

Initiation Tableau – Tableau prep


Niveau I - Découverte

Introduction

Tableau Prep vous aide à nettoyer et à organiser vos données pour l'analyse. La première étape de ce processus consiste à identifier les données avec lesquelles vous allez travailler.

Pour les sections qui viennent téléchargez 'ORDERS_USCA_2015' afin de tester et reproduire les exemples.

Se connecter via des connecteurs intégrés pour les types de données courants

La manière la plus courante de se connecter aux données est d'utiliser les connecteurs intégrés dans Tableau Prep Builder. Ceux-ci sont disponibles pour les types de données les plus courants, et de nouveaux connecteurs sont ajoutés fréquemment avec les nouvelles versions de Tableau Prep Builder. Pour voir la liste des connecteurs disponibles, ouvrez Tableau Prep Builder, puis cliquez sur le bouton **Ajouter une connexion**  pour voir si le volet gauche contient les connecteurs disponibles sous **Connexion**.

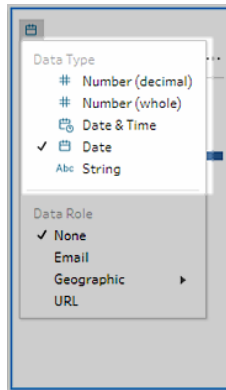
Examiner et Filtrer vos données

Utilisez les options de cette rubrique pour bien appréhender la composition de vos données et mieux comprendre les modifications que vous devez apporter, ainsi que l'effet des opérations que vous incluez dans le flux.

Vérifier les types de données affectés à vos données

Comme Tableau Desktop, Tableau Prep interprète les données de vos champs lorsque vous faites glisser une connexion vers le volet **Flux** et lui attribue automatiquement un type de données. Étant donné que différentes bases de données peuvent gérer les données de différentes manières, il peut arriver que l'interprétation de Tableau Prep ne soit pas correcte.

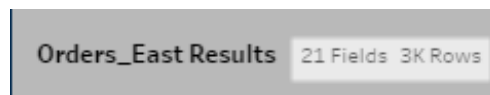
Pour modifier un type de données, cliquez sur l'icône du type de données et sélectionnez le type de données correct dans le menu contextuel. Vous pouvez modifier des types de données « chaîne » ou « entier » en **Date** ou **Date et heure**, et Tableau Prep déclenchera la fonction Auto DateParse (analyse automatique de date) pour modifier ces types de données. Comme Tableau Desktop, si la modification échoue, vous voyez des valeurs null dans les champs à la place, et vous pouvez créer un calcul pour effectuer la modification.



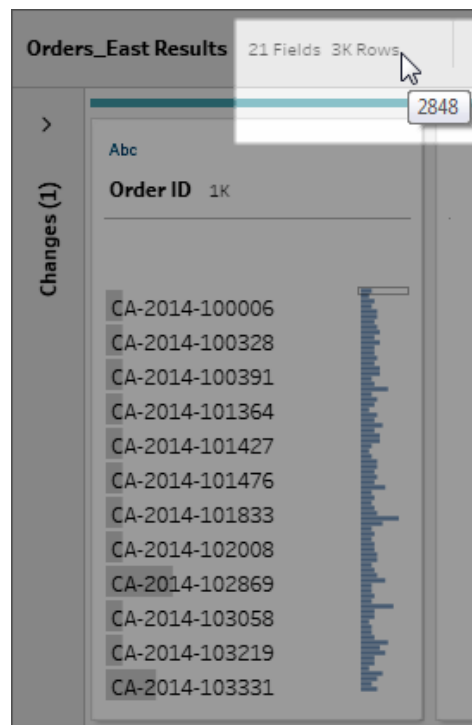
Consulter les informations de taille de vos données

Après vous être connecté à vos données, ajoutez une table au flux, puis ajoutez une étape. Vous pouvez utiliser le volet **Profil** pour voir l'état et la structure actuels de vos données et repérer les valeurs null et les données atypiques.

- **Nombre de fichiers et de lignes** : dans le coin supérieur gauche du volet **Profil**, vous pouvez voir des informations qui résument le nombre de champs et de lignes dans les données à un point particulier du flux. Tableau Prep arrondit au millier le plus proche. Dans l'exemple ci-dessous, l'ensemble de données comporte 21 champs et 3000 lignes.



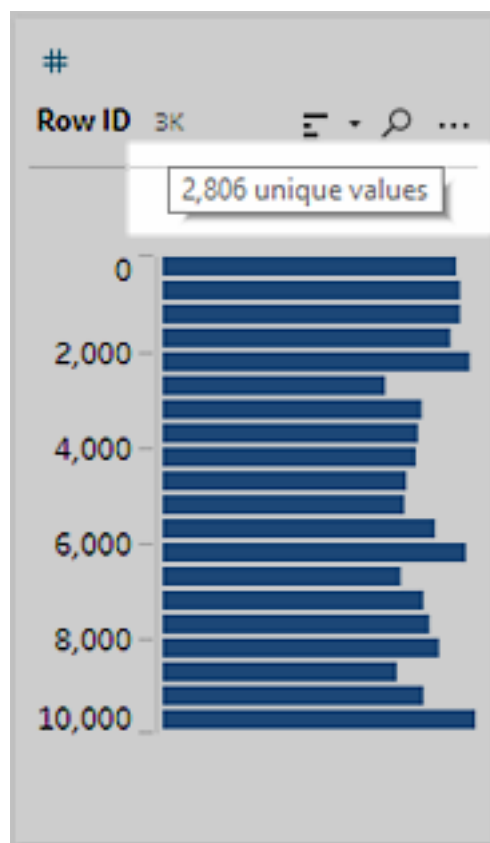
Lorsque vous survolez le nombre de champs et de lignes, vous pouvez voir le nombre exact de lignes (dans cet exemple, 2848).



- **Taille de l'ensemble de données** : utilisez un sous-ensemble de vos données en spécifiant le nombre de lignes à inclure dans l'onglet **Échantillon de données** du volet **Entrée**.
- **Échantillonné** : pour vous permettre d'interagir directement avec vos données, Tableau Prep utilise un sous-ensemble de vos données brutes. Le nombre de lignes est déterminé par les types de données et le nombre de champs rendus. Les champs de chaînes ont plus d'espace de stockage que les entiers, donc si vous avez 10 champs de chaînes dans votre ensemble de données, vous obtiendrez moins de lignes que si vous aviez 10 champs d'entiers.

Un indicateur **Échantillonné** Sampled s'affiche en regard des informations de taille dans le volet **Profil** pour indiquer qu'il s'agit d'un sous-ensemble de votre ensemble de données. Vous pouvez modifier la quantité de données que vous incluez dans votre flux.

- **Nombre de valeurs uniques** : le nombre en regard de chaque en-tête de champ représente les valeurs distinctes contenues dans chaque champ. Tableau Prep arrondit au millier le plus proche. Dans l'exemple ci-dessous, 3000 valeurs distinctes sont représentées dans le champ Description mais si vous survolez le nombre, vous pouvez voir le nombre exact de valeurs uniques.

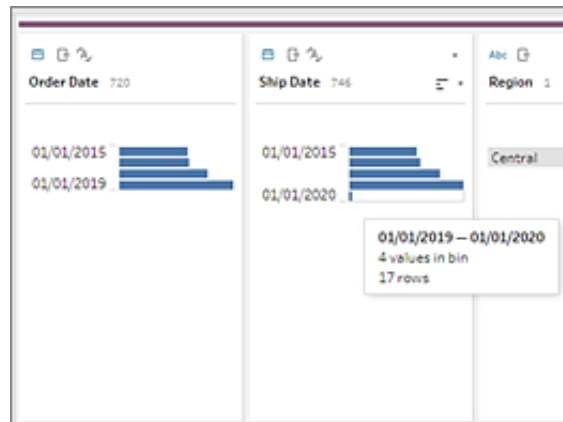


Consulter la distribution des valeurs ou les valeurs uniques

Par défaut, Tableau Prep regroupe les valeurs numériques, de date et de date/heure d'un champ en compartiments. Ces compartiments sont aussi connus sous le nom de classes. Les classes font en sorte que vous puissiez voir la distribution des valeurs dans leur ensemble, et identifier rapidement les

valeurs atypiques et les valeurs null. La taille de la classe est calculée sur les valeurs minimum et maximum dans le champ et les valeurs null s'affichent toujours en haut de la distribution.

Par exemple, les dates de commande et d'expédition sont résumées ou « classées » par année. Chaque classe représente une année depuis janvier de l'année de début à janvier de l'année suivante, et étiquetée en conséquence. Étant donné que certaines dates de ventes et dates d'expédition tombent dans la dernière partie de 2018 et 2019, une classe est créée pour l'année suivante pour ces valeurs.



Si un champ de données discrètes (ou de catégorie) contient de nombreuses lignes ou a une distribution si large qu'elle ne peut être affichée dans le champ sans défilement, vous pouvez voir une distribution résumée à droite du champ. Vous pouvez cliquer et faire défiler la distribution pour cibler des valeurs spécifiques.

Exercice

- ➔ Connectez-vous à 'ORDERS_USCA_2015'
- ➔ Vérifiez que les types de données sont corrects
- ➔ Ajoutez une étape de nettoyage
- ➔ Observez taille des données
- ➔ Observer la distribution des données

Combien y a-t-il de champs dans cette base ? Combien de lignes ?

Combien y a-t-il de villes différentes ?

Quelle est la sous-catégorie ayant le plus de commandes ? Combien de commandes ?

Créer et organiser votre flux

Après vous être connecté aux données que vous souhaitez inclure dans votre flux, vous pouvez commencer le nettoyage et l'organisation de vos données en ajoutant de nouvelles étapes au flux ou en insérant des étapes entre les étapes existantes.

Pour organiser votre flux, vous pouvez modifier les couleurs par défaut des étapes, ajouter des descriptions afin de fournir le contexte de vos étapes ou de vos actions de nettoyage, ou réorganiser la disposition de votre flux afin de faciliter le suivi des flux complexes.


Ajouter ou insérer des étapes

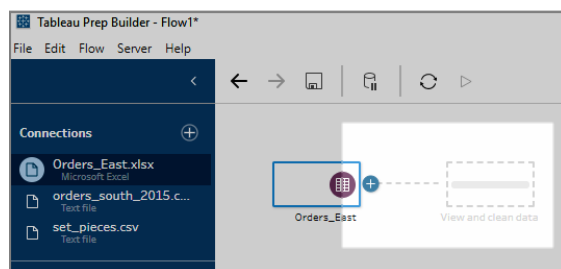
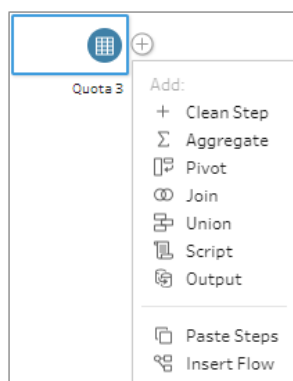
Au fur et à mesure que vous construisez votre flux, vous pouvez ajouter différents types d'étapes pour effectuer les actions dont vous avez besoin. Par exemple, ajoutez une **Étape de nettoyage** pour effectuer des opérations telles que scinder des champs, appliquer des filtres ou effectuer diverses autres opérations de nettoyage de données désordonnées.

À mesure que votre flux commence à prendre forme, vous devrez peut-être revenir aux étapes précédentes de votre flux et insérer différents types d'étapes pour effectuer diverses actions, par exemple ajouter une étape de nettoyage supplémentaire ou agréger vos données afin d'utiliser le même niveau de détail qu'une étape ultérieure.

Ajouter des étapes


Après vous être connecté à vos données et avoir fait glisser une table dans l'espace de travail, cliquez

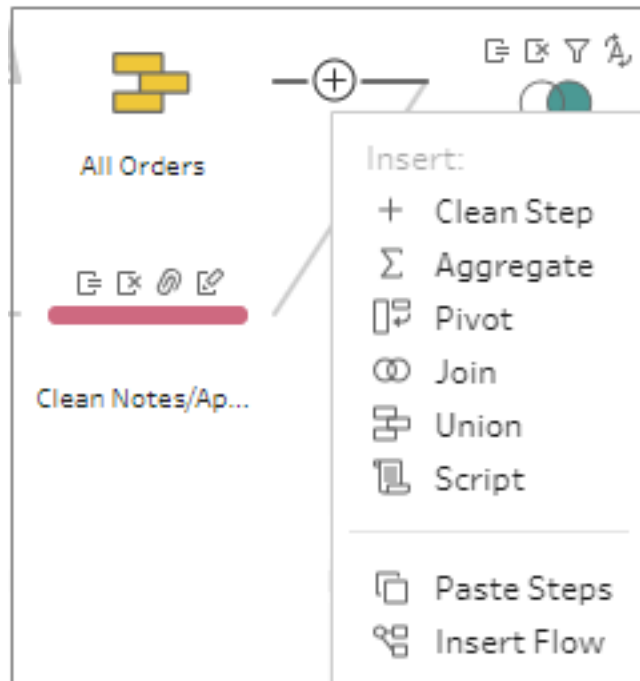
sur le bouton plus  pour sélectionner un type d'étape dans le menu. Sinon, cliquez sur l'étape de nettoyage suggérée pour ajouter automatiquement une étape de nettoyage à votre flux.



Insérer des étapes

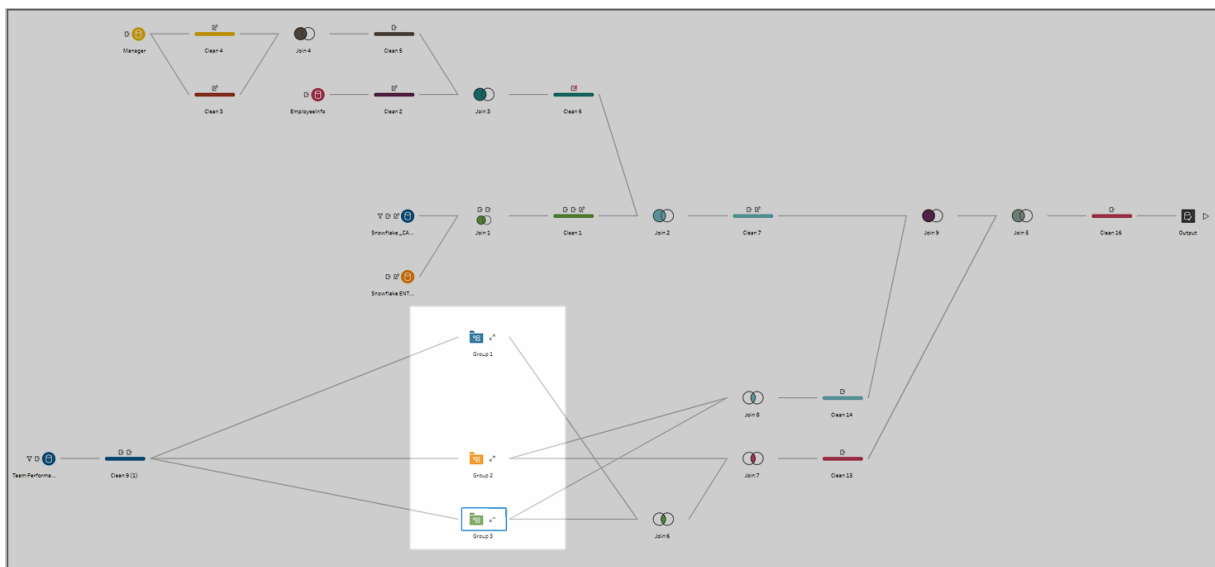
Insérez une étape entre les étapes existantes. Les types d'étapes d'entrée et de sortie ne sont pas disponibles dans ce menu.

Placez le curseur au milieu de la ligne de flux là où vous souhaitez insérer une étape jusqu'à ce que l'icône plus  apparaisse. Cliquez ensuite sur l'icône et sélectionnez un type d'étape.



Regrouper des étapes

Vous pouvez utiliser l'option Regrouper pour compartimenter des sections de flux volumineux complexes dans des dossiers afin de faciliter le suivi, la résolution des problèmes ou le partage de votre flux avec d'autres personnes. Vous pouvez modifier la couleur du groupe, ajouter une description, copier et coller les étapes regroupées dans d'autres zones de votre flux.



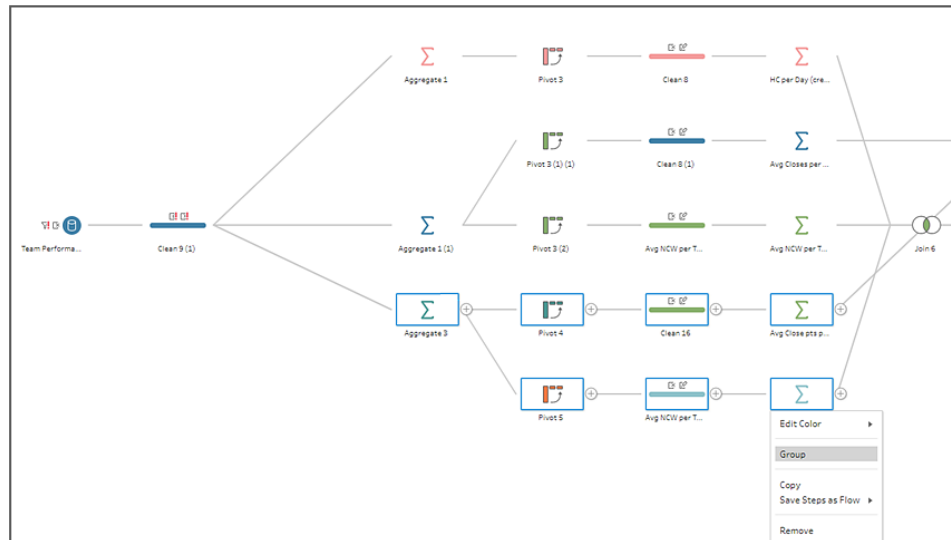
Lors de la création de groupes, les critères suivants s'appliquent :

- Les étapes doivent être directement connectées à une ligne de flux.
- Les étapes ne peuvent être incluses que dans un seul groupe à la fois.
- Les groupes ne peuvent pas être imbriqués.



- Vous pouvez ajouter des étapes à un groupe ou en supprimer à tout moment, tant que vous maintenez les connexions de ligne de flux entre les étapes du groupe. Cela s'applique également à la suppression des étapes d'un flux qui sont déjà incluses dans un groupe. Dans ce scénario, le groupe est automatiquement dégroupé.

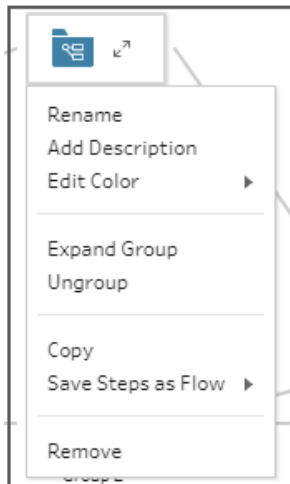
Créer un groupe


Sélectionnez un ensemble d'étapes connectées dans votre flux (vous pouvez également faire glisser des étapes pour sélectionner plusieurs étapes en un seul clic), puis faites un clic droit ou Cmd-clic (MacOS) sur les étapes sélectionnées et sélectionnez **Regrouper** dans le menu.



Après avoir créé le groupe, vous pouvez effectuer l'une des opérations suivantes :

- Cliquez sur les flèches doubles pour développer  ou réduire le groupe à tout moment .
- Ajoutez d'autres étapes au groupe en faisant glisser une étape connectée et en la déposant sur le dossier réduit.
- Supprimez des étapes du groupe. Dans l'état développé, faites un clic droit ou Cntrl-clic (Mac OS) sur une étape et sélectionnez **Supprimer du groupe**.
- Dans l'état réduit, faites un clic droit ou Cntrl-clic (MacOS) pour ouvrir le menu et faites votre choix parmi les options suivantes :



- **Renommer** : modifiez le nom du groupe.
- **Ajouter une description** : entrez une description pour le groupe
- **Modifier la couleur** : modifiez la couleur du dossier de groupe. Cela ne changera pas la couleur des étapes individuelles du groupe.
- **Développer le groupe** : affichez toutes les étapes du groupe. Vous pouvez également cliquer sur les flèches doubles  pour développer le groupe.
- **Dégrouper** : supprimez toutes les étapes du groupe et supprimez le groupe.
- **Copier** : copie le groupe et toutes les étapes du groupe sur votre presse-papiers pour les coller ailleurs dans votre flux.
- **Enregistrer les étapes sous forme de flux** (Tableau Prep Builder uniquement) : enregistrez vos étapes regroupées localement dans un fichier sur votre ordinateur.
- **Supprimer** : supprime le groupe et toutes les étapes du groupe du flux.

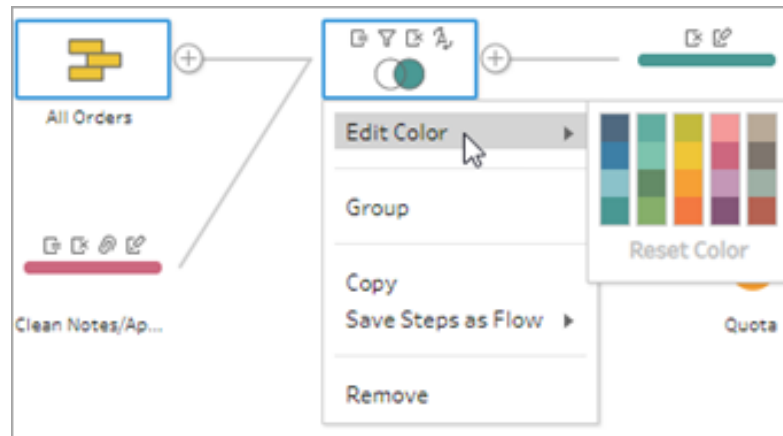


Modifier le thème de couleur du flux

Tableau Prep attribue à chaque étape de votre flux une couleur par défaut. Ce thème de couleur est appliqué au flux entier pour vous aider à suivre vos données tout au long du flux à mesure que vous appliquez les étapes de nettoyage, que vous liez, réunissez ou agrégez les données. Vous savez ainsi quels fichiers sont affectés par vos opérations.

Pour sélectionner un thème de couleur différent pour vos étapes, procédez comme suit :

1. Sélectionnez une ou plusieurs étapes.
2. Faites un clic droit ou Ctrl-clic (MacOS) sur une étape sélectionnée et sélectionnez **Modifier la couleur**.



3. Cliquez sur une couleur dans la palette de couleurs pour l'appliquer.

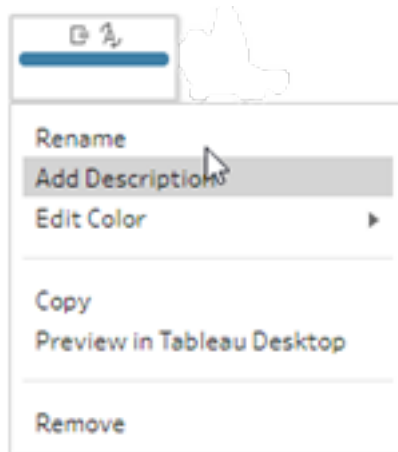
Ajouter des descriptions aux étapes du flux et aux actions de nettoyage

À mesure que vous créez votre flux et effectuez diverses opérations de nettoyage, vous pouvez souhaiter ajouter une description. Elle aidera les autres utilisateurs, qui examineront ou exploiteront votre flux ultérieurement, à mieux comprendre vos étapes. Vous pouvez ajouter une description à n'importe quelle étape individuelle de votre flux directement dans le volet Flux, à tout groupe d'étapes, ou à toute action de nettoyage dans le volet **Modifications** afin de fournir un contexte supplémentaire pour vos modifications. La description peut contenir jusqu'à 200 caractères.

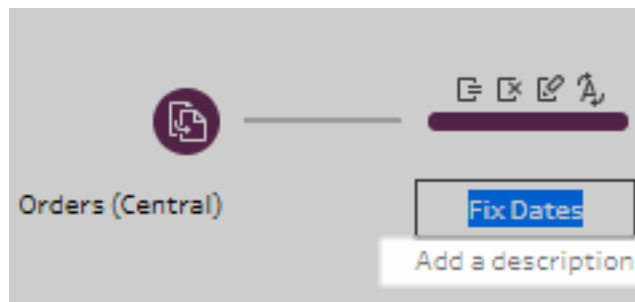
Ajouter une description aux étapes du flux

Lorsque vous ajoutez une description, une icône  de message est ajoutée sous l'étape. Cliquez sur l'icône pour afficher ou masquer le texte de la description dans le volet Flux.

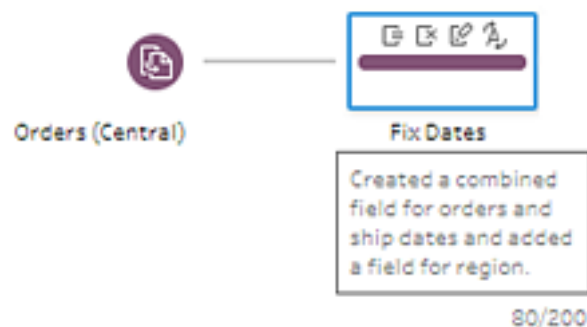
1. Dans le volet Flux, sélectionnez une étape.
2. Effectuez l'une des actions suivantes :
 - Faites un clic droit ou Ctrl-clic (MacOS) sur l'étape, puis sélectionnez **Ajouter une description** dans le menu.




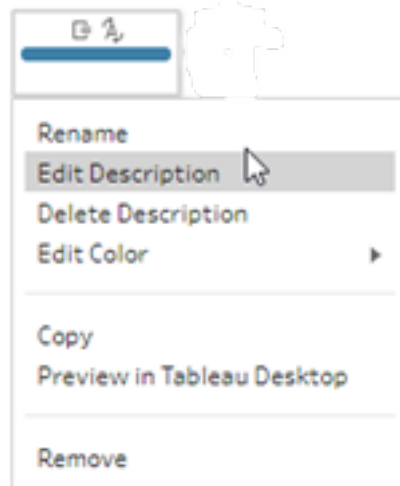
- Double-cliquez sur le champ de nom dans l'étape, puis cliquez sur **Ajouter une description**.



3. Entrez votre description dans la zone de texte.

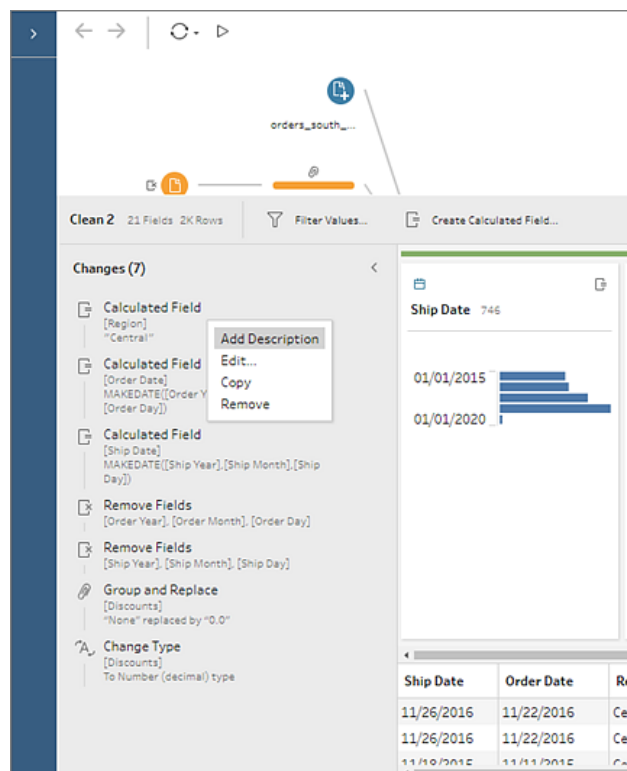


4. Cliquez-en-dehors de la zone de texte ou appuyez sur Entrée pour appliquer vos modifications. Par défaut, la description s'affiche sous l'étape. Pour masquer la description, cliquez sur l'icône de message .
5. Pour modifier ou supprimer la description, faites un clic droit (ou Ctrl+clic sur MacOS) sur l'étape ou la description. Ensuite, à partir du menu, sélectionnez **Modifier la description** ou **Supprimer la description**.

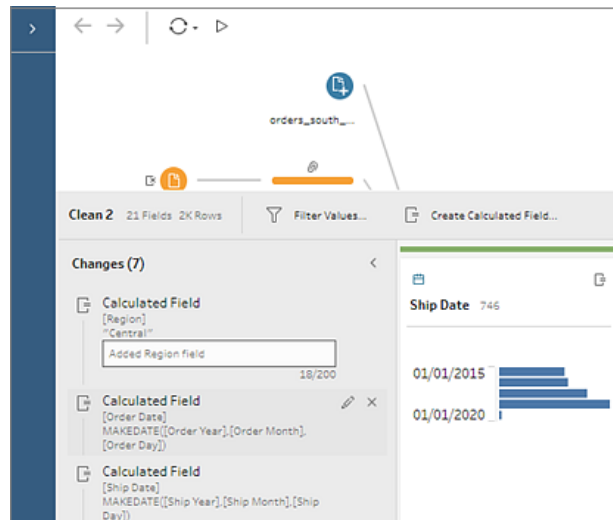



Ajouter une description à une entrée modifiée

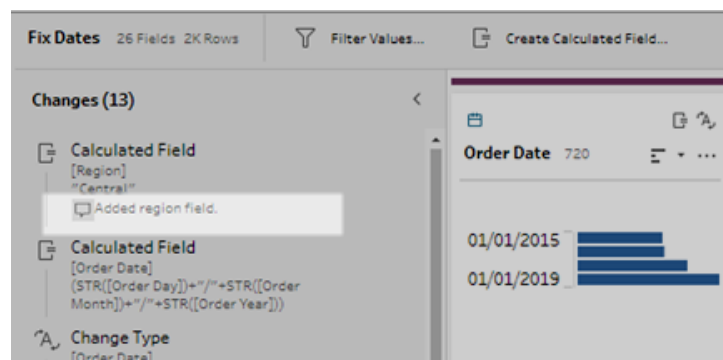
1. Sélectionnez une étape dans le volet de flux.
2. Ouvrez le volet **Modifications** ou l'onglet **Modifications**.
3. Faites un clic droit ou Ctrl-clic (MacOS) sur une entrée dans le volet **Modifications** et sélectionnez **Ajouter une description**.



4. Entrez une description pour l'action de modification.



La description apparaît sous le texte généré pour la modification avec une icône de commentaire .



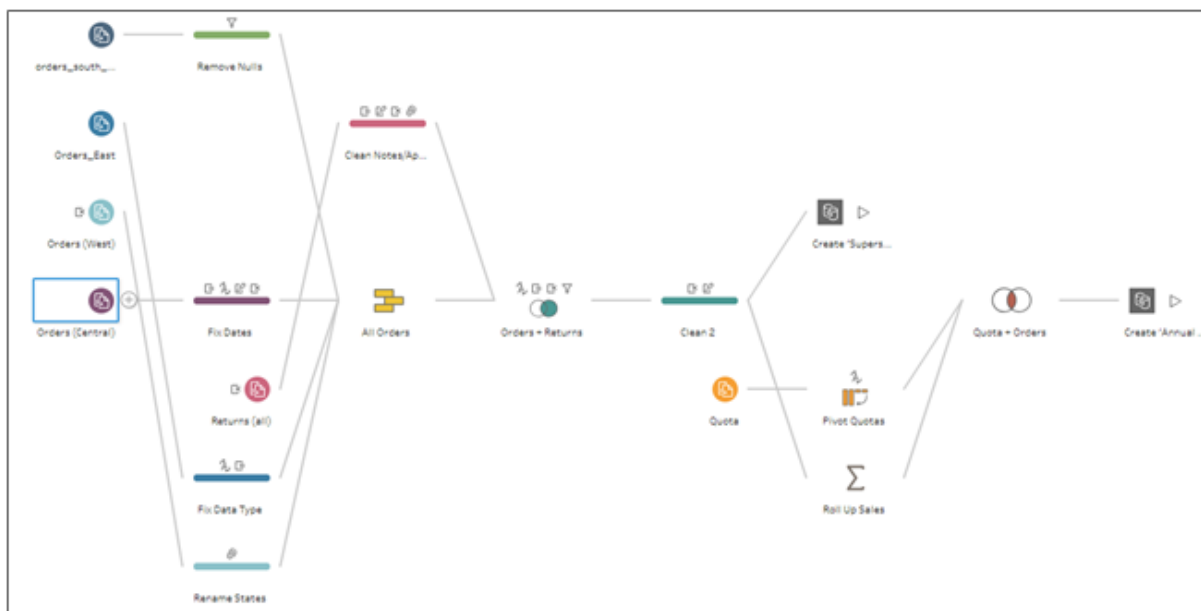
5. Pour modifier ou supprimer la description, faites un clic droit (ou Ctrl+clic sur MacOS) sur l'élément à modifier, puis sélectionnez **Modifier la description** ou **Supprimer la description**.

Réorganiser la disposition de votre flux

Lorsque vous créez un flux, Tableau Prep Builder utilise une disposition par défaut. Chaque flux organise et traite vos données de gauche à droite, avec les étapes des données entrantes commençant à l'extrême gauche de l'espace de travail, et les étapes de sortie se terminant à droite de cet espace. Cependant, si vous créez des flux importants et complexes, ils peuvent rapidement devenir difficiles à suivre.

Vous pouvez nettoyer la disposition de votre flux en sélectionnant et en déplaçant des étapes afin de l'organiser de manière logique pour vous. Par exemple, vous pouvez fixer des lignes de flux croisées, déplacer des étapes de votre flux pour nettoyer l'espace blanc supplémentaire, ou réorganiser les étapes de votre flux afin de montrer une séquence claire d'événements.

Par exemple, le flux suivant est déroutant et difficile à suivre :



Pour nettoyer ce flux, sélectionnez et faites glisser les étapes vers le haut, le bas, la gauche ou la droite, puis déposez-les à un nouvel emplacement dans l'espace de travail. Les étapes du flux ne peuvent pas être déplacées dans une position qui perturbe le flux du processus de gauche à droite. Par exemple, vous ne pouvez pas faire glisser une étape d'union positionnée avant une étape de jointure vers une position qui se trouve après cette étape de jointure dans le flux.

Lorsque vous faites glisser des étapes du flux vers un emplacement autorisé, une zone orange s'affiche. Si l'emplacement n'est pas autorisé, aucune zone orange ne s'affiche et les étapes reviennent à leur emplacement d'origine lorsque vous essayez de les déposer.

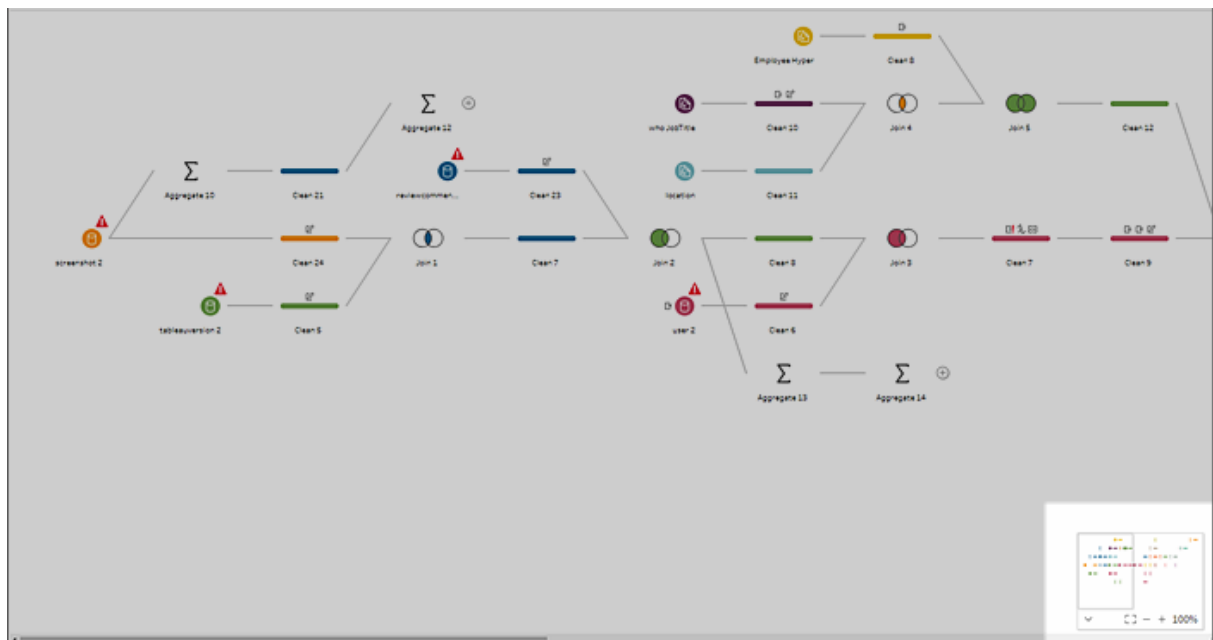
Pour déplacer des étapes dans votre flux, procédez comme suit :

1. Dans le volet Flux, sélectionnez les étapes que vous souhaitez déplacer. Vous pouvez cliquer sur une étape spécifique, la faire glisser pour sélectionner plusieurs étapes, ou faire un Ctrl-clic ou Cmd-clic (MacOS) pour sélectionner des étapes qui ne sont pas adjacentes.
2. Glissez-déposez les étapes vers le nouvel emplacement.

Voir la réorganisation des étapes du flux en action

Utiliser l'outil de navigation de flux

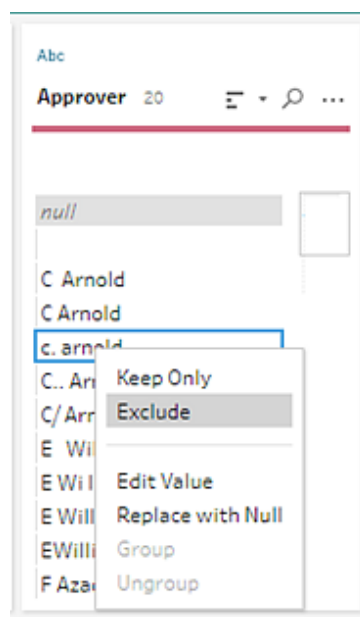
Lorsque vous travaillez avec des flux volumineux, il peut être difficile de faire défiler la liste d'avant en arrière pour rechercher une zone particulière du flux et y revenir. L'outil de navigation de flux facilite cette tâche. Le navigateur de flux est une version miniature de votre flux qui apparaît dans le coin inférieur droit de l'espace de travail.



Filtrer vos données

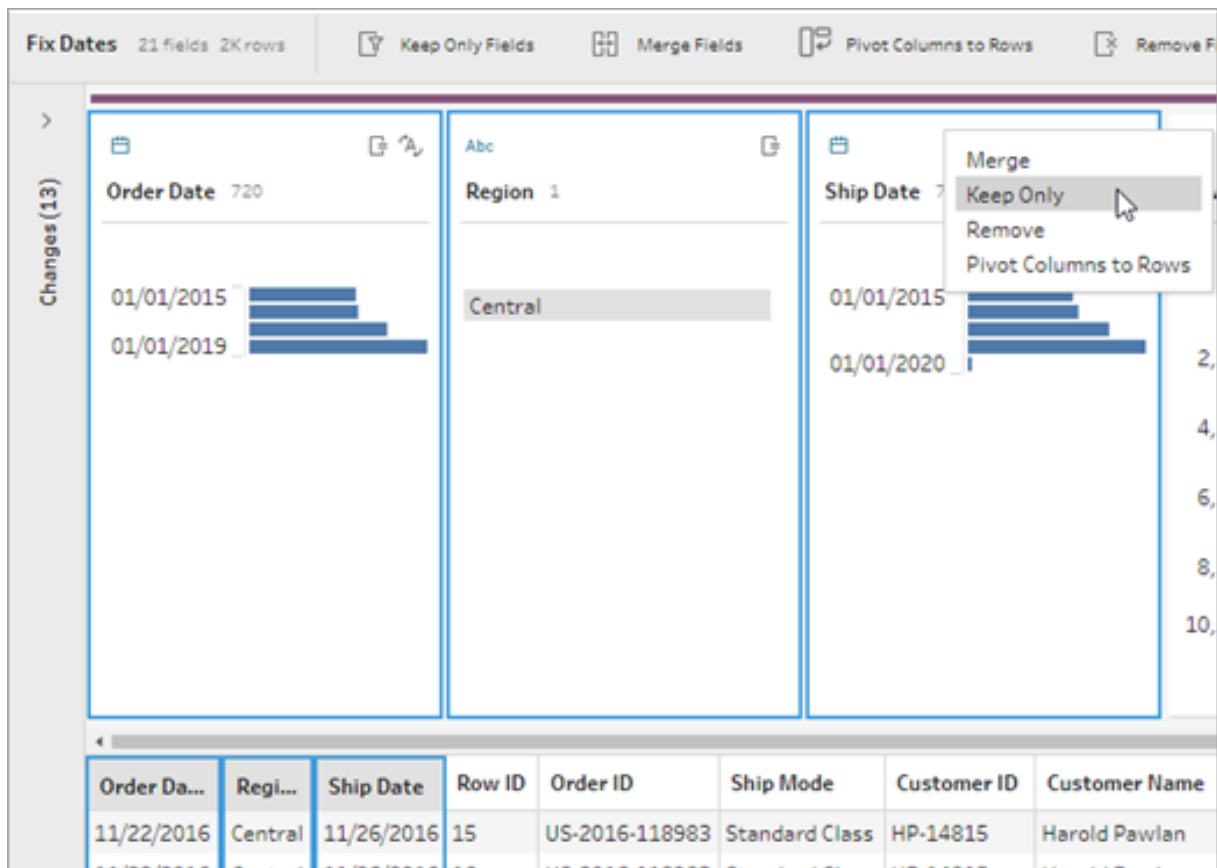
Tableau Prep fournit diverses options que vous pouvez utiliser pour filtrer vos données. Par exemple, utilisez **Conserver uniquement** ou **Exclure** pour effectuer un filtrage en un clic sur une valeur spécifique pour un champ d'une fiche Profil, d'une grille de données ou d'une fiche de résultats, ou sélectionnez parmi diverses options de filtrage pour des besoins de filtrage plus complexes. Vous pouvez également conserver ou supprimer des champs entiers.

Vous pouvez filtrer les données à toute étape du flux. Si vous souhaitez simplement modifier une valeur spécifique, vous pouvez sélectionner **Modifier la valeur** pour modifier la valeur en ligne ou remplacer la valeur par une valeur Null.



Conserver ou supprimer des champs

Au fur et à mesure que vous travaillez avec vos données dans votre flux, vous voudrez peut-être supprimer les champs inutiles. Dans le volet Profil ou dans la grille de données de toute étape de nettoyage ou d'action, sélectionnez un ou plusieurs champs et faites un clic droit ou Ctrl-clic (MacOS) et sélectionnez **Supprimer** pour supprimer les champs sélectionnés, ou sélectionnez **Conserver uniquement** pour ne conserver que les champs sélectionnés et supprimer tous les champs non sélectionnés.



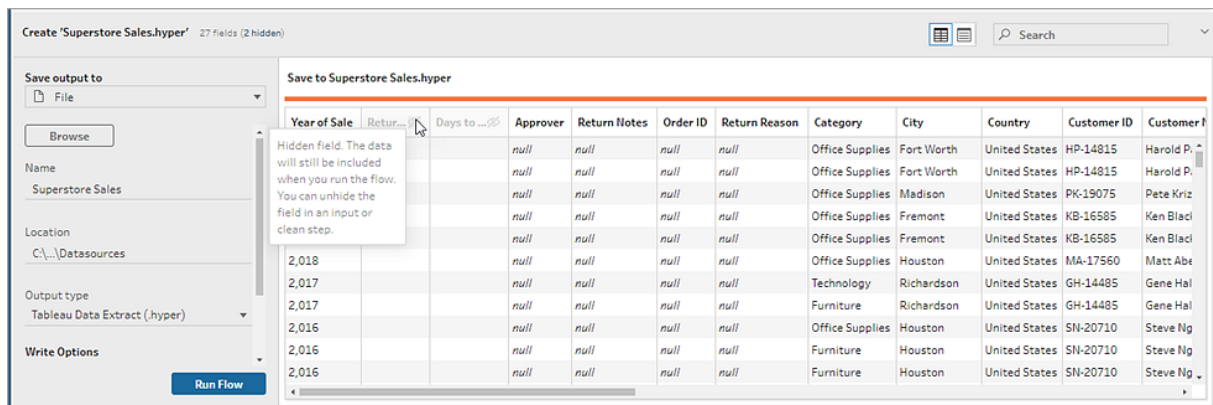
Masquer des champs

Si votre flux comporte des champs qui n'ont pas besoin d'être nettoyés, mais que vous souhaitez malgré tout les inclure dans votre flux, vous pouvez masquer les champs au lieu de les supprimer. Les données de ces champs ne seront pas chargées tant que vous n'aurez pas affiché les champs ou exécuté votre flux pour générer votre sortie.

Lorsque vous masquez des champs, une nouvelle fiche Profil appelée **Champs masqués** est automatiquement ajoutée au volet Profil, ce qui vous permet d'afficher facilement les champs de la liste lorsque vous en avez besoin.

Vous pouvez inclure des champs masqués dans la plupart des opérations, mais dans le cas des jointures, des agrégations et des permutations, le champ doit être affiché pour être utilisé dans l'un de ces types d'étape. Si vous masquez le champ après qu'il a été utilisé dans l'une de ces opérations, le champ s'affichera comme masqué et l'opération ne sera pas affectée.


Tous les champs masqués sont signalés par une icône d'œil .



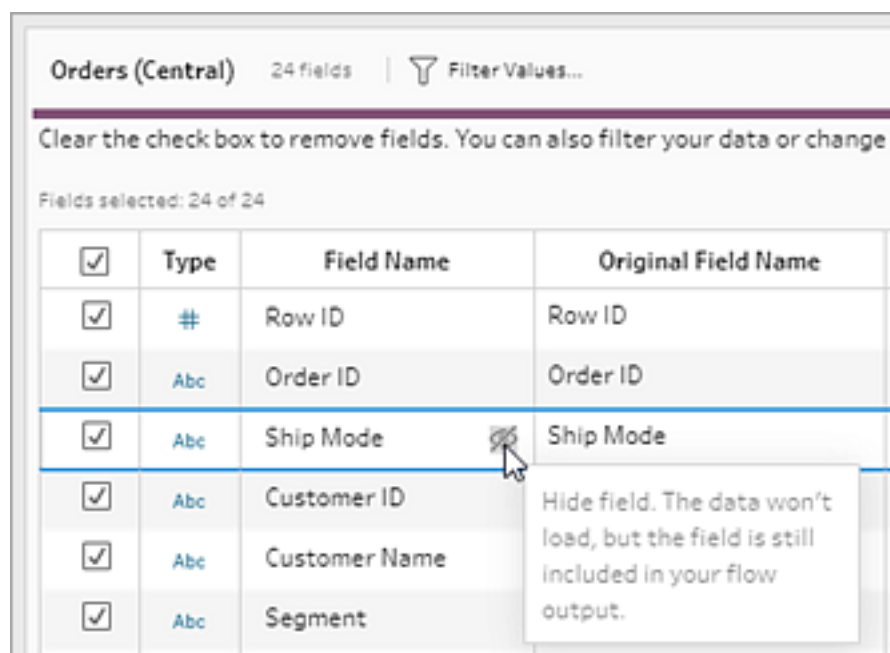
Masquer ou afficher des champs

Pour masquer ou afficher des champs, vous devez être dans une étape des données entrantes ou dans une étape de nettoyage. Dans l'étape de nettoyage, vous pouvez masquer ou afficher les champs du volet Profil, d'une grille de données et d'une liste.

Depuis l'étape des données entrantes

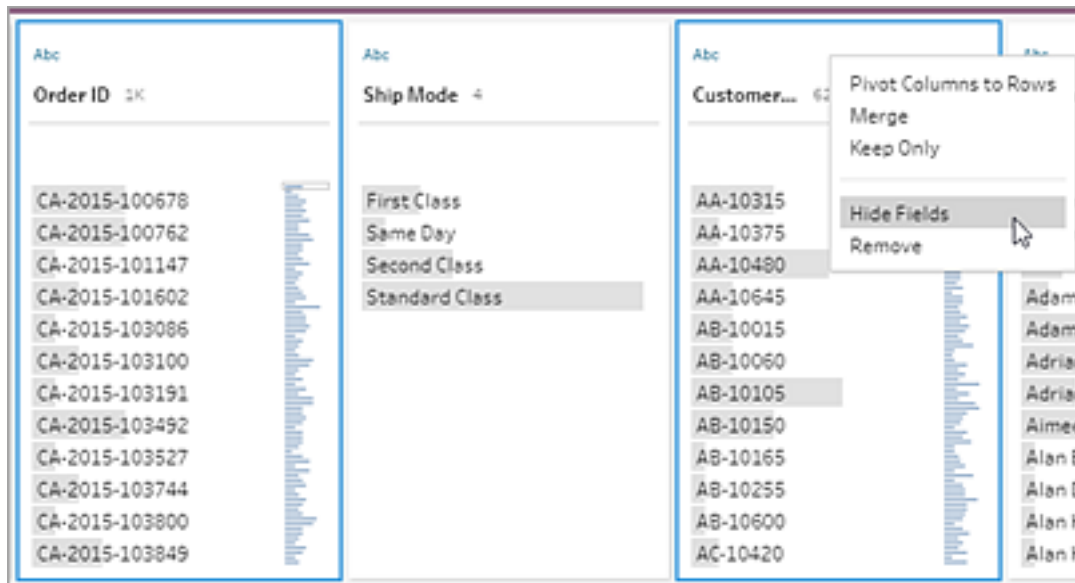
1. Connectez-vous à vos données.
2. Dans l'étape des données entrantes, sélectionnez le champ à masquer ou à afficher.
3. Cliquez sur l'icône d'œil  pour masquer ou afficher le champ.

La sélection de plusieurs champs dans l'étape des données entrantes n'est pas prise en charge actuellement.

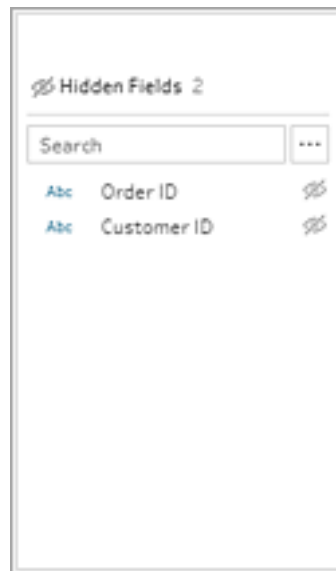



Depuis le volet Profil

1. Sélectionnez les champs à masquer.
2. Cliquez avec le bouton droit, Ctrl-clic (MacOS), dans le menu **Options supplémentaires** *** ou dans le menu de la barre d'outils, sélectionnez **Masquer le champ** ou **Masquer les champs**.





3. Une nouvelle fiche Profil est générée et affiche vos champs masqués.

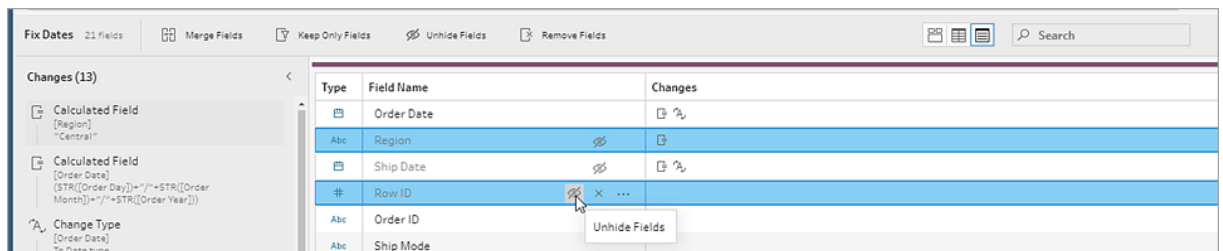


4. Pour afficher les champs, dans la fiche Profil **Champs masqués**, sélectionnez un ou plusieurs champs, puis cliquez sur l'icône d'œil , faites un clic droit ou Ctrl-clic (MacOS) et sélectionnez **Afficher les champs** dans le menu.

Depuis la liste

1. Dans une étape de nettoyage, dans la barre d'outils, cliquez sur l'icône Liste  pour passer à la liste.

- Sélectionnez un ou plusieurs champs à masquer ou à afficher.
- Cliquez sur l'icône d'œil  pour masquer ou afficher les champs.



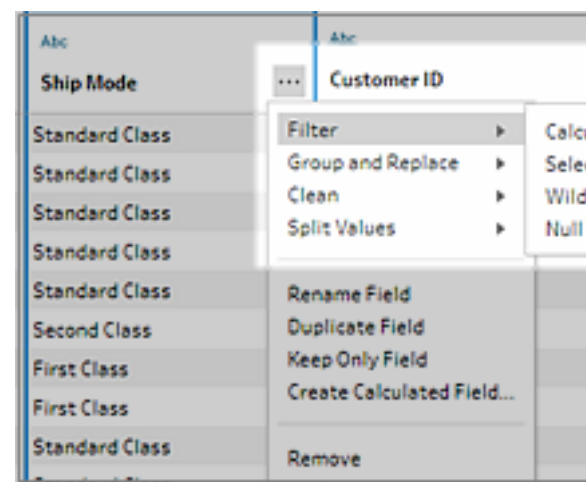
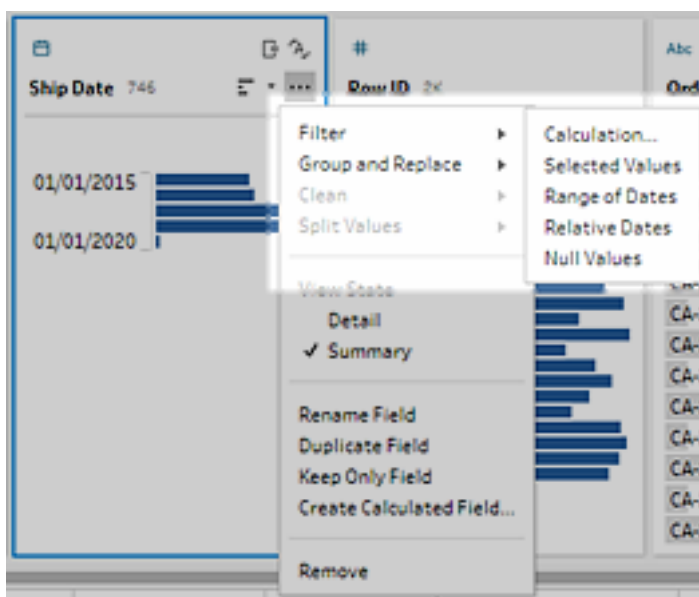
Filtres disponibles pour chaque type de données

Type de données	Filtres disponibles
Chaîne	Calcul, correspondance de caractères génériques, valeurs nulles, valeurs sélectionnées
Nombre	Calcul, plage de valeurs, valeurs null, valeurs sélectionnées
Date et Date et Heure	Calcul, plage de dates, date relative, valeurs nulles, valeurs sélectionnées

Où sont mes options de filtrage ?

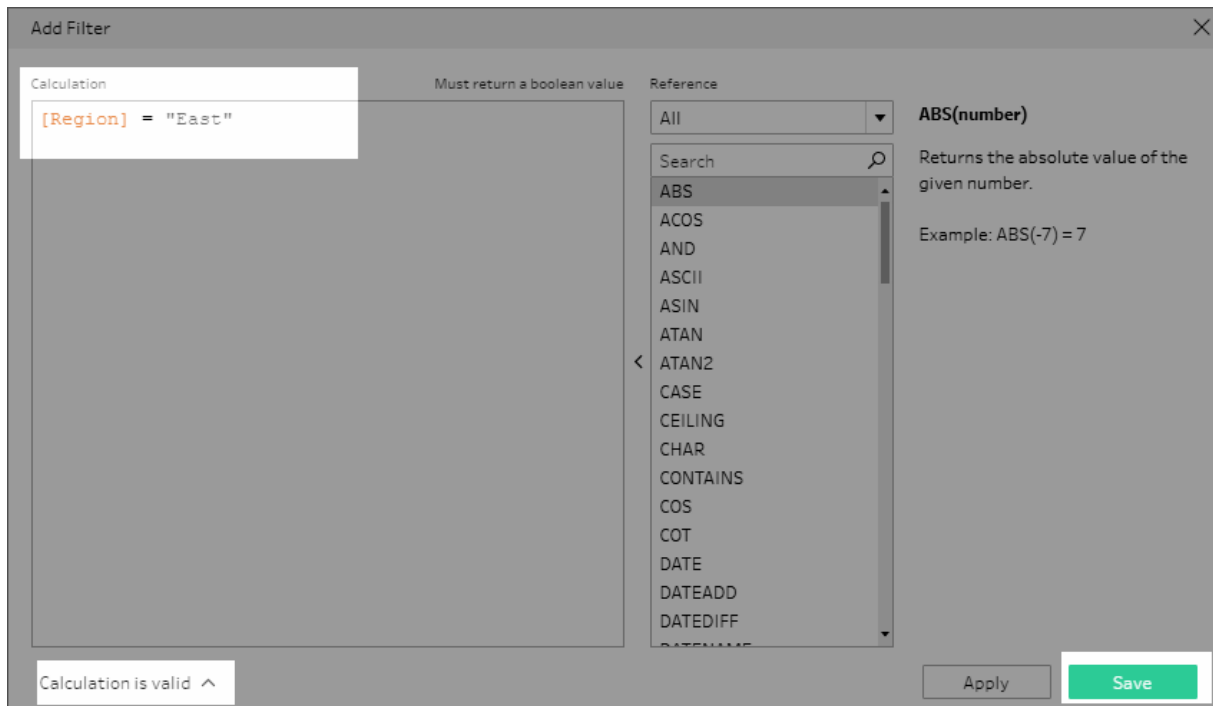
Pour voir les différentes options de filtrage disponibles pour vos champs, sur la fiche Profil, dans la grille de données ou dans le volet de résultats, cliquez sur le menu **Options supplémentaires** ***. Pour voir le menu sur la grille de données, vous devez d'abord cliquer sur le bouton **Masquer le volet**

Profil , puis sur **Options supplémentaires** ***.



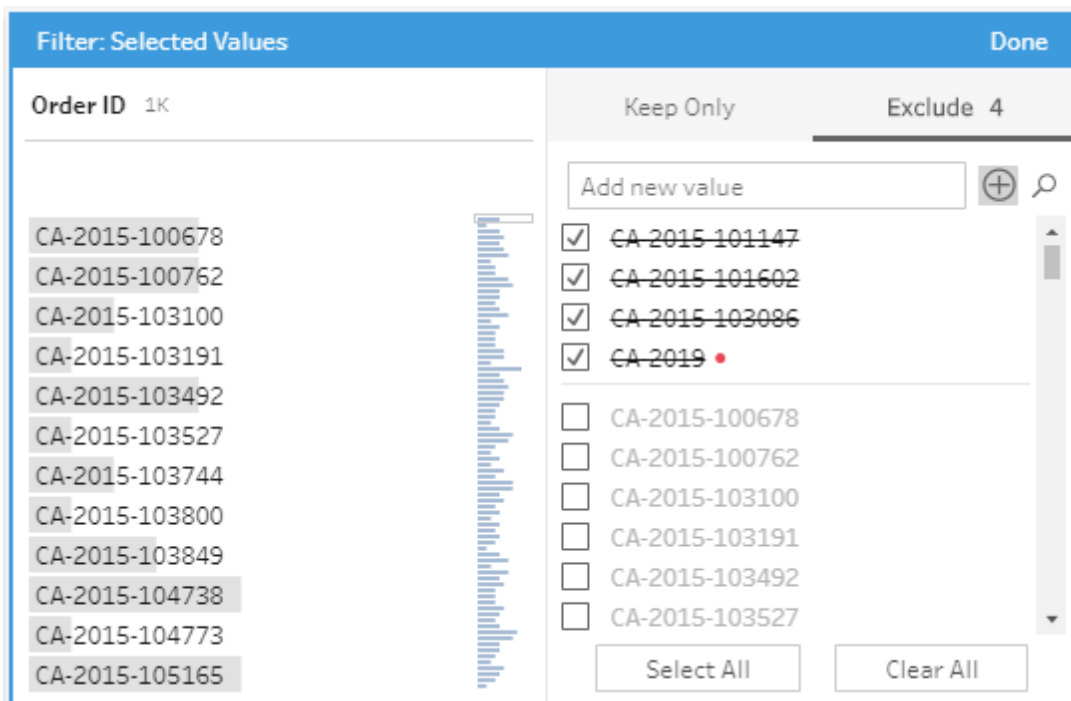
Filtre de calcul

Lorsque vous sélectionnez **Calcul**, la boîte de dialogue **Ajouter un filtre** s'ouvre. Entrez le calcul, vérifiez qu'il est valide et cliquez sur **Enregistrer**.



Filtre Valeurs sélectionnées

Dans Tableau Prep Builder vous pouvez utiliser le filtre **Valeurs sélectionnées** pour sélectionner les valeurs que vous voulez conserver ou exclure pour un champ.



Filtre Plage de valeurs

Filtrez les valeurs qui se situent à l'intérieur d'une plage spécifique. Lorsque vous sélectionnez **Plage de valeurs**, vous pouvez spécifier une plage ou définir des valeurs minimum ou maximum.

Filter: Range of Values

Done

Sales 2K

Range

Minimum

Maximum

0

4,000

8,000

12,000

0

12000

0

12,000

☐ Include null values

Filtre Plage de dates

Filtrez les valeurs qui se situent à l'intérieur d'une plage de dates spécifique. Lorsque vous sélectionnez **Plage de dates**, vous pouvez spécifier une plage de dates ou définir une date minimum ou maximum.

Filter: Range of Dates

Done

Order Date 787

Range

Minimum

Maximum

01/01/2014, 12:00...

01/01/2018, 12:00...

01/01/2014

01/01/2014

01/01/2018

☐ Include null values

Filtre Date relative

Utilisez le filtre **Dates relatives** pour spécifier l'intervalle exact d'années, de trimestres, de mois, de semaines ou de jours que vous voulez voir dans vos données. Vous pouvez également configurer un ancrage relatif à une date spécifique et inclure les valeurs null.

Filter: Relative Dates Done

Order Date 787

01/01/2014, 12:00...
01/01/2018, 12:00...

Years Quarters Months Weeks Days

Select a date range.

☐ Previous year

☐ This year

☐ Next year

☐ Last 3 years

☐ Next 3 years

☐ Year to date

☐ Anchor relative to Today

☐ Include null values

Filtre Correspondance avec des caractères génériques

Lorsque vous sélectionnez **Correspondance avec des caractères génériques**, vous pouvez filtrer les valeurs de champ de manière à conserver ou exclure les valeurs correspondant à une tendance. Dans l'éditeur de filtres, sélectionnez l'onglet **Conserver uniquement** ou **Exclure**, entrez une valeur de correspondance, puis définissez les critères **Options de correspondance** pour renvoyer les valeurs que vous recherchez.

Les résultats filtrés s'affichent dans le volet gauche de l'éditeur de filtre afin que vous puissiez les consulter et expérimenter avec les résultats. Une fois que vous avez obtenu les résultats souhaités, cliquez sur **Terminé** pour appliquer votre modification.

Filter: Wildcard Match

Done

Product Name 98

Keep Only

Exclude

binder

Matching Options

☒ Contains
 ☐ Starts with
 ☐ Ends with
 ☐ Exact match

24 Capacity Maxi Data Binder Rack...

Acco Clips to Go Binder Clips, 24 Cli...

Acco D-Ring Binder w/DubLock

Acco Data Flex Cable Posts For Top ...

Acco Economy Flexible Poly Round ...

Acco Expandable Hanging Binders

Acco Flexible ACCOHIDE Square Rin...

Acco Four Pocket Poly Ring Binder ...

Acco Hanging Data Binders

Acco PRESSTEX Data Binder with S...

Acco Recycled 2" Capacity Laser Pr...

Acco Suede Grain Vinyl Round Ring ...

Filtre Valeurs null

Lorsque vous sélectionnez **Valeurs null**, vous pouvez filtrer les valeurs du champ sélectionné de manière à afficher uniquement les valeurs null ou exclure toutes les valeurs null.

Filter: Null Values

Done

Order ID 183

Keep Only

Select values to keep.

☐ Null values
 ☐ Non-null values

Search

...

CA-2015-103492

CA-2015-104738

CA-2015-105165

CA-2015-105340

CA-2015-105417

CA-2015-108182

CA-2015-118339

CA-2015-119172

CA-2015-120278

CA-2015-123498

CA-2015-126193

CA-2015-126200

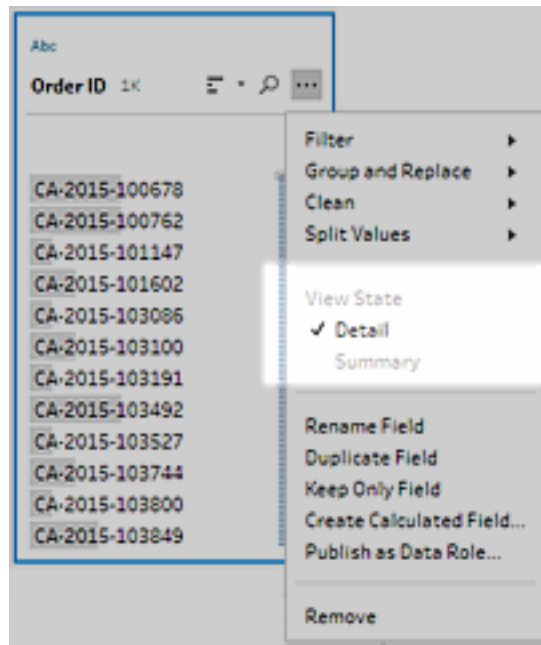


Lorsque vos données contiennent des champs numériques ou de date, vous pouvez basculer sur une version détaillée (discrète) des valeurs ou une version résumée (continue) des valeurs. La vue résumée affiche la plage des valeurs dans un champ et la fréquence avec laquelle certaines valeurs apparaissent.

Ce basculement peut vous aider à isoler des valeurs uniques (comme le nombre d'enregistrements « 3 » dans un champ) ou la distribution de valeurs (comme la somme de tous les enregistrements « 3 » dans un champ)

Pour changer votre vue :

1. Dans le volet Profil, le volet Résultats ou la grille de données, cliquez sur le menu **Options supplémentaires** *** pour un champ numérique ou de date.



2. Dans le menu contextuel, sélectionnez **Détails** pour voir une version détaillée des valeurs, ou **Résumé** pour voir la version distribuée des valeurs.

Exercice

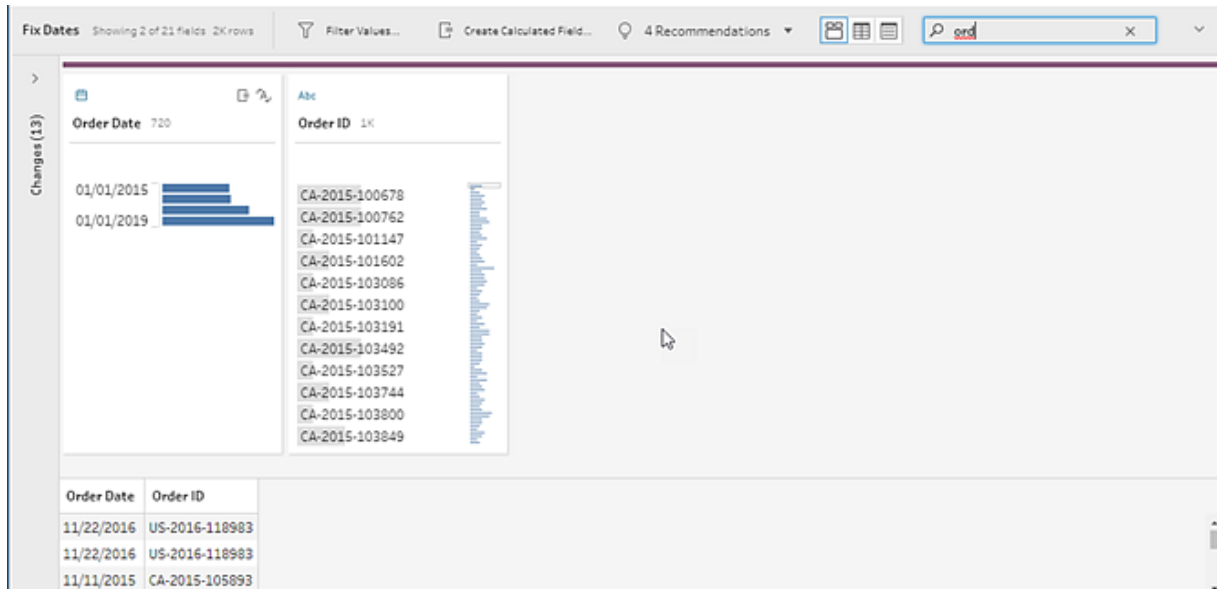
Toujours sur 'ORDER_USCA_2015'

- ➔ Exclure les valeurs null de ID commande
- ➔ Filtrer sur les dates de commande entre janvier et juin
- ➔ Exclure tous les clients dont le nom de famille se termine par 'y'
- ➔ Supprimer le champs 'nom du produit'
- ➔ Masquer le champs 'ID produit'
- ➔ Supprimer toutes ces modifications et dé-masquer 'ID produit'


Rechercher des champs et des valeurs

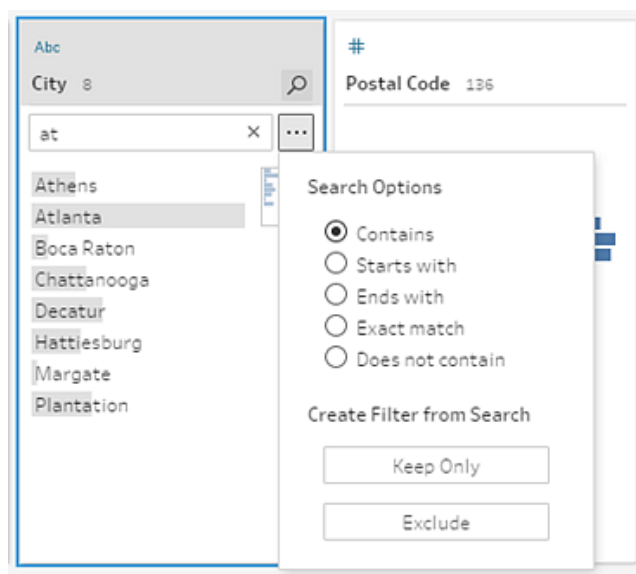
Dans le volet Profil ou le volet Résultats, vous pouvez rechercher des champs ou des valeurs qui vous intéressent particulièrement, et utiliser les résultats de la recherche pour filtrer vos données.

Pour rechercher des champs, entrez un terme de recherche complet ou partiel dans la zone de recherche de la barre d'outils.

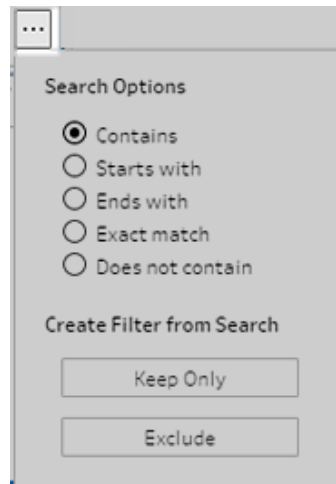


Pour rechercher une valeur dans un champ :

1. Cliquez sur l'icône Rechercher  pour rechercher un champ et entrez une valeur.



2. Pour utiliser des options de recherche avancées, cliquez sur le bouton **Options de recherche...**

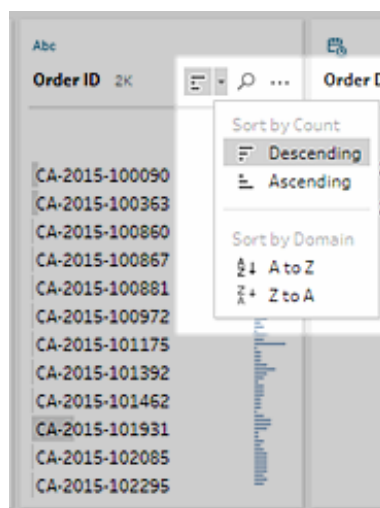


3. Pour utiliser les résultats de la recherche pour filtrer les données, sélectionnez **Conserver uniquement** ou **Exclure**.

Dans le volet **Flux**, une icône de filtre apparaît au-dessus des étapes concernées.

Trier les valeurs et les champs

Vous pouvez utiliser les options de tri d'une fiche Profil pour trier les classes (le nombre de valeurs représentées par les barres de distribution) dans l'ordre croissant ou décroissant ou les valeurs de champs individuelles dans l'ordre alphabétique.



Pour réorganiser l'ordre de vos champs, dans le volet **Profil**, le volet **Résultats** ou la grille de **Données**, il vous suffit de sélectionner une fiche Profil ou un champ dans la grille de données et de le faire glisser jusqu'à ce qu'une ligne noire cible apparaisse. Ensuite, déposez-le sur place. Le volet Profil et la grille de données sont synchronisés afin que le champ apparaisse dans le même ordre aux deux emplacements.

au_id	au_name	phone	address	city	state	zip	contract	pub_id	pub_name	city	state
409-95-7008	Bennet, Abraham	415 650-9932	6223 Baranman St.	Berkeley	CA	94705	true	1309	Algodata Infosystems	Berkeley	CA
239-95-7766	Carson, Cheryl	415 545-7723	589 Devon Ln.	Berkeley	CA	94705	true	1309	Algodata Infosystems	Berkeley	CA

Surligner les champs et les valeurs dans un flux

Tableau Prep vous aide à trouver facilement des champs et des valeurs dans vos données de flux. Dans le volet Flux, tracez le point d'origine d'un champ et là où il est utilisé dans le flux, ou cliquez sur des valeurs individuelles dans une fiche Profil ou dans la grille des données pour surligner les valeurs associées ou identiques.

Tracer des champs dans un flux

Dans Tableau Prep, vous pouvez surligner tout point où un champ est utilisé dans un flux, et même son emplacement d'origine pour vous aider à repérer les valeurs manquantes ou corriger un flux lorsque vous ne voyez pas les résultats attendus.

Cliquez sur un champ dans le volet **Profil** dans une étape de nettoyage, ou dans le volet **Résultats** dans un autre type d'étape. Le volet Flux met en évidence le chemin où ce champ est utilisé.

Clean Notes/Approver	All Orders
Product ID	Product ID
Order ID	Order ID

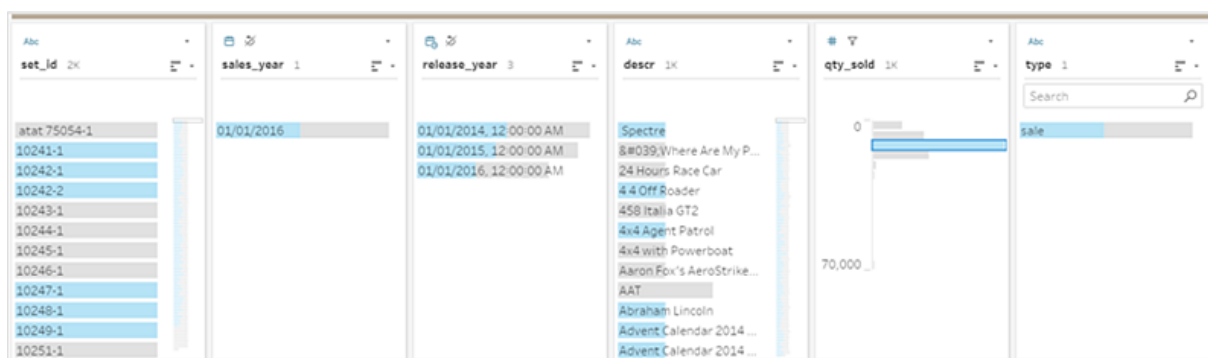
Year of Sale	Returned?	Days to Ship
2015	Yes	0
2016	Yes	1
2017	Yes	2
2018	Yes	3
2019	Yes	4
2020	Yes	5
2021	Yes	6
2022	Yes	7

Consulter les valeurs associées

Vous pouvez utiliser le surlignage pour trouver des valeurs associées à travers les champs. Lorsque vous cliquez sur une valeur dans la fiche **Profil** dans le volet Profil ou le volet Résultats, toutes les

valeurs associées dans les autres champs sont surlignées en bleu. La couleur bleu montre la distribution de la relation entre la valeur que vous avez sélectionnée et les valeurs dans les autres champs.

Par exemple, pour surligner les valeurs associées, dans le volet **Profil**, cliquez sur une valeur dans un champ. Les valeurs associées dans un autre champ deviennent bleues et la proportion de la barre surlignée en bleu représente le degré d'association.



Surligner les valeurs identiques

Lorsque vous sélectionnez une valeur dans la grille de données, toutes les valeurs identiques sont également surlignées. Ces surlignages vous aident à identifier les tendances ou les irrégularités dans vos données.

Type	Customer	Purchases	Date
Cash	Wei	5	08/18/2016
Cash	Jim	7	07/15/2016
Credit	Arnold	5	06/29/2016
Credit	Lee	1	08/07/2016
Cash	Maria	2	08/30/2016
Cash	Wendy	1	07/21/2016
Credit	Max	2	07/02/2016
Credit	Juan	1	05/10/2016
Cash	Isaac	4	06/28/2016
Credit	Philip	1	08/09/2016
Credit	Lane	5	05/04/2016

Nettoyer et organiser les données

Tableau Prep fournit diverses opérations de nettoyage que vous pouvez utiliser directement pour nettoyer et organiser vos données. Le nettoyage des données désordonnées facilite la combinaison et l'analyse de vos données ou permet à d'autres personnes de comprendre plus facilement vos données lorsque vous partagez vos ensembles de données.

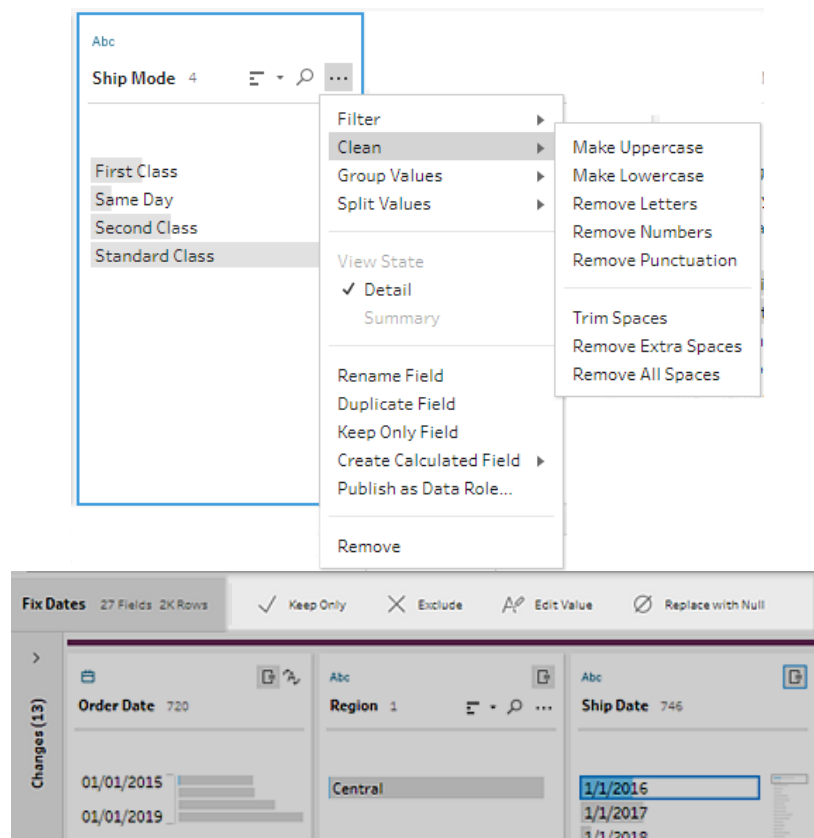
À propos des opérations de nettoyage

Vous nettoyez les données en appliquant des opérations de nettoyage telles que le filtrage, l'ajout, le changement de nom, la scission, le regroupement ou la suppression de champs. Vous pouvez effectuer des opérations de nettoyage dans la plupart des étapes de votre flux. Vous pouvez également effectuer des opérations de nettoyage dans la grille de données lors d'une étape de nettoyage.

Appliquer des opérations de nettoyage

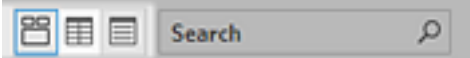
Pour appliquer des opérations de nettoyage à des champs, utilisez les options de la barre d'outils ou cliquez sur **Options supplémentaires** *** dans la fiche Profil du champ, la grille de données ou le volet Résultats pour ouvrir le menu.


Dans les types d'étape Agrégation, Permutation, Jointure et Union, le menu **Options supplémentaires** *** est disponible dans les fiches Profil du volet Résultats et la grille de données correspondante. Si vous effectuez les mêmes opérations de nettoyage ou actions à plusieurs reprises tout au long de votre flux, vous pouvez copier et coller vos étapes, actions ou même champs.

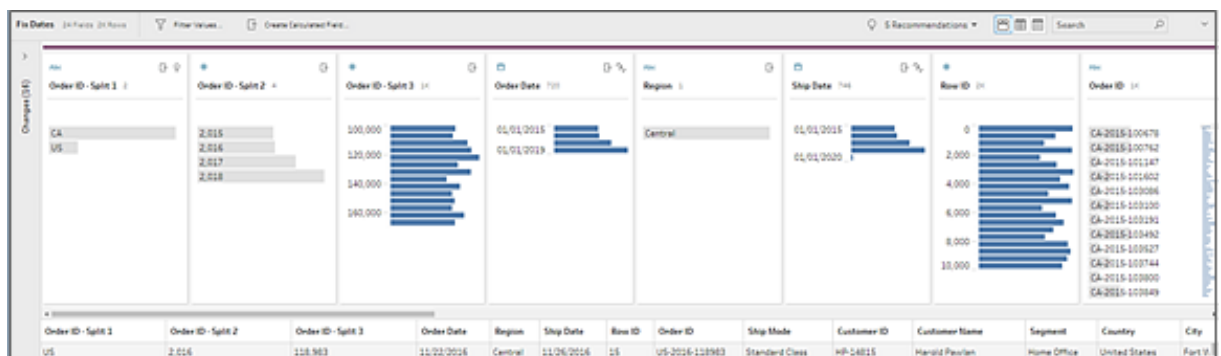



Sélectionner votre vue

Vous pouvez effectuer des opérations de nettoyage en dehors des volets Profil ou Résultats dans la grille de données ou dans la liste. Utilisez la barre d'outils de la


vue  (Tableau Prep Builder à partir de la version 2019.3.2 et sur le Web) pour modifier votre vue, puis cliquez sur **Options supplémentaires** *** sur un champ pour ouvrir le menu de nettoyage.

- **Afficher le volet Profil**  : il s'agit de la vue par défaut. Sélectionnez ce bouton pour revenir à la vue du volet Profil ou du volet Résultats.



- **Afficher la grille de données**  : Réduisez le volet Profil ou Résultats pour développer et afficher uniquement la grille de données. Cette vue fournit une vue plus détaillée de vos données et peut être utile lorsque vous devez travailler avec des valeurs de champ spécifiques. Une fois que vous avez sélectionné cette option, cet état de vue persiste à travers toutes les étapes de votre flux, mais vous pouvez le modifier à tout moment.

Order ID - Split 1	Order ID - Split 2	Order ID - Split 3	Order Date	Region	Ship Date	Row ID	Order ID	Ship Mode
US	2.016	118.963	11/22/2016	Central	11/26/2016	15	US-2016-118963	Standard Class
US	2.016	118.963	11/22/2016	Central	11/26/2016	16	US-2016-118963	Standard Class
CA	2.015	105.893	11/11/2016	Central	11/26/2016	17	CA-2016-105893	Standard Class
CA	2.017	137.330	12/09/2017	Central	12/13/2017	22	CA-2017-137330	Standard Class
CA	2.017	137.330	12/09/2017	Central	12/13/2017	23	CA-2017-137330	Standard Class
CA	2.018	107.727	10/19/2018	Central	10/23/2018	35	CA-2018-107727	Second Class
CA	2.017	117.590	12/08/2017	Central	12/10/2017	36	CA-2017-117590	First Class
CA	2.017	117.590	12/08/2017	Central	12/10/2017	37	CA-2017-117590	First Class
CA	2.016	117.415	12/27/2016	Central	12/31/2016	38	CA-2016-117415	Standard Class
CA	2.016	117.415	12/27/2016	Central	12/31/2016	39	CA-2016-117415	Standard Class
CA	2.016	117.415	12/27/2016	Central	12/31/2016	40	CA-2016-117415	Standard Class
CA	2.016	117.415	12/27/2016	Central	12/31/2016	41	CA-2016-117415	Standard Class
CA	2.018	120.999	09/15/2018	Central	09/15/2018	42	CA-2018-120999	Standard Class
CA	2.017	118.255	03/11/2017	Central	03/13/2017	45	CA-2017-118255	First Class
CA	2.017	118.255	03/11/2017	Central	03/13/2017	46	CA-2017-118255	First Class

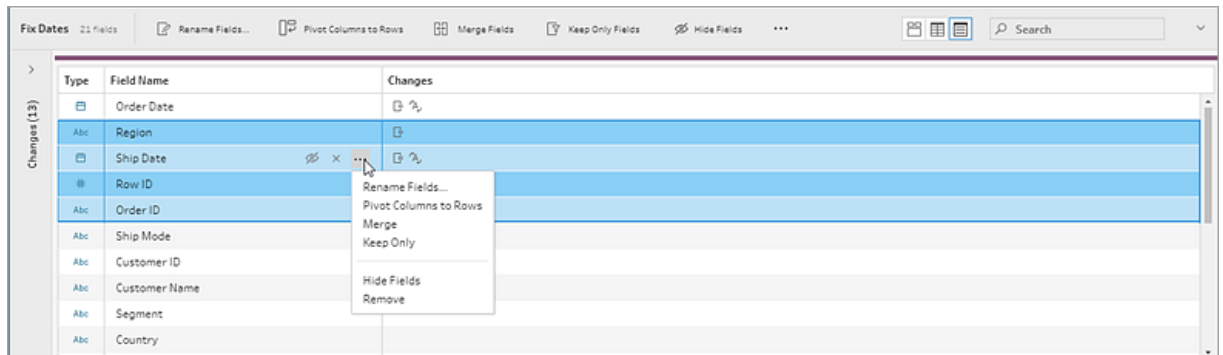
- **Afficher la liste**  (Tableau Prep Builder à partir de la version 2019.3.2 et sur le Web) : Convertissez le volet Profil ou le volet Résultats en une liste. Une fois que vous avez sélectionné cette option, cet état de vue persiste à travers toutes les étapes de votre flux, mais vous pouvez le modifier à tout moment.

Dans cette vue, vous pouvez :

- Sélectionner et supprimer plusieurs lignes à l'aide de l'option X.

- Utilisez le menu **Options supplémentaires** *** pour appliquer des opérations aux champs sélectionnés.

Si vous affectez un type de données au champ ou sélectionnez **Filtrer**, **Regrouper les valeurs**, **Nettoyer** ou **Scinder les valeurs**, vous êtes renvoyé à la vue Profil ou Résultats pour terminer ces actions. Toutes les autres options peuvent être exécutées dans la liste.



Appliquer des opérations de nettoyage

Pour appliquer des opérations de nettoyage à un champ, procédez comme suit :

1. Dans le volet **Profil**, dans la grille de données, dans le volet Résultats ou dans la liste, sélectionnez le champ auquel vous souhaitez apporter des modifications.
2. Dans la barre d'outils ou le menu **Options supplémentaires** *** du champ, faites votre choix parmi les options suivantes :
 - **Filtrer** ou **Filtrer les valeurs** : sélectionnez l'une des options de filtre, faites un clic droit (ou Ctrl+clic sur MacOS) sur une valeur de champ pour conserver ou exclure des valeurs. Vous pouvez également utiliser le filtre **Valeurs sélectionnées** pour choisir et sélectionner les valeurs à filtrer, y compris les valeurs non comprises dans votre échantillon de flux.
 - **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) : sélectionnez manuellement les valeurs ou utilisez le regroupement automatique. Vous pouvez aussi sélectionner plusieurs valeurs dans la fiche Profil et faire un clic droit (ou Ctrl+clic sur MacOS) pour regrouper ou dégroupier des valeurs ou encore modifier la valeur de regroupement.
 - **Nettoyer** : faites votre choix dans la liste d'opérations de nettoyage rapides à appliquer à toutes les valeurs du champ.
 - **Scinder les valeurs** : Scindez les valeurs automatiquement en fonction d'un séparateur commun ou utilisez la scission personnalisée pour spécifier la façon dont vous souhaitez scinder les valeurs de champ.
 - **Renommer le champ** : modifiez le nom du champ.

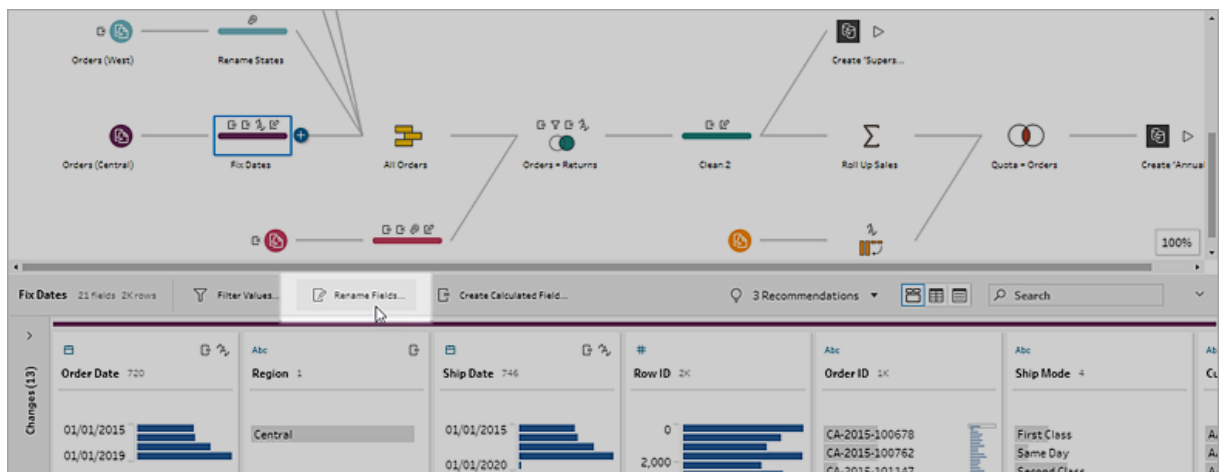
- **Créer un champ calculé** : écrivez un calcul personnalisé dans l'éditeur de calcul ou utilisez l'éditeur de calcul visuel pour créer des calculs de niveau de détail (LOD), de classement ou de nombre de lignes.
3. Pour modifier une valeur, faites un clic droit ou Ctrl+clic (MacOS) sur une ou plusieurs valeurs, et sélectionnez **Modifier la valeur** puis entrez une nouvelle valeur. Vous pouvez également sélectionner **Remplacer par une valeur null** pour remplacer les valeurs par une valeur null ou double-cliquer sur un champ pour le modifier directement.
 4. Vérifiez le résultat de ces opérations dans le volet **Profil**, les volets **Résumé** ou la grille de données.

Renommer les champs en bloc

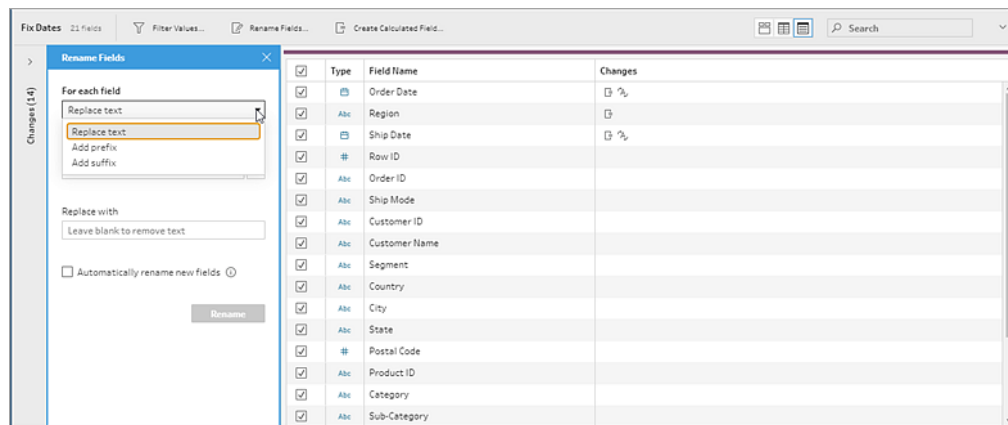
Utilisez l'option **Renommer les champs** pour renommer plusieurs champs en bloc. Recherchez des parties d'un nom de champ pour le remplacer ou le supprimer, ou ajoutez des préfixes ou des suffixes à tous les champs ou à certains champs de votre ensemble de données.

Vous pouvez également appliquer automatiquement la même modification à tous les champs ajoutés par la suite qui correspondent à vos critères en cochant la case **Renommer automatiquement les nouveaux champs** lors de vos modifications.

1. Dans une étape de **Nettoyage**, dans la barre d'outils, sélectionnez **Renommer les champs**.



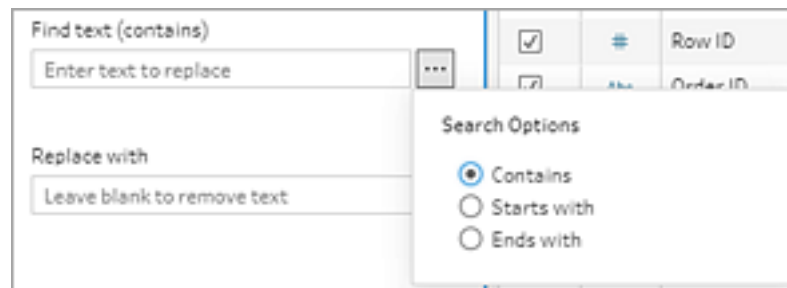
Votre vue est automatiquement convertie en **Liste** et affiche tous les champs de votre flux. Vous pouvez utiliser l'option **Rechercher** dans la barre d'outils pour affiner vos résultats.



Tous les champs sont sélectionnés par défaut. Désélectionnez la case supérieure pour effacer la sélection de tous les champs et sélectionner manuellement uniquement les champs que vous souhaitez modifier.

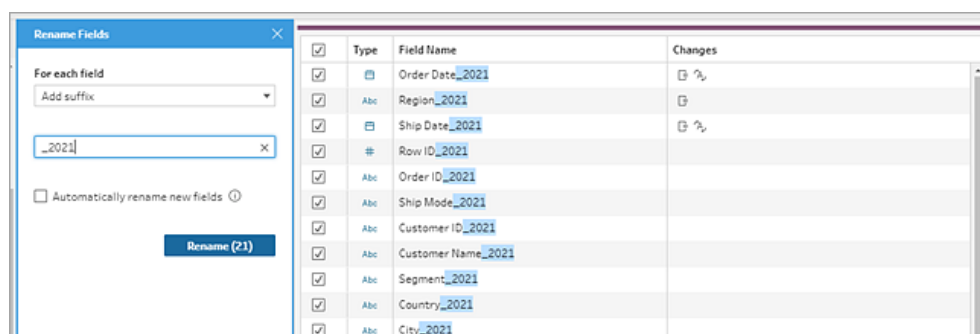
2. Dans le volet **Renommer les champs**, vous avez le choix entre les options suivantes :

- **Remplacer le texte** : dans le champ **Trouver le texte**, recherchez le texte correspondant à l'aide des **Options de recherche**, puis saisissez le texte de remplacement dans le champ **Remplacer par**. Pour rechercher des espaces vides, appuyez sur la barre d'espace dans le champ **Trouver le texte**.

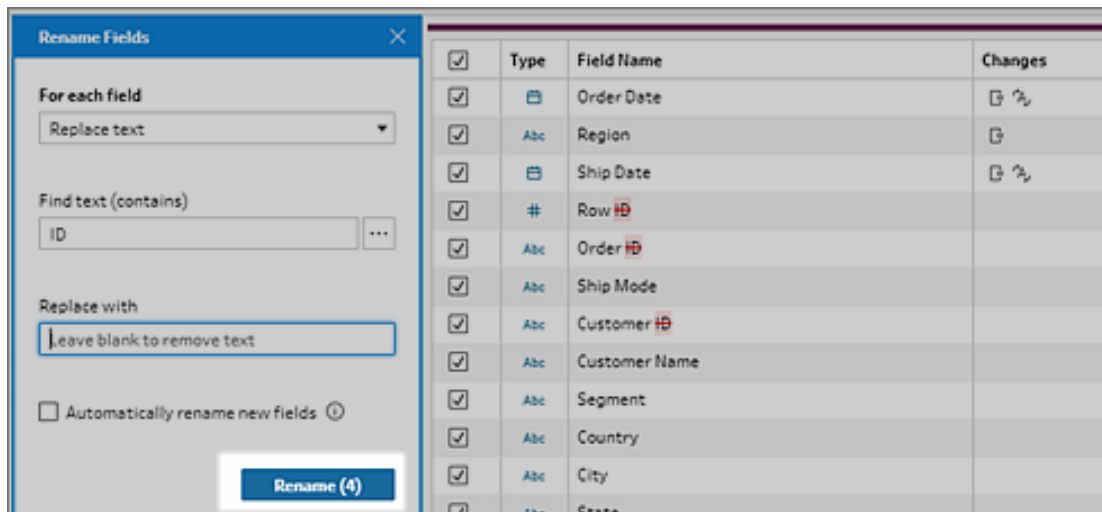


- **Ajouter un préfixe** : ajoutez un texte au début de tous les noms de champs sélectionnés.
- **Ajouter un suffixe** : ajoutez un texte à la fin de tous les noms de champs sélectionnés.

Au fur et à mesure que vous effectuez vos entrées, vos résultats s'affichent dans le volet **Liste**.



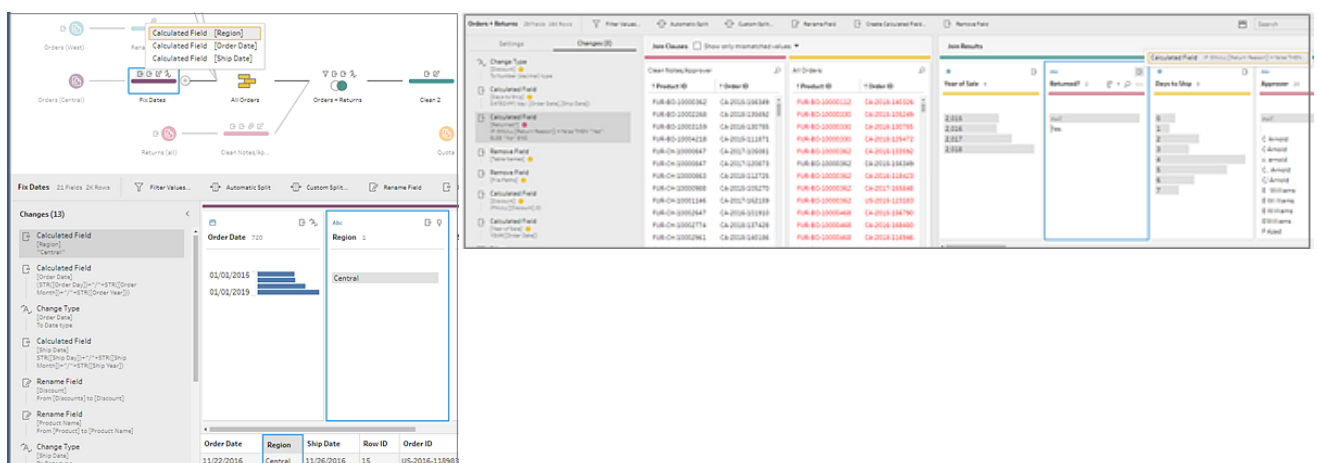
- (facultatif) Sélectionnez **Renommer automatiquement les nouveaux champs** pour appliquer automatiquement ces mêmes modifications aux nouveaux champs qui correspondent à vos critères de remplacement lorsque vos données sont actualisées.
- Cliquez sur **Renommer** pour Appliquer les modifications et fermer le volet. Le bouton **Renommer** affiche le nombre de champs affectés par vos modifications.



Afficher vos modifications

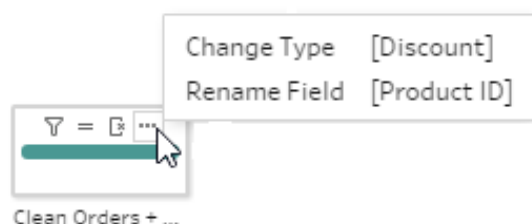
Les différents types d'opérations de nettoyage sont représentés par des icônes à travers les différentes étapes de votre flux. Si plus de quatre types d'opérations sont appliqués à une étape, trois points de suspension apparaissent sur l'étape. Survolez ces icônes pour voir les annotations indiquant les opérations appliquées, ainsi que leur ordre d'exécution.

Dans Tableau Prep Builder vous pouvez cliquer sur une annotation sur l'icône de modification dans une étape du volet Flux ou sur une fiche Profil dans le volet Profil ou Résultats. La modification et le champ affecté seront surlignés dans le volet **Modifications** ainsi que dans le volet **Profil** ou **Résultats**.

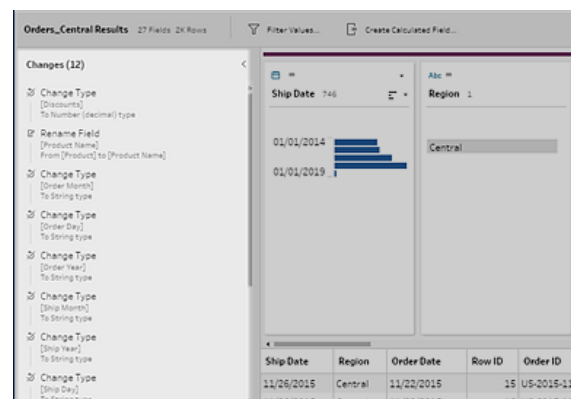


Vous pouvez également sélectionner une étape, puis développer le volet **Modifications** pour consulter les détails de chaque modification, éditer ou supprimer vos modifications, ou faire glisser des modifications vers le haut ou vers le bas pour changer leur ordre d'application. Vous pouvez aussi ajouter une description pour fournir un contexte aux autres utilisateurs.

Annotation sur le nettoyage



Suivi des modifications



Fusionner des champs

Si des champs contiennent les mêmes valeurs mais ont un nom différent, vous pouvez facilement les fusionner en un seul champ pour les combiner en faisant glisser un champ sur un autre. Lorsque vous fusionnez des champs, le champ cible devient le champ principal et le nom du champ cible persiste. Le champ que vous fusionnez sur le champ cible est supprimé.

Lorsque vous fusionnez des champs, Tableau Prep conserve tous les champs du champ cible et remplace toutes les valeurs null de ce champ par les valeurs issues des champs source que vous fusionnez avec le champ cible. Les champs source sont supprimés.

Exemple

Nom	Téléphone_Contact	Téléphone_Professionnel	Téléphone_Mobile	Téléphone_Domicile
Bob	123-4567	123-4567	null	null
Sally	null	null	456-7890	789-0123
Fred	null	null	null	567-8901
Emma	null	234-5678	345-6789	null

Si vous fusionnez **Téléphone_Professionnel**, **Téléphone_Mobile** et **Téléphone_Domicile** avec le champ **Téléphone_Contact**, les autres champs sont supprimés et le résultat se présente ainsi :

Nom	Téléphone_Contact
Bob	123-4567
Sally	456-7890
Fred	567-8901
Emma	234-5678

Pour fusionner des champs, procédez comme suit :


- Faites glisser et déposez un champ sur un autre. Un indicateur **Déposer pour fusionner les champs** s'affiche.
- Sélectionnez plusieurs champs et faites un clic droit dans la sélection pour ouvrir le menu contextuel, puis cliquez sur **Fusionner les champs**.
- Sélectionnez plusieurs champs, puis cliquez sur **Fusionner les champs** dans la barre d'outils.

Appliquer des opérations de nettoyage à l'aide de recommandations

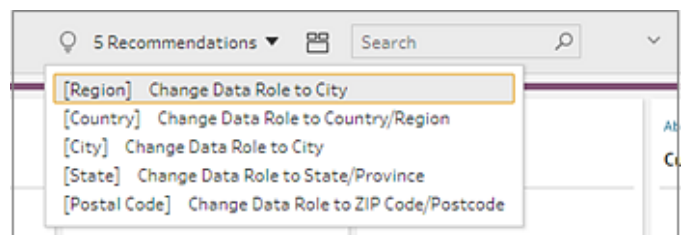
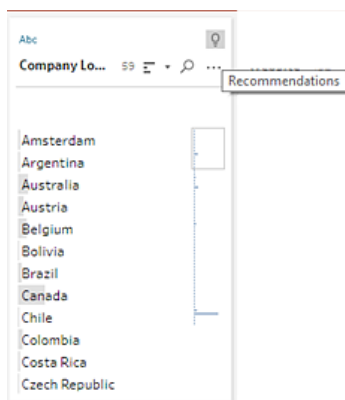
Il peut parfois être difficile d'identifier l'opération de nettoyage nécessaire pour corriger les problèmes de vos données. Tableau Prep peut analyser vos données et recommander les opérations de nettoyage à appliquer automatiquement pour corriger rapidement les problèmes de vos champs de données ou identifier les problèmes afin que vous puissiez les corriger. Cette fonctionnalité est disponible dans tous les types d'étape, sauf les types d'étape Entrée, Sortie et Jointure.

Appliquer les recommandations

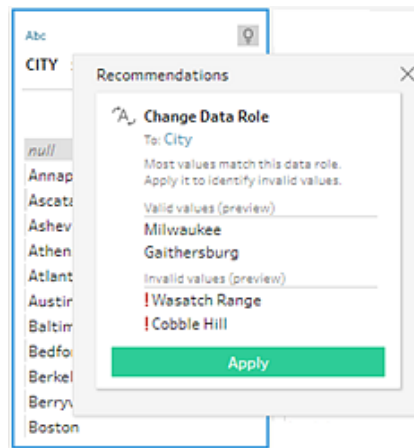
1. Effectuez l'une des actions suivantes :

- Cliquez sur l'icône d'ampoule  en haut à droite de la fiche Profil.
- Dans la barre d'outils, cliquez sur la flèche déroulante **Recommandations** pour afficher toutes les recommandations concernant votre ensemble de données et sélectionnez une recommandation dans la liste.

Cette option apparaît uniquement lorsque les modifications recommandées sont identifiées par Tableau Prep.



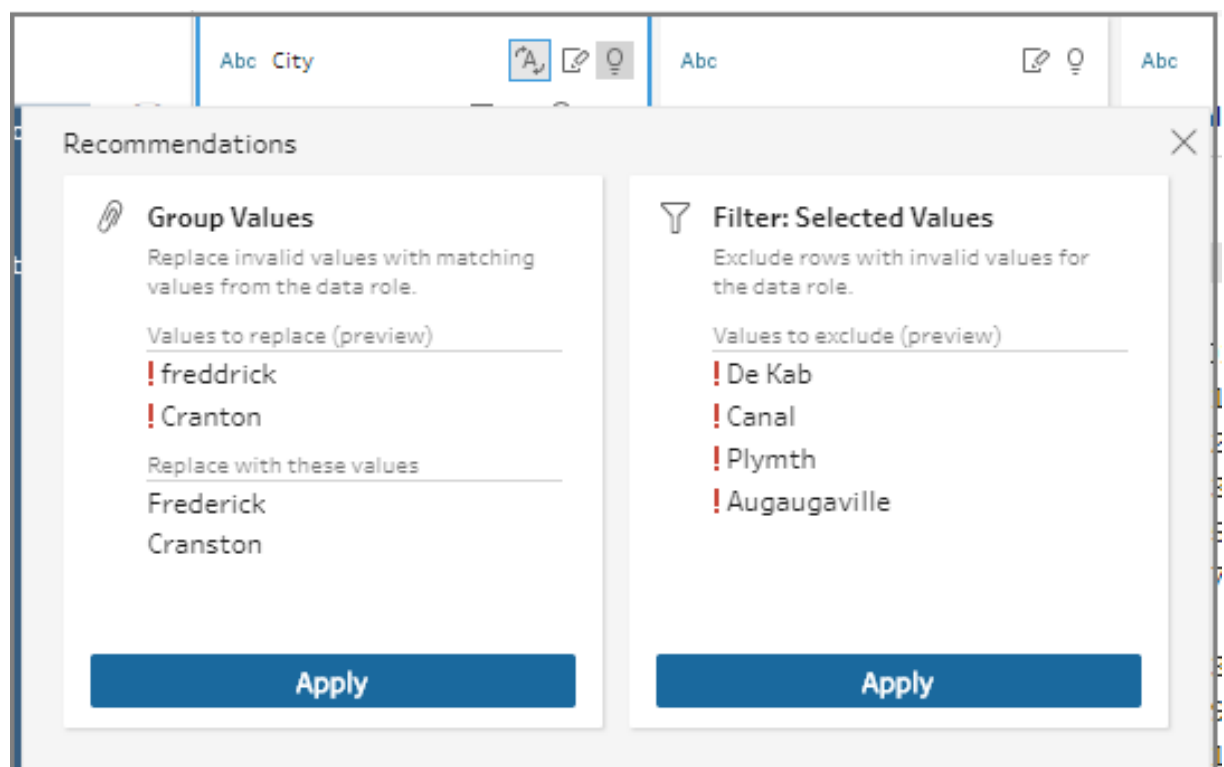
2. Pour appliquer la recommandation, survolez la fiche Recommendations puis cliquez sur **Appliquer**.



La modification est automatiquement appliquée et une entrée est ajoutée au volet **Modifications**. Pour supprimer la modification, cliquez sur **Annuler** dans le menu supérieur ou survolez la modification dans le volet **Modifications** puis cliquez sur le signe **X** pour la supprimer.

3. Si Tableau Prep identifie d'autres recommandations suite à la modification, l'icône d'ampoule reste sur la fiche Profil jusqu'à ce qu'il n'y ait plus d'autres recommandations.


Répétez les étapes ci-dessus pour appliquer des modifications supplémentaires ou ignorez les modifications suggérées et utilisez les autres outils de nettoyage pour résoudre les problèmes de données.



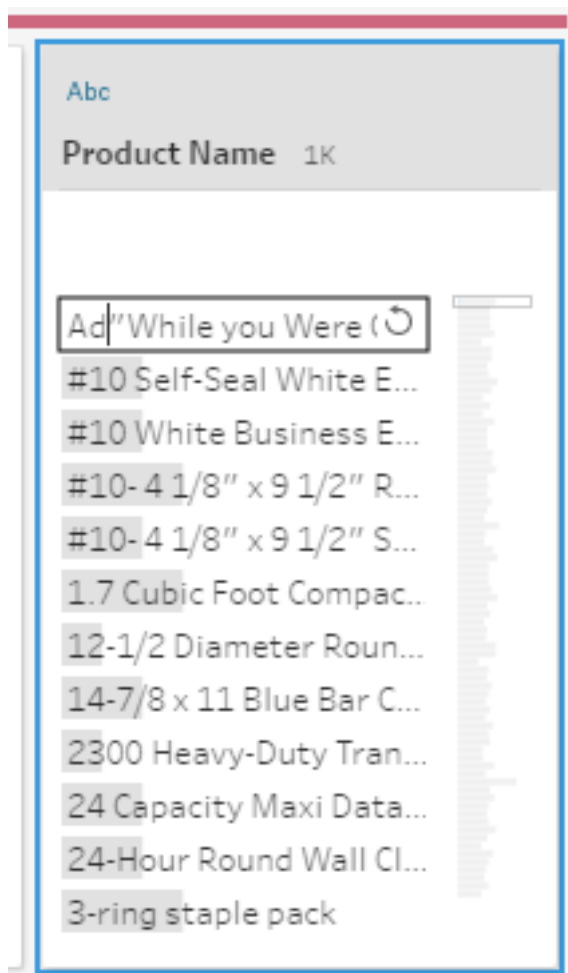
Modifier des valeurs de champ

Plusieurs variations de la même valeur peuvent vous empêcher de résumer correctement vos données. Vous pouvez corriger ces variations rapidement et aisément en appliquant les options suivantes.

Modifier une seule valeur

1. Dans la fiche **Profil**, cliquez sur la valeur à modifier puis entrez la nouvelle valeur. Une icône de groupe  s'affiche à côté de la valeur.

Sinon, faites un clic droit sur une valeur et cliquez sur **Modifier la valeur**. La modification est enregistrée dans **Suivi des modifications** sur le côté gauche de l'écran.



2. Affichez les résultats dans le volet **Profil** et la grille de données.

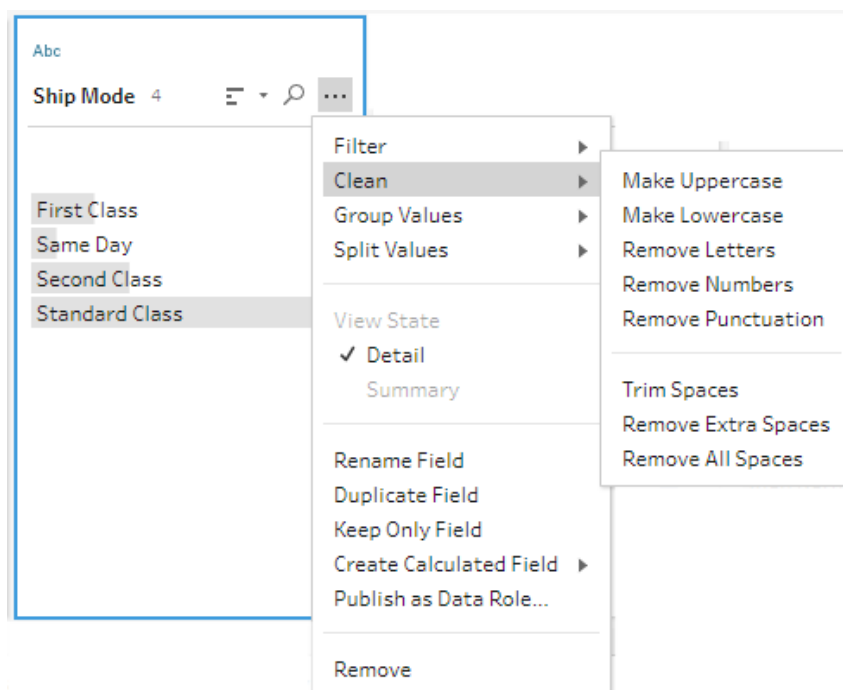
Modifier plusieurs valeurs

Vous avez plusieurs possibilités pour modifier plusieurs valeurs à la fois. Par exemple, utilisez des opérations de nettoyage rapide pour supprimer la ponctuation de toutes les valeurs du champ, regrouper manuellement des valeurs en utilisant la sélection multiple ou regrouper automatiquement ensemble des valeurs à l'aide d'algorithmes de correspondance approximative qui trouvent les valeurs similaires, ou encore sélectionner plusieurs valeurs et les remplacer par des valeurs null.

Modifier plusieurs valeurs en utilisant les opérations de nettoyage rapides

Cette option s'applique uniquement aux champs textuels.

1. Dans le volet **Profil**, le volet Résultats ou la grille de données, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** **...**, sélectionnez **Nettoyer**, puis sélectionnez l'une des options suivantes:
 - **Mettre en majuscule** : modifiez toutes les valeurs en texte en majuscule.
 - **Mettre en minuscule** : modifiez toutes les valeurs en texte en minuscule.
 - **Supprimer les lettres** : supprimez toutes les lettres et laissez uniquement les autres caractères.
 - **Supprimer les nombres** : supprimez tous les nombres et laissez les lettres et les autres caractères.
 - **Supprimer la ponctuation** : supprimez toutes la ponctuation.
 - **Supprimer les espaces** : supprimez les espaces de début et de fin.
 - **Supprimer les espaces supplémentaires** : supprimez les espaces blancs de début et de fin, et remplacez les espaces blancs supplémentaires entre des caractères par un seul espace.
 - **Supprimer tous les espaces** : supprimez tous les espaces blancs, y compris les espaces blancs de début et de fin, et tout espace blanc entre les caractères.



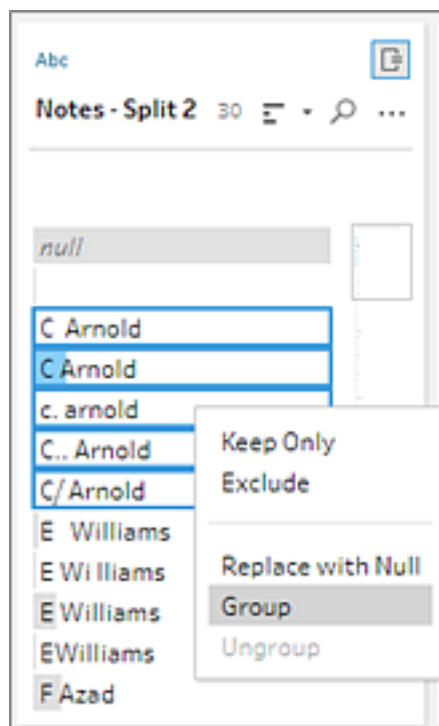
Vous pouvez empiler les opérations pour appliquer plusieurs opérations de nettoyage aux champs. Par exemple, commencez par sélectionner **Supprimer** > **Supprimer les nombres** puis sélectionnez **Nettoyer** > **Supprimer la ponctuation** pour supprimer tous les nombres et la ponctuation des valeurs de champ.

3. Pour annuler vos modifications, cliquez sur la flèche **Annuler** en haut du volet **Flux** ou supprimez la modification de la liste des modifications.

Regrouper et modifier plusieurs valeurs en ligne

Utilisez cette option pour sélectionner manuellement plusieurs valeurs et les regrouper sous une valeur standard dans la fiche Profil.


1. Dans la fiche **Profil**, sélectionnez le champ à modifier.
2. Appuyez sur Ctrl ou Maj+clic, ou Command ou Maj+clic (MacOS), et sélectionnez les valeurs que vous souhaitez regrouper.
3. Faites un clic droit et sélectionnez **Groupe** dans le menu contextuel. La valeur de la sélection sur laquelle vous faites un clic droit devient le nom par défaut du nouveau groupe, mais vous pouvez le modifier en ligne.

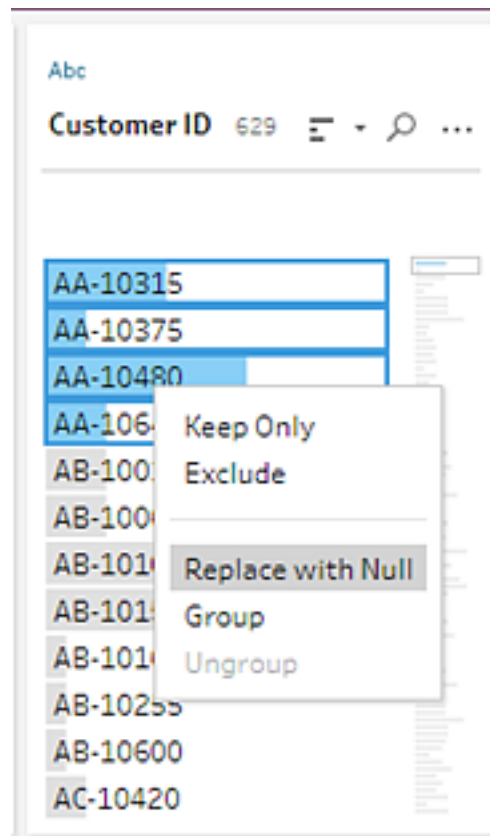


4. Pour modifier le nom du groupe, sélectionnez le champ groupé et modifiez la valeur ou faites un clic droit (ou Ctrl+clic sur un Mac) sur le champ groupé et sélectionnez **Modifier la valeur** dans le menu contextuel.
5. Pour dégroupier les valeurs d'un champ groupé, faites un clic droit sur le champ groupé, puis sélectionnez **Dégroupier** dans le menu contextuel.

Remplacer une ou plusieurs valeurs par une valeur null

Si vous souhaitez inclure des lignes de données dans votre analyse, mais exclure certaines valeurs de champ, vous pouvez les transformer en une valeur null.

1. Dans la fiche **Profil**, appuyez sur Ctrl ou Maj+clic, ou Command ou Maj+clic (sur Mac), et sélectionnez les valeurs que vous souhaitez modifier
2. Faites un clic droit ou Ctrl+clic (Mac), et sélectionnez **Remplacer par une valeur null** dans le menu. Les valeurs sont transformées en valeurs null et l'icône de groupe  apparaît à côté de la valeur.



Associer manuellement plusieurs valeurs à une valeur standard

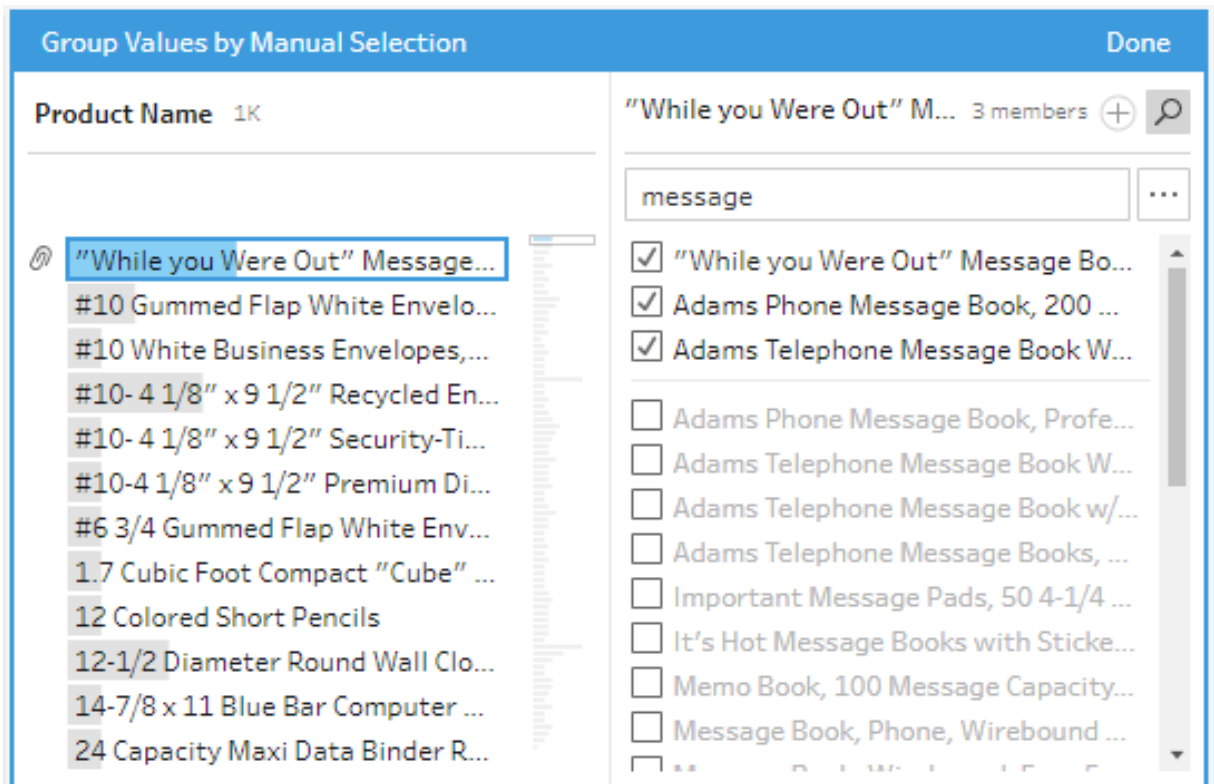
Utilisez l'option **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) pour associer la valeur d'un champ à une autre valeur ou sélectionner manuellement plusieurs valeurs à regrouper. Vous pouvez même ajouter de nouvelles valeurs pour configurer des relations de mappage afin d'organiser vos données.

Par exemple, imaginons qu'un champ comporte trois valeurs : Ma société, Ma société Incorporated et Ma société Inc. Ces valeurs représentent toutes trois la même entreprise, Ma société. Vous pouvez utiliser **Regrouper les valeurs** pour associer les valeurs Ma société Incorporated et Ma société Inc. à Ma société afin que les valeurs s'affichent toutes trois comme Ma société dans le champ.

Associer plusieurs valeurs à un champ sélectionné unique

1. Dans le volet **Profil** ou le volet Résultats, sélectionnez le champ à modifier.

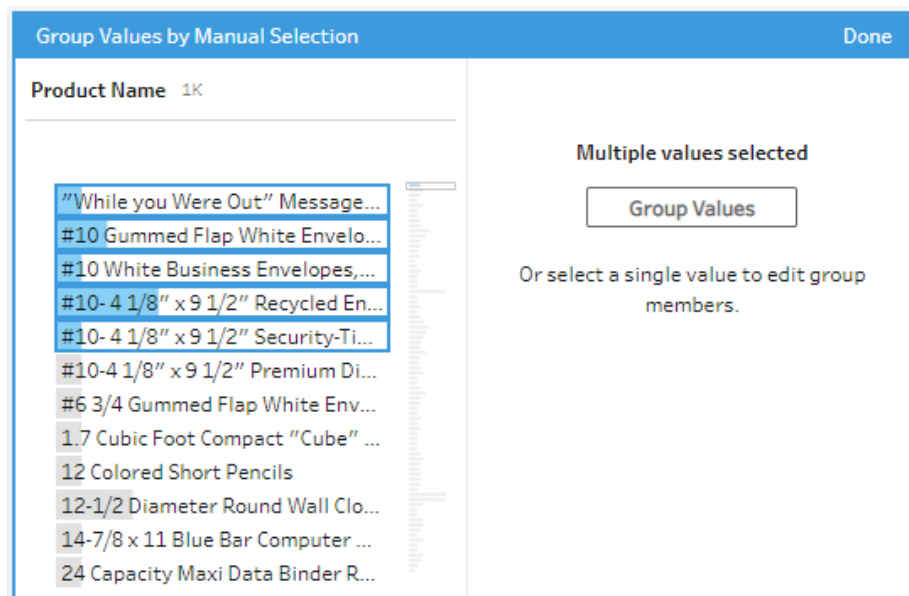
2. Cliquez sur **Options supplémentaires** *** et sélectionnez **Regrouper les valeurs (Regrouper et remplacer dans les versions précédentes)** > **Sélection manuelle** dans le menu.
3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, sélectionnez la valeur de champ que vous souhaitez utiliser comme valeur de regroupement. Cette valeur s'affiche maintenant en haut du volet droit.
4. Dans la section inférieure du volet droit dans l'éditeur **Regrouper les valeurs**, sélectionnez les valeurs à ajouter au groupe.



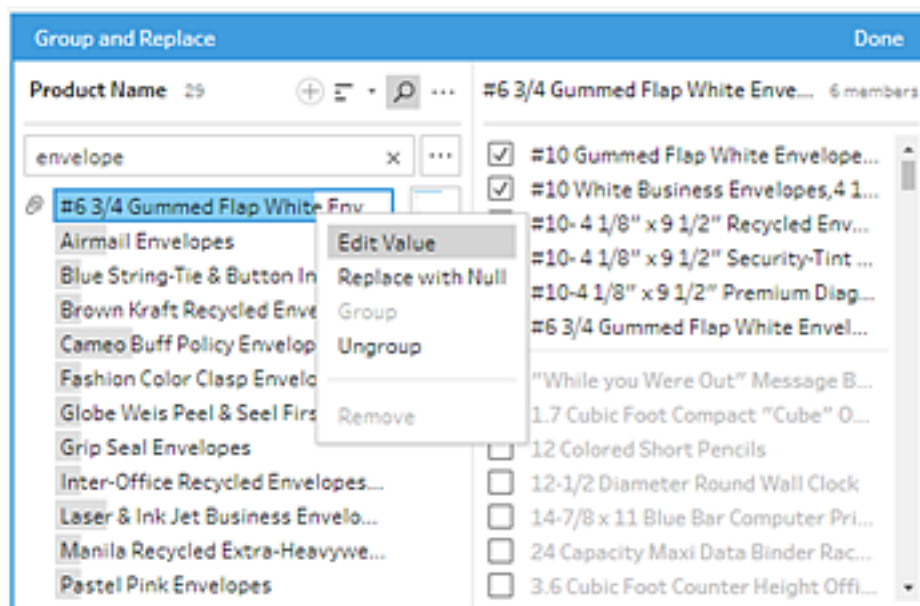
Pour supprimer des valeurs du groupe, dans la section supérieure du volet droit, dans l'éditeur **Regrouper les valeurs**, désactivez la case à cocher à côté des valeurs.

Créer un groupe en sélectionnant plusieurs valeurs

1. Dans le volet **Profil** ou le volet **Résultats**, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** *** et sélectionnez **Regrouper les valeurs (Regrouper et remplacer dans les versions précédentes)** > **Sélection manuelle** dans le menu.
3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, sélectionnez plusieurs valeurs que vous souhaitez regrouper.
4. Dans le volet droit de l'éditeur **Regrouper les valeurs**, cliquez sur **Regrouper les valeurs**.



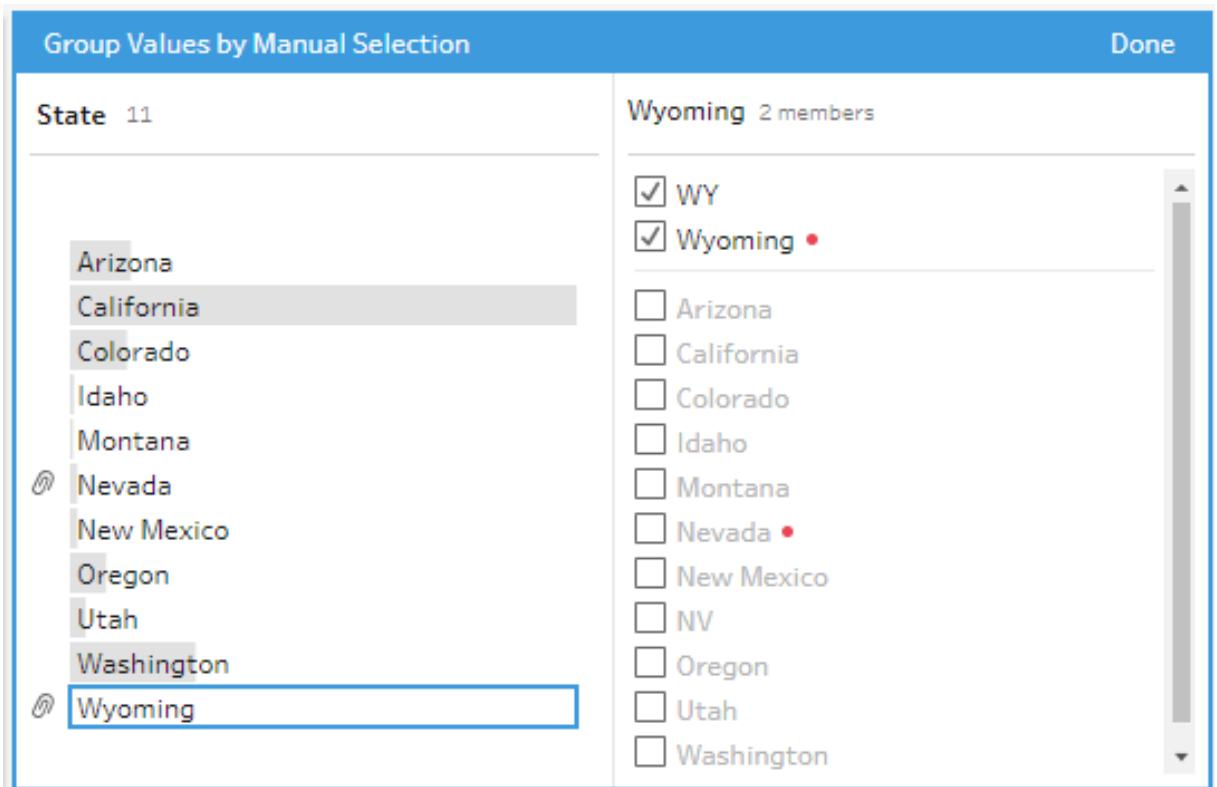
Un nouveau groupe est créé à l'aide de la dernière valeur sélectionnée comme nom de groupe. Pour modifier le nom du groupe, sélectionnez le champ groupé et modifiez la valeur ou faites un clic droit ou Ctrl+clic (MacOS) sur le champ groupé et sélectionnez **Modifier la valeur** dans le menu.



Ajouter et identifier des valeurs qui ne sont pas présentes dans l'ensemble de données

Si vous souhaitez associer des valeurs de votre ensemble de données à une nouvelle valeur qui n'existe pas, vous pouvez les ajouter à l'aide de **Regrouper les valeurs (Regrouper et remplacer)** dans les versions précédentes). Afin que vous puissiez identifier rapidement les valeurs qui ne sont pas dans l'ensemble de données, ces valeurs sont marquées avec un point rouge à côté du nom de la valeur dans l'éditeur **Regrouper les valeurs**.


Par exemple, dans l'image ci-dessous, le Wyoming et le Nevada ne figurent pas dans l'ensemble de données.

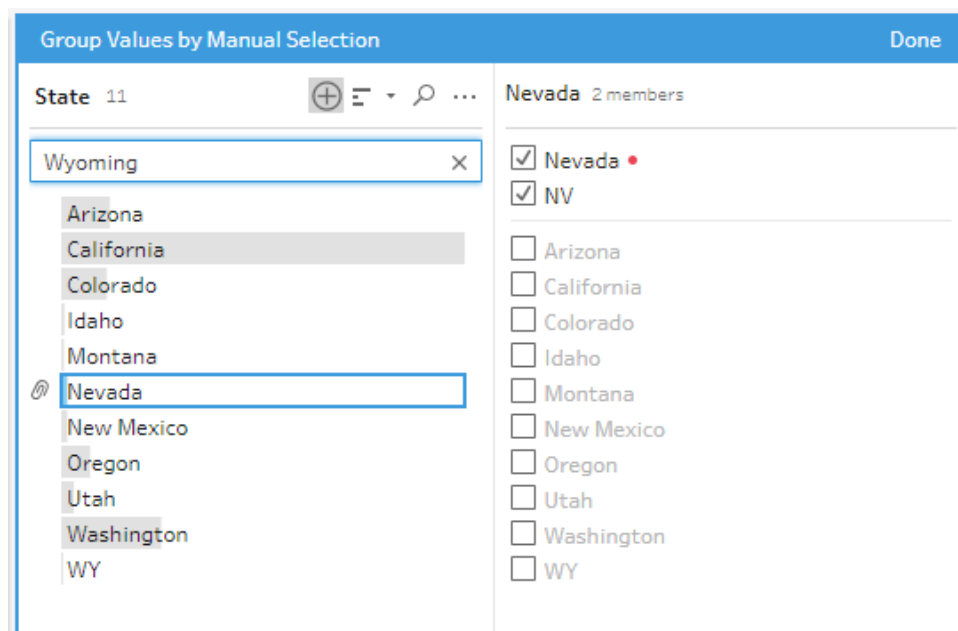


Voici quelques raisons pour lesquelles une valeur peut ne pas être dans l'ensemble de données :

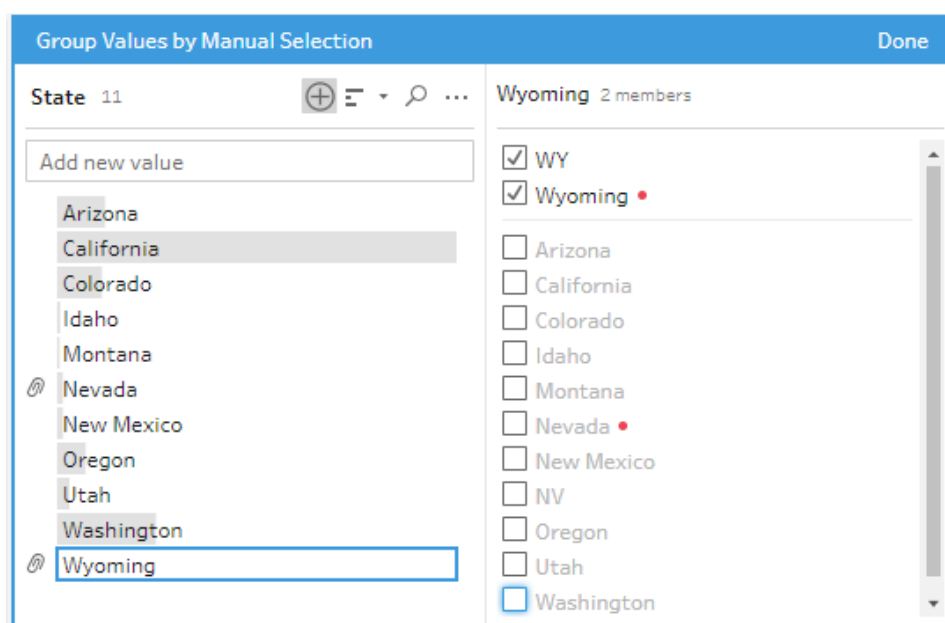
- Vous venez d'ajouter la nouvelle valeur manuellement.
- La valeur n'est plus dans les données.
- La valeur est dans les données mais n'est pas dans l'ensemble de données échantillonné.

Pour ajouter une nouvelle valeur :

0. Dans le volet **Profil** ou le volet **Résultats**, sélectionnez le champ à modifier.
1. Cliquez sur **Options supplémentaires** *** et sélectionnez **Regrouper les valeurs (Regrouper et remplacer dans les versions précédentes) > Sélection manuelle** dans le menu contextuel.
2. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, cliquez sur l'icône plus  pour ajouter une nouvelle valeur.
3. Entrez une nouvelle valeur dans le champ et appuyez sur Entrée pour l'ajouter.



4. Dans le volet droit, sélectionnez les valeurs à associer à la nouvelle valeur.



5. (Facultatif) Pour ajouter de nouvelles valeurs supplémentaires à votre valeur associée, cliquez sur le bouton plus (+) dans le volet droit de l'éditeur **Regrouper les valeurs**.

Associer automatiquement des valeurs à une valeur standard à l'aide de la correspondance approximative

Pour rechercher et regrouper automatiquement des valeurs similaires, utilisez l'un des algorithmes de correspondance approximative. Les valeurs de champ sont regroupées sous la valeur qui apparaît le plus fréquemment. Consultez les valeurs regroupées, et ajoutez ou supprimez des valeurs dans le groupe comme nécessaire.

Si vous utilisez des types de données pour valider vos valeurs de champ, vous pouvez utiliser l'option **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) pour faire correspondre des valeurs non valides avec des valeurs valides.

Choisissez l'une des options suivantes pour regrouper les valeurs :

- **Prononciation** : trouvez et regroupez les valeurs ayant la même sonorité. Cette option utilise l'algorithme Metaphone 3 qui indexe les mots selon leur prononciation et est plus spécifiquement adapté aux mots anglais. Ce type d'algorithme est utilisé par de nombreux correcteurs orthographiques.
- **Caractères communs** : trouvez et regroupez les valeurs ayant des lettres et des nombres en commun. Cette option utilise l'algorithme d'empreinte digitale ngram qui indexe les mots par leurs caractères uniques après avoir supprimé la ponctuation, les doublons et les espaces blancs. Cet algorithme fonctionne quelle que soit la langue prise en charge. Cette option n'est pas disponible pour les rôles de données.

Par exemple, cet algorithme associerait les noms représentés sous la forme « John Smith » et « Smith, John » parce qu'ils génèrent tous deux la clé « hijmnost ». Cet algorithme ne prenant pas en considération la prononciation, la valeur « Tom Jhinois » aurait la même clé « hijmnost » et serait également inclus dans le groupe.

- **Orthographe** : trouvez et regroupez les valeurs ayant la même orthographe. Cette option utilise l'algorithme de distance Levenshtein qui calcule une distance de modification entre deux valeurs de texte utilisant un seuil fixe par défaut. Elle les regroupe ensuite lorsque la distance de modification est inférieure à la valeur de seuil. Cet algorithme fonctionne quelle que soit la langue prise en charge.
- **Prononciation et orthographe** : si vous affectez un type de données à vos champs, vous pouvez utiliser ce type de données pour associer et regrouper des valeurs avec la valeur standard définie par votre type de données. Cette option associe ensuite des valeurs non valides à la valeur valide la plus proche en fonction de l'orthographe et de la prononciation. Si la valeur standard ne se trouve pas dans votre exemple d'ensemble de données, Tableau Prep l'ajoute automatiquement et marque la valeur comme n'appartenant pas à l'ensemble de données d'origine. Cette option est celle qui convient le mieux aux mots anglais.

Regrouper des valeurs similaires à l'aide de la correspondance floue

1. Dans le volet **Profil** ou le volet **Résultats**, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** *** et sélectionnez **Regrouper les valeurs** puis sélectionnez l'une de ces options :
 - **Prononciation**
 - **Caractères communs**
 - **Orthographe**

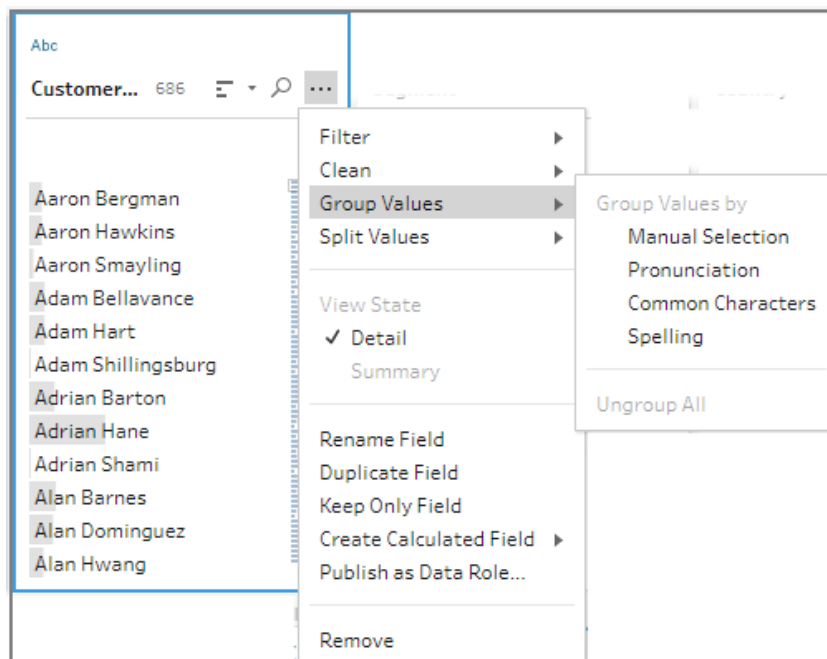
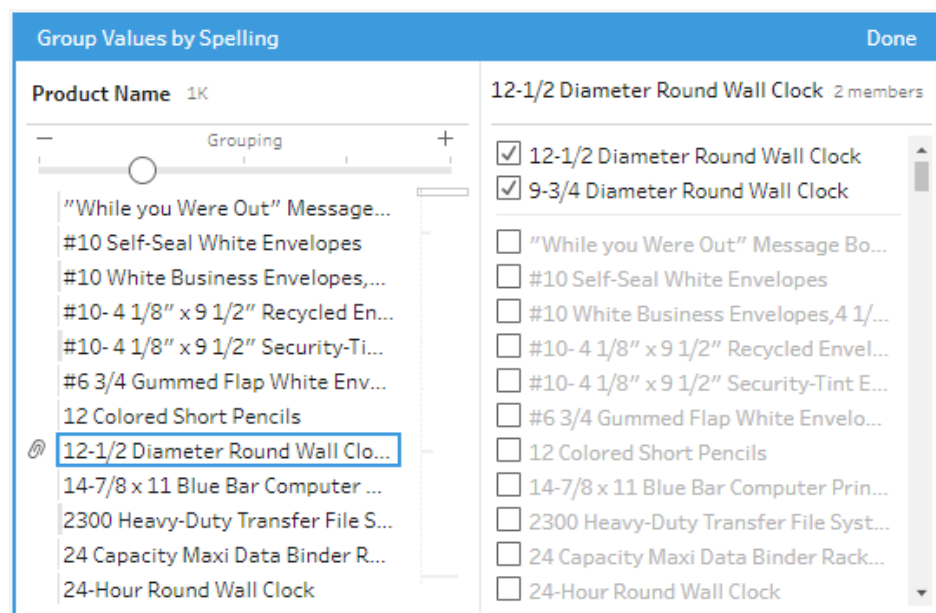


Tableau Prep Builder trouve et regroupe les valeurs qui correspondent et les remplace par la valeur qui se présente le plus fréquemment dans le groupe.

3. Consultez les regroupements, et ajoutez ou supprimez manuellement des valeurs, ou modifiez-le comme nécessaire. Cliquez ensuite sur **Terminer**.



Ajuster vos résultats lors du regroupement des valeurs de champ

Si vous regroupez des valeurs similaires par **Orthographe** ou **Prononciation**, vous pouvez modifier vos résultats en utilisant le curseur sur le champ afin d'ajuster le degré de rigidité des paramètres de regroupement.

Selon la manière dont vous définissez le curseur, vous pouvez avoir davantage de contrôle sur le nombre de valeurs incluses dans un groupe et le nombre de groupes créés. Par défaut, Tableau Prep détecte le paramètre de regroupement optimal et affiche le curseur dans cette position.

Lorsque vous modifiez le seuil, Tableau Prep analyse un échantillon des valeurs pour déterminer le nouveau regroupement. Les groupes générés à partir du paramètre sont sauvegardés et enregistrés dans le volet **Modifications**, mais le paramètre de seuil n'est pas enregistré. La prochaine fois que vous ouvrez l'éditeur **Regrouper les valeurs**, soit pour modifier votre changement existant, soit pour effectuer une nouvelle modification, le curseur de seuil est affiché dans la position par défaut, ce qui vous permet d'effectuer des ajustements en fonction de votre ensemble de données actuel.

1. Dans le volet **Profil** ou le volet **Résultats**, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** *** et sélectionnez **Regrouper les valeurs (Regrouper et remplacer** dans les versions précédentes) puis sélectionnez l'une de ces options :
 - **Prononciation**
 - **Orthographe**

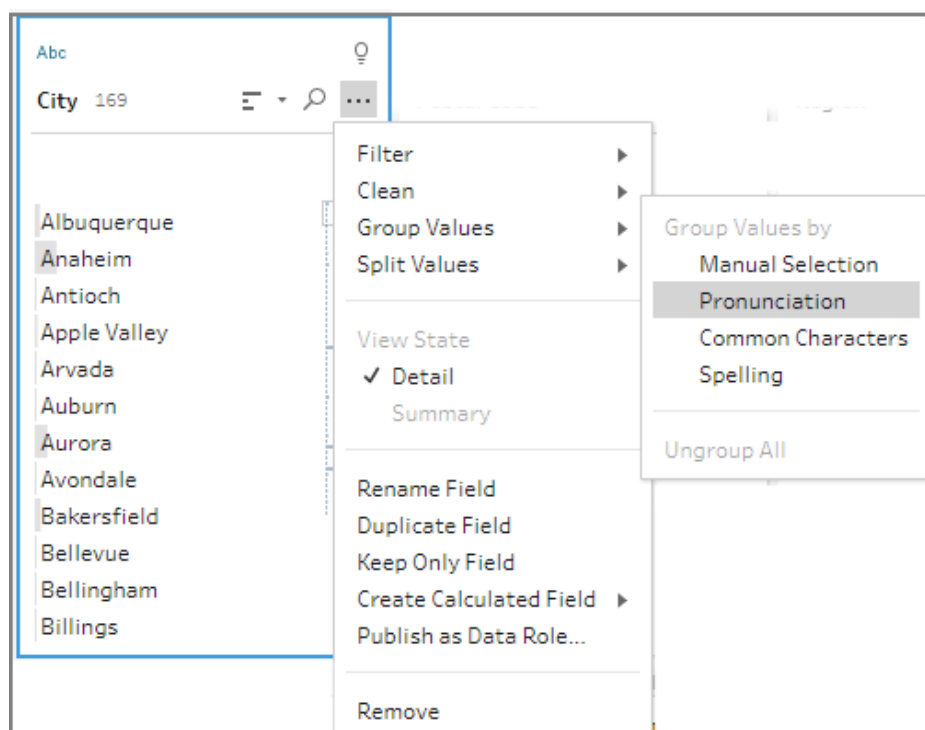
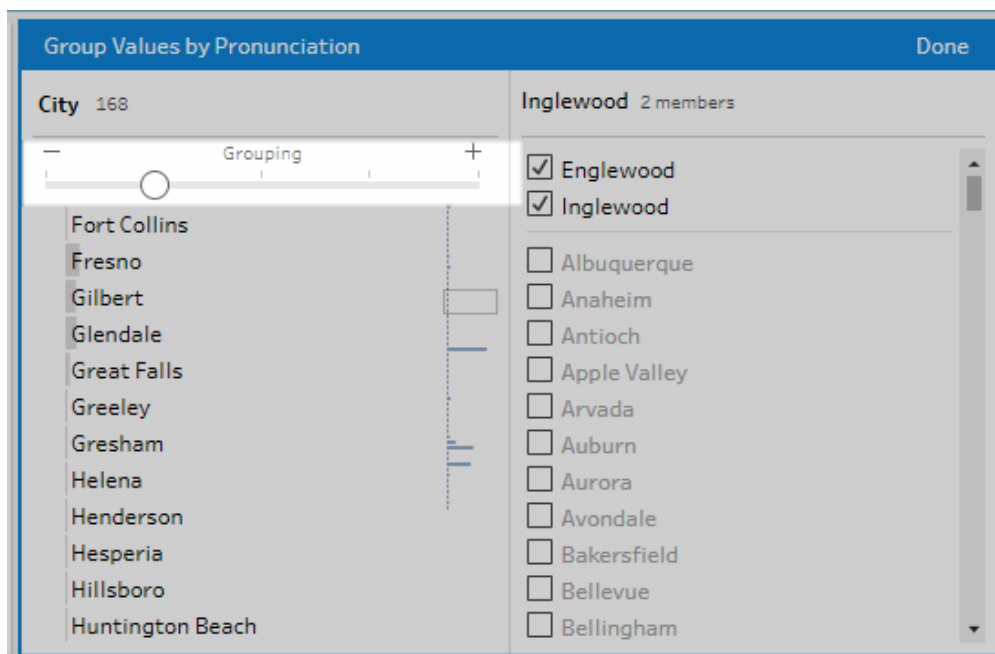


Tableau Prep trouve et regroupe les valeurs qui correspondent et les remplace par la valeur qui se présente le plus fréquemment dans le groupe.

3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, faites glisser le curseur vers l'un des 5 niveaux de seuil pour modifier vos résultats.



Pour définir un seuil plus strict, déplacez le curseur vers la gauche. Il en résulte moins de correspondances et la création de moins de groupes inférieurs. Pour définir un seuil moins strict, déplacez le curseur vers la droite. Il en résulte plus de correspondances et la création de plus de groupes.

4. Cliquez sur **Terminé** pour enregistrer vos modifications.

Exercice

L'objectif est de nettoyer les différents fichiers afin de les regrouper ultérieurement, pour un regroupement efficace il faut que les champs et les valeurs qu'ils contiennent soient de même nature. Dans un premier temps sans la suite détaillée essayer de nettoyer les fichiers USCA, APAC, EMEA et LATAM dans l'objectif de les rendre compatibles.

- ➔ Sur le champ 'région de vente', regrouper les 4 régions de vente en 'USCA' uniquement
- ➔ Ajouter aux flux les données :

APAC – nettoyage

- ➔ Supprimer les lettres de montant des ventes
- ➔ Mettre montant des ventes en décimale
- ➔ Sur le champ 'région de vente' regrouper en 'APAC'

EMEA – nettoyage :

- ➔ Ne pas prendre les champs Droit_XXX dans la base
- ➔ Dans le champ 'État/province', grouper correctement les provinces (ex : Danemark du S. et Danemark du Sud)
- ➔ Sur le champ 'région de vente' regrouper en 'EMEA'

LATAM – nettoyage :

- ➔ Sur le champ 'Région de vente' regrouper en 'LATAM'
- ➔ Créer un champ calculé pour reconstruire la date de commande complète à partir de l'année, du mois et du jour sous la forme 'j/m/a'
- ➔ Créer un champ calculé pour reconstruire la date de livraison complète à partir de l'année, du mois et du jour sous la forme 'j/m/a'
- ➔ Modifier le type de données de ces champs calculés en date
- ➔ Renommer le champ 'Pays\région' en 'Pays/région'
- ➔ Remplacer les nulls par des 0 dans le champ Remise
- ➔ Exclure les valeurs nulls de ID commande
- ➔ Renommer Produit en 'Nom du Produit'

Sauvegarder ce flux pour la suite

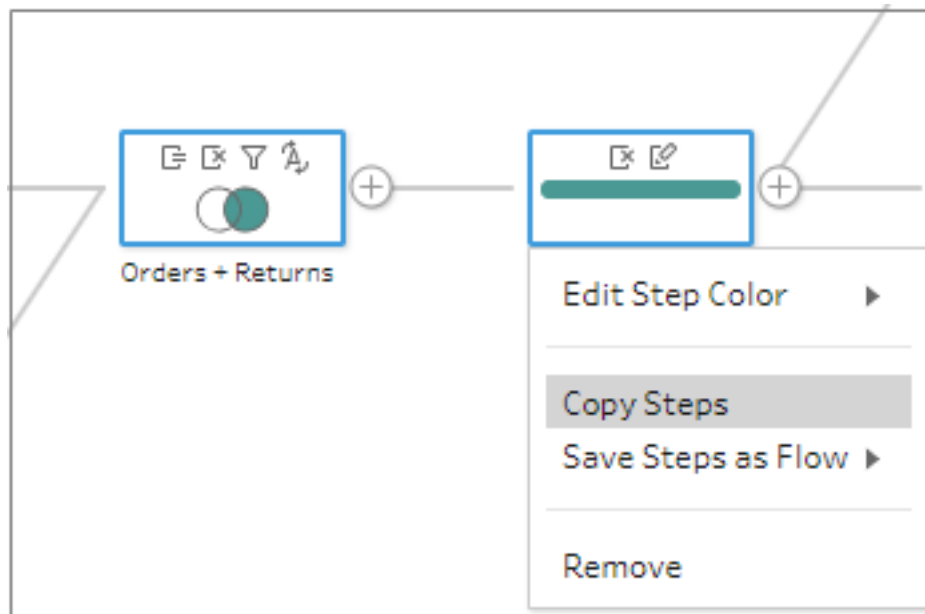
Copier des étapes, des actions et des champs

Lors du nettoyage de vos données, vous répétez souvent les mêmes actions ou opérations de nettoyage tout au long de votre flux. Pour optimiser le nettoyage et l'organisation de vos données, vous pouvez copier-coller les opérations ou les actions tout au long de votre flux, ou même copier des étapes ou des groupes sélectionnés et les enregistrer. Vous pouvez ainsi effectuer des actions ou des opérations de nettoyage une seule fois, puis les réutiliser là où vous en avez besoin. Vous pouvez même dupliquer des champs pour expérimenter différentes opérations de nettoyage.


Copier et coller des étapes

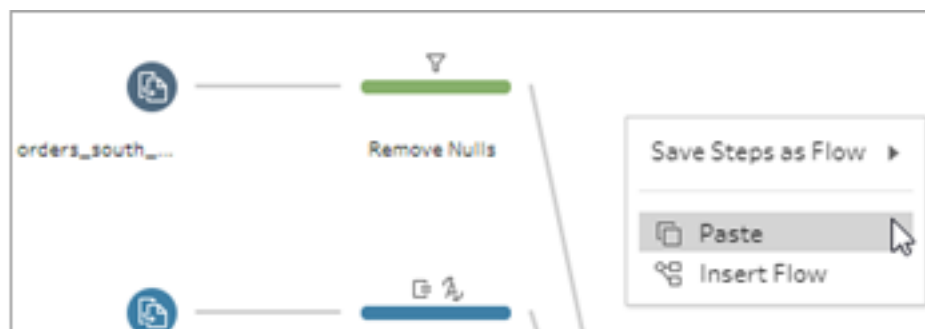
Copiez une ou plusieurs étapes de manière à les utiliser dans une autre partie du même flux. Cette option n'est pas disponible pour les étapes des données entrantes qui incluent une union avec caractère générique.

1. Dans le volet Flux, sélectionnez une ou plusieurs étapes ou groupes dans le flux.
2. Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur l'étape sélectionnée et sélectionnez **Copier**.



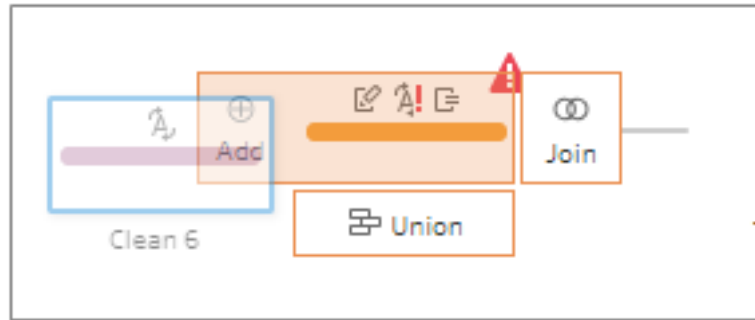
3. Pour coller les étapes copiées, effectuez l'une des opérations suivantes :

- Survolez une étape jusqu'à ce que l'icône plus  apparaisse, puis cliquez sur l'icône et sélectionnez **Coller** dans le menu.
- Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) dans n'importe quel espace blanc sur l'espace de travail et cliquez sur **Coller**.



4. Si vous avez collé les étapes dans l'espace blanc du flux, faites glisser les étapes là où vous les souhaitez dans le flux et déposez-les. Si vous ajoutez des étapes à la fin d'une étape de flux, les étapes sont automatiquement ajoutées à la fin de l'étape. Si vous insérez des étapes entre les étapes de flux existantes, déplacez les étapes là où vous les souhaitez dans le flux et corrigez les erreurs éventuelles.

Vous pouvez supprimer des lignes du flux ou déplacer des étapes selon vos besoins. Par exemple, pour connecter une étape aux étapes copiées, supprimez la ligne de flux existante s'il y en a une, puis faites glisser l'étape existante vers la nouvelle étape et déposez-la sur **Ajouter**.



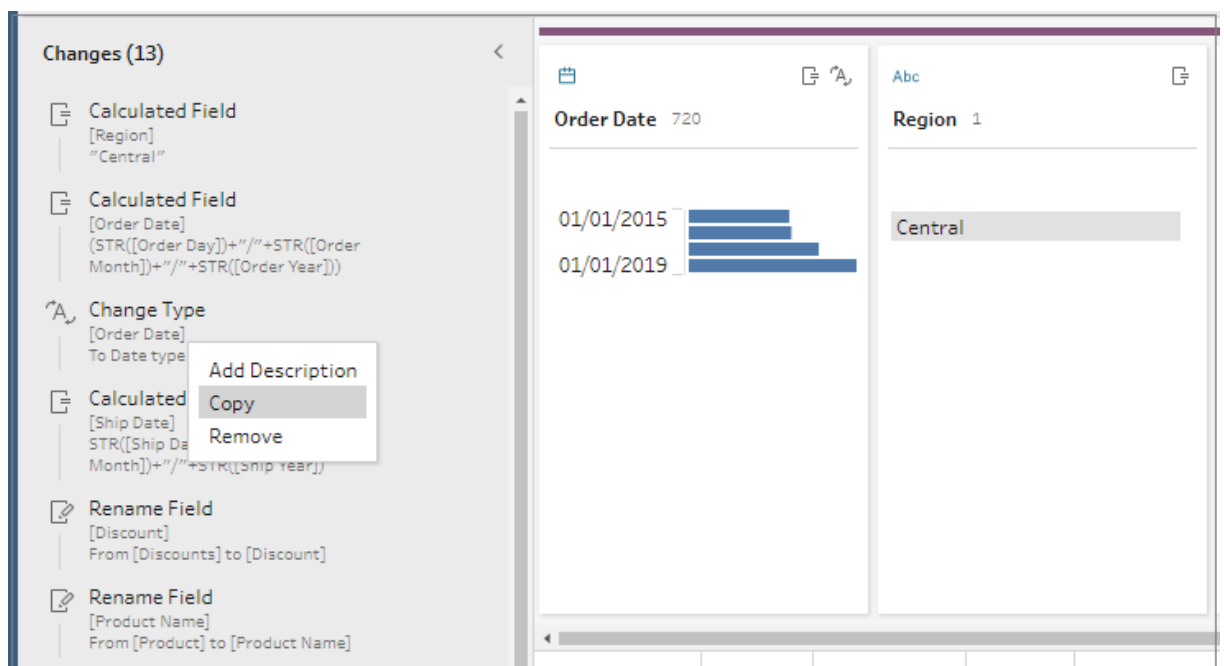
Copier et coller des opérations de nettoyage

Vous pouvez copier et coller des opérations de nettoyage dans le même flux pour réutiliser vos actions à l'aide de l'une des options suivantes :

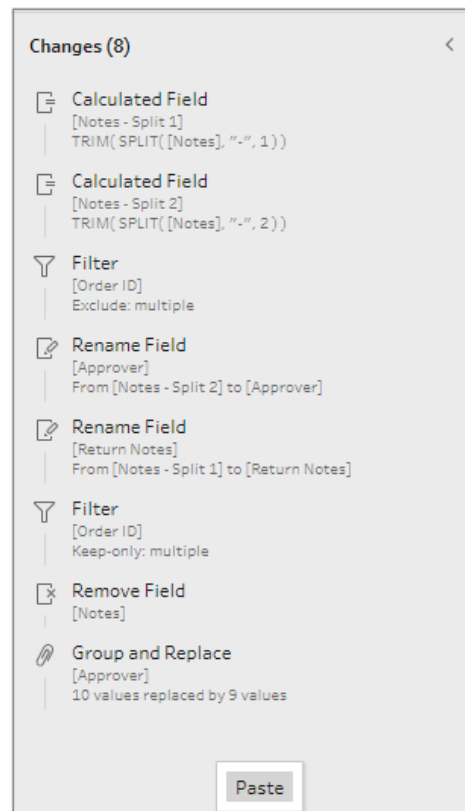
- Copiez une opération depuis le volet **Modifications** dans une étape et collez-la dans le volet **Modifications** de la même étape ou d'une autre étape pour appliquer la même opération dans cette étape.
- Faites glisser et déposez une opération depuis le volet **Modifications** dans d'autres champs du volet Profil de cette étape pour appliquer cette opération à plusieurs champs. Cette option n'est pas disponible pour les opérations affectant plusieurs champs, par exemple les champs calculés.

Pour copier et coller une modification apportée à une étape dans la même étape ou dans une autre étape, procédez comme suit :

1. Dans le volet **Modifications**, sélectionnez la modification que vous souhaitez copier.
2. Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur l'élément de modification, puis sélectionnez **Copier** dans le menu.

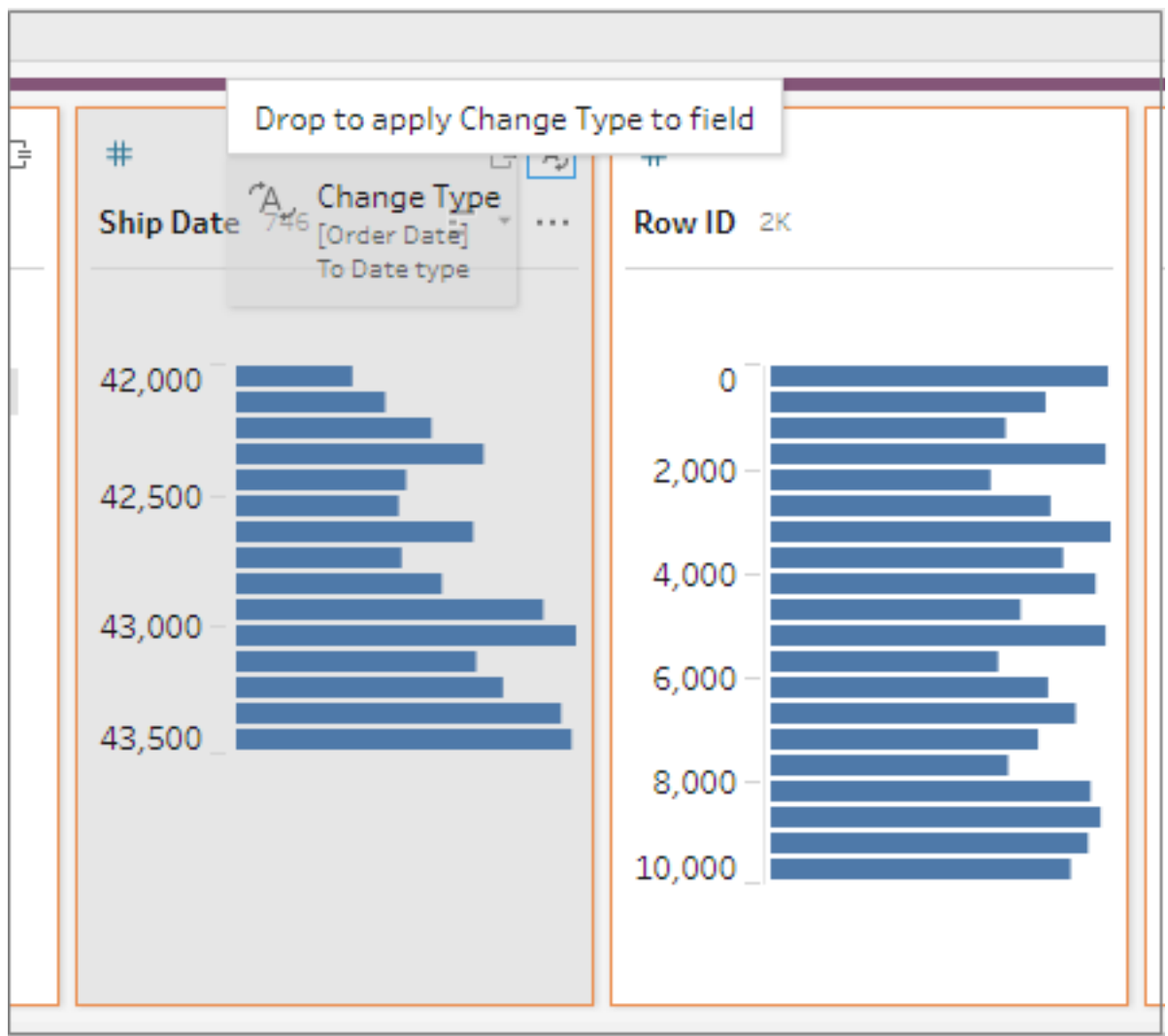


3. Dans le volet **Modifications** où vous voulez coller la modification, cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) et sélectionnez **Coller**. Sélectionnez la modification et cliquez sur **Modifier** pour effectuer les ajustements nécessaires.



Pour faire glisser et déposer une modification dans d'autres zones de l'étape, procédez comme suit :

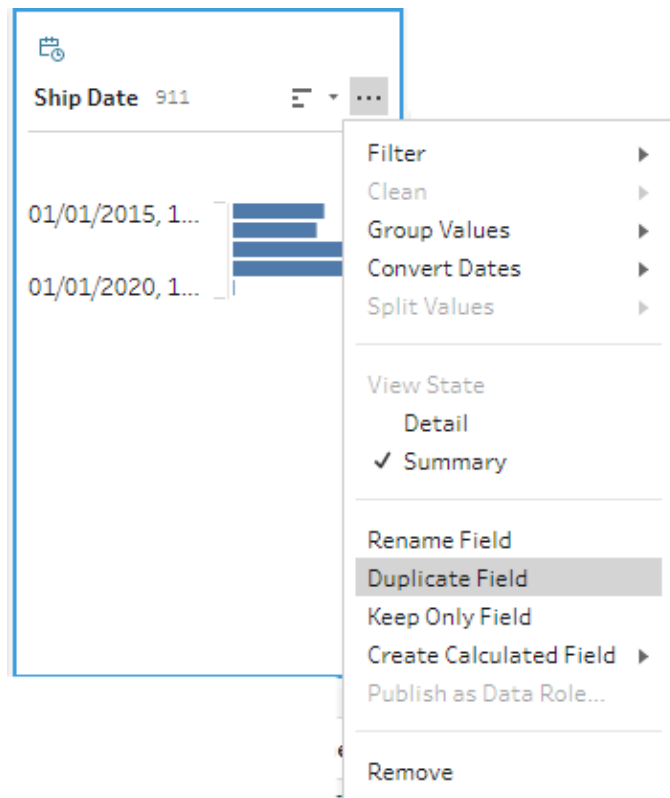
1. Dans le volet **Modifications**, sélectionnez la modification que vous souhaitez copier.
2. Faites glisser la modification sur le champ où vous voulez l'appliquer et déposez-la. Répétez cette action si nécessaire.



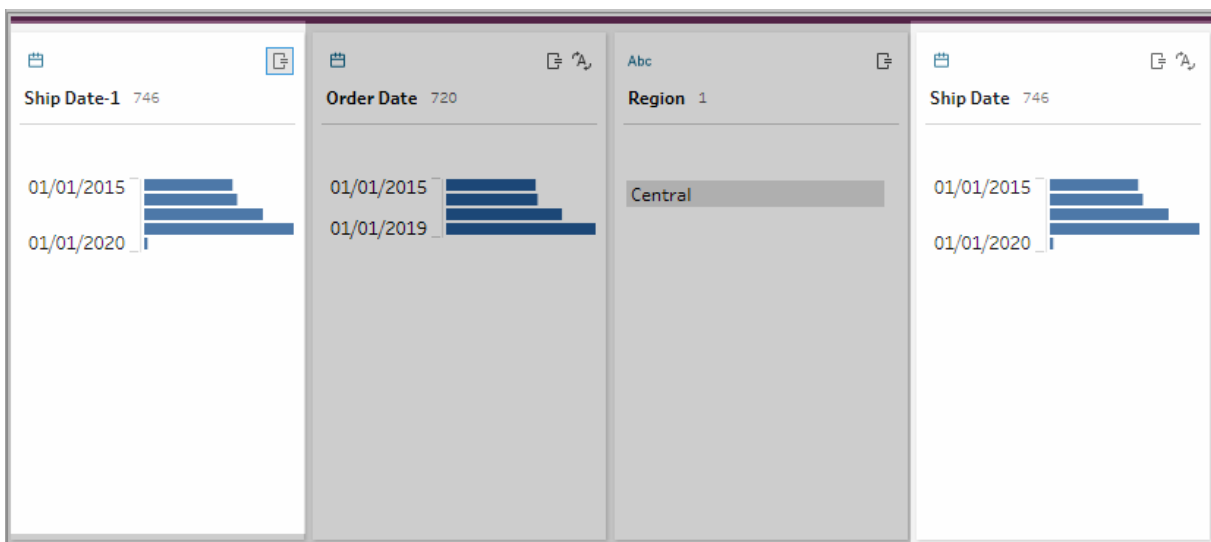
Copier des champs

Dans Tableau Prep Builder, si vous souhaitez expérimenter avec vos opérations de nettoyage sur un champ mais sans modifier les données d'origine, vous pouvez copier vos champs.

1. Dans le volet Profil, dans la grille de données, dans le volet Résultats ou dans la liste, sélectionnez le champ que vous souhaitez copier.
2. Dans le menu **Options supplémentaires**, sélectionnez **Dupliquer le champ**.



Un nouveau champ est créé avec le même nom et un modificateur. Par exemple, « Ship Date -1 ».



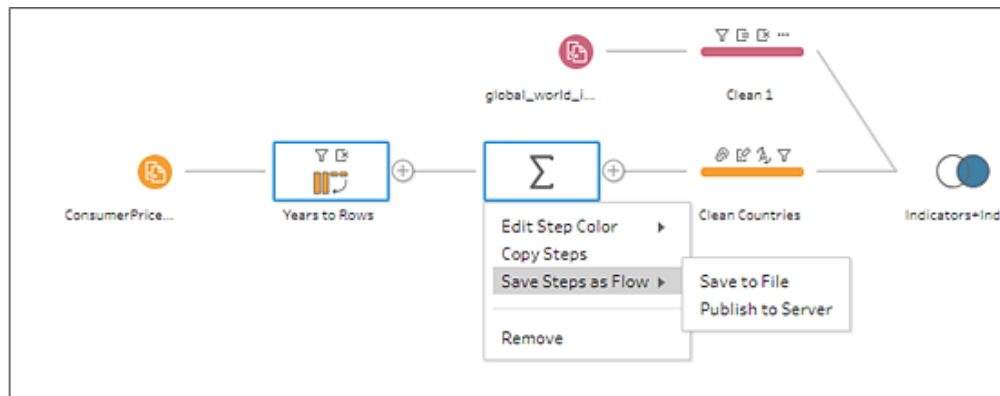
Créer des étapes de flux réutilisables

Si vous répétez souvent les mêmes actions et que vous souhaitez appliquer ces mêmes étapes dans d'autres flux, Tableau Prep Builder version 2019.3.2 et versions ultérieures vous permet de sélectionner un(e) ou plusieurs étapes ou groupes de flux et leurs actions associées ou le flux tout entier et de les enregistrer localement dans un fichier sur votre ordinateur.

Créer des étapes réutilisables

1. Sélectionnez une ou plusieurs étapes.

2. Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur une étape sélectionnée et sélectionnez **Enregistrer les étapes en tant que flux**.



4. Sélectionne **Enregistrer dans le fichier** pour enregistrer le flux localement