

Informations supplémentaires sur la fusion des données

Bienvenue dans cette vidéo présentant des aspects supplémentaires de la fusion de données.

Vous pouvez télécharger le classeur d'exercices pour suivre les étapes dans votre propre logiciel Tableau.

Fusion des données et calculs

Effectuer des calculs avec des données provenant de deux sources distinctes et créer un champ calculé ordinaire peuvent être des actions légèrement différentes.

Cette vue présente une comparaison État par État des données de vente d'Office City et de Coffee Chain.

Que faire pour afficher les ventes combinées pour chaque État ?

Afin de combiner nos ventes provenant des deux sources de données, nous devons créer un calcul.

- Vérifiez que la source de données Office City est sélectionnée,
- puis cliquez avec le bouton droit dans le volet Données et sélectionnez « Créer un champ calculé ».
- Ici est indiquée la source de données avec laquelle nous créons le calcul.
- Faites glisser les ventes d'Office City
- et saisissez le signe plus (+).

Passez maintenant à la source de données Coffee Chain.

- Faites glisser les ventes de Coffee Chain.

Utilisation de données agrégées dans les calculs

Comme vous pouvez le remarquer, les ventes de Coffee Chain ne se présentent pas de la même manière que les ventes d'Office City dans la formule.

Tout d'abord, les ventes de Coffee Chain contiennent également le nom de la source de données. Cela nous permet de distinguer les différentes sources de données dans un calcul, et de toujours savoir d'où viennent les champs.

Ensuite, les ventes de Coffee Chain sont automatiquement englobées dans une fonction SUM. Les données provenant de la source secondaire doivent toujours être agrégées. La fonction SUM correspond à l'agrégation par défaut.

Comme vous pouvez le constater, le calcul comporte une erreur, car il n'est pas possible de mélanger les arguments agrégés et non agrégés.

Pour corriger cette erreur, il suffit d'englober également les ventes d'Office City dans une fonction SUM.

Notre calcul fonctionne, nous allons donc le renommer « Ventes combinées », puis cliquer sur OK.

Notre calcul n'apparaît pas dans le volet Données car nous utilisons une autre source de données que celle sur laquelle le calcul se base. Lorsque nous passons à Office City, le champ calculé apparaît.

Utilisation des valeurs null

Nous sommes désormais prêts à tester notre calcul.

- Faites glisser le champ Ventes combinées sur l'étagère Colonnes.
 - Hum. Les données des ventes combinées ne sont pas affichées pour les États où la valeur des ventes de Coffee Chain est null.

Une valeur null (de Coffee Chain) combinée à une valeur réelle (d'Office City) entraîne toujours une valeur null. Le calcul de Tableau est donc correct. Cependant, nous savons que nos valeurs null de Coffee Chain correspondent en fait à zéro. Nous devons donc ajuster notre calcul en conséquence.

- Cliquez avec le bouton droit sur Ventes combinées, puis sur Modifier.
- Nous allons englober les ventes de Coffee Chain dans une fonction ZN, qui signifie « zéro-null » et qui indique à Tableau que les valeurs null doivent être traitées comme des zéros.

Notre colonne Ventes combinées affiche désormais correctement les informations de ventes combinées pour Office City et Coffee Chain.

Astérisques lors de la fusion

L'apparition d'astérisques est un autre problème fréquent. Lorsque les données sont fusionnées, l'ajout d'une autre dimension à une vue peut entraîner l'affichage d'un astérisque à la place des valeurs attendues. Pourquoi ?

Dans notre vue actuelle, nous utilisons Office City en tant que source de données principale (en bleu) et nous effectuons une fusion sur les États.

Le champ Indicatif régional n'est présent que dans la source de données secondaire. Puisqu'il existe plusieurs indicatifs régionaux pour la plupart des États, Tableau Desktop affiche un astérisque afin d'indiquer qu'il y a plusieurs valeurs.

Les valeurs pour New Hampshire et Nouveau-Mexique s'affichent normalement, car ces États n'ont qu'un seul indicatif régional.

Solution concernant les astérisques : intervertir la source principale et la source secondaire

Comment résoudre le problème des astérisques lors de la fusion de données ? Essayez d'intervertir les sources de données principale et secondaire dans une nouvelle vue.

Nous utilisons désormais Coffee Chain en tant que source de données principale ; elle s'affiche en bleu.

Vérifiez qu'elle est sélectionnée, puis déposez le champ Indicatif régional dans la vue. Cela correspond au comportement attendu du champ Indicatif régional. Nous avons désormais une liste des indicatifs régionaux par État.

Cependant, la source de données Office City ne contient pas de champ Indicatif régional. Ainsi, les ventes sont agrégées au niveau de la fusion, État, et cette valeur se répète pour chaque indicatif régional.

Une jointure entre bases de données convient peut-être mieux à votre analyse. Pour en savoir plus, consultez les vidéos consacrées aux jointures entre bases de données.

CONCLUSION

Merci d'avoir suivi cette vidéo présentant des informations supplémentaires sur la fusion de données. Nous vous invitons à découvrir les autres vidéos gratuites pour en apprendre davantage sur l'utilisation des produits Tableau.