

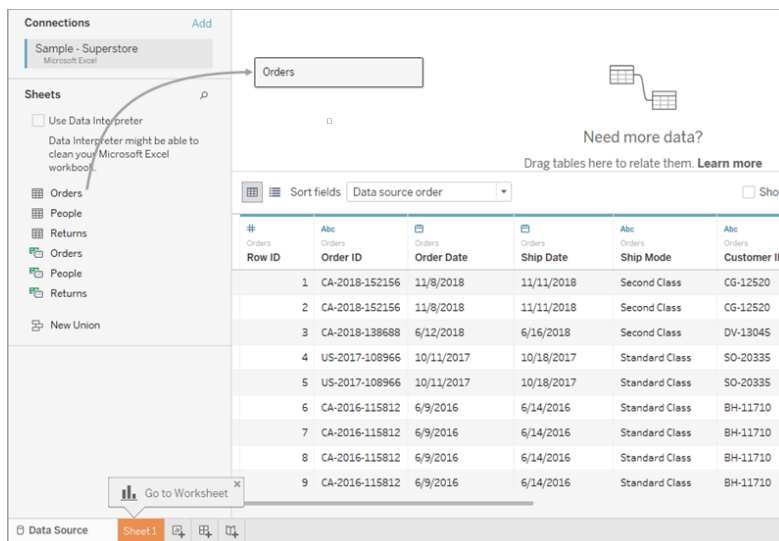
Initiation Tableau

Niveau I - Découverte

Connexion aux données

La première étape consiste à vous connecter aux données que vous souhaitez explorer. Cet exemple montre la procédure permettant de se connecter à des données Sample - Superstore dans Tableau Desktop.

1. Ouvrez Tableau. Sur la page de démarrage, sous **Connexion**, cliquez sur Excel. Dans la boîte de dialogue Ouvrir, accédez au fichier Excel Sample - Superstore sur votre ordinateur.
2. Une fois que vous êtes connecté aux données Excel, le volet Source de données affiche les feuilles ou les tables de vos données. Faites glisser la table « Orders » vers l'espace de travail pour commencer à explorer ces données.

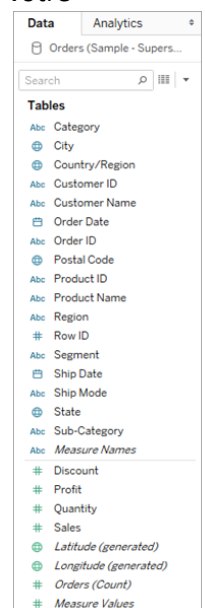


3. Cliquez sur l'onglet de la feuille pour atteindre la nouvelle feuille de calculs et démarrer votre analyse.

Dans la feuille de calcul, les colonnes de votre source de données sont affichées sous forme de champs sur la partie gauche dans le volet **Données**. Le volet **Données** contient une variété de champs organisés par table. Pour chaque table ou dossier d'une source de données, les champs de dimensions apparaissent au-dessus de la ligne grise, et les champs de mesures apparaissent sous la ligne grise.

En général, **les champs de dimensions contiennent des données de catégorie**, telles que les dates et les types de produits, **tandis que les champs de mesures contiennent des données numériques**, telles que les ventes et les profits.

Si vous avez des champs de dimensions, il peut arriver que vous souhaitiez les regrouper dans un dossier ou en tant que **hiérarchie**. Par exemple, dans cette source de données, Country, State, City et Postal Code sont regroupés dans une hiérarchie appelée Location.



Créer une vue

Une **vue** est une visualisation que vous créez dans Tableau. Il peut s'agir d'un graphique, d'une carte, d'un tracé ou même d'un tableau de texte.

Chaque vue que vous créez dans Tableau devrait commencer par une question. **Que voulez-vous savoir ?**

Différentes manières de commencer à créer une vue :

Lorsque vous créez une vue, vous ajoutez des champs dans le volet **Données**. Vous pouvez faire ceci de différentes manières.

Ces étapes montrent comment créer une vue de base affichant le profit année par année.

1. Depuis la zone **Dimensions** dans le volet **Données**, faites glisser le champ **Order Date** vers l'étagère **Colonnes**. Vous devez peut-être développer la hiérarchie Order pour voir Order Date.

Lorsque vous faites glisser un champ sur une étagère, un signe plus indique que l'étagère peut accepter ce champ.

La table obtenue contient quatre colonnes et une ligne. Chaque entête de colonne représente un membre du champ Order Date (le niveau de date par défaut est YEAR). Chaque cellule contient une étiquette "Abc", qui indique que le type de repère actuel pour cette vue est Texte.

Le niveau de date par défaut est déterminé par le plus haut niveau contenant plusieurs valeurs discrètes (par exemple, plusieurs années, plusieurs mois, etc.).

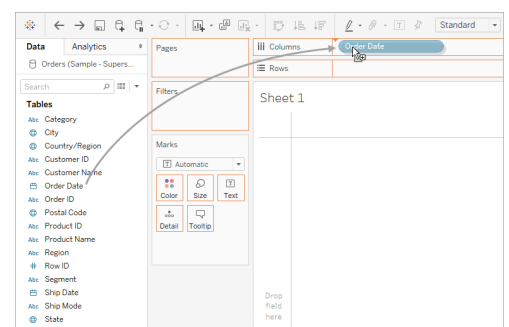
Conseil : pour masquer le titre sur la feuille, sélectionnez la liste déroulante à droite du titre (Feuille 1), et sélectionnez **Masquer le titre**.

2. Dans le volet **Données**, faites glisser le champ **Profit** vers l'étagère **Lignes**.

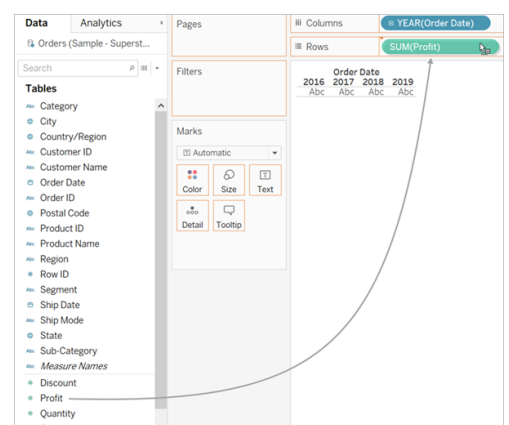
Le champ **Profit** s'affiche en vert sur l'étagère **Lignes**, ce qui signifie qu'il est continu. De même, le nom du champ change en **SUM(Profit)**. C'est parce que les mesures sont agrégées automatiquement lorsque vous les ajoutez à la vue et l'agrégation par défaut pour cette mesure est SUM. (Pour plus d'information lire la partie **Dimension et mesure**)

Le graphique en courbes affiche les profits sur une période définie. Chaque point de la courbe représente la somme des profits pour l'année correspondante.

L'étape suivante consiste à approfondir l'analyse de la dimension temporelle.




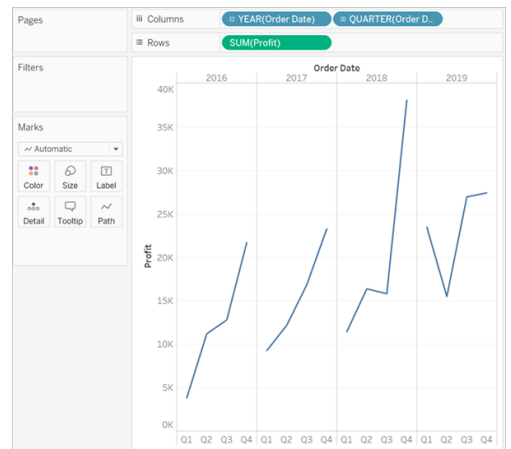
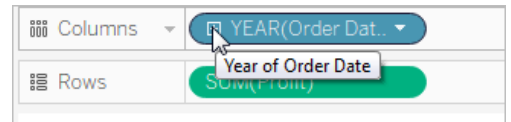
Columns				
YEAR(Order Date)				
Rows				
Order Date				
2016	2017	2018	2019	
Abc	Abc	Abc	Abc	



Explorer les données en détail

Vous pouvez afficher le contenu du champ **[Order Date]** par trimestre en utilisant l'une des méthodes suivantes :

- Cliquez sur le bouton plus  sur la partie gauche du champ **YEAR (Order Date)** sur Colonnes.
- Faites (à nouveau) glisser le champ **Order Date** du volet Données vers l'étagère Colonnes à droite du champ **Year (Order Date)**.

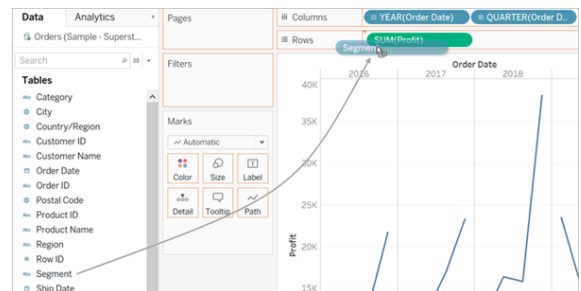


Augmenter le niveau de détail avec les petits multiples

Cette étape montre comment modifier la vue de table imbriquée pour ajouter un segment client. Une vue de petits multiples est créée.

Faites glisser la dimension **Segment** depuis le volet Données. Déposez-le juste à gauche de **SUM(Profit)** sur Lignes.

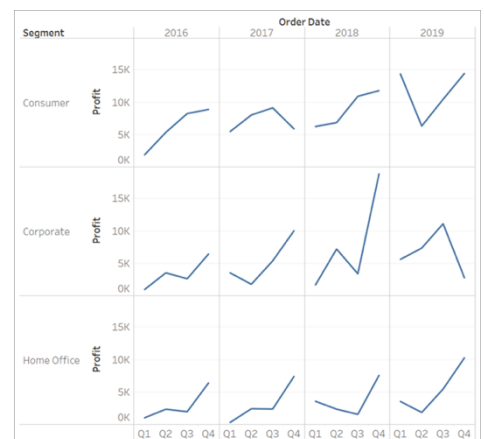
Le champ est ajouté à l'étagère Lignes et des en-têtes de lignes sont créés. Chaque en-tête représente un membre du champ Segment.



Remarque : Tableau ne permet pas de placer une dimension à droite d'une mesure, que ce soit dans les étagères des colonnes ou des lignes parce que cette structure visuelle ne ferait pas sens dans la vue.

La nouvelle dimension divise la vue en 12 volets : un pour chaque combinaison d'année et de segment. Cette vue est un exemple plus complexe d'une table imbriquée. Toute vue qui contient cette sorte de grille de graphiques individuels est appelée une vue **Petits multiples**.

À ce stade, vous ne souhaitez probablement pas rendre la vue plus granulaire ; en fait, vous souhaitez peut-être simplifier la vue en supprimant (c'est-à-dire, filtrant) certaines données.

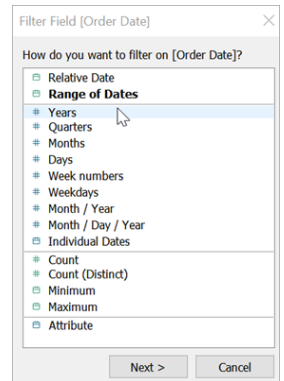


Filtrer la vue pour centrer votre exploration

Cette section explique comment vous pouvez centrer votre exploration de manière à n'afficher qu'un sous-ensemble de données, par exemple modifier la vue pour n'afficher que les données de commandes de 2012 et 2013.

1. Faites glisser la mesure **Order Date** du volet Données vers l'étagère Filtres.

2. Dans la boîte de dialogue Filtrer, choisissez le niveau de dates sur lequel vous souhaitez effectuer le filtrage—**Années**. Cliquez ensuite sur **Suivant**.
3. Dans le volet suivant, supprimez deux années que vous ne souhaitez pas inclure dans la vue.
4. Une fois que vous avez terminé, cliquez sur **OK**.



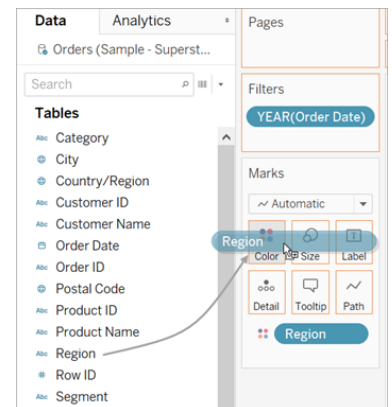
Vous allez ensuite augmenter le niveau de détail dans votre exploration de données en faisant glisser un champ vers Couleur sur la fiche Repères.

Utiliser la fiche Repères pour approfondir votre analyse

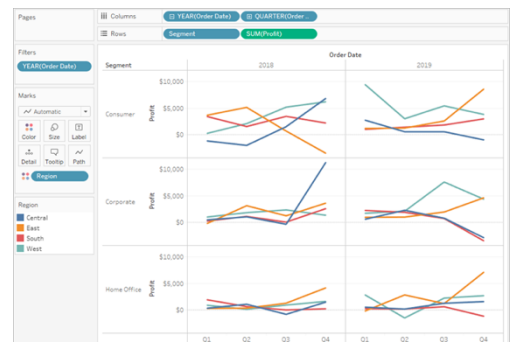
Cette étape vous montre comment modifier la vue pour colorer les repères basés sur la région.

Faites glisser la dimension **Region** du volet **Données** vers l'étagère **Couleurs**.

Le placement d'une dimension sur l'étagère des **couleurs** permet de séparer les repères en fonction des membres présents dans la dimension et d'attribuer une couleur unique à chacun d'entre eux. La légende des couleurs affiche chaque nom de membre et la couleur qui lui est associée.



Chaque volet contient désormais quatre lignes, soit une par région. La visualisation affiche désormais les données de bénéfices résumées au niveau de détail Région.



Bilan intermédiaire :

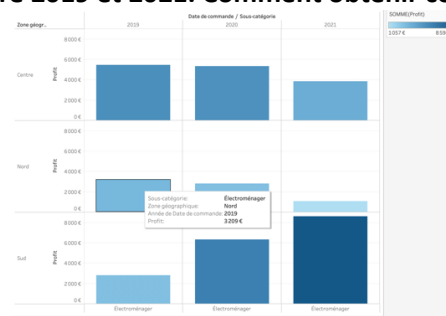
Vous devez maintenant comprendre les notions :

- Connexion aux données
- Création d'une vue
- Dimension et mesure / Discret et continue
- Hiérarchie
- **Agrégation sum() > approfondissement**
- Exploration de donnée (petits multiples)
- Repère (couleurs)
- Filtre

Exercice :

Connectez-vous à la BDD Supermarché, nous nous intéresserons à la table 'Achat'

1 - Vous cherchez à comparer les profits réalisés entre 2020 et 2021 par segment et zone géographique afin d'avoir une première vue de vos données. Comment obtenir cette vue ?



Niveau I – Approfondissement

Structure des données

Tableau Desktop fonctionne de manière optimale avec les données qui sont dans des tables formatées en tant que feuilles de calcul, c'est-à-dire des données stockées en lignes et en colonnes, avec des entêtes de colonne dans la première ligne. Comment devrait se présenter une ligne ou une colonne ?

Qu'est-ce qu'une ligne ?

Une ligne, ou un enregistrement, peut désigner tout élément depuis des d'informations sur une transaction dans un magasin de détail, des mesures météorologiques à un lieu précis jusqu'aux statistiques d'une publication sur les réseaux sociaux.

Il est important de savoir ce que représente un enregistrement (une ligne) dans les données. Il s'agit de la granularité des données.

Ici, chaque enregistrement est un jour

Date	Max TemperatureF	Mean TemperatureF	Min TemperatureF
1/1/2015	42	34	26
1/2/2015	42	37	32
1/3/2015	41	38	35
1/4/2015	51	45	38
1/5/2015	54	52	49
1/6/2015	54	49	43
1/7/2015	46	44	42
1/8/2015	46	41	35
1/9/2015	50	44	38
1/10/2015	46	45	43

Ici, chaque enregistrement est un mois

Date	Max TemperatureF	Mean TemperatureF	Min TemperatureF
January	63	45	26
February	62	49	33
March	69	51	31
April	77	52	37
May	82	59	43
June	92	68	49
July	95	71	54
August	92	69	54
September	81	61	45
October	74	58	45
November	60	44	25

Conseil : une bonne pratique consiste à avoir un identifiant unique, une valeur qui identifie chaque ligne comme un élément de données unique. Dans Superstore, ce serait Row ID (ID de ligne). Notez que tous les ensembles de données ne disposent pas d'un identifiant unique, mais cela ne peut pas faire de mal d'en avoir un.

Essayez de vérifier que vous pouvez répondre à la question « **Que représente une ligne dans l'ensemble de données ?** ». Si vous n'arrivez pas à articuler votre réponse, les données risquent d'être mal structurées pour l'analyse.

La granularité fait référence au degré de détail des données.

Qu'est-ce qu'un champ ou une colonne ?

Une *colonne* de données Tableau apparaît dans Tableau Desktop sous forme de *champ* dans le volet de données, mais il s'agit essentiellement de termes interchangeables. (Nous conservons le terme *colonne* dans Tableau Desktop pour l'utiliser dans l'étagère Colonnes et Lignes et pour décrire certaines visualisations). Un champ de données doit contenir des éléments qui peuvent être regroupés dans une relation plus large. Les éléments eux-mêmes sont appelés *valeurs* ou *membres* (seules les dimensions discrètes contiennent des membres).

Les valeurs autorisées dans un champ donné sont déterminées par le *domaine* du champ. Par exemple, une colonne pour les « rayons d'épicerie » pourrait contenir les membres « charcuterie », « boulangerie », « produits », etc., mais elle n'inclurait pas « pain » ou « salami » car ce sont des

articles et non des rayons. En d'autres termes, le domaine du champ des rayons est limité aux éventuels rayons d'épicerie.

En outre, un ensemble de données bien structuré comporterait une colonne pour « Ventes » et une colonne pour « Bénéfices », et non une seule colonne pour « Argent », car les bénéfices sont un concept distinct des ventes.

- Le domaine du champ « Ventes » intégrerait les valeurs ≥ 0 , puisque les ventes ne peuvent pas être négatives.
- Le domaine du champ « Bénéfices », cependant, comprendrait toutes les valeurs, puisque les bénéfices peuvent être négatifs.

Dimension et mesure

À propos des rôles et types de champs de données

Les champs de données sont constitués des colonnes de votre source de données. Chaque champ est automatiquement affecté à un type de données (par exemple entier, chaîne, date) et un rôle : dimension discrète ou mesure continue (plus courant) ou dimension continue ou mesure discrète (moins courant).

- Les *Dimensions* contiennent des valeurs qualitatives (par exemple noms, dates ou données géographiques). Vous pouvez utiliser des dimensions pour catégoriser, segmenter et révéler les détails dans vos données. Les dimensions affectent le niveau de détail dans la vue.
- Les *Mesures* contiennent des valeurs numériques, quantitatives que vous pouvez mesurer. Les mesures peuvent être agrégées. Lorsque vous faites glisser une mesure dans la vue, Tableau applique une agrégation à cette mesure (par défaut).

Champs bleus vs. Champs verts

Tableau représente les données différemment dans la vue selon que le champ est discret (bleu) ou continu (vert). *Continu* et *Discret* sont des termes mathématiques. Continu signifie « formant un tout continu, sans interruption ». Discret signifie « séparés individuellement et distincts ».

- Les mesures **SUM(Profit)** et les dimensions **YEAR(Order Date)** vertes sont continues. Les valeurs de champs continus sont traitées comme une plage infinie. Généralement, les champs continus ajoutent des axes à la vue.
- Les mesures **SUM(Profit)** et les dimensions **Product Name** bleues sont discrètes. Les valeurs discrètes sont traitées comme finies. Généralement, les champs discrets ajoutent des en-têtes à la vue.

Agrégation

Agrégation : désigne la manière dont plusieurs valeurs de données sont combinées en une seule valeur, comme prendre la moyenne de toutes les températures relevées autour de Seattle un jour donné.

Dans Tableau, vous pouvez agréger des mesures ou une dimension, bien qu'il soit plus courant d'agréger des mesures. Dès que vous ajoutez une mesure à votre vue, une agrégation est appliquée à cette mesure par défaut. Le type d'agrégation appliquée varie selon le contexte de la vue.

Modifier l'agrégation d'une mesure dans la vue

Lorsque vous ajoutez une mesure à la vue, Tableau agrège automatiquement ces données. La somme, la moyenne et la médiane sont des agrégations courantes.

L'agrégation active apparaît en tant que nom de la mesure dans la vue. Par exemple, **Sales** devient **SUM(Sales)**.

Vous pouvez modifier l'agrégation pour une mesure de la vue à partir du menu contextuel

