

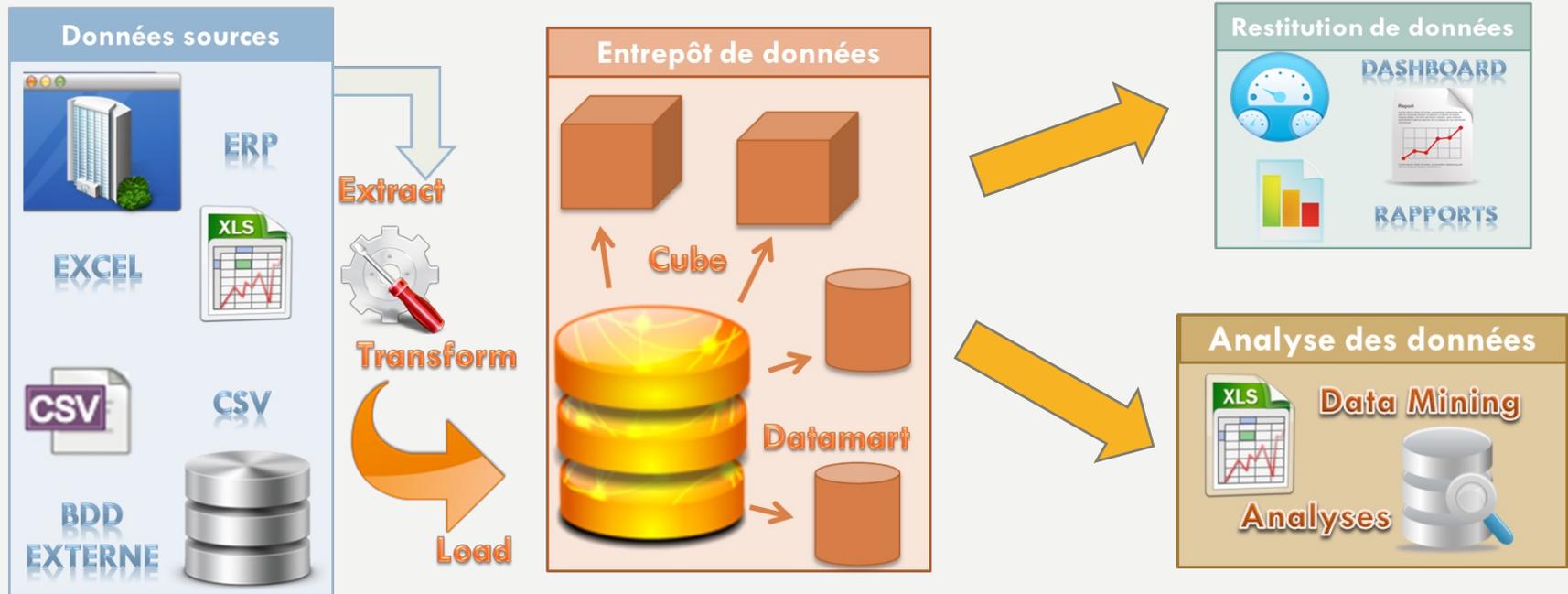
TP DE L'ETL À LA RESTITUTION

AUTEUR : JONATHAN FONTANEL

DATE : 29/11/2022

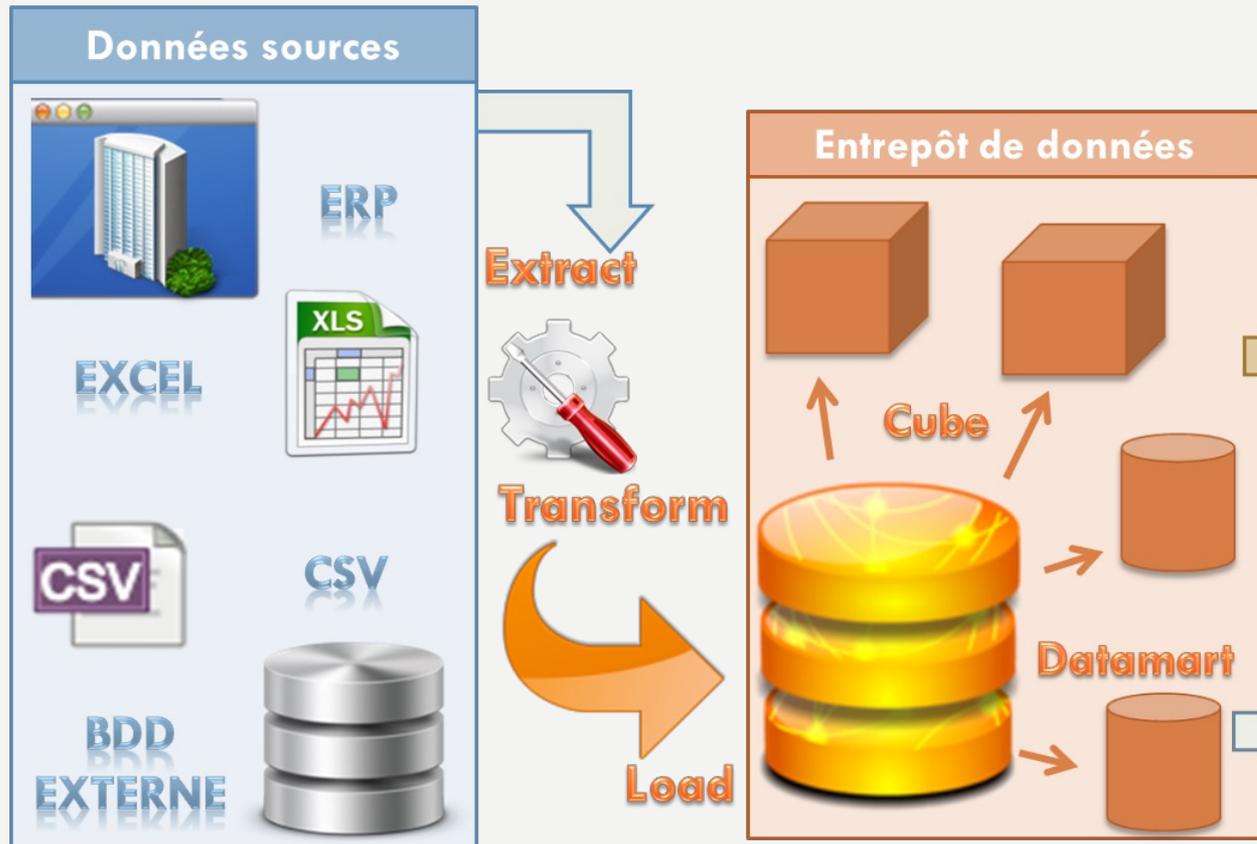
TP ISIMA ZZ2 F3 FONDEMENT DES BASES DE DONNÉES

DE L'ETL À LA RESTITUTION



LES DONNÉES SOURCES

- Des données provenant du SI de l'entreprise ou de l'extérieur
 - Interne : ERP, fichiers XLS, Systèmes Legacy, BDD spécifiques, ...
 - Externe : Flux EDI Fournisseurs, clients, ..



BUT DES 4 HEURES DE TP

- TP sur Tableau Desktop :
 - Se connecter et manipuler un datawarehouse
 - Comprendre le principe des tables de dimensions et de faits
 - Créations de visualisations
 - Création d'un tableau de bord
 - Option : Introduction au data storytelling
 - Débuter avec l'étude de cas sur l'entrepôt de données adventure work
- TP2 sur Tableau Desktop :
 - Terminer l'étude de cas sur l'entrepôt de données adventure work
 - Etude de cas sur l'implantation de magasins (Implémentation des starbucks dans le monde)
 - Etude de cas sur l'analyse des informations nutritionnelles (Cas des informations nutritionnelles du menu de McDonald's)



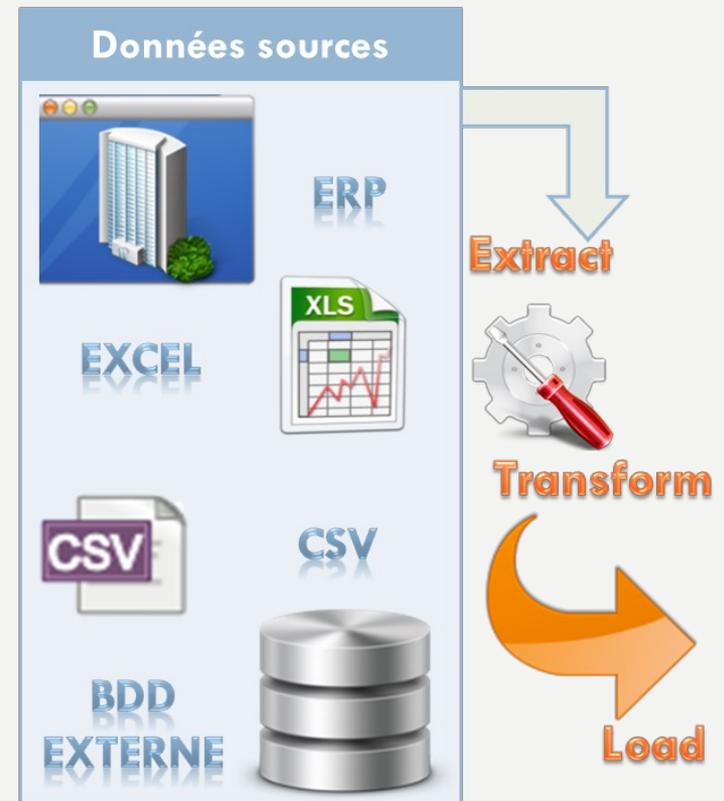
PRÉSENTATION DES OUTILS

ETL : EXTRACT, TRANSFORM LOAD

- Extraction de données de tous types (Excel, fichier plat ou CSV, BDD, etc...)
- Transformation de données (Création de colonnes, changement de type, Jointure de tables)
- Chargement de données (Chargement dans l'entrepôt de données ou des fichiers)

PRÉSENTATION DE TALEND OPEN STUDIO

- But d'un ETL : l'intégration de données
 - Intégration opérationnelle :
 - Partage de données entre applications avec transformation ou non
 - Flux temps réels entre applications ;
 - Réplication de base de données ;
 - Migration de données ;
 - Lien avec des partenaires (fournisseurs clients, ...)
 - ...
 - Intégration décisionnelle :
 - Charger des schémas du datawarehouse
 - Charger des datalakes
 - Transformation des données entre les différentes couches du datawarehouse
 - ...



MAGIC QUADRANT FOR DATA INTEGRATION TOOLS



As of July 2019 © Gartner, Inc



As of July 2020 © Gartner, Inc

PRÉSENTATION DE TALEND OPEN STUDIO

- Talend Open Studio for Data Integration : Outil Open Source pour l'intégration de données

Talend Open Source Solution GRATUITE pour tous les utilisateurs TÉLÉCHARGER MAINTENANT	Stitch Data Loader Version payante : De 100 à 1 000 USD par mois ESSAI GRATUIT 14 JOUR	Talend Pipeline Designer Tarif horaire en fonction de l'utilisation En savoir plus ESSAI GRATUIT 14 JOUR	Talend Cloud Data Integration* Abonnement mensuel : 1 170 USD par utilisateur Abonnement annuel : 12 000 USD par utilisateur (réduction de 15 % par rapport à l'abonnement mensuel) *Comprend des licences Talend Studio et Pipeline Designer. ESSAI GRATUIT	Talend Data Fabric* Inclut les aspects suivants : intégration des big data, gouvernance des données, intégration des applications, services de support client Platinum. *Comprend des licences Talend Studio et Pipeline Designer. CONTACTER LE SERVICE COMMERCIAL
--	--	---	--	---

PRÉSENTATION DE TABLEAU

Tableau Desktop

Tableau Server

Tableau Online

Tableau Mobile

Tableau Public



Analysez vos données
et créez vos tableaux de bords interactifs

Tableau Desktop



Tableau Reader

Partagez vos travaux
avec n'importe quels
utilisateurs



Tableau Server

Diffusez vos travaux
vers vos collaborateurs
et clients en mode web



Tableau Online

Profitez des capacités de
Tableau Server, sans
aucune installation



tableau public

Publiez vos travaux
vers le grand public

PRÉSENTATION DE TABLEAU

Tableau Desktop

Tableau Desktop vous permet de créer en quelques clics des **tableaux de bord interactifs** à partir de n'importe quelles sources de données. Le produit exploite au maximum la capacité naturelle de chacun à penser de manière visuelle. Pour répondre à des questions complexes il vous suffit de créer une série de vues simples que **vous assemblerez sous forme de tableaux de bord en quelques clics**. Vous suivez ainsi votre raisonnement naturel et n'êtes jamais coincé par des assistants d'exécution ni obligé de rédiger des scripts. Et tout cela 10 à 100 x plus rapidement que les solutions concurrentes.

Tableau Server

Tableau Server vous permet de **diffuser très facilement en mode web** vers vos collaborateurs, clients ou partenaires les analyses et tableaux de bords que vous aurez préalablement créés via Tableau Desktop. Pour cela, il vous suffit de publier en quelques clics vos classeurs sur le serveur en choisissant les destinataires et le niveau de sécurité souhaité. Les documents sont alors immédiatement disponibles et consultables via n'importe quel navigateur web sur leur PC et via une **application gratuite sur iPad et tablette Android**. Aucune installation n'est nécessaire sur les postes des utilisateurs.

Tableau Reader

Tableau Reader est une **application gratuite** à installer sur PC ou Mac permettant de partager facilement des visualisations de données. Les utilisateurs **ne disposant pas de la version complète** de Tableau Desktop peuvent néanmoins lire des classeurs existants, et **intéragir avec les données**.

Tableau Online

Tableau Online est la **version Cloud** de Tableau Server. Elle en offre les mêmes fonctions, tout en y ajoutant la souplesse et la flexibilité d'une version hébergée. Ainsi, les données et tableaux de bords peuvent être instantanément partagés et lus sur PC et tablettes, tout ceci de façon sécurisée. Optez pour la puissance de Tableau sans aucune gestion d'infrastructure.

+tableau⁺public

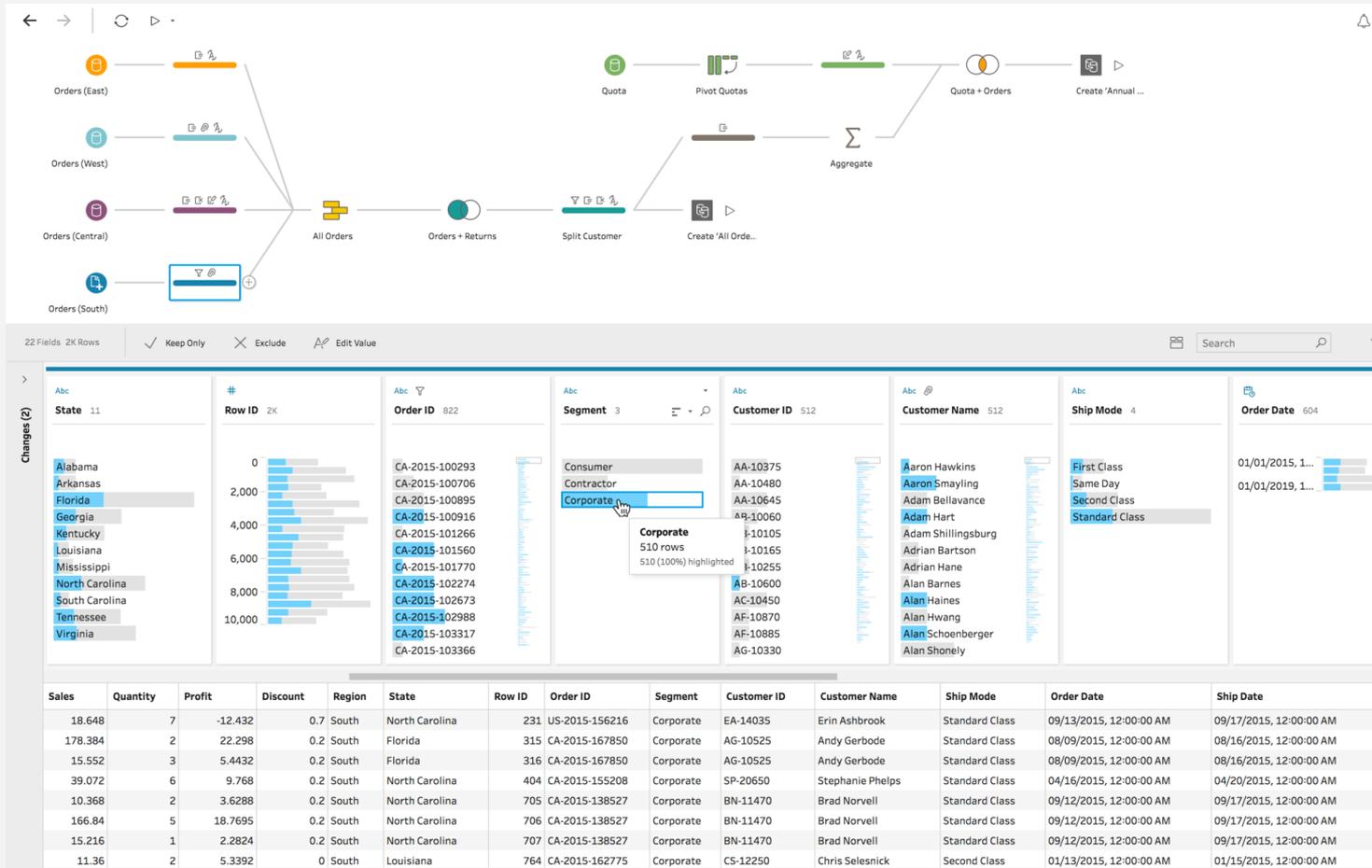
Tableau Public est gratuit et rend accessible à tous l'**analyse et la visualisation des données**. Il permet de créer facilement des contenus interactifs pour faire raconter des histoires aux données, et **publier facilement ces visualisations sur le web**.

Remarque : tout ce qui est publié sur Tableau Public est accessible et libre d'accès sur le web (classeurs et données) !

- Source : <http://www.olap-partner.fr/tableau-software/tableau-software--logiciels-de-business-intelligence/>

PRÉSENTATION DE TABLEAU

- Nouveau produit :



- Source : <https://www.tableau.com/products/prep>

PRÉSENTATION DE TABLEAU



PRÉSENTATION DE TABLEAU



Source: Gartner (February 2019)



Source: Gartner (February 2020)

PRÉSENTATION DE TABLEAU

- Un extrait du Gartner 2019 :
- By 2020, augmented analytics — a paradigm that includes natural language query and narration, augmented data preparation, automated advanced analytics and visual-based data discovery capabilities — will be a dominant driver of new purchases of business intelligence, analytics and data science and machine learning platforms and of embedded analytics.
- By 2020, the number of users of modern business intelligence and analytics platforms that are differentiated by augmented data discovery capabilities will grow at twice the rate — and deliver twice the business value — of those that are not.
- By 2020, natural-language generation and artificial intelligence will be a standard feature of 90% of modern business intelligence platforms.
- By 2020, 50% of analytical queries will be generated via search, natural-language processing or voice, or will be automatically generated.
- By 2020, organizations that offer users access to a curated catalog of internal and external data will derive twice as much business value from analytics investments as those that do not.
- Through 2020, the number of citizen data scientists will grow five times faster than the number of expert data scientists.
- Sources : <https://www.gartner.com/doc/reprints?id=I-4JT7ECY&ct=I71102&st=sb>

PRÉSENTATION DE TABLEAU

- Un extrait du Gartner 2020 :
- By 2022, augmented analytics technology will be ubiquitous, but only 10% of analysts will use its full potential.
- By 2022, 40% of machine learning model development and scoring will be done in products that do not have machine learning as their primary goal.
- By 2023, 90% the world's top 500 companies will have converged analytics governance into broader data and analytics governance initiatives.
- By 2025, 80% of consumer or industrial products containing electronics will incorporate on-device analytics.
- By 2025, data stories will be the most widespread way of consuming analytics, and 75% of stories will be automatically generated using augmented analytics techniques.

PRÉSENTATION DE TABLEAU

	Year Ended December 31,				
	2016	2015	2014	2013	2012
	(in thousands, except per share data)				
Consolidated Statements of Operations Data:					
Revenues					
License	\$ 481,659	\$ 423,766	\$ 279,944	\$ 159,930	\$ 89,883
Maintenance and services	345,284	229,821	132,672	72,510	37,850
Total revenues	826,943	653,587	412,616	232,440	127,733
Cost of revenues					
License	7,003	3,852	1,211	740	305
Maintenance and services	92,087	69,833	35,774	17,784	10,057
Total cost of revenues (1)	99,090	73,685	36,985	18,524	10,362
Gross profit	727,853	579,902	375,631	213,916	117,371
Operating expenses					
Sales and marketing (1)	476,506	356,723	216,672	123,573	62,333
Research and development (1)	302,759	204,131	110,923	60,769	33,065
General and administrative (1)	88,149	71,078	41,712	25,905	17,715
Total operating expenses	867,414	631,932	369,307	210,247	113,113
Operating income (loss)	(139,561)	(52,030)	6,324	3,669	4,258
Other income (expense), net	2,134	1,223	858	(804)	(54)
Income (loss) before income tax expense (benefit)	(137,427)	(50,807)	7,182	2,865	4,204
Income tax expense (benefit)	7,022	32,893	1,309	(4,211)	2,777
Net income (loss)	\$ (144,449)	\$ (83,700)	\$ 5,873	\$ 7,076	\$ 1,427

- Source :

http://sl.q4cdn.com/149179428/files/doc_financials/2016/FY2016-Annual-Report.pdf

PRÉSENTATION DE TABLEAU

- Salesforce rachète Tableau Software, leader dans la visualisation des données. Montant : 15,7 milliards \$ par échange d'actions.
- Grosse opération de concentration dans le domaine de l'analytics : Salesforce achète Tableau Software, par échange d'actions, pour un montant de 15,7 milliards \$. Soit une prime de 42 % sur le cours moyen de l'action de Tableau Software coté au New York Stock Exchange.
- Selon les chiffres communiqués, l'opération devrait augmenter le chiffre d'affaires de Salesforce d'environ 350 à 400 millions \$ pour l'exercice 2020. En 2018, Tableau a réalisé un chiffre d'affaires proche de 1,2 milliards \$ sans atteindre la rentabilité opérationnelle.
- Source :

<https://www.silicon.fr/analytics-salesforce-achete-tableau-software-243615.html>



PLACE À LA PRATIQUE

TABLEAU DESKTOP

TODO

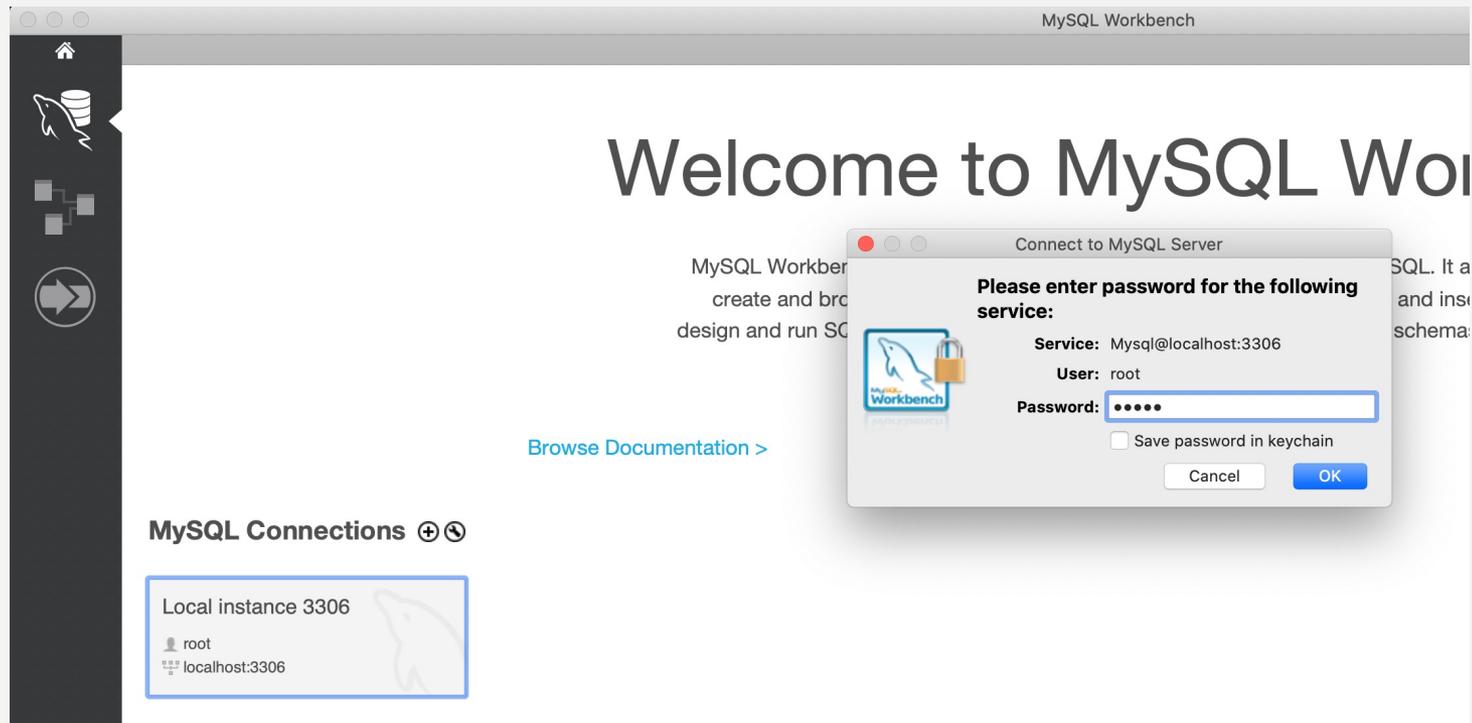
- Installation de Tableau Desktop (lien par mail de la part du support Tableau)
- Restaurer le datawarehouse AdventureWorkDW
- Connexion à votre entrepôt de données
- Création d'un premier graphique
- Drill dans les données
- Rajouter un premier filtre
- Rajouter de la couleur en fonction d'une autre métrique
- Rajouter une info bulle
- Créer une combinaison
- Créer une combinaison de graphique
- Filtrer les données
- Utiliser les noms et valeurs des mesures
- Créer une bande de référence
- Création d'un paramètre et d'un champ calculé

TO DO

- Etude de cas avec questions et création de visualisations, de tableau de bord
- Découverte du data storytelling
- Bonnes pratiques sur les visualisations

LANCER MYSQL

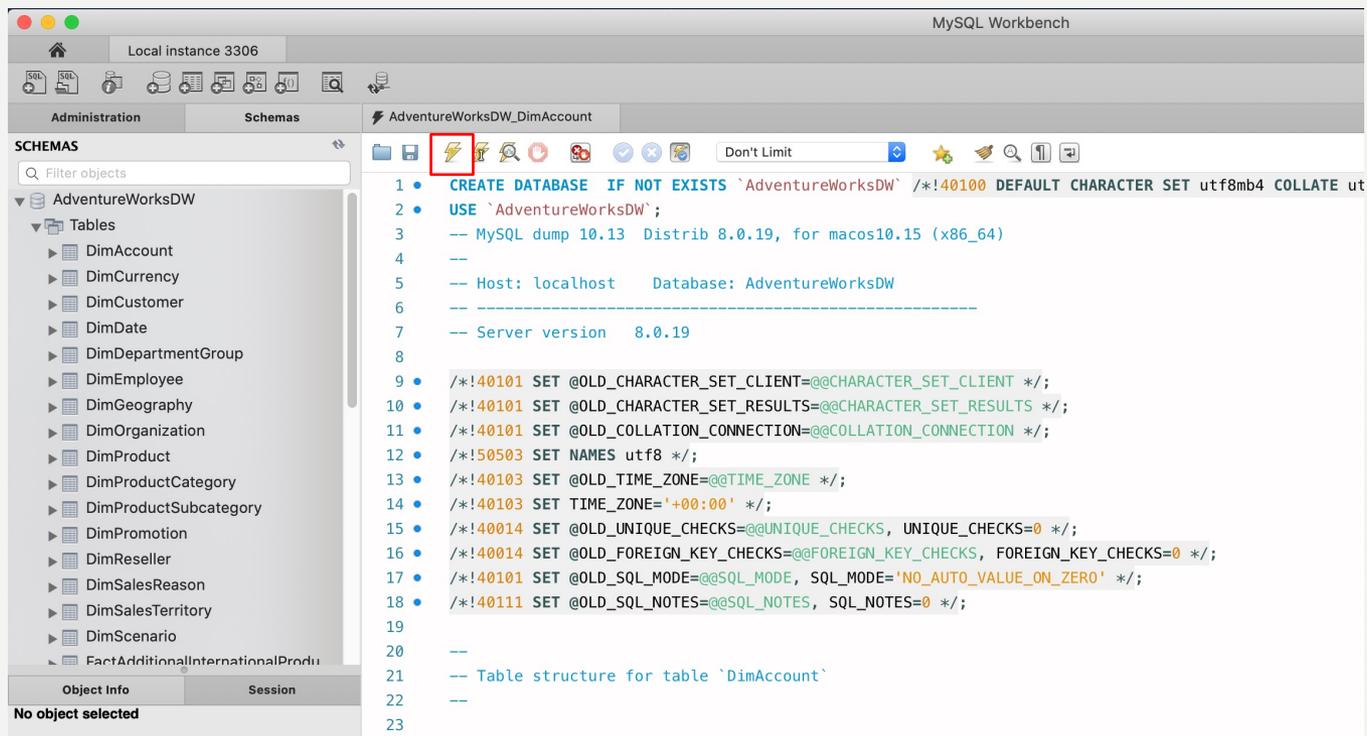
- Se rendre dans services.msc via le menu démarrer
- Puis démarrer le service : MySQL si pas démarré
- Ouvrir MySQLWorkbench et se connecter en tant que « root »



RESTAURER LE DATAWAREHOUSE

23

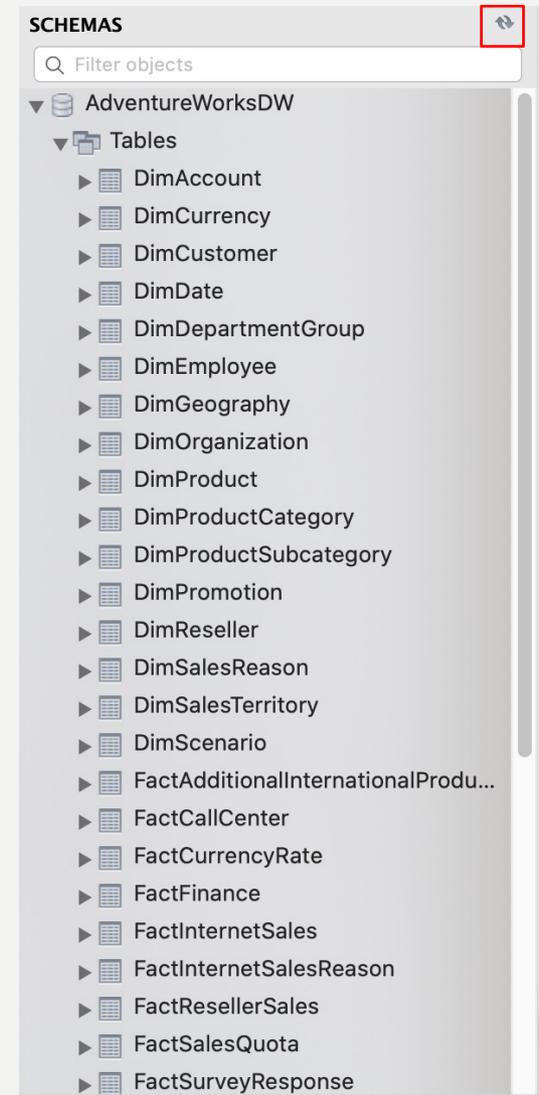
- Restauration d'un entrepôt de données à partir des scripts MySQL :
 - Télécharger le dossier zip contenant les scripts sur mon perso.isima.fr
 - Puis ouvrir chaque script avec MySQL et cliquer sur éclair (cf. carré rouge sur la figure ci-dessous)



RESTAURER LE DATAWAREHOUSE

24

- Contrôle en actualisant les schémas à gauche (cf. carré rouge sur la figure ci-dessous)
- Vous devez voir toutes les tables du datawarehouse
- Vous pouvez consulter le schéma PDF pour comprendre les relations entre les tables



CONNEXION À VOTRE ENTREPÔT DE DONNÉES

- Ouvrir Tableau Desktop
- Installer votre licence étudiant
- Cliquer sur MySQL

Tableau - Classeur2
Fichier Données Serveur Aide

Connexion

Dans un fichier

- Microsoft Excel
- Fichier texte
- Fichier JSON
- Microsoft Access
- Fichier PDF
- Fichier de données spatiales
- Fichier de statistiques
- Plus...

Sur un serveur

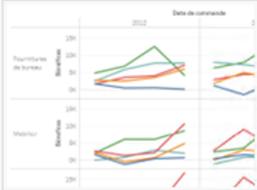
- Tableau Server
- MySQL**
- Oracle
- Amazon Redshift
- Microsoft SQL Server
- Plus... >

Ouvrir

Prep_ZZ2F3

Ballard		
M		72
R		18
S		83
T		13
Barstow	R	4
Basingstoke Hants	M	10
R		30
S		40
T		20

TP5_2018



Starbuck



MCdo

Inclus	Breakfast	Big Breakfast with
		Big Breakfast with
Chicken & Fish		Chicken McNugget
Smoothies & Shakes		Chocolate Shake (L)
		McFlurry with M&M
		Strawberry Shake

CONNEXION À VOTRE ENTREPÔT DE DONNÉES

- Télécharger et installer les pilotes si cela n'a pas déjà été fait en cliquant sur le lien en bas de l'écran de connexion

my

Alibaba AnalyticDB for MySQL

Amazon Aurora pour MySQL

MySQL

MySQL

Serveur : localhost Port : 3306

Base de données : AdventureWorksDW

Entrez les informations de connexion à la base de données :

Nom d'utilisateur : root

Mot de passe :

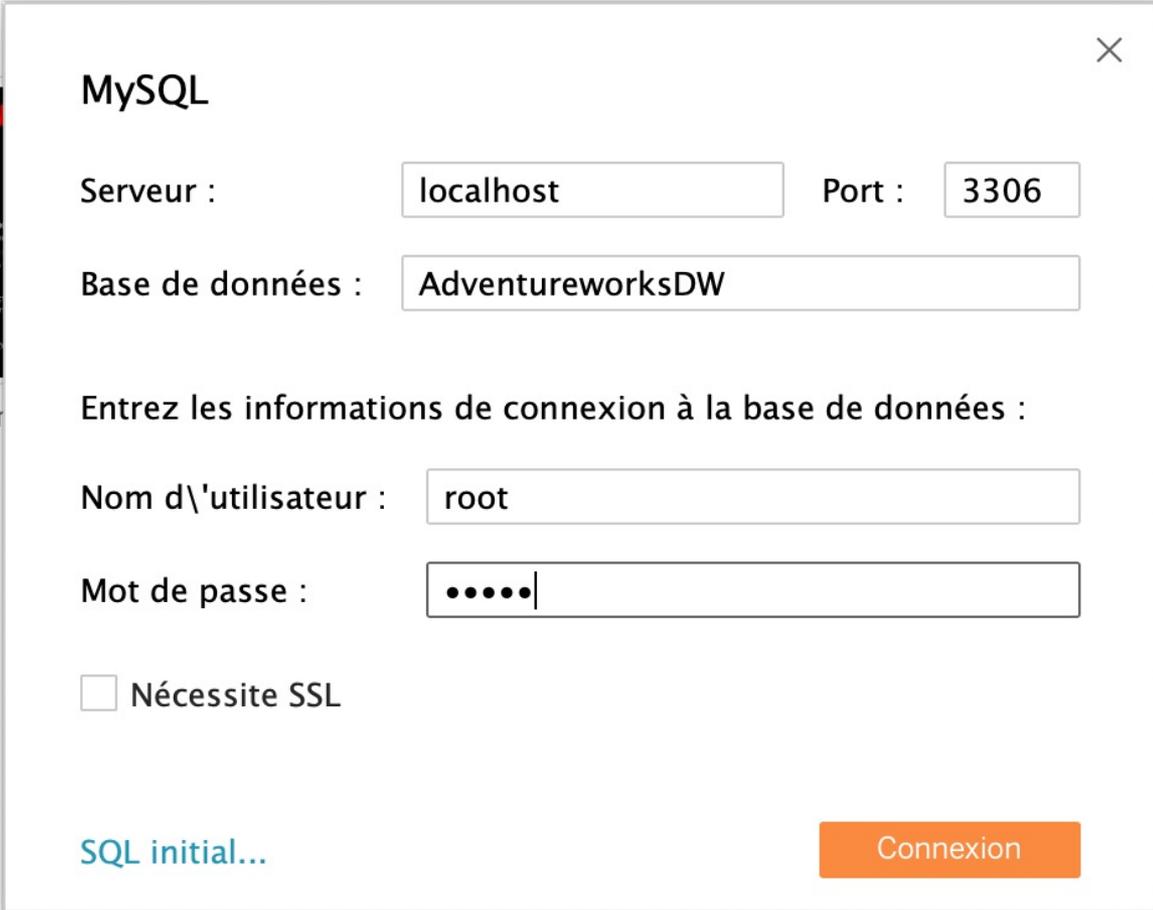
Nécessite SSL

[SQL initial...](#)

[Téléchargez et installez les pilotes](#), puis connectez-vous.

CONNEXION À VOTRE ENTREPÔT DE DONNÉES

- Puis saisir les paramètres suivants en fonction du nom de votre base de données et du user/mdp que vous avez utilisé



A screenshot of a MySQL connection dialog box. The window title is "MySQL" with a close button (X) in the top right corner. The dialog contains several input fields and a checkbox. The "Serveur" field is set to "localhost" and the "Port" field is set to "3306". The "Base de données" field is set to "AdventureworksDW". Below these fields, there is a prompt "Entrez les informations de connexion à la base de données :". The "Nom d'utilisateur" field is set to "root" and the "Mot de passe" field is masked with five dots. At the bottom left, there is a checkbox labeled "Nécessite SSL" which is currently unchecked. At the bottom right, there is an orange button labeled "Connexion". At the bottom left, there is a blue link labeled "SQL initial...".

MySQL

Serveur : localhost Port : 3306

Base de données : AdventureworksDW

Entrez les informations de connexion à la base de données :

Nom d'utilisateur : root

Mot de passe : ●●●●●

Nécessite SSL

SQL initial... Connexion

CONNEXION À VOTRE ENTREPÔT DE DONNÉES

- Faire glisser les tables du côté gauche vers le haut de la page

The screenshot shows the Tableau interface with the following elements:

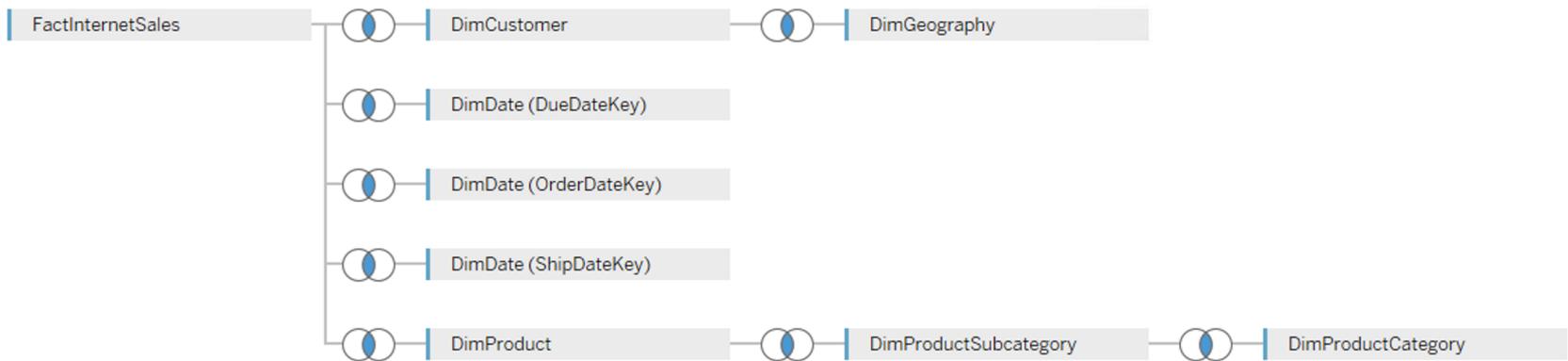
- Tableau - Classeur2** (Title bar)
- Fichier Données Serveur Fenêtre Aide** (Menu bar)
- Connexions**: A210PC03 (Microsoft SQL Server)
- Base de données**: AdventureWorksDW2008R2
- Table** (List of tables):
 - AdventureWor...BuildVersion
 - DatabaseLog
 - DimAccount
 - DimCurrency
 - DimCustomer
 - DimDate
 - DimDepartmentGroup
 - DimEmployee
 - DimGeography
 - DimOrganization
 - DimProduct
 - DimProductCategory
 - DimProductSubcategory
 - DimPromotion
 - DimReseller
 - DimSalesReason
 - DimSalesTerritory
 - DimScenario
 - FactAdditiona...ctDescription
 - FactCallCenter
 - FactCurrencyRate
 - FactFinance
 - FactInternetSales
 - FactInternetSalesReason
 - Nouvelle requête SQL personnalisée
 - Nouvelle union
- Workspace**:
 - AdventureWorksDW2008R2 (View name)
 - Faites glisser des tables ici* (Instruction)
 - Trier les champs (Sort fields)
 - Ordre de la source de données (Data source order)
 - Afficher les alias (Show aliases)
 - Afficher les champs masqués (Show hidden fields)
 - lignes (Rows)

CONNEXION À VOTRE ENTREPÔT DE DONNÉES

- Glisser les tables afin d'obtenir ce schéma

DimProduct+ (AdventureWorksDW2008R2)

Connexion
 En direct Extrait



- Les relations sont « internes »
- Essayer de trouver les jointures entre les tables en naviguant avec SQL Server Management Studio

CREATION D'UN PREMIER GRAPHIQUE

- Cliquer sur Feuille 1 en bas dans les onglets, vous avez les dimensions et les mesures qui sont dans la première colonne à gauche

The screenshot displays the Tableau interface. On the left, the 'Données' pane is active, showing a list of dimensions and measures. The dimensions include 'Catégorie', 'Code postal', 'Date d'expédition', 'Date de commande', 'Pays, Ville' (expanded to 'Pays' and 'Ville'), 'État', 'ID de client', 'ID de commande', 'ID produit', 'Marché', 'Mode d'expédition', 'Nom du client', and 'Nom du produit'. The measures include 'Bénéfices', 'Frais d'expédition', 'ID de ligne', 'Quantité', 'Remise', 'Ventes', and two generated fields: 'Latitude (générée)' and 'Longitude (générée)'. The central workspace shows a 'Feuille 1' tab and a large area with the text 'Déposer champ ici'. The right-hand side features a 'Montre-moi' pane with various visualization options and a legend: 'Sélectionner ou faire glisser des données' and 'Utilisez les touches Maj ou Cmd pour sélectionner plusieurs champs'.

CREATION D'UN PREMIER GRAPHIQUE

- Faire glisser dans Lignes le champ « Full Date Alternate Key » de la dimension DimDate pour la jointure « OrderDateKey ».
- Faire glisser la mesure « Sales Amount » sur les « Abc » de la feuille

Feuille 10

Année de F..	
2005	Abc
2006	Abc
2007	Abc
2008	Abc

The screenshot shows the Tableau Desktop interface for a workbook named 'Tableau - Prep_ZZ2F3'. The main view is a pivot table on 'Feuille 10'. The 'Lignes' (Rows) shelf contains the field 'ANNÉE(Full Date ..)'. The 'Colonnes' (Columns) shelf is empty. The 'Filtres' (Filters) shelf is empty. The 'Repères' (Marks) shelf is set to 'Automatique'. The 'Mesures' (Measures) shelf contains the field 'SOMME(Sales ..)'. The 'Dimensions' pane on the left shows the hierarchy: DimCustomer, DimDate (DueDateKey), and DimDate (OrderDateKey). The 'Mesures' pane on the right lists various measures, with 'Sales Amount' highlighted. The pivot table data is as follows:

Année de F..	
2005	3 266 374
2006	6 530 344
2007	9 791 060
2008	9 770 900

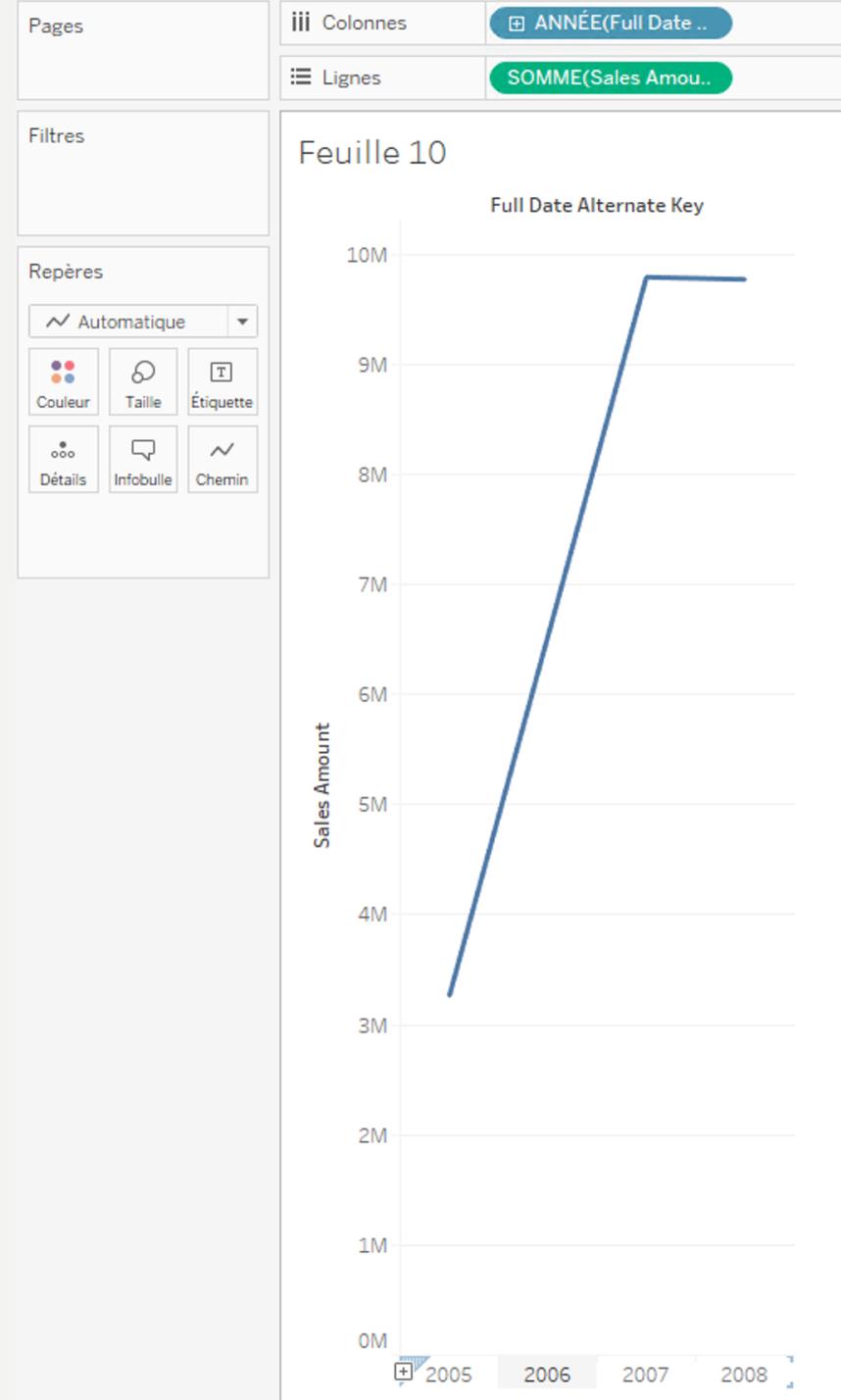
CREATION D'UN PREMIER GRAPHIQUE

- Il est possible de pouvoir choisir plusieurs ruptures pour le tableau.

Pages		iii Colones								
Filtres		Lignes	ANNÉE(Full Date ..	TRIMESTRE(Full D..	MOIS(Full Date Alt..					
Repères		Feuille 10								
Automatique		Année de F..	Trimestre d..	Mois de Full..						
Couleur	Taille	Texte	Détails	Infobulle	SOMME(Sales ..	^	2005	T3	juillet	473 388
							août	506 192		
							septembre	473 943		
							T4	octobre	513 329	
							novembre	543 993		
							décembre	755 528		
							2006	T1	janvier	596 747
							février	550 817		
							mars	644 135		
							T2	avril	663 692	
							mai	673 556		
							juin	676 764		
T3	juillet	500 365								
août	546 001									
septembre	350 467									
T4	octobre	415 390								
novembre	335 095									
décembre	577 314									
2007	T1	janvier	438 865							
février	489 090									
mars	485 575									
T2	avril	506 399								
mai	562 773									
juin	554 799									
T3	juillet	886 669								
août	847 414									
septembre	1 010 258									
T4	octobre	1 080 450								
novembre	1 196 981									
décembre	1 731 788									

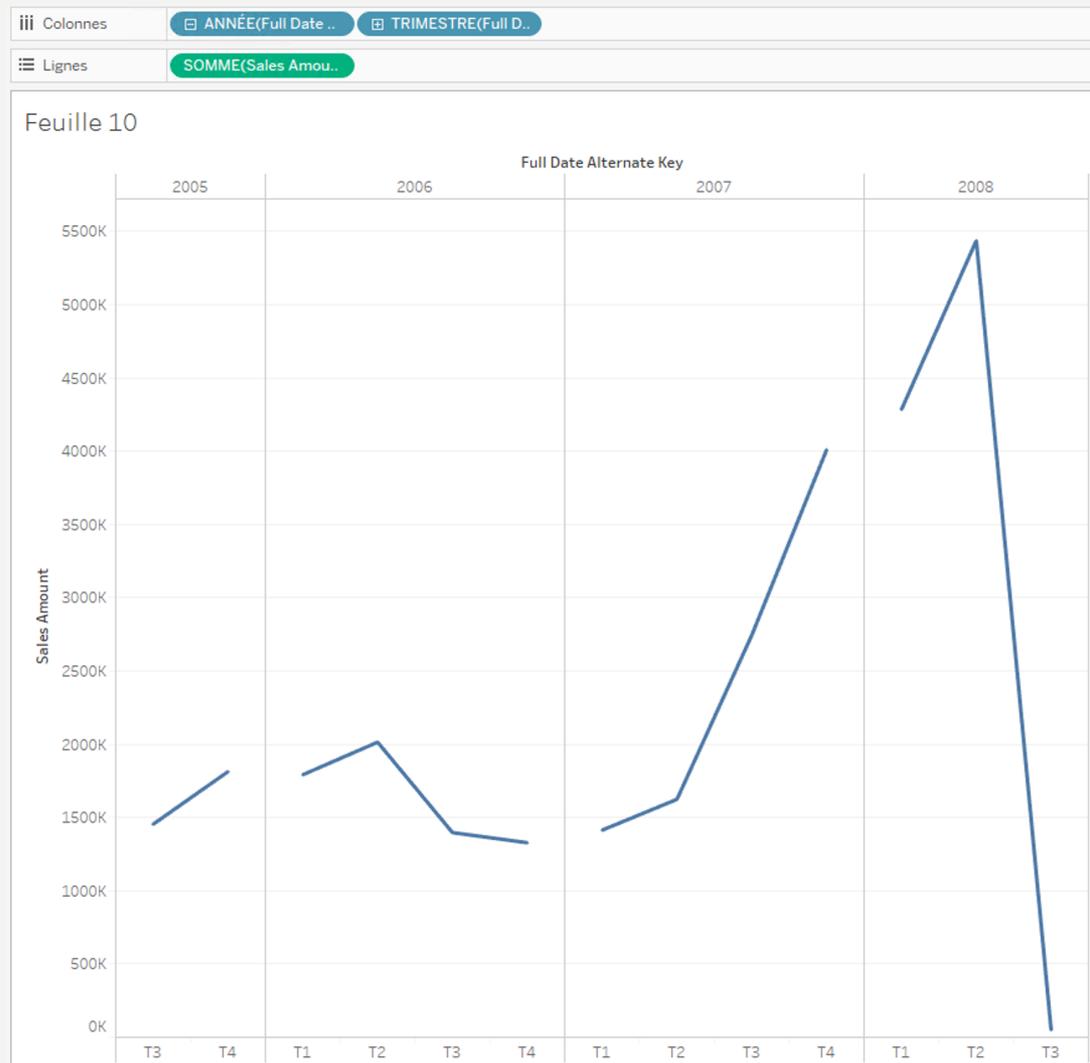
CREATION D'UN PREMIER GRAPHIQUE

- Créer une nouvelle feuille
- Faire glisser le champ « Sales Amount" dans Lignes et FullDate Alternate Key dans Colonnes
- Qu'est ce que l'on obtient ?
- Tableau interprète automatiquement la mesure avec la fonction agrégat "Somme".



DRILL DANS LES DONNÉES

- En cliquant sur le + à gauche de "ANNEE", vous allez faire apparaître les trimestres



RAJOUTER UN PREMIER FILTRE

- Faire glisser ANNEE dans Filtres sur la partie gauche
- Sélectionner Années puis cliquer sur Suivant
- Sélectionner uniquement les commandes de 2007 et 2008

The screenshot shows an Excel PivotTable with the following configuration:

- Pages:** (empty)
- Colonnes:** ANNEE(Full Date ..), TRIMESTRE(Full D..)
- Lignes:** SOMME(Sales Amou..)
- Feuille 10:** PivotTable with columns for years (2005, 2006, 2007) and rows for Sales Amount (2000K to 5500K).

Two dialog boxes are overlaid on the PivotTable:

Dialog 1: Filtrer [Année de Full Date Alternate Key]

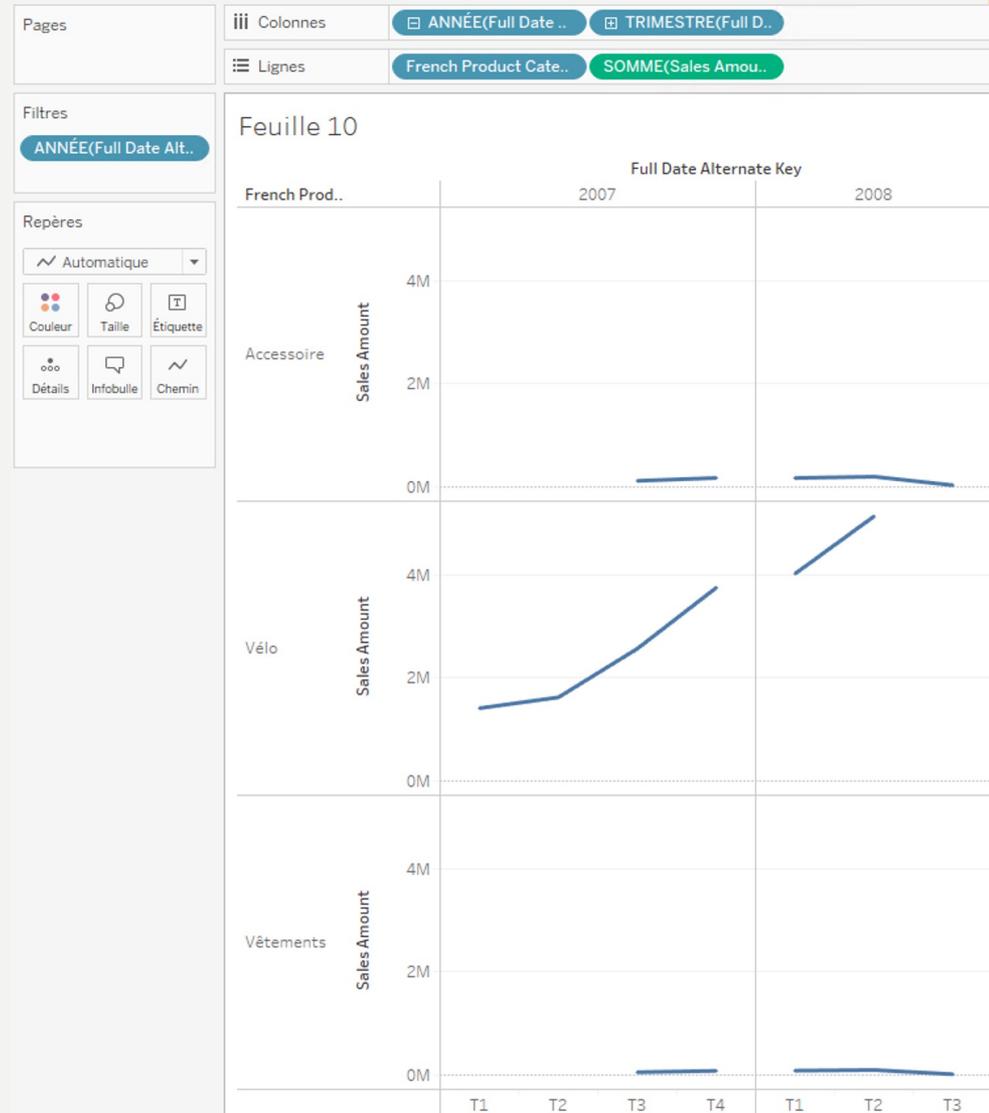
- Tab: Général
- Radio buttons: Sélectionner dans la liste, Liste de valeurs
- Search box: Saisir le txt de recherche
- List of years: 2005, 2006, 2007, 2008. 2007 and 2008 are checked.

Dialog 2: Filtrer le champ [Full Date Alternate Key]

- Question: Quel filtre voulez-vous appliquer à [Full Date Alternate Key] ?
- Categories: Date relative, Plage de dates, Dates individuelles, Total, Minimum, Maximum, Attribut.
- Selected options: Années, Total (Distinct), Minimum.
- Buttons: Suivant >, Annuler.

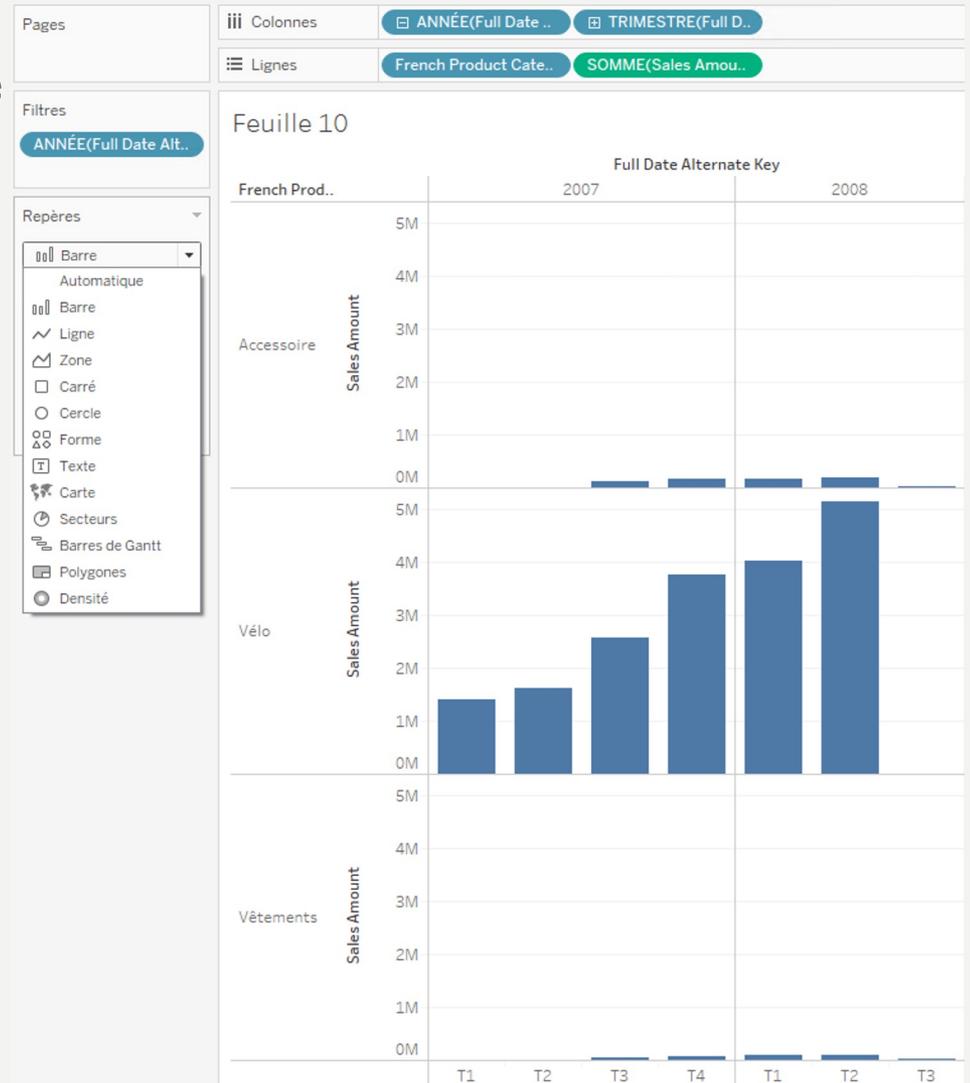
DRILL DANS LES DONNÉES

- Ajouter des données en faisant glisser « French Product Category Name » à gauche de « Sales Amount »
- On peut voir que le graphique s'est découpé en 6 panneaux



DRILL DANS LES DONNÉES

- Changer le type de graphique en transformant en histogramme
- Quelle catégorie a connu la plus belle progression ?



RAJOUTER DE LA COULEUR EN FONCTION D'UNE AUTRE METRIQUE

- Faire glisser « French Country Region Name » de la dimension « DimGeography » sur Couleur et observer ce qui se passe
- A quoi correspond cet attribut de cette dimension ? Que peut-on voir ? Quelle information peut-on rajouter sur la hausse ?



RAJOUTER UNE INFO-BULLE

- Cliquer sur "Info Bulle" afin d'ajouter une infobulle sur les données affichées

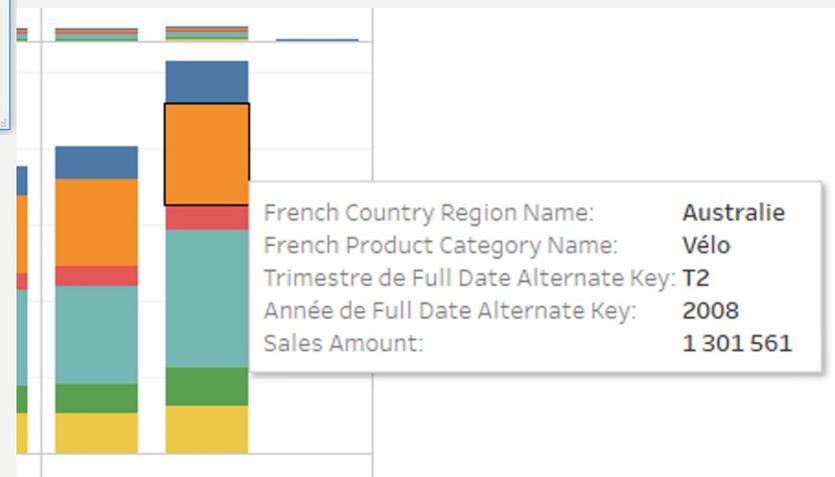
The screenshot shows the Tableau interface with a pivot table. The columns are 'Full Date Alternate Key' (2007, 2008) and the rows are 'French Product Category Name' (Accessoire, Vélo). A 'Modifier l'infobulle' dialog box is open, displaying the following text:

Tableau Book | 10 | B | I | U | [Color] | [Align] | [Format] | Insérer | X

French Country Region Name: <French Country Region Name>
French Product Category Name: <French Product Category Name>
Trimestre de Full Date Alternate Key: <TRIMESTRE(Full Date Alternate Key)>
Année de Full Date Alternate Key: <ANNÉE(Full Date Alternate Key)>
Sales Amount: <SOMME(Sales Amount)>

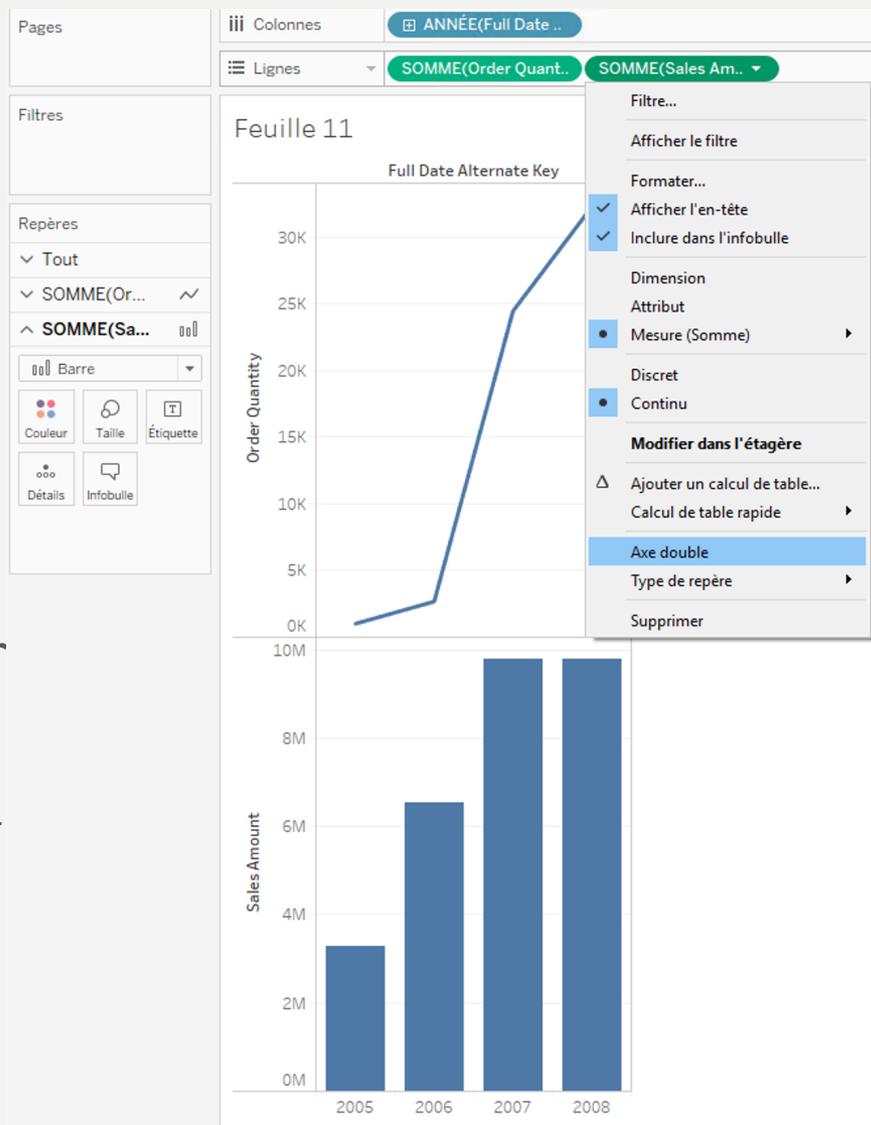
Afficher les infobulles Réactif - Afficher les infobulles instantanément
 Inclure les boutons de commande
 Autoriser la sélection par catégorie

Réinitialiser Aperçu OK Annuler



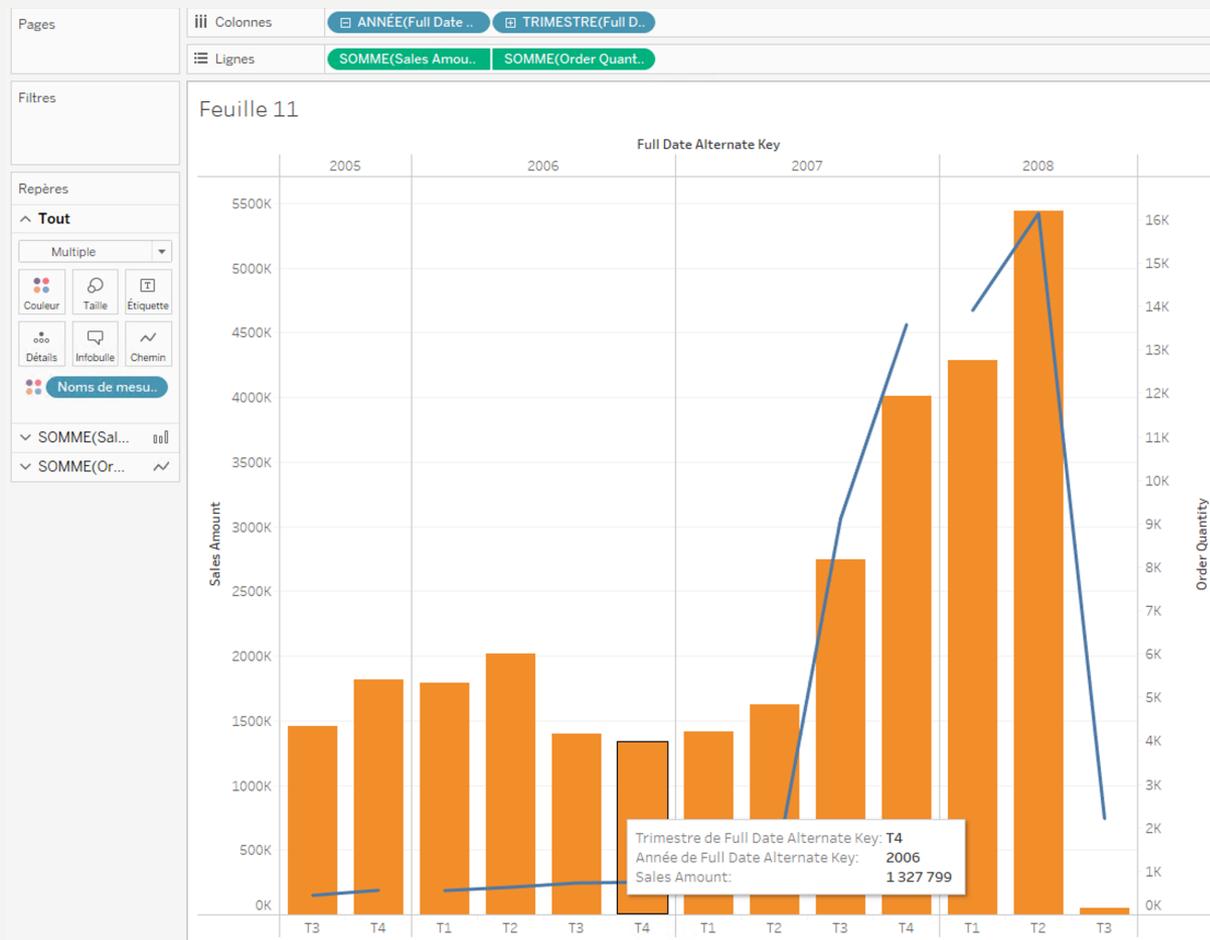
CRÉER UNE COMBINAISON DE GRAPHIQUE

- Créer un nouvelle feuille puis mettre en colonne l'attribut FullDate Alternate Key de la dimension DimDate (OrderDateKey) et en ligne les mesures suivantes « Order Quantity» et « Sales Amount ».
- Cliquer dans la colonne de droite sur "somme Sales Amount" puis changer le type de graphique pour mettre Barre
- Cliquer sur la flèche à droite de la métrique Bénéfice dans les lignes pour sélectionner "Axe Double"



CRÉER UNE COMBINAISON DE GRAPHIQUE

- Echanger les deux métriques afin d'avoir la courbe « Order Quantity » devant les barres.
- Descendre dans les trimestres pour comprendre l'évolution



FILTRE LES DONNÉES

- Créer le tableau suivant permettant d'analyser les ventes (Sales Amount) par pays/région et en fonction de la catégorie et de la sous catégorie

Pages

iii Colonnes **French Country Regi..** **State Province Name**

Lignes **French Product Cate..** **French Product Subc..**

Filtres

Feuille 12

French Country Region Name / State Province Name

French Product Cat..	French Product Subcategory Name	Allemagne						Australie					
		Bayern	Brandenbu..	Hamburg	Hessen	Nordrhein-..	Saarland	New South Wales	Queensland	South Australia	Tasmania	Victoria	Alberta
Accessoire	Bidon et porte-bid...	670	107	1 041	1 193	1 229	1 176	4 972	2 554	666	282	2 771	28
	Casque	3 359	455	3 779	5 563	4 899	5 703	19 734	9 692	3 079	1 540	10 917	35
	Garde-boue	528	22	637	571	835	945	2 967	1 714	462	242	1 758	22
	Nettoyant	40		95	111	151	111	819	310	103	48	429	
	Pneu et chambre à ...	2 569	197	3 337	4 365	5 068	5 325	20 314	10 229	2 566	1 606	12 078	90
	Porte-vélo	360		360	600	720	840	2 040	1 560	600		1 800	240
	Range-vélo	159		318	477	1 113	1 113	5 088	3 021	318		159	1 749
Sac d'hydratation	165		220	440	825	440	3 904	2 475	1 045	165	2 914		
Vélo	Vélo de randonnée	56 300	8 367	75 372	100 316	73 395	110 620	420 648	238 081	59 964	27 283	247 706	5 510
	Vélo de route	191 253	32 420	229 436	330 525	257 367	339 341	2 203 910	1 062 537	344 305	135 212	1 258 584	14 022
	VTT	141 783	15 720	160 819	213 590	215 009	256 880	1 219 303	642 065	200 786	70 702	720 963	2 295
Vêtements	Casquette	369	72	333	530	494	692	1 762	818	288	117	827	18
	Chaussettes	54	27	45	72	63	54	450	207	108	63	180	9
	Cuissards	70		140	70	70	280	4 129	2 520	840	280	3 010	70
	Gants	171	49	318	539	563	686	3 478	1 714	416	171	1 984	24
	Maillot	1 864	420	2 368	2 568	3 741	4 177	17 983	7 585	2 266	1 688	10 140	104
	Veste	254	64	508	572	572	699	2 985	1 334	445	381	2 096	

Repères

Automatique

Couleur Taille Texte

Détails Infobulle

SOMME(Sales ..)

FILTRE LES DONNÉES

- Puis cliquer en tenant « ctrl » sur la ligne Vélo et la colonne South Australia

Pages

iii Colonnes **French Country Regi..** **State Province Name**

Lignes **French Product Cate..** **French Product Subc..**

Filtres

Feuille 12

Repères

Automatique

Couleur Taille Texte

Détails Infobulle

SOMME(Sales ..)

		French Country Region Name / State Province Name														
		Allemagne						New South Wales			Australie			Canada		
French Product Cat..	French Product Subcategory Name	Bayern	Brandenbu..	Hamburg	Hessen	Nordrhein..	Saarland	Wales	Queensland	South Australia	Tasmania	Victoria	Alberta	Columbia	Ontar	
Accessoire	Bidon et porte-bid...	670	107	1041	1193	1229	1176	4972	2554	666	282	2771	28	7645		
	Casque	3359	455	3779	5563	4899	5703	19734	9692	3079	1540	10917	35	28062		
	Garde-boue	528	22	637	571	835	945	2967	1714	462	242	1758	22	8748		
	Nettoyant	40		95	111	151	111	819	310	103	48	429		1161		
	Pneu et chambre à ..	2569	197	3337	4365	5068	5325	20314	10229	2566	1606	12078	90	39224		
	Porte-vélo	360		360	600	720	840	2040	1560	600		1800	240	6840		
	Range-vélo	159		318	477	1113	1113	5088	3021	318	159	1749		5088		
Sac d'hydratation	165		220	440	825	440	3904	2475	1045	165	2914		6159			
Vélo	Vélo de randonnée	56300	8367	75372	100316	73395	110620	420648	238081	59964	27283	247706	5510	264735		
	Vélo de route	31253	32420	229436	330525	257367	339341	2203910	1062537	344305	135212	1258584	14022	921594		
	VTT	41783	15720	160819	213590	215009	256880	1219303	642065	200786	70702	720963	2295	613145		
Vêtements	Casquette	369	72	333	530	494	692	1762	818	288	117	827	18	2158		
	Chaussettes	54	27	45	72	63	54	450	207	108	63	180	9	791		
	Cuissards	70		140	70	70	280	4129	2520	840	280	3010	70	16238		
	Gants	171	49	318	539	563	686	3478	1714	416	171	1984	24	6074		
	Maillot	1864	420	2368	2568	3741	4177	17983	7585	2266	1688	10140	104	21774		
	Veste	254	64	508	572	572	699	2985	1334	445	381	2096		5906		

- Puis faire une clic droit et sélectionner Conserver uniquement

FILTRE LES DONNÉES

- Faire un filtre en conservant uniquement la région de Washington

The screenshot shows a data visualization interface with a sidebar on the left and a main table area on the right. The sidebar contains sections for 'Pages', 'Filtres', and 'Repères'. The 'Filtres' section has a button for 'State Province Nam..'. The 'Repères' section includes a dropdown menu set to 'Automatique', icons for 'Couleur', 'Taille', and 'Texte', and buttons for 'Détails', 'Infobulle', and a green button labeled 'SOMME(Sales ..)'. The main table area is titled 'Feuille 12' and displays a table with columns for 'French Prod..', 'French Product Su..', and 'French Count.. États-Unis Washington'. The table lists various bicycle accessories and clothing items with their respective counts.

French Prod..	French Product Su..	French Count.. États-Unis Washington
Accessoire	Bidon et porte-bid..	5 866
	Casque	23 303
	Garde-boue	6 045
	Nettoyant	882
	Pneu et chambre à ..	26 833
	Porte-vélo	4 560
	Range-vélo	4 770
	Sac d'hydratation	3 849
Vélo	Vélo de randonnée	300 725
	Vélo de route	1 146 761
	VTT	905 906
Vêtements	Casquette	1 591
	Chaussettes	593
	Cuissards	10 149
	Gants	3 576
	Maillot	17 777
	Veste	4 064

FILTRE LES DONNÉES

- "Afficher le filtre" sur « French Product SubCategory » en faisant un clic droit dessus
- Puis se rendre dans la zone "filtres" et cliquer sur "afficher le filtre"
- Enfin, sélectionner les sous-catégories suivantes : Casque', Vélo de route' et 'Maillot'

Pages

Colonnes: French Country Regi... State Province Name

Lignes: French Product Cate... French Product Subc...

Feuille 12

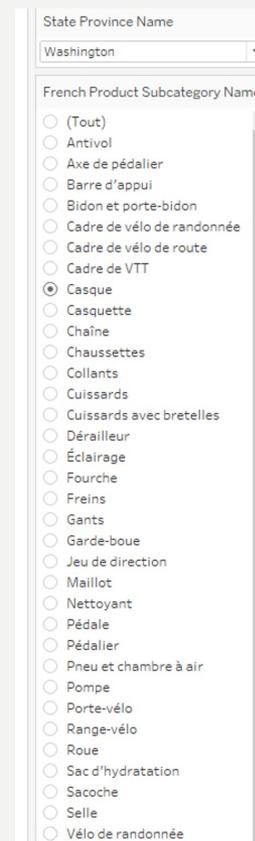
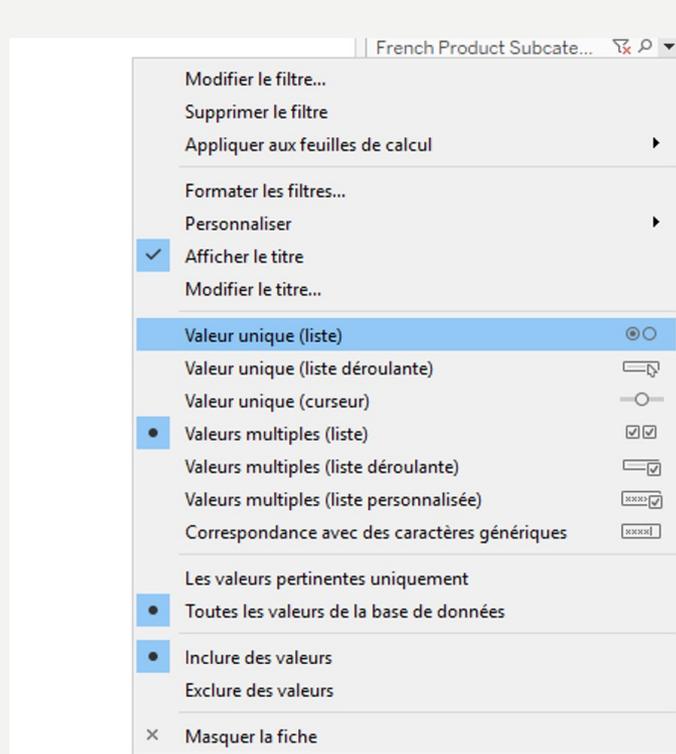
French Prod..	French Product Su..	Washington	French Count..
Accessoire	Casque		23 903
Vélo	Vélo de route		1 146 761
Vêtements	Maillot		17 777

French Product Subcategory Name

- (Tout)
- Antivol
- Axe de pédalier
- Barre d'appui
- Bidon et porte-bidon
- Cadre de vélo de randonnée
- Cadre de vélo de route
- Cadre de VTT
- Casque
- Casquette
- Chaîne
- Chaussettes
- Collants
- Cuissards
- Cuissards avec bretelles
- Dérailleur
- Éclairage
- Fourche
- Freins
- Gants
- Garde-boue
- Jeu de direction
- Maillot
- Nettoyant
- Pédale
- Pédalier
- Pneu et chambre à air
- Pompe
- Porte-vélo
- Range-vélo
- Roue
- Sac d'hydratation
- Sacoche
- Selle
- Vélo de randonnée
- Vélo de route
- Veste
- VTT

FILTRE LES DONNÉES

- Modifier le filtre dans la card de droite pour passer la sous-catégorie en radio bouton
- Faire la même chose sur Région en passant la card en mode Valeur unique (Liste déroulante)



UTILISER LES NOMS ET VALEURS DES MESURES

- Créer une nouvelle feuille en mettant en ligne "Noms de mesures" et glisser "Valeurs de mesures" dans "Repères" puis glisser valeurs de mesures dans le tableau
- Ajouter un filtre sur le nom des mesures

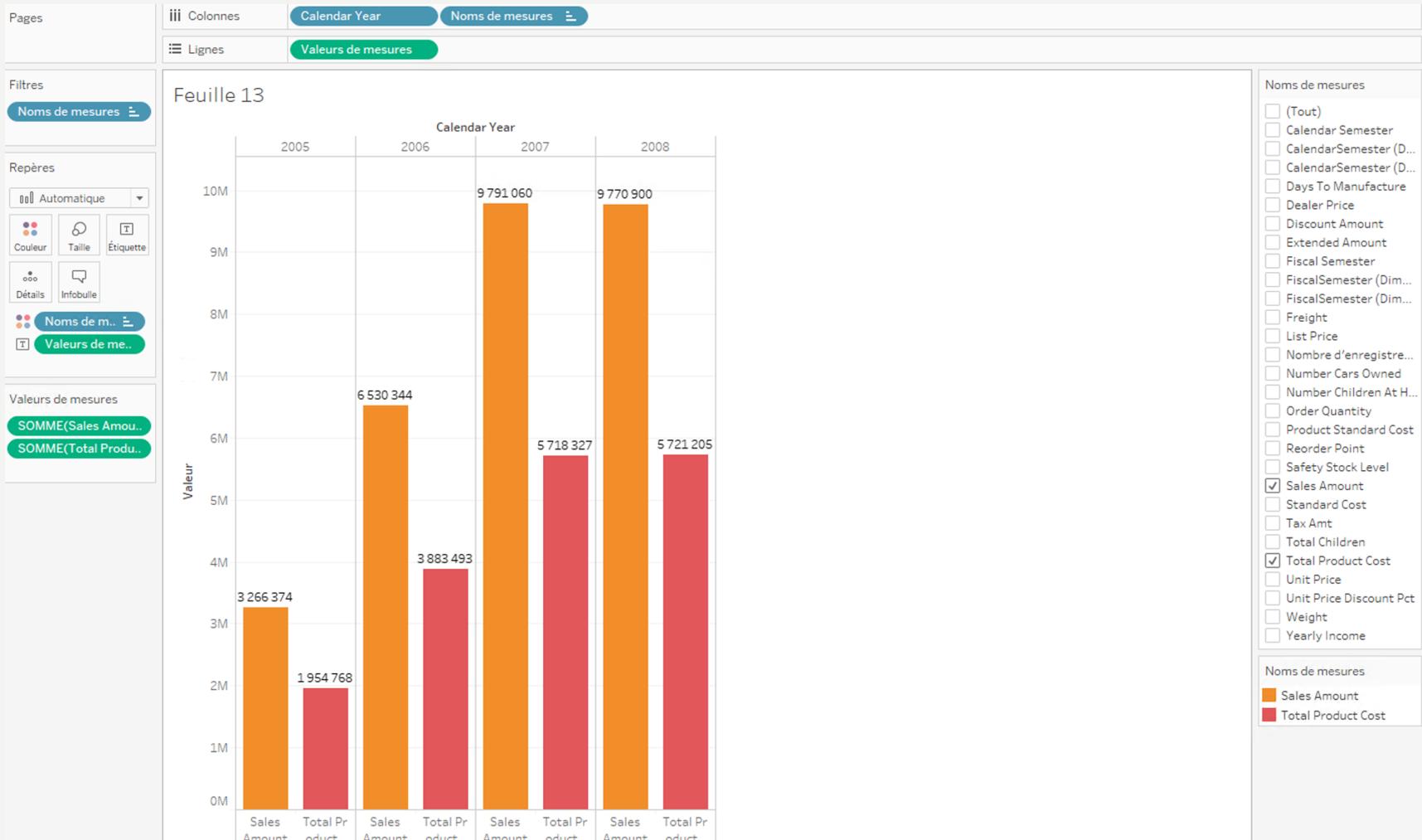
The screenshot displays a Tableau interface with a table of measures and their values. The table is titled "Feuille 13" and contains the following data:

Calendar Semester	87 809
CalendarSemester (DimD..	87 906
CalendarSemester (DimD..	88 109
Days To Manufacture	60 820
Dealer Price	17 615 206
Discount Amount	0
Extended Amount	29 358 677
Fiscal Semester	93 385
FiscalSemester (DimDate)	93 288
FiscalSemester (DimDate1)	93 085
Freight	733 970
List Price	29 358 677
Nombre d'enregistrements	60 398
Number Cars Owned	91 152
Number Children At Home	65 249
Order Quantity	60 398
Product Standard Cost	17 277 794
Reorder Point	6 536 406
Safety Stock Level	8 715 208
Sales Amount	29 358 677

The interface also shows a sidebar with filters and a list of measures. The "Noms de mesures" filter is active, and the "Valeurs de me..." filter is also active. The "Repères" section shows the "Automatique" filter selected. The "Valeurs de mesures" section shows a list of measures, including "SOMME(Calendar S..", "SOMME(CalendarSe..", "SOMME(CalendarSe..", "SOMME(Days To Ma..", "SOMME(Dealer Price)", "SOMME(Discount A..", "SOMME(Extended A..", "SOMME(Fiscal Sem..", "SOMME(FiscalSeme..", "SOMME(FiscalSeme..", "SOMME(Freight)", "SOMME(List Price)", "SOMME(Nombre d'e..", "SOMME(Number Ca..", "SOMME(Number Ch..", "SOMME(Order Quan..", "SOMME(Product St..", "SOMME(Reorder Poi..", "SOMME(Safety Stoc..", and "SOMME(Sales Amou..".

UTILISER LES NOMS ET VALEURS DES MESURES

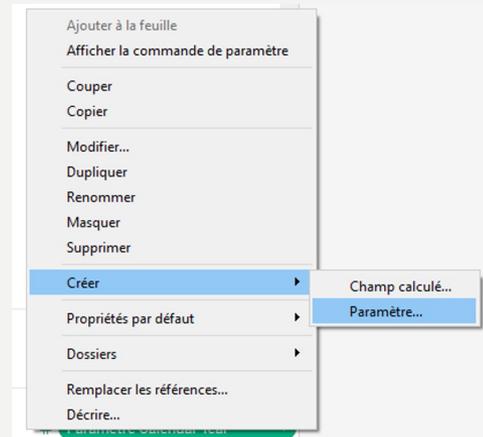
- Créer une nouvelle feuille et essayer de reproduire ce graphique



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Créer un paramètre avec les données suivantes :

- Nom : TOP Customers
- Type de donnée : integer
- Valeur courante : 15
- Régler le range :
 - Minimum : 5
 - Maximum : 20
 - Taille du pas : 5

A screenshot of a dialog box titled 'Modifier le paramètre [TOP Customer]'. It contains the following fields and options:

- Nom :
- Propriétés:
 - Type de données :
 - Valeur actuelle :
 - Format d'affichage :
 - Valeurs autorisées : Tout Liste Plage
- Plage de valeurs:
 - Minimum :
 - Maximum :
 - Taille d'incrément :
- Buttons:

CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Créer un ensemble de données sur le « Last Name » comme sur la 1ère capture
- Saisir les données comme sur la deuxième capture :

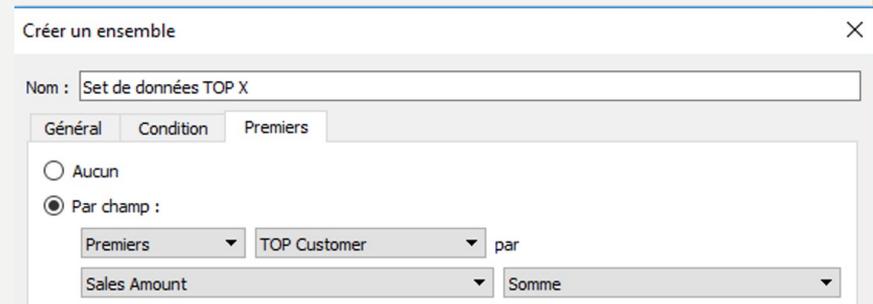
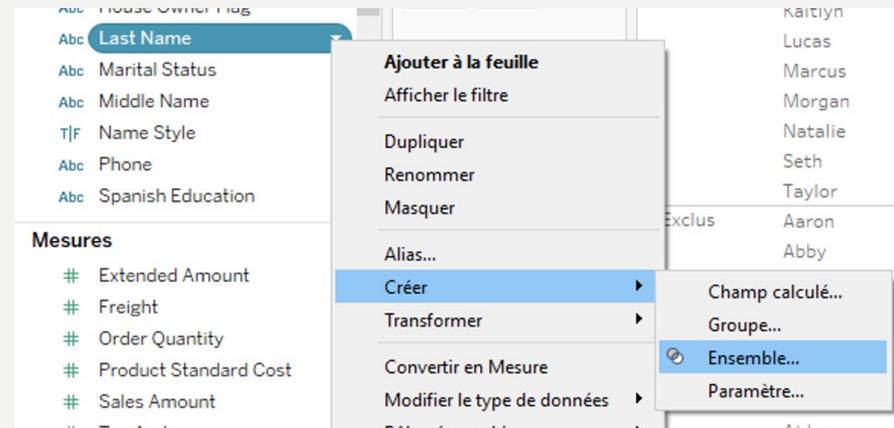
- Nom : Set de données TOP X
- Sélectionner « Premiers »
- Puis le radio bouton "Par Champ" avec les paramètres

suivants :

Type :TOP

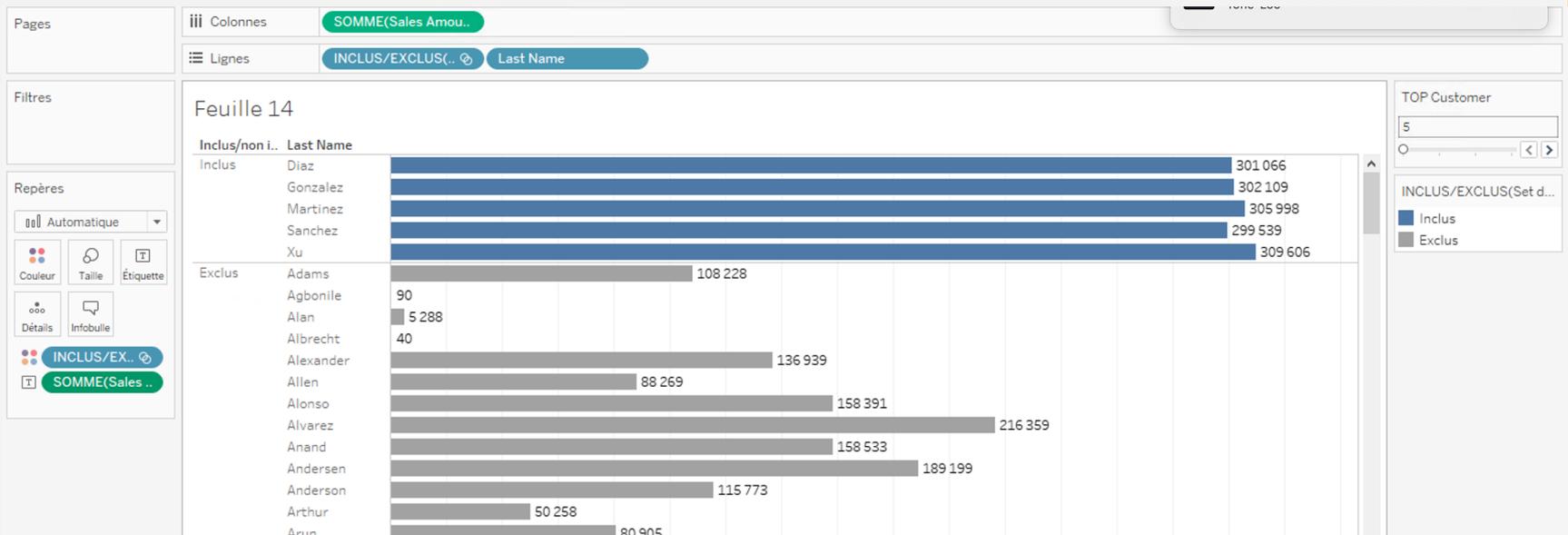
Champ paramètre : "Top Customer"

Par : Somme des Sales Amount



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Créer un graphique permettant de montrer le top 10 des meilleurs clients en fonction des chiffres de ventes réalisés



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Il est possible de reprendre les valeurs des paramètres dans les tableaux. Pour cela, il est nécessaire de procéder de la manière suivante :

- Créer un champ calculé nommé : Nom du TOP

- Puis insérer la formule suivante

IF [Set de données TOP X]

THEN "TOP " + str([TOP Customers]) + " Clients »

ELSE "Autres »

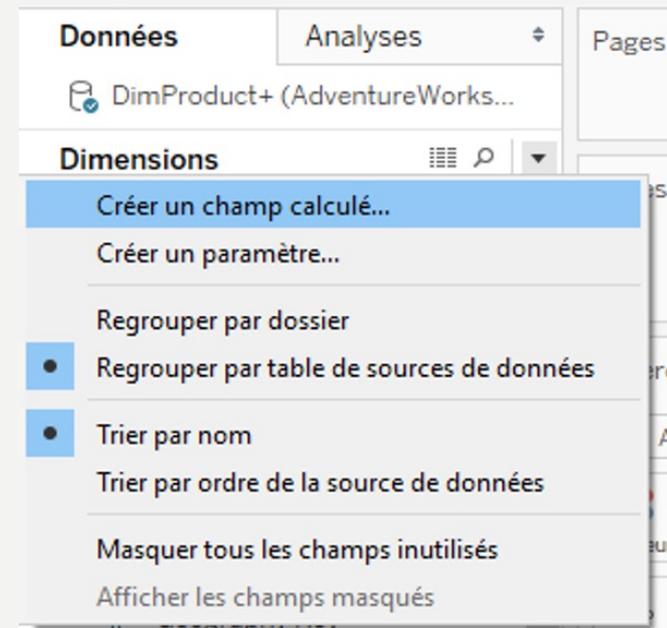
END

- Expliquer cette formule ?

- Cacher le champ "IN/OUT

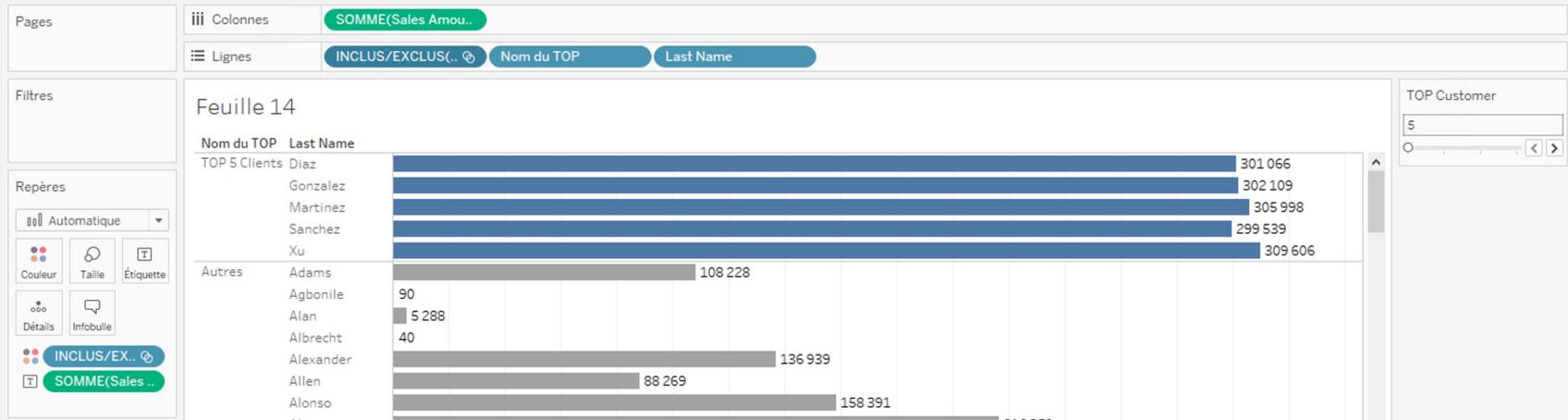
(Set de données) puis ajouter le

nouveau champ dans le graphique



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Il est possible de reprendre les valeurs des paramètres dans les tableaux.



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Il est possible de cacher les données en créant dans un premier temps un paramètre (Etendre/Cacher) puis dans un second temps un champ calculé qui affiche seulement les données nécessaires.

– Créer le paramètre "Réduire/Elargir" de type String

Avec deux valeurs choisies dans une liste :

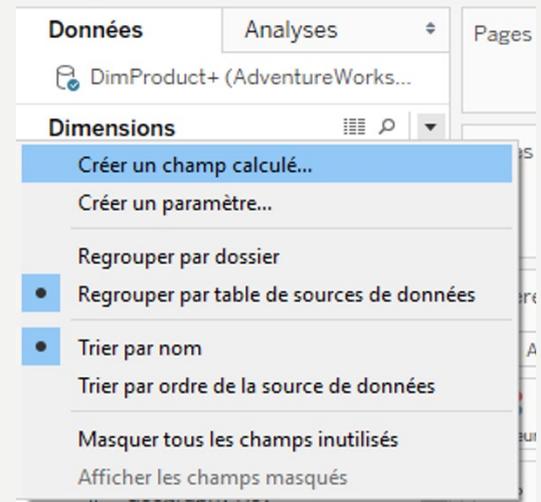
- Elargir
- Réduire

– Créer un champ calculé avec le nom suivant

"Nom Client Dynamique" et la formule suivante :

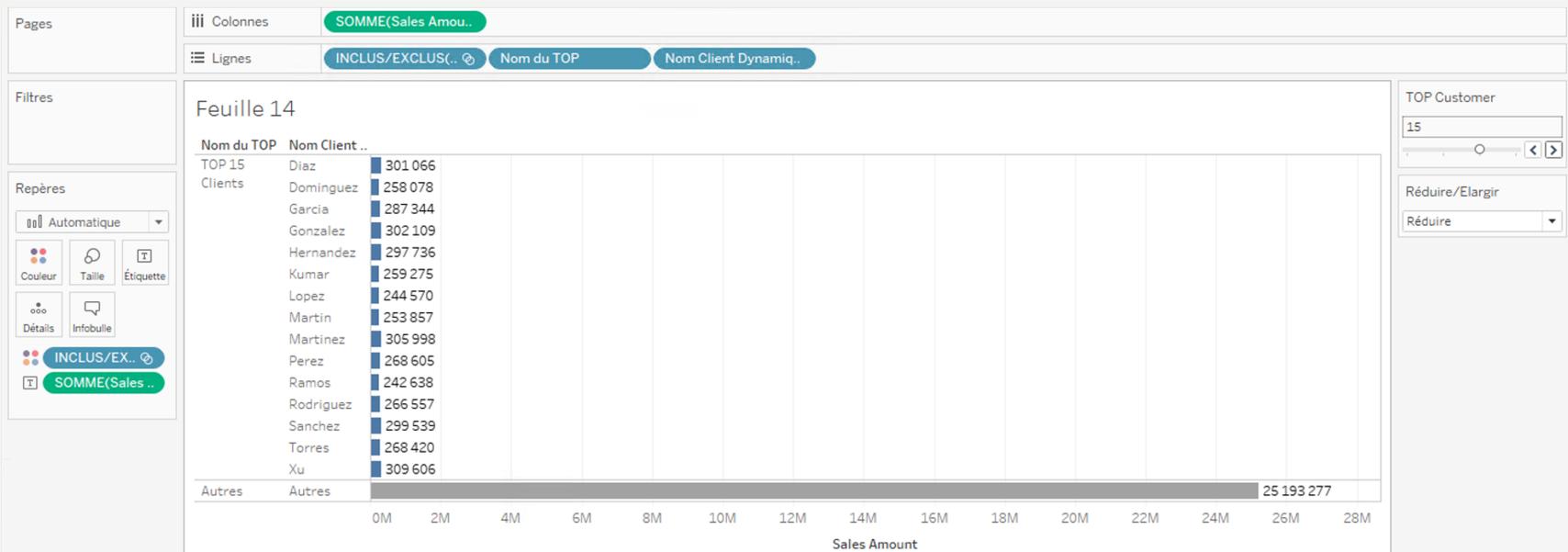
```
IF [Réduire/Elargir]="Réduire" THEN  
IF [Set de données TOP X Last Name]  
THEN [Last Name]  
ELSE "Autres" END  
ELSE [Last Name] END
```

– Expliquer cette formule ?



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Il est possible de cacher les données en créant dans un premier temps un paramètre (Etendre/Cacher) puis dans un second temps un champ calculé qui affiche seulement les données nécessaires.



- Est ce que vous trouvez ces résultats corrects ?
- Afficher le « Customer Alternate Key ». Que voyez vous ?

CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- A présent, on va modifier l'ensemble pour que celui-ci soit plus juste
- Créer un ensemble de données sur le « CustomerKey (DimCustomer) » comme sur la 1ère capture
- Saisir les données comme sur la deuxième capture :

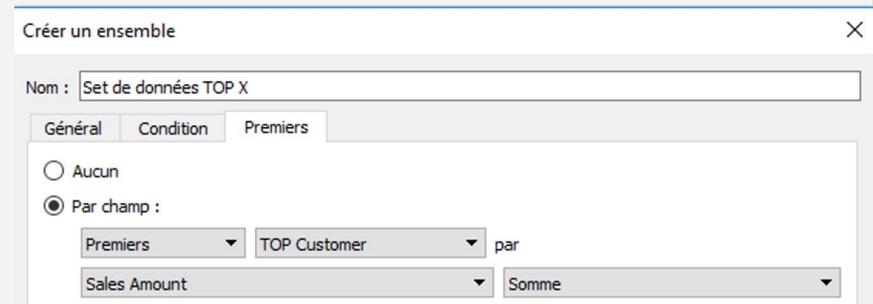
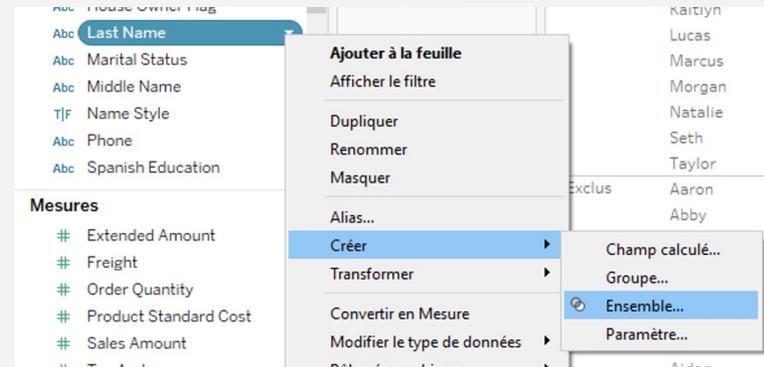
- Nom : Set de données TOP X CustID
- Sélectionner « Premiers »
- Puis le radio bouton "Par Champ" avec les paramètres

suivants :

Type : TOP

Champ paramètre : "Top Customer"

Par : Somme des Sales Amount



CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Reproduire la visualisation suivante :



- Pensez vous que ce nouveau TOP est plus juste ? Pourquoi ?

CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Rajouter le pays du client dans le graphique (French Country Region Name).



- Pourquoi tous les clients sont français ?
- Changer le paramètre TOP Customer pour afficher les 30 meilleurs clients
- Que voyez-vous ?

CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Pour aller plus loin, on va modifier l'ensemble des meilleurs clients pour que celui-ci prenne en compte le pays du client.
- Pour cela, on va créer un nouveau paramètre permettant de sélectionner le pays

The screenshot shows a dialog box titled "Modifier le paramètre [Pays]". It contains the following elements:

- Nom :** A text input field containing "Pays".
- Commentaire >>** A button to the right of the name field.
- Propriétés** section:
 - Type de données :** A dropdown menu set to "Chaîne de caractères".
 - Valeur actuelle :** A dropdown menu set to "Allemagne".
 - Format d'affichage :** A dropdown menu.
 - Valeurs autorisées :** Radio buttons for "Tout", "Liste" (selected), and "Plage".
- Liste de valeurs** section:
 - A table with two columns: "Valeur" and "Afficher en tant que".
 - Buttons: "Ajouter à partir du paramètre", "Ajouter à partir du champ", "Coller à partir du Presse-papiers", and "Tout supprimer".
- OK** and **Annuler** buttons at the bottom.

Valeur	Afficher en tant que
Allemagne	Allemagne
Australie	Australie
Canada	Canada
États-Unis	États-Unis
France	France
Royaume-Uni	Royaume-Uni
Ajouter	

CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- On va créer un champ calculé pour que celui-ci prenne en compte que les montant de commande du pays sélectionné
- Son nom : Sales Amount/Pays

×

```
IF [Pays] = [French Country Region Name] THEN
[Sales Amount]
ELSE
0
END
```

Le calcul est valide.

4 dépendances ▾

CRÉATION D'UN PARAMÈTRE, UN SET DE DONNÉES ET D'UN CHAMP CALCULÉ

- Puis on va modifier l'ensemble Set de données TOP X CustID pour que celui-ci somme ce nouveau champ calculé

Général Condition Premiers

Aucun

Par champ :

Premiers TOP Customer par

Sales Amount /Pays Somme

- On va modifier le champ calculé « Nom Client Dynamique » pour que celui-ci affiche plus d'informations sur le client

```
THEN [Last Name]+" "+[First Name]+" "+[French Country Region Name]+" "+[Customer Alternate Key]
ELSE "Autres" END
ELSE [Last Name]+" "+[First Name]+" "+[French Country Region Name]+" "+[Customer Alternate Key] END
```

- Ajouter le paramètre Pays dans la feuille

CRÉER UNE CARTE

- Créer une nouvelle feuille et double cliquer sur « French Country Region Name »

The screenshot shows a software interface with a sidebar on the left and a main content area on the right. The sidebar contains the following sections:

- Pages**: A section with a grid icon and the text 'Pages'.
- Filtres**: A section with the text 'Filtres'.
- Repères**: A section with a dropdown menu set to 'Automatique' and several icons for 'Couleur', 'Taille', 'Texte', 'Détails', and 'Infobulle'.

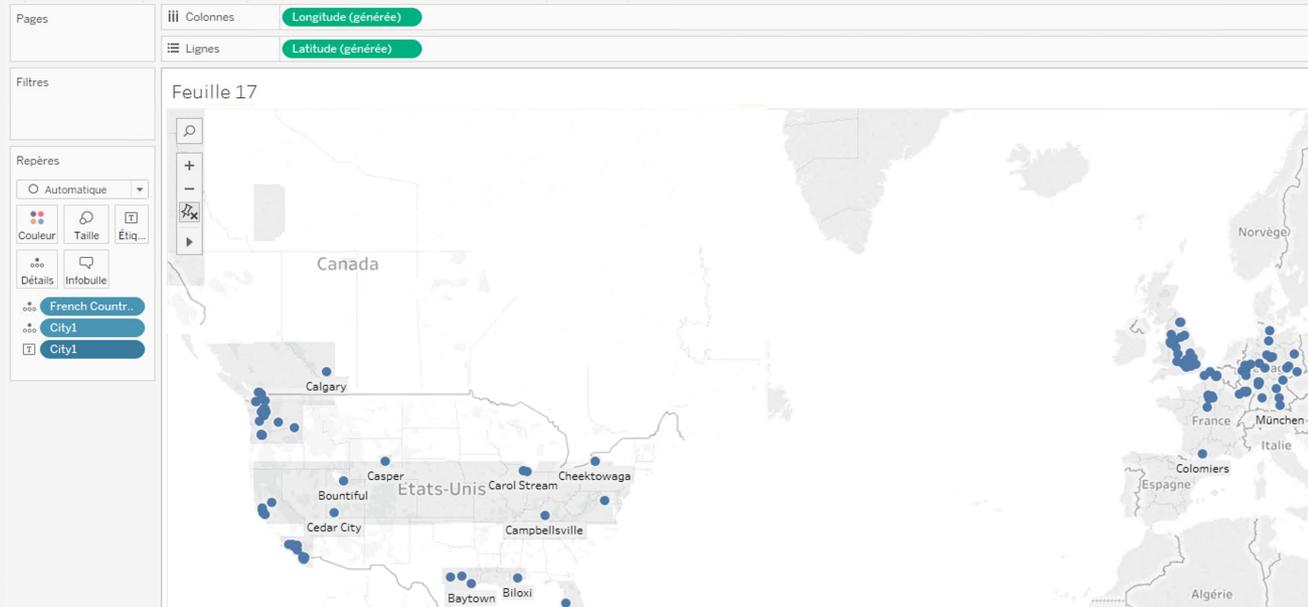
The main content area shows a table titled 'Feuille 17' with the following data:

French Country..	
Allemagne	Abc
Australie	Abc
Canada	Abc
États-Unis	Abc
France	Abc
Royaume-Uni	Abc

- Puis cliquer sur montre-moi et choisir Cartes de symboles

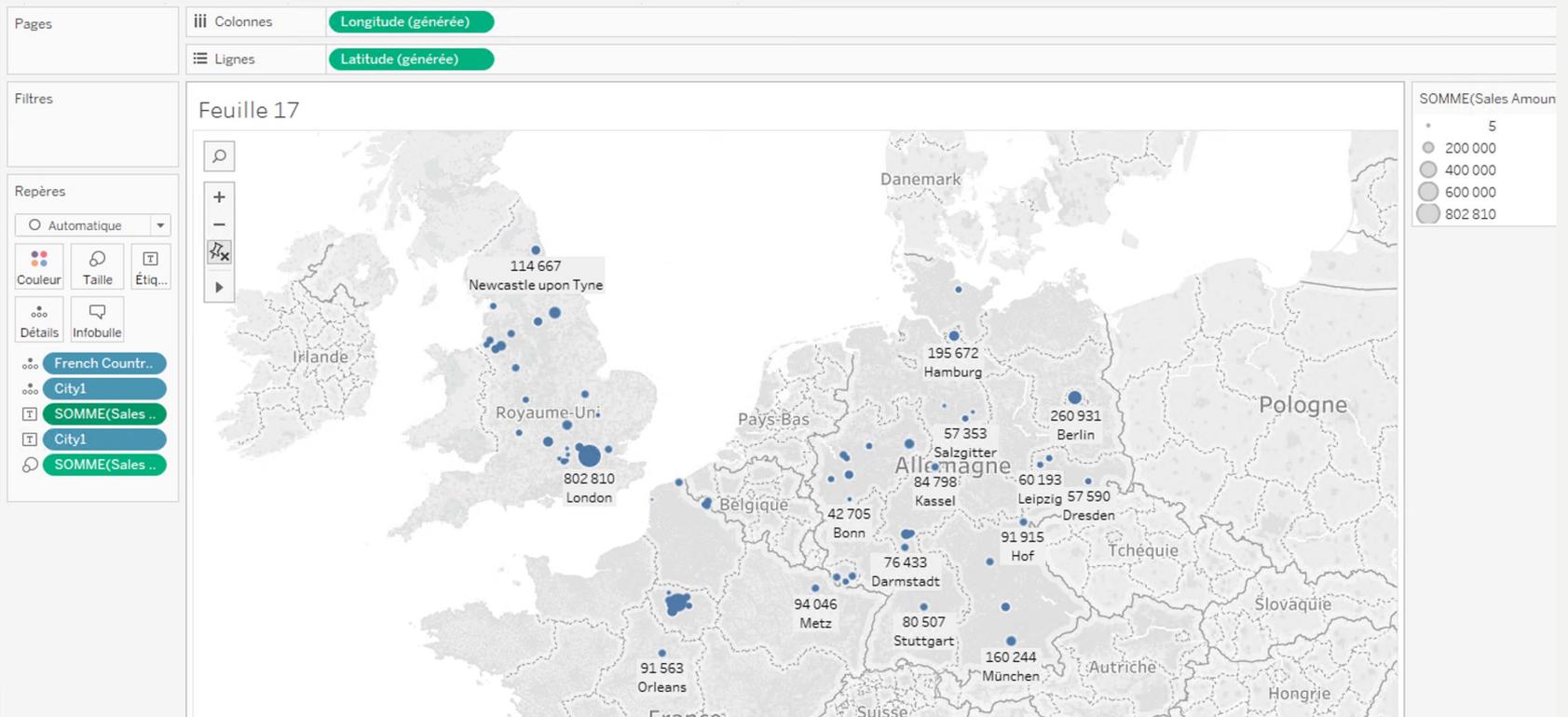
CRÉER UNE CARTE

- Ajouter les villes en double cliquant sur l'attribut



CRÉER UNE CARTE

- Puis ajouter les « Sales Amount » pour chaque ville avec un rapport de taille en fonction du CA





ETUDE DE CAS SUR L'ENTREPÔT DE DONNÉES ADVENTURE WORK

CAS PORTANT SUR L'ANALYSE DES VENTES

ETUDE DE CAS SUR L'ENTREPÔT DE DONNÉES ADVENTURE WORK

- Contexte :

Adventure Works Cycles est un fabricant international de vélos avec un siège social à Bothell, Washington. L'entreprise compte 300 employés et 29 commerciaux. Il possède plusieurs magasins dans le monde et un réseau de revendeur. Ils sont localisés en Australie, Canada, France, Allemagne, Royaume-Unis et USA. Ils vendent aussi à travers leur site web.

- Objectif principal :

Le but de cette étude de cas est de répondre à plusieurs points :

- Comprendre l'incroyable croissance de l'entreprise et aussi ses transformations durant les derniers années
- Connaître les meilleurs produits et les moins bons afin de rendre l'entreprise plus en adéquation avec son marché
- Connaître les clients en fonction des pays et des informations que l'on possède sur lui afin de déterminer ses habitudes d'achats dans le but de lui proposer des produits les plus proches de ses besoins

ETUDE DE CAS SUR L'ENTREPÔT DE DONNÉES ADVENTURE WORK

- Cahier des charges du tableau de bord sur l'analyse générale de la société :
 - Le board de l'entreprise souhaite voir l'évolution du CA (Sales Amount) dans le temps par trimestres
 - Une carte permettant de voir les villes où il y a le plus de ventes de vélo. Des informations seront présentes dans les info-bulles.
 - Afficher une treemap (carte proportionnelle) permettant de montrer les meilleures ventes de produits
 - Créer une visualisation avec la marge pour les produits les plus vendus
 - En fonction des chiffres et des analyses selon vous, quels conseils pourriez vous donner à cette société pour que son CA et sa marge évoluent encore plus positivement ? Démontrez le avec des visualisations
 - [BONUS] Créer une carte avec les villes en décrochage d'un point de vue CA
- Pour chaque demande, créer une ou plusieurs feuilles spécifiques. Puis les rassembler dans un tableau de bord ergonomique

ETUDE DE CAS SUR L'ENTREPÔT DE DONNÉES ADVENTURE WORK

- Cahier des charges du tableau de bord sur une analyse précise du marché français d'un point de vue de la clientèle et des produits :
 - Quels sont les critères communs qui sont possédés par les meilleurs clients français (25 meilleurs clients) ?
 - Est ce qu'ils sont représentatifs du client moyen en terme de segmentation de clientèle ou de choix de produits (vélos) ?
 - Analyser les ventes des clients français afin de déterminer les meilleurs ventes de vélos ?
 - Quels sont les informations importantes que l'on peut analyser sur l'évolution du marché français (nombre de clients, évolution des ventes, segmentation de clientèle la plus représentative, ...). Démontrez le avec des visualisations.
 - Quels sont les pays qui ont le même profil de croissance que la France ? Quels sont à contrario les pays qui évoluent différemment ?
- Pour chaque demande, créer une ou plusieurs feuilles spécifiques. Puis les rassembler dans un tableau de bord ergonomique
- Pour aller plus loin, créer à partir de vos feuilles, une histoire (data storytelling). Cette histoire reprendra les principales informations du marché français en 10 points clés.

ETUDE DE CAS SUR L'ENTREPÔT DE DONNÉES ADVENTURE WORK

- Cahier des charges du tableau de bord sur le marché du vélo :
 - Reprendre la treemap (carte proportionnelle) permettant de montrer les meilleurs ventes de produits pour repartir sur les produits phares. A partir de ces produits, expliquez la stratégie de renouvellement de vélos appliquée par la société ?
 - Est ce que celle-ci a été porteuse à chaque fois pour la société (relais de croissance positive) ?
 - Comment évolue le marché du vélo si on se base sur les ventes réalisées par la société ?
 - Est ce que tous les vélos ont le même profil de vente dans tous les pays ?
 - Peut-on déterminer des caractéristiques sur le profil de l'acheteur pour chaque type de vélos (Randonnées, VTT, Course) ?
- Pour chaque demande, créer une ou plusieurs feuilles spécifiques. Puis les rassembler dans un tableau de bord ergonomique
- Pour aller plus loin, créer à partir de vos feuilles, une histoire (data storytelling). Cette histoire reprendra les principales informations du marché du vélo en 10 points clés.



ETUDE DE CAS SUR L'IMPLANTATION DE MAGASINS

**CAS DE L'IMPLÉMENTATION DES STARBUCKS
DANS LE MONDE**

ETUDE DE CAS SUR L'IMPLANTATION DE MAGASINS

- Contexte :

Ce jeu de données fournit les informations d'implantation des magasins Starbucks. Actuellement, plus de 25600 magasins sont présents dans le monde.

- Objectif principal :

Le but de cette étude de cas est de comprendre l'implantation des magasins Starbucks à travers le monde.

- Nom du fichier : « Implantation Starbucks.csv »

- Logiciel à utiliser : Tableau Desktop

ETUDE DE CAS SUR L'IMPLANTATION DE MAGASINS

- Cahier des charges du tableau de bord :
 - Un paramètre doit permettre de sélectionner les TOP (de 5 à 20 par pas de 5)
 - Une carte permettant de voir les villes où il y a le plus de magasins Starbucks avec le paramètre qui permet de changer dynamiquement. Des informations seront présentes dans les info-bulles.
 - Créer un champ calculé montrant la proportion de magasins dans les pays par rapport à l'implantation globale de Starbuck (25600 magasins).
 - Afficher une treemap (carte proportionnelle) permettant de montrer les 10 pays ayant le plus de magasins Starbuck
 - Afficher dans un coin du rapport le nombre de villes et le nombre de pays où il y a des magasins Starbucks (utiliser le changement de type de champ).

ETUDE DE CAS SUR L'IMPLANTATION DE MAGASINS

- Cahier des charges du 2^{ème} tableau de bord :
 - Créer une visualisation permettant de voir la répartition des magasins en fonction du type de partenariat avec Starbuck (Franchise, Licensed, ...). Que peut-on observer ?
 - Montrer les 10 pays qui ont le plus de magasins en Joint Venture ? Que peut-on remarquer sur la localisation ? Tentative d'explications ?
 - Quels sont les différentes stratégies de déploiements des magasins en Europe ? (Pour filtrer sur les pays d'Europe, faire une sélection à main levée sur une carte puis basculer vers une autre visualisation). Créer une visualisation permettant de montrer les différentes stratégies.
 - Quels sont les 5 premiers états américains à avoir le plus de magasins ? Avec quelles données externes, il pourrait être intéressant de recouper ces informations ?



ETUDE DE CAS SUR L'ANALYSE DES INFORMATIONS NUTRITIONNELLES

**CAS DES INFORMATIONS NUTRITIONNELLES
DU MENU DE MCDONALD'S**

ETUDE DE CAS SUR L'ANALYSE DES INFORMATIONS NUTRITIONNELLES

- Contexte :

Ce jeu de données fournit les informations nutritionnelles de tous les menus McDonald's américain du petit déjeuner aux desserts. Ces informations ont été récupérées à partir du site web américain de McDonald's.

- Objectif principal :

Le but de cette étude de cas est de créer un tableau de bord permettant d'aider les clients McDonald's à choisir leur menu en respectant les apports de référence en énergie et en nutriments.

- Nom du fichier : « Informations nutritionnelles McDo.csv »

- Logiciel à utiliser : Tableau Desktop

ETUDE DE CAS SUR L'ANALYSE DES INFORMATIONS NUTRITIONNELLES

- RÈGLEMENT (UE) No 1169/2011 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 25 octobre 2011 concernant l'information des consommateurs sur les denrées alimentaires

PARTIE B – APPORTS DE RÉFÉRENCE EN ÉNERGIE ET EN CERTAINS NUTRIMENTS À L'EXCLUSION DES VITAMINES ET DES SELS MINÉRAUX (ADULTES)

Énergie ou nutriment	Apport de référence
Énergie	8 400 kJ (2 000 kcal)
Graisses totales	70 g
Acides gras saturés	20 g
Glucides	260 g
Sucres	90 g
Protéines	50 g
Sel	6 g

- Source : https://www.senat.fr/europe/textes_europeens/ue0120.pdf

ETUDE DE CAS SUR L'ANALYSE DES INFORMATIONS NUTRITIONNELLES

- Cahier des charges du tableau de bord :
 - Au moins 4 paramètres permettront de choisir les aliments sur le tableau de bord (choix du burger, choix de l'accompagnement, choix des boissons, choix des desserts, ...)
 - Dans un gramme de sel courant (NaCl), il y a un peu moins de 400 mg de sodium (40% de Na). Créer un champ calculé permettant de déterminer le sel dans les aliments sélectionnés pour le comparer aux apports de référence.
 - Une ou plusieurs visualisations permettra(ont) de comparer les sept valeurs des métriques du slide précédent avec les valeurs des éléments sélectionnés dans les paramètres.
 - Une alerte sur un ou plusieurs surplus par rapport aux apports de référence devra être prévue.
 - Une visualisation montrera les 10 éléments les moins « riches » à prendre pour aider aux choix.
 - Créer une visualisation libre donnant du sens à ce tableau de bord
 - Ajouter ce tableau de bord dans votre rapport final

ETUDE DE CAS SUR L'ANALYSE DES INFORMATIONS NUTRITIONNELLES

- Répondre aux questions suivantes et insérer les réponses dans votre rapport :
 - Combien de calories sont contenues dans les burgers de McDonald ?
 - Quelles sont les métriques qui sont le plus impactées par les boissons ?
 - Quel est la différence entre le poulet grillé et le poulet croustillant d'un point de vue valeur nutritionnelle ?
 - Quels sont les articles les plus équilibrés à commander pour un petit déjeuner (boisson chaude, boisson froide et plat) ?
 - Quels sont les métriques qui sont corrélées entre-elles sur ce jeu de données ? Quel est la meilleure visualisation permettant de le montrer ? Justifier ?



**BONNES
PRATIQUES SUR
LES
VISUALISATIONS**

BONNES PRATIQUES SUR LES VISUALISATIONS

- Hiérarchie d'informations dans vos tableaux de bord (Du plus important vers le moins important)
- Adapter les couleurs aux contenus pour les comparer (importance, nombres négatifs, données différentes, ...)
- Attention à mettre trop de couleurs qui fatiguent et complexifient au lieu de simplifier et mettre en valeur
- Utiliser des polices lisibles par tous (Arial, Trebuchet MS, Verdana, Times New Roman, Lucida sans)
- Utiliser des bulles d'informations pour montrer des informations plus précises et secondaires. Cela allège votre graphique au premier regard mais donne toutes les informations si cela est nécessaire.
- Travailler vos axes de graphique pour qu'ils soient bien calibrés à vos données et renseignent facilement vos utilisateurs
- Source : http://onlinehelp.tableau.com/current/pro/desktop/en-us/help.htm#visual_best_practices.html

BONNES PRATIQUES SUR LES VISUALISATIONS

- The representation and presentation of data that exploits our visual perception abilities in order to amplify cognition.

De Andy Kirk in *Data Visualization: A Successful Design Process*.

<http://www.visualisingdata.com/>

- Explications :
 - Representation : Formes physiques (lignes, barres, nuages de points, ...)
 - Presentation : Couleurs, Interaction, bulle d'aide, ...
 - Visual perception abilities : les yeux, le principe d'interprétation du cerveau, le processus de reconnaissance de formes, le raisonnement spatial, ...
 - Amplify cognition : transformer l'information en perceptions et connaissances

POUR ALLER PLUS LOIN

- Lire le document "Tableau Software Visual Analytics Best Practices" pour comprendre pourquoi et comment construire des visualisations qui permettent d'améliorer la transformation de l'informations en connaissances
- Questions :
 - Quels sont les différents types de données ?
 - Veuillez expliquer pourquoi on trouve des représentations différentes en fonction du type de données dans les "Best Practices" ?