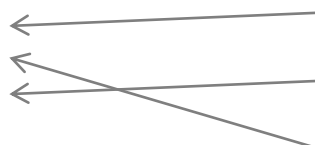
Nom	Prénom	Date
DUPONT	Hervé	01/01/2000
MARIE	Marie	31/12/2012
...	...	...

- **La première ligne est réservée au nom des colonnes.** Un nom bien choisi est précis, concis et unique (pas de doublon). Il ne faut pas laisser de colonne comportant des informations sans y mettre un nom.

Nom	Prénom	Date	Date	...
...	...	...	...	...

- **[FACULTATIF] la première colonne peut être réservée à un identifiant unique et propre à chaque ligne (appelé « ID » ou « Clé »).** Elle permet ensuite de retrouver facilement les informations, et notamment de joindre des bases de données distinctes faisant référence à un élément commun. Par exemple : mettre la date de naissance provenant d'une base de données B en face d'une base de données A comportant un nom et prénom.

Clé	Nom	Prénom	Clé DatNaiss
1	DUPONT	Hervé	21
2	DURAND	Jeanne	52
3	MARTIN	Antoine	33



Clé	Date
...	...
21	05/05/1973
...	...
33	01/02/1997
...	...
52	03/09/1967

- **Plusieurs enregistrements peuvent faire référence à un même élément, mais au moins l'un des champs (=colonne) doit être différent des autres.** Par exemple, DUPONT Hervé a commandé 3 articles différents, j'aurais donc 3 lignes distinctes dans ma base de données, faisant référence à cette même personne :

Clé	Nom	Prénom	Date	Article	Quantité
1	DUPONT	Hervé	11/11/2011	R010203	3
2	DUPONT	Hervé	11/11/2011	B0AA203	1
3	DUPONT	Hervé	11/11/2011	F67CD43	1

- **L'utilisation de plusieurs colonnes pour y stocker un même type d'information est à proscrire.** Par exemple (basé du l'exemple précédent) :

Nom	Prénom	Date	Article1	Article2	Article3
DUPONT	Hervé	11/11/2011	R010203	B0AA203	F67CD43

- **L'exploitation des données se fait à part.** La base de données ne doit comporter aucune ligne de sous total, total, etc. De même, on évite d'utiliser le même onglet pour y mettre d'autres tableaux, données, graphiques, etc.

Clé	Nom	Prénom	Date	Article	Quantité
1	DUPONT	Hervé	11/11/2011	R010203	3
2	DUPONT	Hervé	11/11/2011	B0AA203	1
3	DUPONT	Hervé	11/11/2011	F67CD43	1
TOTAL	DUPONT	Hervé	11/11/2011	3	5

- **Utiliser un format de donnée uniformisé.** Utiliser une saisie standardisée est primordial, on saisit donc des données dans la même unité (ex : une longueur en mètres), une date au même format (ex : JJ/MM/AAAA), etc.

Tirer profit des outils Excel

Excel dispose d'outils facilitant la gestion de données. En voici quelques-uns :

- **Le filtre automatique** : simple et utile le filtre automatique permet de limiter les lignes affichées selon différents critères, ou encore de trier une colonne par ordre croissant/décroissant (pour les valeurs numériques et dates) ou par ordre alphabétique.

Filtrage (Ctrl+Shift+L)

Active le filtrage des cellules sélectionnées.

Lorsque le filtrage est activé, cliquez sur la flèche dans l'en-tête de la colonne pour sélectionner le filtre à appliquer à la colonne.

Appuyez sur F1 pour obtenir de l'aide.

	A	B	C	D	E	F
1	NOM	Prénom	Nom complet	Date	Article	Valeur unitaire
2	ANDRE	Jules	ANDRE Jules	43735	A088M3	1,29
3	BERTRAND	Julien	BERTRAND Ju	43327	A088M3	1,29
4	MOREL	Marie	MOREL Marie	42980	A088M3	1,29
5	THOMAS	Marie	THOMAS Ma	43217	A476B5	2,99
6	DAVID	Laure	DAVID Laure	42759	A994T4	5,92
7	DAVID	Léonie	DAVID Léoni	43106	B179G1	7,84
8	RICHARD	Jérémy	RICHARD Jér	43351	B506O8	8,12

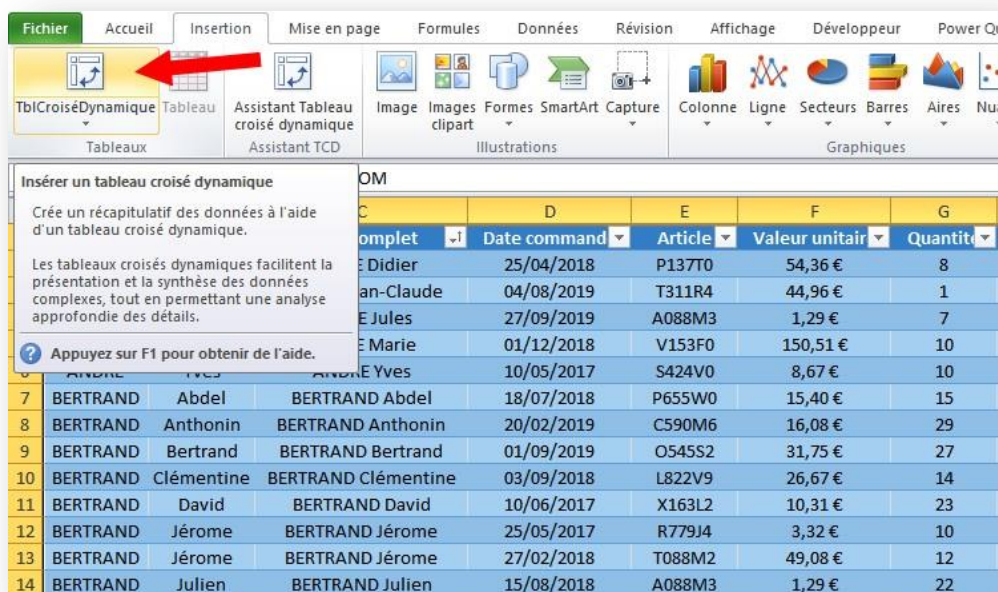
- **Le tableau structuré** : il se redimensionne seul à l'ajout/suppression de données et facilite l'actualisation des fonctionnalités qui dépendent de ce tableau (formules, tableaux croisés dynamiques, graphiques...). Autre avantage, les colonnes comportant des formules s'incrémentent automatiquement, et la syntaxe des formules change en faisant directement référence au nom des colonnes (ce qui facilite souvent la compréhension des formules).

Tableau (Ctrl+L)
Crée un tableau pour gérer et analyser des données associées.
Les tableaux facilitent le tri, le filtrage et la mise en forme des données dans une feuille.
Microsoft Power Query pour Excel
Appuyez sur F1 pour obtenir de l'aide sur le complément.

Formulaire de transition: =A2+B2 devient =[@[Champ1]]+ @[@[Champ1]]

	A	B	C	D
1	Champ1	Champ2	Champ3	Champ4
2	0,61	0,66	Nom1	1,08
3	1,00	1,41	Nom2	0,78
4	0,36	1,87	Nom3	0,96
5	0,25	0,97	Nom4	0,71
6	0,73	1,38	Nom5	0,13
7	0,66	1,04	Nom6	0,43
8	0,62	0,84	Nom7	3,56
9	0,79	1,86	Nom8	3,69
10	0,12	0,59	Nom9	3,32
11	0,34	0,65	Nom10	1,13
12	0,23	0,32	Nom11	1,69
13	0,12	1,92	Nom12	3,70
14	0,26	1,14	Nom13	0,01

- **Le tableau croisé dynamique (TCD) :** est un outil très fiable et performant pour exploiter une base de données. Il permet de synthétiser l'information rapidement, sans avoir à écrire de formule et sans doublon.



Liste sans doublon

Liste des variables du tableau

Glisser-déposer

Variable(s) à afficher en colonne

Champ des filtres (pour restreindre l'affichage à certaines modalités d'une variable)

Variable(s) à afficher en ligne

Variable(s) à synthétiser

Liste de champs de tableau croisé dynamique

Choisissez les champs à inclure dans le rapport :

☐ NOM

☐ Prénom

☒ Nom complet

☒ Date commande

☐ Article

☐ Valeur unitaire

☐ Quantité

☒ Valeur totale

Faites glisser les champs dans les zones ci-dessous :

☒ Filtre du rapport

Étiquettes de colon...

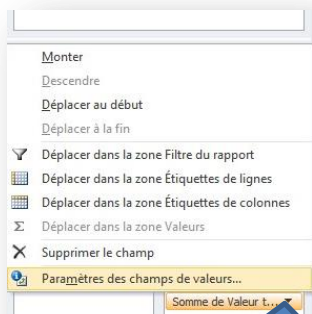
Σ Valeurs

Étiquettes de lignes

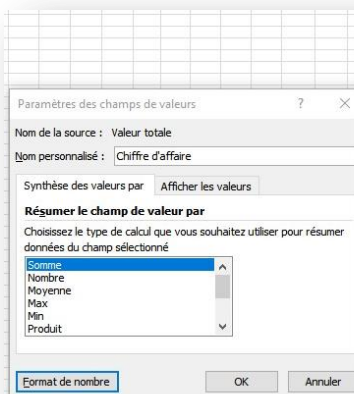
Nom complet

Nombre de Date c...

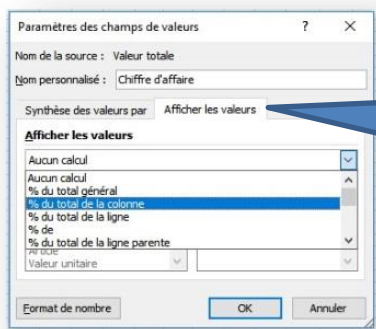
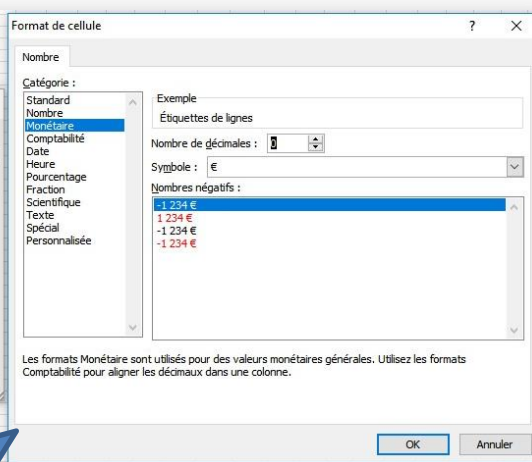
Somme de Valeur t...



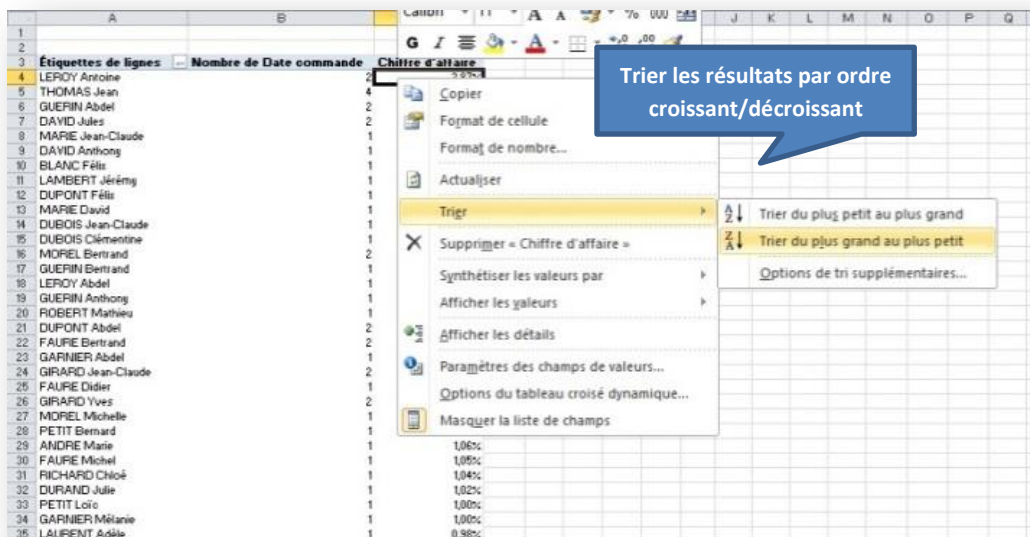
Paramétrer le calcul à réaliser
(nombre, somme, moyenne,
écart-type...)



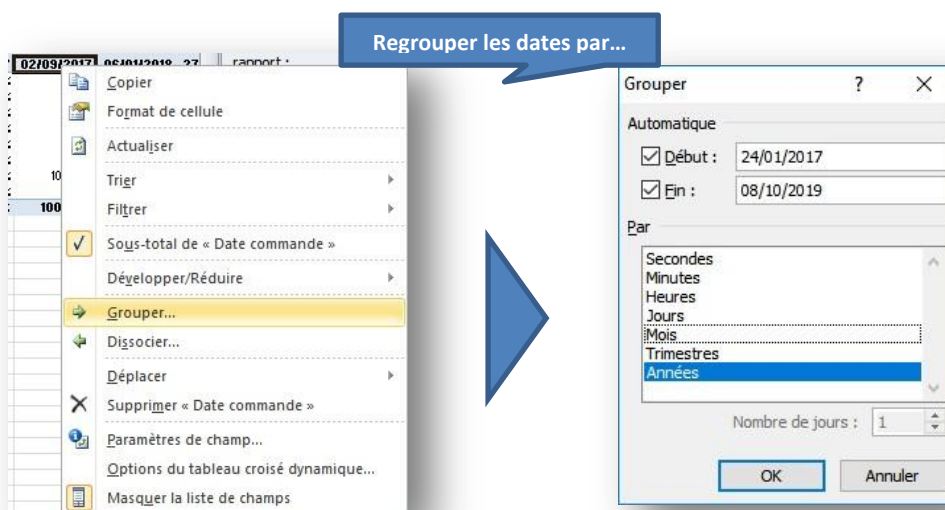
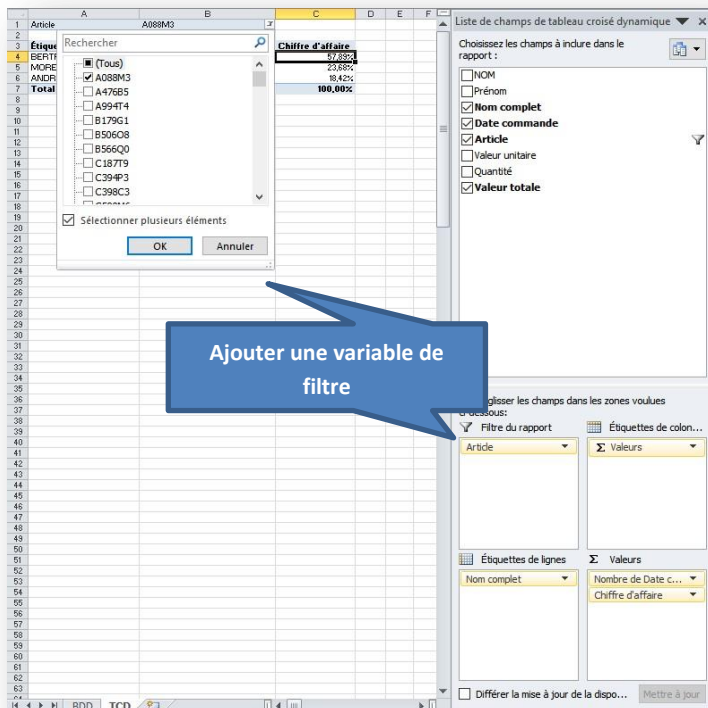
Format du résultat



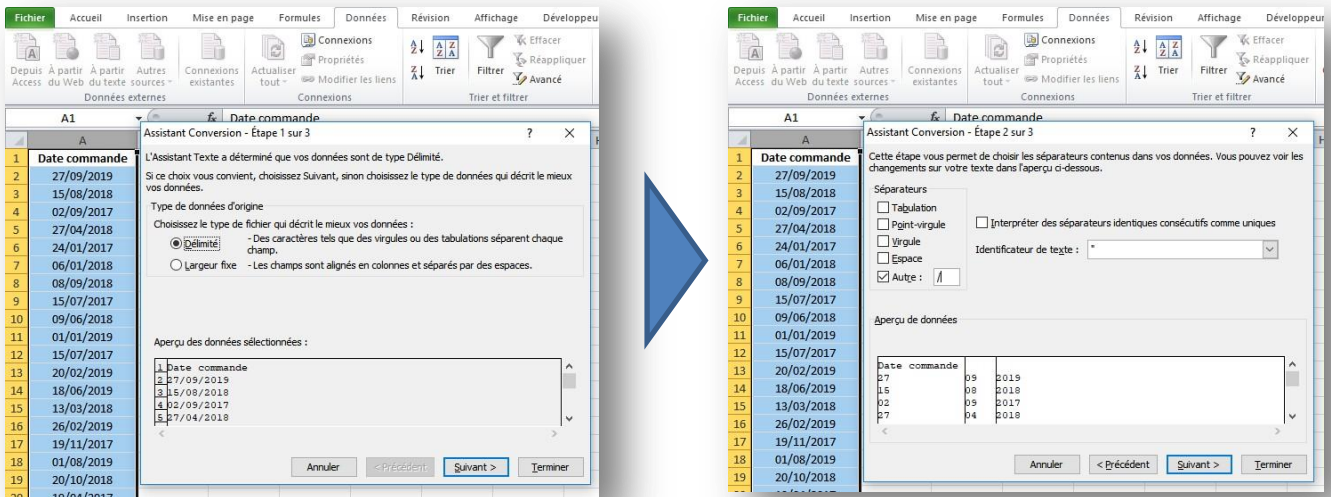
Mode d'expression du résultat
(brut, exprimé en % du total de la
ligne et/ou de la colonne, etc)



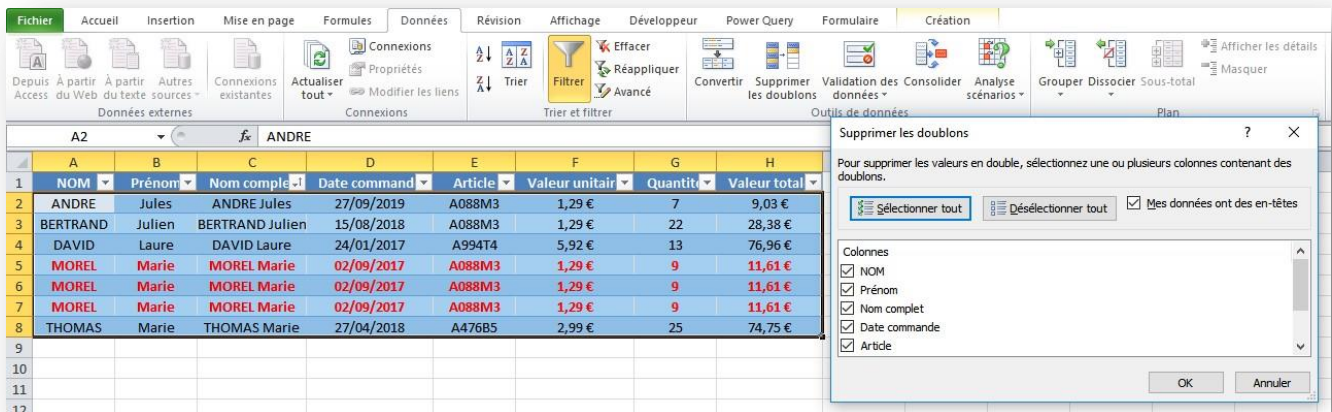
Trier les résultats par ordre
croissant/décroissant



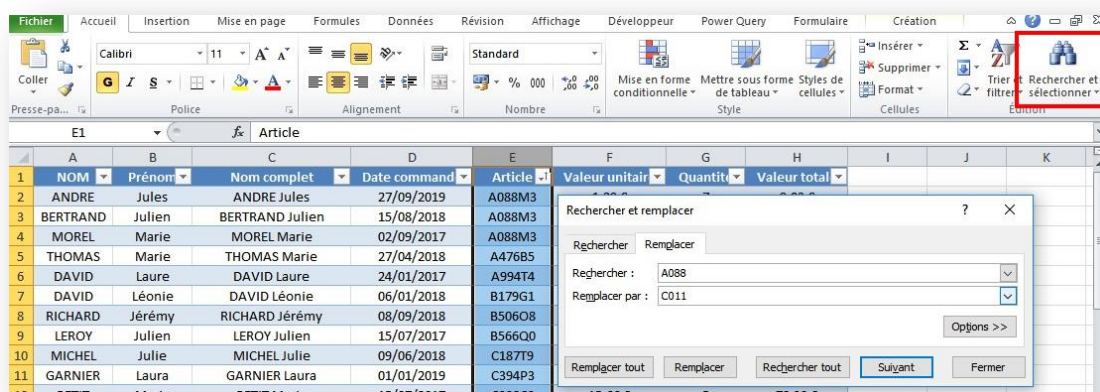
- **La fonctionnalité « Convertir »** : très utile pour séparer les informations d'une même colonne (courant à l'ouverture de fichiers CSV). La fonctionnalité convertir permet aussi de convertir automatiquement des données mal reconnues (nombres ou dates qui ne sont pas reconnus comme tels).



- **La fonctionnalité « Supprimer les doublons »** : le problème de nombreuses bases de données est la présence de doublons, qui peuvent fausser leur exploitation et analyse. Cet outil permet de supprimer les doublons sur la base d'une ou de plusieurs colonnes. Dans le dernier cas, seules les lignes qui sont strictement identiques sur toutes les colonnes indiquées sont supprimées.



- **La fonctionnalité « Rechercher et remplacer »** : elle permet de remplacer un ou plusieurs caractères dans une plage, y compris dans des formules. Cette fonctionnalité s'avère utile pour convertir des données numériques comportant un séparateur décimal « . » en nombres reconnaissables par Excel (séparateur « , »).



Autres astuces

- **Les données de date et heure dans un tableur** : une donnée valide de date est un nombre de jours écoulé depuis la référence du 00/01/1900. La partie décimale de ce nombre représente l'heure ($0,5 = 12/24 = 12$ heures). Le format JJ/MM/AAAA n'est qu'une histoire d'affichage, mais ne change rien à la valeur réellement stockée dans la cellule. Pour vérifier si une date est valide, il suffit de changer le format vers « Nombre », si la date est bien reconnue par Excel, elle apparaîtra sous une forme du type 43466,00 (correspond pour cet exemple à la date du 1^{er} janvier 2019).
- **La concaténation** : elle permet de combiner des informations sous forme de texte, par exemple créer un champ unique « Nom prénom » à partir d'un champ « Nom » (colonne A) et d'un champ « prénom » (colonne B). La formule s'écrit ainsi `=A2&« »&B2`, ou `=CONCATENER(A2;« »;B2)`.
- **La validation de données** : elle permet de restreindre le type de données qui peuvent être saisies dans une plage (par exemple seulement des nombres entre 0 et 100) ou d'imposer une liste de valeurs. Pour ce dernier cas, une liste déroulante apparaît dans chaque cellule de la plage concernée, au moment de la saisie.

The screenshot shows the Excel interface with the 'Validation des données' (Data Validation) dialog box open. The dialog box has three tabs: 'Options', 'Message de saisie', and 'Alerte d'erreur'. The 'Options' tab is selected. Under 'Critères de validation' (Validation Criteria), the 'Autoriser' (Allow) dropdown is set to 'Liste' (List). The 'Source de données' (Data Source) is set to '=Feuille1!\$A\$2:\$A\$12'. The 'Ignorer si vide' (Ignore blank) checkbox is checked. The 'Appliquer ces modifications aux cellules de paramètres identiques' (Apply these changes to identical cells) checkbox is unchecked. The 'Effacer tout' (Clear all) button is visible at the bottom left of the dialog box.

	A	B	C	D	E	F	G	H
	NOM	Prénom	Nom complet	Date commande	Article	Valeur unitaire	Quantité	Valeur totale
1	ANDRE	Jules	ANDRE Jules	27/09/2019	A088M3	1,29 €	7	9,03 €
2	BERTRAND	Julien	BERTRAND Julien	15/08/2018	A088M3	1,29 €	22	28,38 €
3	MOREL	Marie	MOREL Marie	02/09/2017	A088M3	1,29 €	9	11,61 €
4	THOMAS	Marie	THOMAS Marie	27/04/2018	A476B5	2,99 €	25	74,75 €
5	DAVID	Laure	DAVID Laure	24/01/2017	A994T4	5,92 €	13	76,96 €
6	DAVID	Léonie	DAVID Léonie	06/01/2018	B179G1	7,84 €	23	180,32 €
7	RICHARD	Jérémy	RICHARD Jérémy	08/09/2018	B506O8	8,12 €	11	89,32 €
8	LEROY	Julien	LEROY Julien	15/07/2017	B566Q0	13,47 €	22	296,34 €
9	MICHEL	Julie	MICHEL Julie	09/06/2018	C187T9	15,76 €	1	15,76 €
10	GARNIER	Laura	GARNIER Laura	01/01/2019	C394P3	11,91 €	29	345,39 €
11	PETIT	Marie	PETIT Marie	15/07/2017	C398C3	15,66 €	5	78,30 €