TP DE L'ETL À LA Restitution

AUTEUR : JONATHAN FONTANEL DATE : 15/11/2022 TP ISIMA ZZ2 F4 BASES DE DONNÉES ET FOUILLES 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



PRÉSENTATION DE TALEND OPEN STUDIO

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

Talend Open Source	Stitch Data Loader	Talend Pipeline Designer	Talend Cloud Data	Talend Data Fabric*
Solution GRATUITE pour tous les utilisateurs	Version payante : De 100 à 1 000 USD par mois	Tarif horaire en fonction de l'utilisation En savoir plus	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.	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.
TÉLÉCHARGER MAINTENANT	ESSAI GRATUIT 14 JOUR	ESSAI GRATUIT 14 JOUR	ESSAI GRATUIT	CONTACTER LE SERVICE COMMERCIAL

Tableau Desktop

Tableau Server

Tableau Online

Tableau Mobile

Tableau Public



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.

+ableau++public

Tableau Public est gratuit et rend accessible à tous l'analyse et la <u>visualisation des données</u>. 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) !

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

• Nouveau produit :



• Source : https://www.tableau.com/products/prep

• 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=1-4JT7ECY&ct=171102&st=sb

- 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.

	Year Ended December 31,									
		2016		2015 2014		2014	2013			2012
			_	(in thousar	nds,	except per	sha	re 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

- 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 <u>réalisé</u> 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 »

○○○			MySQL	Workbench		
		Welcor MySQL Workber create and bro design and run SC	Connect to Connect to Please enter service: User: User: Password:	MySQL Server password for the fo Mysql@localhost:3306 root 	_ V Ilowing	SQL. It a and inse schema:
	MySQL Connections ⊕ Local instance 3306 root root isocalhost:3306	browse bocumentation >		Cancel	OK	

RESTAURER LE DATAWAREHOUSE

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

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

47 SCHEMAS Q Filter objects AdventureWorksDW ▼ Tables DimAccount DimCurrency DimCustomer DimDate DimDepartmentGroup DimEmployee DimGeography DimOrganization DimProduct DimProductCategory DimProductSubcategory DimPromotion DimReseller DimSalesReason DimSalesTerritory DimScenario FactAdditionalInternationalProdu... FactCallCenter FactCurrencyRate FactFinance FactInternetSales FactInternetSalesReason FactResellerSales FactSalesQuota

- FactSurveyResponse

VOTRE ENTREPÔT DE DONNÉES

• Identifier les dimensions et les tables de fait

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

Tableau - Classeur2

Fichier Données Serveur Aide

曓

Connexion

Microsoft Excel

Fichier texte

Fichier JSON

Microsoft Access

Fichier PDF

Fichier de données spatiales

Fichier de statistiques

Plus...

Ouvrir

TP5 2018

Inclus	Breakfast	Big Breakfast with			
		Big Breakfest with			
		Big Breakfast with			
		Big Breakfest with			
	Chicken &	Chicken McNuppet			
	Fish	Chicken McNugget			
	Smoothies &	Chocolate Shake (L			
	Shakes	McFlurry with M81			
		Strawberry Shake			

MCdo

Tableau Server

MySQL

Oracle

Amazon Redshift

Microsoft SQL Server

Plus...

• 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 : 330	5
		Base de données :	AdventureWorksDW		
		Entrez les informatio	ns de connexion à la bas	se de données :	
		Nom d\'utilisateur :	root		
		Mot de passe :	••••		
		Nécessite SSL			
					_
		SQL initial		Connexion	
		Téléchargez et install	ez les pilotes, puis conn	ectez-vous.	

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

MySQL				×
Serveur :	localhost	Port :	3306	
Base de données :	AdventureworksDW			
Entrez les informati	ons de connexion à la ba	se de don	nées :	
Nom d\'utilisateur :	root			
Mot de passe :	••••			
Nécessite SSL				
SQL initial		Con	nexion	

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

• Glisser les tables afin d'obtenir ce schéma

- 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 I en bas dans les onglets, vous avez les dimensions et les mesures qui sont dans la première colonne à gauche

• • •			Tableau - Classeur1				
* ← → □ द G	• 🖓 • 📕 • 🖬 🖓	· · · · · · · · · · · · · · · · · · ·	<u>/</u> • Ø • I 4	Standard -			ntre-moi
Données Analyse +	Pages	iii Colonnes				- 1214 1078	
Achats (global_superstor global_superstore_20		⊞ Lignes				- 5/8 (914) - 5/8	- 348 - 2428 Stano
Dimensions Ⅲ P Abc Catégorie # Code postal	Filtres	Feuille 1		Déposer char	np ici		
 Date d'expédition Date de commande Pays, Ville Pays Ville Abc État Abc ID de client Abc ID de client Abc ID roduit Abc Marché Abc Mode d'expédition Abc Nom du client Abc Nom du produit 	Repères	Déposer champ iri		Déposer chan	np ici	Sélectionner ou faire	Solution Sol
Mesures # Bénéfices # Frais d'expédition # ID de ligne # Quantité # Remise # Ventes @ Latitude (générée) @ Longitude (générée)		- Mari				Utilisez les touches M Cmd pour sélectionne plusieurs champs	aj ou r

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

Feu	il	le	1	0

Année de F	
2005	Abc
2006	Abc
2007	Abc
2008	Abc

辞		Ð			₽, ↓	ê ↓ <u></u> ₽	0,	7 - 0	* [<u>T</u>]	\$	Standar
Donné	Analyses	\$	Pages		iii C	olonnes					
🌏 Dir	mProduct+ (AdventureWorks	-			⊞ L	ignes			NNÉE(F	ull Dat	e
Dimer	nsions 🏢 ຂ	-	F -11		_						
D 🖩	imCustomer	^	Filtres		Fe	uille 1	0				
) 🖩 D	imDate (DueDateKey)										
/ 🆩 D	imDate (OrderDateKey)				Ani 200	nee de F	-	2266 27/	1		
#	Calendar Quarter		Papàras		200	05	6	5200374	•		
#	Calendar Year		Reperes		200	07		9 791 060	·		
#	Date Key		T Automatique	Ŧ	200	08	9	770 900)		
#	Day Number Of Month										
#	Day Number Of Week		Couleur Taille Text	e							
#	Day Number Of Year										
Abc	English Day Name Of Week										
Abc	English Month Name		Details								
#	Fiscal Quarter		T SOMME(Sales .								
#	Fiscal Year										
Abc	French Day Name Of Week										
Abc	French Month Name										
Ë	Full Date Alternate Key										
#	Month Number Of Year	~									
Mesur	res										
#	Extended Amount	^									
#	Freight										
#	Order Quantity										
#	Product Standard Cost										
#	Sales Amount										
#	Tax Amt										
#	Total Product Cost										
#	Unit Price										
#	Unit Price Discount Pct										

CREATION D'UN PREMIER GRAPHIQUE

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

Pages	iii Colonnes						
	⊞ Lignes		NNÉE(Full Date .	. 🗆 TRIN	IESTRE(Full D	. MOIS(Full Date Alt
Filtres	Feuille 1	.0					
	Année de F	Trimestre d.	Mois de Full				
	2005	T3	juillet	473 388	^		
Repères			août	506 192			
-			septembre	473 943			
T Automatique 🔻		T4	octobre	513 329			
			novembre	543 993			
Couleur Taille Texte			décembre	755 528			
	2006	T1	janvier	596 747			
000 Ly			février	550 817			
Détails Infobulle			mars	644 135			
T SOMME(Sales		T2	avril	663 692			
			mai	673 556			
	1		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	1010258			
		T4	octobre	1080450			
			novembre	1 196 981			
			décembre	1731788			

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".

Pages

Filtres

...

ē ē Couleur

000

Détails

DRILL DANS LES DONNÉES

• En cliquant sur le + à gauche de "ANNEE", vous allez faire apparaitre les

trimestres

RAJOUTER UN PREMIER FILTRE

Pages

Filtres

Repères

Couleur

ooo Détails

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

Filtrer [Anné	e de Full Date	e Alternate k	(ey]
Général	Condition	Premiers	
Sélection	onner dans la l	iste 🔿 Liste	e de vale
	xt de recherch	ie	
 ✓ 2007 ✓ 2008 			

	iii Colonnes	ANNÉE	Full Date 🗵 TRIMESTRE(Full D
	Ξ Lignes	SOMME(Sa	ales Amou
	Feuille 10		
			Full Date Alternate Key
		2005	2006 2007
natique 🔻	5500K		
ÐI			Filtrer le champ [Full Date Alternate Key]
Taille Étiquette	5000K		Quel filtre voulez-vous appliquer à [Full Date Alternate Key] ?
Fobulle Chemin	4500K		 Date relative Plage de dates
	HOUR		# Années # Trimestres # Mois
	4000K		# Jours # Numéros de semaine
	3500K		Hois / Année Jour / Mois / Année Jour / Mois / Année
	3000K		# Total # Total (Distinct)
	Sale		Image: Minimum Image: Minimum Image: Minimum
	2500K		🛱 Attribut
le	2000K	/	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



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



 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 Count	ry Regi St	ate Province I	Name									
			🗄 Lignes	French Produ	ct Cate Fr	ench Product	Subc									
Filtres			Feuille 1	.2												
				I					Frencl	h Country Reg	ion Name / St	tate Province N	ame		1	
Repère	s		French	French Product			Allema	gne			New South		Australie			
			Product Cat.	. Subcategory Name	Bayern 8	Brandenbu	Hamburg	Hessen	Nordrhein	Saarland	Wales	Queensland	Australia	Tasmania	Victoria	Alberta
T AL	utomatique	•	Accessoire	Bidon et porte-bid	670	107	1041	1193	1 2 2 9	1176	4972	2 5 5 4	666	282	2771	28
::	Ð	Т		Casque	3 3 5 9	455	3 779	5 563	4 899	5 703	19734	9 692	3 0 7 9	1540	10 917	35
Couleur	Taille	Texte		Garde-boue	528	22	637	571	835	945	2 967	1714	462	242	1758	22
				Nettoyant	40		95	111	151	111	819	310	103	48	429	
000				Pneu et chambre à	2 569	197	3 3 3 7	4 365	5 0 6 8	5 3 2 5	20 314	10 229	2 566	1606	12078	90
Details	Intobulie			Porte-vélo	360		360	600	720	840	2 0 4 0	1560	600		1800	240
T	SOMME(Sa	ales		Range-vélo	159		318	477	1113	1113	5 0 8 8	3 0 2 1	318	159	1749	
				Sac d'hydratation	165		220	440	825	440	3 904	2 475	1045	165	2914	
			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	1062537	344 305	135 212	1 258 584	14 022
				VTT	141 783	15720	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	1762	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 1 2 9	2 520	840	280	3 010	70
				Gants	171	49	318	539	563	686	3 478	1714	416	171	1984	24
				Maillot	1864	420	2 368	2 568	3741	4177	17 983	7 585	2 266	1688	10 140	104
				Veste	254	64	508	572	572	699	2 985	1334	445	381	2 0 9 6	
					<											

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

Pages			iii Colonnes	French Cou	ntry Reg	i State Pro	ovince Name										
			∷ Lignes	French Proc	duct Cate	e French P	roduct Subc										
Filtres			Feuille 1	12													
							Allema	one		French Coun	try Region Nar	me / State Pro	ovince Name				Canada
Repère	s		French Product Cat.	French Product . Subcategory Name	Bayern	Brandenbu	Hamburg	Hessen	Nordrhein	Saarland	New South Wales	Queensland	South	Tasmania	Victoria	Alberta	British Columbia Ontar
T A	utomatique	•	Accessoire	Bidon et porte-bid	670						4 972		666		2771		
::	Ø	Т		Casque									3 0 7 9				
Couleur	Taille	Texte		Garde-boue									462				
				Nettoyant									103				
000 Détails	Infobulle			Pneu et chambre à .				4365			20 314		2 566				39 224
Details	Intoballe			Porte-vélo									600				
T	SOMME(S	ales		Range-vélo				477					318				
				Sac d'hydratation									1045				
			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	264 735
				Vélo de route	91 253	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	921 594
				VTT	41 783	15 720	160 819	213 590	215 009	256 880	1 219 303	642 065	200 786	70 702	720 963	2 295	613 145
			Vêtements	Casquette									288				
				Chaussettes	54			72					108				
				Cuissards									840				
				Gants							3 478		416			24	6074
				Maillot									2 266				21774
				Veste									445				

- Puis faire une clic droit et sélectionner Conserver uniquement

- Faire un filtre en conservant uniquement la région de Washington

Pages	iii Colonnes	French Cour	ntry Regi State Province Name)								
	⊞ Lignes	French Prod	uct Cate French Product Subc									
Filtres	Feuille 1	.2										
State Province Nam.		French Count										
		États-Unis										
Repères	French Prod.	. French Product Su	Washington									
	Accessoire	Bidon et porte-bid	5 866									
T Automatique		Casque	23 303									
: 0 I		Garde-boue	6 045									
Couleur Taille Texte		Nettoyant	882									
•		Pneu et chambre à	. 26 833									
Détails Infobulle		Porte-vélo	4 560									
Details		Range-vélo	4 770									
T SOMME(Sales		Sac d'hydratation	3 849									
	Vélo	Vélo de randonnée	300 725									
		Vélo de route	1 146 761									
		VTT	905 906									
	Vêtements	Casquette	1591									
		Chaussettes	593									
		Cuissards	10 149									
		Gants	3 576									
		Maillot	17 777									
		Veste	4 064									

- "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	iii Colonnes French Country Regi. State Province Name	
	E Lignes French Product Cate. French Product Subc.	
Filtres	Fauille 12	French Product Subcategory Name
State Province Nam	Tedille 12	(Tout)
French Product Sub	French Count	Antivol
Trendit Todader bub.	États-Unis	Axe de pédalier
	French Prod French Product Su Washington	Barre d'appui
Repères	Accessoire Casque 23 303	Bidon et porte-bidon
	Vélo Vélo de route 1146761	Cadre de vélo de randonnée
T Automatique 🔻	Vêtements Maillot 17777	Cadre de vélo de route
•• • •		Cadre de VTT
Coulour Taile Texts		✓ Casque
Couleur Taille Texte		Casquette
		Chaîne
Détails Infobulle		Chaussettes
		Collants
E COMMECONCS		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
		U VTT

- 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



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 •

		Décrire	ai icai	
Aodifier le paramètre [T0	OP Customer]			>
Iom: TOP Customer				Commentaire >>
Propriétés				
Type de données : E	intier	•		
Valeur actuelle : 1	5			
Format d'affichage :	lutomatique	~		
Valeurs autorisées :) Tout 🔵 Liste 💿 Plag	e		
Plage de valeurs				
Minimum :	5		Définir à partir	du paramètre 🗼
Maximum :	20		Définir à part	ir du champ 🕨
✓ Taille d'incrément :	5			
_			OK	Appular

Afficher la commande de parame	etre	
Couper		
Copier		
Modifier		
Dupliquer		
Renommer		
Masquer		
Supprimer		
Créer	- F	Champ calculé.
Propriétés par défaut	•	Paramètre
Dossiers	•	
Remplacer les références		

- Créer un ensemble de données sur le « Last Name » comme sur la l'ère 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

AUC	House owner Hag					Kaitiyn	
Abc	Last Name			6 - E		Lucas	
Abc	Marital Status	Ajouter à la feuille				Marcus	
Abc	Middle Name	Afficher le filtre				Morgan	
T F	Name Style	Dupliquer				Natalie	
Abc	Phone	Benommer				Seth	
Abc	Spanish Education	Nenoniniei M				Taylor	
		Masquer		Exclu	us	Aaron	
Mesur	es	Alias		L .		Abby	
#	Extended Amount	Créer	•	-	Champ	calculá	
#	Freight	Transformer			Champ	calcule	
#	Order Quantity	Transformer	·		Groupe	····	
#	Product Standard Cost	Convertir en Mesure		Ø	Ensem	ble	
	Salas Amount	Modifier le type de données			Parame	ètre	
++	Sales Amount	noumer le type de données		-		A - 1	-
			-	-			

Créer u	r un ensemble	>
Nom :	: Set de données TOP X	
Géne	énéral Condition Premiers	
0) Aucun	
• F) Par champ :	
	Premiers TOP Customer par	
	Sales Amount	•

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

Pages	iii Colonnes	SOMME	(Sales Amou														
	⊞ Lignes	INCLUS	/EXCLUS(📎	Last Name													
Filtres	Feuille	14															TOP Customer
	Inclus/non i	Last Name										 					•
	Inclus	Diaz					_					 	3	301066		^	
Repères		Gonzalez										 	-	302 109			INCLUS/EXCLUS(Set d
00 Automatique		Martinez										 	2	305 998	5		Inclus
		Sanchez					-					 	2	200 6	nc		Exclus
	Exclus	Adams	_			108 228								303 00	00		
Couleur Taille Etiquette	Excitab	Adonile	90			100 220											
		Alan	5 288														
Détails Infobulle		Albrecht	40														
🚼 INCLUS/EX 📎		Alexander					136 939										
T SOMME(Sales		Allen			88 269												
		Alonso						158 391									
		Alvarez									216 359						
		Anand						158 533									
		Andersen							189	199							
		Anderson				115 773											
		Arthur		50 258													
		Arun			80 905												

- 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

D	onnées	Analyses	nalyses \$								
6	Control DimProduct+ (AdventureWorks										
D	imensions	م 🏢	•								
	Créer un champ	calculé		IS IS							
	Créer un param	ètre									
_	Regrouper par dossier										
٠	Regrouper par t	able de sources de c	lonné	ies pre							
•	Trier par nom			А							
	Trier par ordre d	e la source de donn	ées								
	Masquer tous le	s champs inutilisés		u							
	Afficher les cha	mps masqués									

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



- 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 ?



 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.

Pages	iii Colonnes	SOM	IME(Sales Amou														
	⊞ Lignes	INCL	.US/EXCLUS(📎	Nom du 1	ГОР	Noi	m Client Dyr	namiq									
Filtres	Feuille 1 Nom du TOP	4 Nom Client	301.055														TOP Customer
Repères	Clients	Garcia Gonzalez Hernandez Kumar Lopez Martin Martinez Perez Ramos Rodriguez Sanchez Torres Xu Autres	258 078 287 344 302 109 297 736 259 275 244 570 253 857 305 998 268 605 242 638 266 557 299 539 268 420 309 606	4M	6M	8M	10M	12M	14M	16M	18M	20M	22M	24M	25 193 27 26M	7728M	Réduire/Elargir
									Sales Amo	ount							

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

Abc

Mesur

- 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 l'ère 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

Last Name				Kaitiyn Lucas			
Marital Status Middle Name	Ajouter à la feuille Afficher le filtre		Marcus Morgan				
Name Style Phone Spanish Education	Dupliquer Renommer Masquer	Dupliquer Renommer Masquer					
es .	Alias		_ EXClus	Abby			
Extended Amount	Créer	+	Char	mp calculé			
Preight Order Quantity	Transformer	•	Grou	ipe			
Product Standard Cost	Convertir en Mesure		Ensemble				
Sales Amount	Modifier le type de données	•	Para	mètre			
T A1	DAL - 4			A : 4 = -			

Créer	r un ensemble	×
Nom :	: Set de données TOP X	
Gér	néral Condition Premiers	
0	Aucun	
۲	Par champ :	
	Premiers TOP Customer par	
	Sales Amount	•

• Reproduire la visualisation suivante :

Pages	iii Colonnes SO	MME(Sales Amou					
	E Lignes	CLUS/EXCLUS(📎 🗍	Last Name	Customer Alternate			
Filtres	Feuille 14 (2)						
	Inclus/non i Last Name	Customer Alter					
	Inclus Domingues	z AW00012131				13 266	^
Repères	Gonzalez	AW00012300				13 243	
	He	AW00012308				13 269	
00 Automatique	Henderson	AW00012132				13 294	
21 A T	Nara	AW00012301				13 295	
Couleur Taille Étiquette	Exclus Adams	AW00011172	129				
		AW00011217			5 953		
		AW00011659	235				
Details		AW00011869	269				
T SOMME(Sales		AW00011931	234				
		AW/00011071	19/				

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

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

Pages	iii Colonnes	SOMME(Sales Amou.									
	⊟ Lignes	INCLUS/EXCLUS(@	French Country	Regi Last Name	Customer Alternat	e					
Filtres	Feuille 14 (2	2)								TOP Customer	r
	Inclus/non i Fren	ich Country Last Name	Customer Alter							0	< < >
	Inclus Fran	ice Cai	AW00011241				11 469		^		
Repères		Dominguez	AW00012131					13 266		Réduire/Elarg	ir
all Automatique		Gill	AW00012124					13 196		Páduira	-
100 Automatique		Gonzalez	AW00012300					13 243		Reduire	•
		He	AW00012308					13 269		Pays	
Couleur Taille Étiquette		Henderson	AW00012132					13 294			
		Hu	AW00012321					13 216		France	•
		Munoz	AW00011242				L1 068				
Details			AW00011439				12	2 489			
T SOMME(Sales		Nara	AW00012301					13 295			
		Sara	AW00012296					13 165			
		Shan	AW00011433					12910			
		She	AW00012307					13 173			
		Turner	AW00011420				11 201				
		Zheng	AW00011417				11 248				

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

- 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

Type de données :		
Valeur actuelle :		
Format d'affichage :	~	
Valeurs autorisées :	○ Tout	
Liste de valeurs		
Valeur	Afficher en tant que	Ajouter à partir du paramètre
Valeur Allemagne	Afficher en tant que Allemagne	Ajouter à partir du paramètre
Valeur Allemagne Australie	Afficher en tant que Allemagne Australie	Ajouter à partir du paramètre Ajouter à partir du champ
Valeur Allemagne Australie Canada	Afficher en tant que Allemagne Australie Canada	Ajouter à partir du paramètre Ajouter à partir du champ Coller à partir du Presse-papiers
Valeur Allemagne Australie Canada États-Unis	Afficher en tant que Allemagne Australie Canada États-Unis	Ajouter à partir du paramètre Ajouter à partir du champ Coller à partir du Presse-papiers
Valeur Allemagne Australie Canada États-Unis France	Afficher en tant que Allemagne Australie Canada États-Unis France	Ajouter à partir du paramètre Ajouter à partir du champ Coller à partir du Presse-papiers
Valeur Allemagne Australie Canada États-Unis France Royaume-Uni	Afficher en tant que Allemagne Australie Canada États-Unis France Royaume-Uni	Ajouter à partir du paramètre Ajouter à partir du champ Coller à partir du Presse-papiers
Valeur Allemagne Australie Canada États-Unis France Royaume-Uni Ajouter	Afficher en tant que Allemagne Australie Canada États-Unis France Royaume-Uni	Ajouter à partir du paramètre Ajouter à partir du champ Coller à partir du Presse-papiers Tout supprimer

- 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

Sales Amount /Pays			×
<pre>IF [Pays] = [French Country Region Name] THEN [Sales Amount] ELSE 0 END</pre>			
Le calcul est valide.	4 dépendances 🕶	Appliquer	ОК

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

Général	Condition	Premiers		
	n			
Par d	namp:			
Pre	miers	TOP Customer 🔹	par	
Sale	es 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

• Puis afficher les 25 meilleurs clients du pays : Allemagne

Pages	iii Colonnes	SOMME(Sales Amou																	
	🗄 Lignes	INCLUS/EXCL ⊗ ≩↓ Nom du TC	P	No	m Client	Dynami	iq												
Filtres	Fouillo	14																	TOP Customer
	reume.	14																	25
	Nom du TOP	Nom Client Dynamique																	
	TOP 25	Alvarez Wendy Allemagne AW00012129	8 5 5 7																
Repères	Clients	Anand Clarence Allemagne AW00011237	10 566																Réduire/Elargir
		Andersen Cassie Allemagne AW00014428	8 321																
000 Automatique 🔻		Bryant Sydney Allemagne AW00011412	9880																Réduire
		Carlson Victor Allemagne AW00014425	8 3 3 2																Paus
Couleur Taille Étiquette		Chander Jon Allemagne AW00013257	8941) ays
		Chandra Adriana Allemagne AW00014181	8 3 2 6																Allemagne 🔻
		Dominguez Anne Allemagne AW00013258	8 8 8 7																
Details Infobulle		Dominguez Desiree Allemagne AW00011427	9918																INCLUS/EXCLUS(Ense
👬 (INCLUS 🛞 👬		Garcia Bryant Allemagne AW00011431	9913																Inclus
T SOMME(Sales		Hughes Madison Allemagne AW00014426	8 3 1 9																Exclus
		Kapoor Cory Allemagne AW00014939	8 367																
		Martin Krista Allemagne AW00014189	8 3 1 7																
		Munoz Kristy Allemagne AW00012128	8681																
		Ortega Diana Allemagne AW00012125	8 5 9 5																
		Patel Cindy Allemagne AW00011249	9 8 9 0																
		Perez Deanna Allemagne AW00011428	9 9 5 4																
		Rana Carmen Allemagne AW00014940	8 368																
		Rubio Latasha Allemagne AW00011246	10 575																
		Saraiva José Allemagne AW00014207	8 361																
		Stewart Jasmine Allemagne AW00011423	9 905																
		Sun Amy Allemagne AW00011421	9780																
		Vazquez Ricky Allemagne AW00011245	10 580																
		Ward Isabella Allemagne AW00014830	8 3 4 6																
		Xie Bonnie Allemagne AW00014775	8 4 0 4																
	Autres	Autres															29 13	0 590	
	-		0M 2M	414	6M	8M	10M	12M	14M	16M	18M	2014	22M	24M	26M	28M	30M	32M	
			2101	-+101	0101	OIVI	10101	12111		C-Low	10101	2011	22111	2-+IVI	20101	20101	5010	JZ IVI	
										Sales A	Amount								

CRÉER UNE CARTE

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

Pages	iii Colonnes	
	🗮 Lignes	French Country Regi
Filtres	Feuille 17	
	French Country	
	Allemagne	Abc
Repères	Australie	Abc
	Canada	Abc
T Automatique	États-Unis	Abc
	France	Abc
Couleur Taille Texte	Royaume-Uni	Abc
ošo 🖓 Détails Infobulle		

• 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



CAS PORTANT SUR L'ANALYSE DES VENTES

• Contexte :

Adventure Works Cycles est un fabriquant 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

- 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 meilleurs 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

- 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.

- 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 ?

CAS DES INFORMATIONS NUTRITIONNELLES DU MENU DE MCDONALD'S

• 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

 RÈGLEMENT (UE) No II69/2011 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 : <u>https://www.senat.fr/europe/textes_europeens/ue0120.pdf</u>

- 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

- 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 : <u>http://onlinehelp.tableau.com/current/pro/desktop/en-us/help.htm#visual_best_practices.html</u>

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 : Forme 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"
 ?