Thème n°5 – Données structurées Cours – Thème: Données structurées

Voir page 90 de votre livre et suivantes



Objectif

"Triturer" des données structurées avec un tableurgrapheur. C'est fou ce que l'on peut faire avec des cellules et nos petites cellules grises;)

Nous allons réaliser le graphique de l'évolution démographique de votre commune !

Indiquez votre nom : Votre réponse

Table des matières

1-Source des données	2
2-Comprendre la structure des données avec le bloc-note	2
3-Données et structures et tableur-grapheur	3
3.1-Mise en place pour le tableur	3
4-Des problèmes des solutions	3
4.1-Liste des problèmes	4
4.2-R&R - La solution à beaucoup de problèmes	4
4.3-Bilan	5
5-Les outils pour la gestion des données	5
5.1-L' Autofiltre	5
5.2-Le tri des données	5
5.3-Trouver votre commune	6
6-L'assistant graphique	6
6.1-Mise en place pour le graphique	6
6.2-Les 4 étapes de l'assistant graphique	7
6.3-Bilan et remargue	8

1- Source des données

Le csv provient du site <u>data.gouv.fr</u> coordonné par <u>Etalab</u> qui est un département de la direction interministérielle du numérique (DINUM), dont les missions et l'organisation sont fixées par le décret du 30 Octobre 2019. Il coordonne notamment la conception et la mise en œuvre de la stratégie de l'État dans le domaine de la donnée.

L'objectif est de développer la politique d'ouverture et de partage des données publiques (open data), c'est pour cela que les contributeurs sont nombreux : services régionaux, municipalités, etc. Pour la région PdL, il y a aussi <u>https://data.paysdelaloire.fr/pages/home/</u>

En quelques lignes, selon vous qu'elle est l'intérêt d'ouvrir au monde des données publiques ? Votre réponse

2- Comprendre la structure des données avec le blocnote

- <u>Télécharger le fichier de données au format csv</u>
- Ouvrir le csv avec le bloc note puis répondre aux questions

Quel est le séparateur de champs ou d'étiquette Votre réponse

A quel étiquette appartiennent les valeurs "52" ? Votre réponse

Que représente selon vous la valeur "44" ? Votre réponse

Quelle est la population de Basse-Goulaine en 2018 (pnum18) ? Votre réponse

Quel est le séparateur de décimale dans ce fichier ? Votre réponse

Que représente l'avant dernière étiquette ? Votre réponse

Impressionnant non ? A quel siècle appartient l'année de cette avant-dernière étiquette ? Votre réponse

Thème n°5 – Données structurées 3- Données et structures et tableur-grapheur

3.1- Mise en place pour le tableur

Lancer le tableur puis ouvrir le fichier csv. Entourer les deux paramètres importants dans la copie d'écran de droite :

(vous pouvez utiliser les polygones rouge du dessous)



mporter Jeu de caracté Langue :	eres : U								
<u>J</u> eu de caracte <u>L</u> angue :	eres : U								
<u>L</u> angue :		nicode (UTI	F-8)	~					
Eangaci	- Dar défaut - Françair (França)								
			rançais (France)						
À partir de la	ligne : 1	-							
ptions de sépa	arateur								
O Largeur fix	e		Séparé	par					
 	on 🔲	<u>V</u> irgule	Point-virgule Espace	Autre					
Eusionn	er les sépar	rateurs	Espaces superflus	Séparateur de	e chaîne de o	aractères :	" ~		
-	· · ·			1 - C	-				
utres options									
Formater I	es champs	entre guille	emets comme texte 🗹 Detec	ter les nom <u>b</u> res	speciaux				
hamps									
			1						
Type de <u>c</u> olor	ine:	\sim							
Standard	Standard	Standard	Standard	Standard	Standard	Standard	S ^		
1 codgeo	reg	dep	libgeo	pmun18	pmun17	pmun16	р		
2 44005	52	44	Chaumes-en-Retz	6827.0	6759.0	6691.0	6		
3 44009	52	44	Basse-Goulaine	9093.0	9036.0	8867.0	8		
- 4404C	52	44	La Boissière-du-Doré	1049.0	1045.0	1040.0			
4 44016	E 2	44	Campbon	3979.0	3998.0	4018.0	1		
5 44016	p2		Casson	2331.0	2303.0	2217.0	3		
4 44016 5 44025 6 44027	52 52	44					1 3 2		
4 44018 5 44025 6 44027 7 44039	52 52 52	44 44	Cheix-en-Retz	1086.0	1067.0	1047.0	1 3 2 1		
4 44016 5 44025 6 44027 7 44039 8 44073	52 52 52 52	44 44 44	Cheix-en-Retz Héric	1086.0 6168.0	1067.0 6049.0	1047.0 5930.0	1 3 2 1 5		
4 44016 5 44025 6 44027 7 44039 8 44073	52 52 52 52	44 44 44	Cheix-en-Retz Héric	1086.0	1067.0 6049.0	1047.0 5930.0	1 3 2 1 5		
4 44016 5 44025 6 44027 7 44039 8 44073 «	52 52 52	44 44	Cheix-en-Retz Héric	1086.0 6168.0	1067.0 6049.0	1047.0 5930.0	1 3 2 1 5		
4 44016 5 44025 6 44027 7 44039 8 44073 <	52 52 52	44 44 44	Cheix-en-Retz Héric	1086.0	1067.0 6049.0	1047.0	1 3 2 1 5 ~		

4- Des problèmes... des solutions

Rappel de l'objectif :

Réaliser un graphique de l'évolution de la population de votre commune du XIX^{ième1} à nos jours.

Si problème de lecture de dates en chiffre romaine, voir la note de bas de page ! <u>Il semblerait qu'il y ait</u> <u>débat au XXI^{ième} siècle dans les musées;</u>)



Population de la commune d'INDRE

¹ XIX = (10) + (-1+10) = 19 donc 19ième - La numération romaine est un système de numération additive <u>https://fr.wikipedia.org/wiki/Num%C3%A9ration_romaine</u>

4	.1-	Liste de	es pr	Format Texte				
	A1		\sim	∱x ∑	codgeo	 -		
		A	В	С	D	E	F	
	1	codgeo	reg	dep	libgeo	pmun18	pmun17	pm
	2	44005	52	44	Chau Format Numérique	6827.0	6759.0	669
	3	44009	52	44		9093.0	9036.0	886
	4	44016	52	44	La Boissière-du-Doré	1049.0	1045.0	104
	5	11005	50	11	Comphon	2070 0	2000 0	40

Puisque votre objectif est de faire un graphique comparable à celui du dessus, au regard des données du classeur je liste les problèmes suivants :

- Les étiquettes ne sont pas numériques ex format texte pnum18 => format numérique 2018
- Le séparateur de décimal est un point et le type de données dans la cellule est alors du texte ex 6827.0 =>6827,0
- Je ne sais pas où se trouve ma commune, il y a trop de communes (au moins 1238;)

4.2- R&R - La solution à beaucoup de problèmes

En utilisant la fonction **Recherche et remplacer** du tableur, remplacer les données d'une manière subtile. Je précise que vous ne cherchez pas dans le contenu d'une cellule entière. Pour vous aider, appuyez-vous sur le tableau des tâches du dessous :

Via le menu
Edition / Rechercher et Remplacer
Ctrl+F (AOO) ou Ctrl+H (LibreOffice)

Kechercher & remplacer	
Respecter la casse Mots er	ntiers uniquement
· · · · · · · · · · · · · · · · · · ·	
Tout re <u>c</u> hercher	égédent Rechercher le suigant Remplacer Jout remplacer
Autres options	
Sélection active se <u>u</u> lement	Remplacer à re <u>b</u> ours
L Expressions régulières	Styles de paragraphe
Rechercher des similarités	<u>S</u> imilarités
	<u>Commentaires</u>
Sensible à la diacritique	
	Format Aucun format
A <u>t</u> tributs	

Votre tableau des tâches (vous compléterez les XX!)

Tâches	Pourquoi	Plage concernée
Remplacer pmun par 20	Pour ne plus avoir pnum18 mais 2018 et avoir une valeur numérique.	E1:Q1
Remplacer psdc par 19	Pour ne plus avoir par exemple psdc99 mais <mark>xxxx</mark> et avoir une valeur numérique.	XX:Y1 ²
Remplacer ptot par (rien)	Pour ne plus avoir par exemple ptot1931 mais xxxx	Z1:AJ1

² Note à l'attention des enseignants : Pour l'année 1954 et 1936 j'ai modifié le csv source

Tâches	Pourquoi	Plage concernée
	et avoir une valeur numérique.	
Remplacer le . par la ,	Pour ne plus avoir par exemple 6827.0 format texte aligné à gauche MAIS 6827 format numérique aligné à droite	E1:AJ1239

4.3- Bilan

Votre classeur maintenant est fonctionnel :

- Données numériques alignées à droite
- Données textes alignées à gauche

D5		\sim	fx 🗵	g = Campbon				
	Α	В	С	D	E	F	G	Н
1	codgeo	reg	dep	libgeo	2018	2017	2016	20
2	44005	52	44	Chaumes-en-Retz	6827	6759	6691	66
3	44009	52	44	Basse-Goulaine	9093	9036	8867	86
4	44016	52	44	La Boissière-du-Doré	1049	1045	1040	10
5	44025	52	44	Campbon	3979	3998	4018	39
6	44027	52	44	Casson	2331	2303	2217	21
7	44039	52	44	Cheix-en-Retz	1086	1067	1047	10
8	44073	52	44	Héric	6168	6049	5930	58
9	44076	52	44	Jans	1378	1370	1342	13
10	44096	52	44	Mésanger	4697	4687	4686	46
11	44098	52	44	Missillac	5367	5342	5271	51
12	44105	52	44	Mouais	378	385	384	9

5- Les outils pour la gestion des données

Voir le menu Données.

5.1- L' Autofiltre

Mettre en place l'Autofiltre pour toutes les données et filtrer les données du 44.

Ce filtre supprime-t-il les données ? Votre réponse

Que se passe-t-il dans votre tableur en dessous de la ligne 24 ? Votre réponse doit contenir le mot « masquée »

Supprimez l'autofiltre

5.2- Le tri des données

Trier vos données par département PUIS par libgeo sans oublier de déclarer votre première ligne comme étiquette. inal_evolution-de-la-population-des-pays-de-la-loire-de-1876-a-2017.csv - OpenOffice Calc

chier E <u>d</u> ition <u>A</u> ff	ichage	Insertion F	orma <u>t</u>	<u>O</u> utils	Donnée <u>s</u> Fe <u>n</u> être Aid <u>e</u>						
🗟 • 볼 • 🔛 🛛	🎭 🛃	? 🗟 🖴	<u>R</u> 1	ABC ABC	<u>D</u> éfinir la plage	1 v	💩 🐉 🧸 🏙 2	2 H	🤌 💼 🗟 🤇	🔪 🛛 💽	Rechercl
Arial		✓	10	~ G	Trier		 % 🐉 號 😹	(🔄 🍕	= 🗆 🝷 🖄	· <u>A</u> - ,	
🗉 🗅 🦄 🐿 🖻	D5		\sim	5 2	Filtre +	*	<u>A</u> utoFiltre				
Par défaut		Α	В	С	<u>S</u> ous-totaux	V	Eiltre standard		F	G	н
Résultat Résultat2	1	codge(-	re	de	<u>V</u> alidité	4	Filtre spécial	18	2017 -	2016 -	201
Titre	2	44005	52	44	Opérations <u>m</u> ultiples		Companying on In filters	582 Z	6750	6601	- 6
Titre1	3	44009	52	44	Texte en colonnes		Masquer l'AutoFiltre	9093	9036	8867	8
	4	44016	52	44	<u>C</u> onsolider	-	masquerrautornate	1049	1045	1040	1
	5	44025	52	44	Plan et groupe			3979	3998	4018	3
	6	44027	52	44	Tableau croisé +			2331	2303	2217	2
	7	44039	52	44	Actualis <u>e</u> r la plage			1086	1067	1047	1
	8	44073	52	44	Héric			6168	6049	5930	E
	9	44076	52	44	Jans			1378	1370	1342	1
	10	44096	52	44	Mésanger			4697	4687	4686	2
	11	44098	52	44	Missillac			5367	5342	5271	ŧ
	10	44405						070	205	204	

final_evolution-de-la-population-des-pays-de-la-loire-de-1876-a-2017.csv - OpenOffice Calc Fichier Édition Affichage Insertion Format Outils Données Fenêtre Aide

i 🗟 • 😕 • 🔒 🖂	🖌 📝 🗟 🖴 🖎 🏜 😖	Définir la plage 🔟 👻 🎯
Arial	✓ 10 ✓ G	Selectionner la glage
		Irier
💷 🖻 🖏 🛅 🕻	E6 🗸 🏹 🛣	F <u>i</u> ltre
Par défaut		Sous-totaux
Dí de la		V-11-11-4
Resultat	1 codgeo reg dep l	validite
Kesultat2		
Titre	2 44001 52 44	Operations <u>multiples</u>
Ti41		- · ·

Tri	×
Critères de tri Options	
Respecter la casse	
✓ La plage contient des étiquettes de colonne	
✓ Inclure les <u>f</u> ormats	
<u>C</u> opier le résultat du tri vers :	
- indéfini -	
Ordre de tri <u>p</u> ersonnalisé	
lun.,mar.,mer.,jeu.,ven.,sam.,dim.	\sim
Langue Op <u>t</u> i	ons
Par défaut - Français (France) 🗸 🗸	~
Direction	
De <u>h</u> aut en bas (tri de lignes)	
O De gauche à droite (tri de colonnes)	
Zone de données : \$A\$1:\$AK\$1239 (sans nom)	
	OK Annuler <u>A</u> ide <u>R</u> établir

5.3- Trouver votre commune

Une recherche devrait vous permettre de la trouver via Edition / Rechercher

6- L'assistant graphique

6.1- Mise en place pour le graphique

Conseil :

Créer une nouvelle feuille puis coller la ligne d'étiquette puis en dessous la ligne concernant votre commune.

Comme nous n'avons pas besoin des colonnes « codreg », ni des « reg », ni des « dep », ni des « dep », ni des « departement », vous pouvez les supprimer.

(Le clic droit contextuel est votre ami! Enfin cela dépend du contexte ;)

Mais qu'est ce donc que le clic droit contextuel ? Votre réponse

Nombre d'habitants

3000

2000

1000 0

1860 1880 1900 1920 1940 1960 1980 2000 2020 2040 Année

6.2- Les 4 étapes de l'assistant graphique

Après avoir sélectionner les données, lancer l'assistant graphique via le menu Insertion / Diagramme...



4. É

<u>A</u>ide

Indre

Ax<u>e</u> X

Axe <u>Y</u>

Axe Z

Afficher les g Axe X

Année

Nombre d'habitants

🖂 A<u>x</u>e Y

Ax<u>e</u> Z

<< Précédent Suivant >> Terminer

◯ <u>H</u>aut

OBas

Annuler

6.3- Bilan et remarque

Faire un graphique c'est simple si l'on dispose de données numériques !



Evolution démographique

Type diagramme YX et axe des X du graphique, que constatez-vous ? Votre réponse

Ligne lissée, quel sens lui donnez-vous ? <mark>Votre réponse</mark>

Les deux séries permettant de faire ce graphique sont en ligne ou en colonne ? Votre réponse

Vos deux étiquettes sont en lignes ou en colonne ? Votre réponse