

Cours - Thème: Données structurées

Voir page 90 de votre livre et suivantes



5

Données et traitements

Objectif

"Triturer" des données structurées avec un tableur-grapheur. 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 remarque.....	8

Thème n°5 – Données structurées

1- Source des données

Le csv provient du site data.gouv.fr coordonné par [Etalab](https://etalab.gouv.fr) 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 <https://data.paysdelaloire.fr/pages/home/>

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 bloc-note

- [Télécharger le fichier de données au format csv](#)
- 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 ?

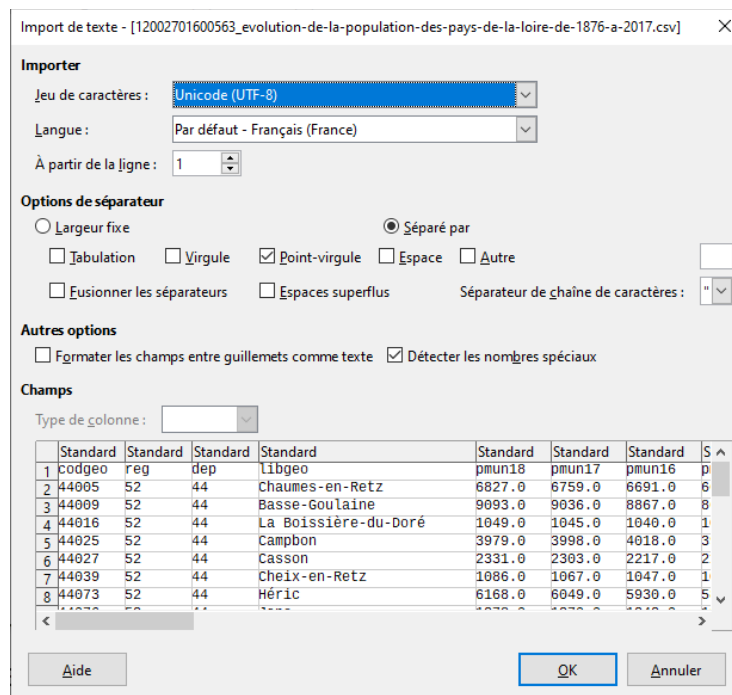
Votre réponse

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)

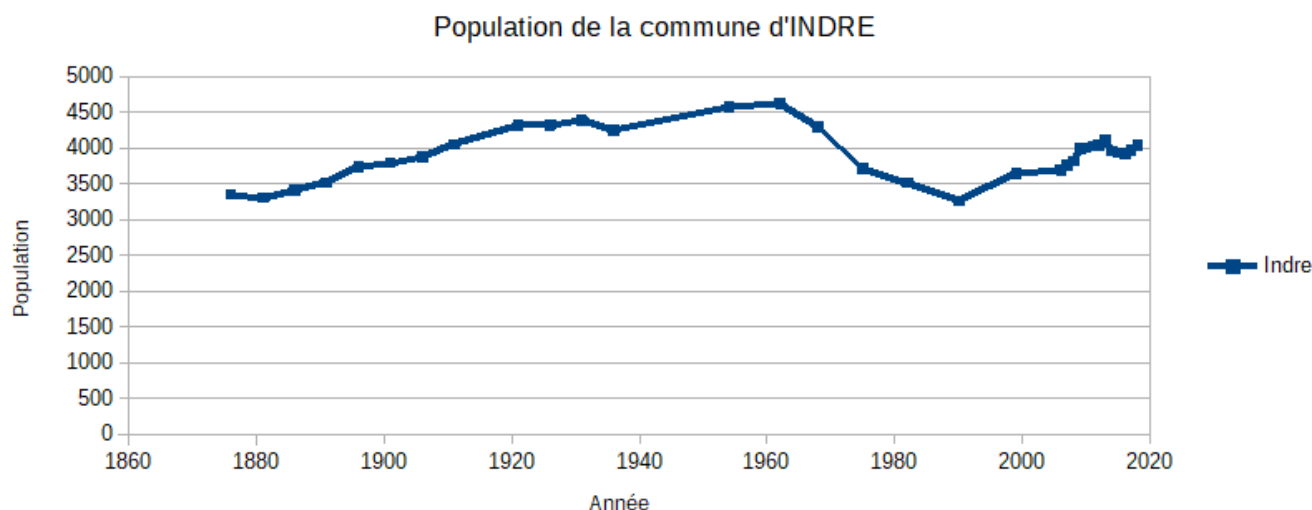


4- Des problèmes... des solutions

Rappel de l'objectif :

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

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



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

Thème n°5 – Données structurées

4.1- Liste des problèmes

The screenshot shows a spreadsheet with columns A through F. Column A contains codes like 'codgeo', '44005', '44009', '44016'. Column B contains 'reg', '52', '52', '52'. Column C contains 'dep', '44', '44', '44'. Column D contains 'libgeo', 'Chau', 'P', 'La Boissière-du-Doré'. Column E contains 'pmun18', '6827.0', '9093.0', '1049.0'. Column F contains 'pmun17', '6759.0', '9036.0', '1045.0'. A callout 'Format Texte' points to 'pmun18' and 'Format Numérique' points to '6827.0'.

	A	B	C	D	E	F
1	codgeo	reg	dep	libgeo	pmun18	pmun17
2	44005	52	44	Chau	6827.0	6759.0
3	44009	52	44	P	9093.0	9036.0
4	44016	52	44	La Boissière-du-Doré	1049.0	1045.0

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 **pnun18** => 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;)

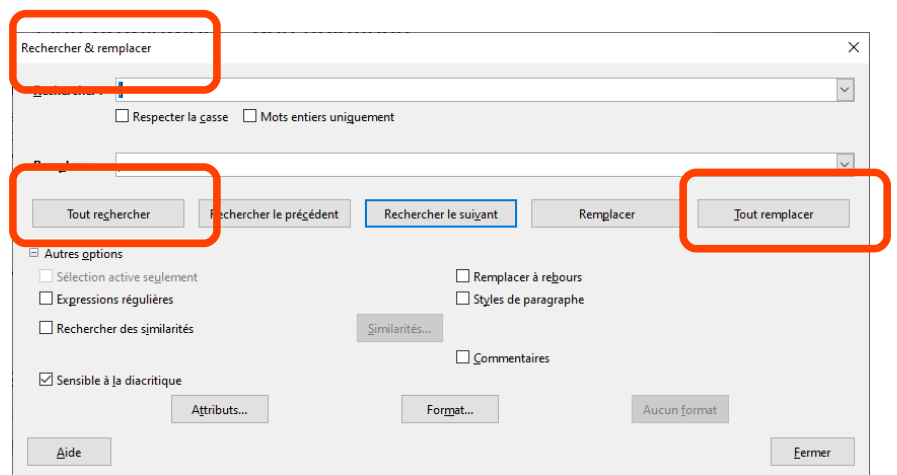
4.2- R&R - La solution à beaucoup de problèmes

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)



Votre tableau des tâches (vous complétez les **XX!**)

Tâches	Pourquoi	Plage concernée
Remplacer pmun par 20	Pour ne plus avoir pnun18 mais 2018 et avoir une valeur numérique.	E1:Q1
Remplacer psdc par 19	Pour ne plus avoir par exemple psdc99 mais XXXX 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

Thème n°5 – Données structurées

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

	A	B	C	D	E	F	G	H
1	codegeo	reg	dep	libgeo	2018	2017	2016	2015
2	44005	52	44	Chaumes-en-Retz	6827	6759	6691	6623
3	44009	52	44	Basse-Goulaine	9093	9036	8867	8800
4	44016	52	44	La Boissière-du-Doré	1049	1045	1040	1035
5	44025	52	44	Campbon	3979	3998	4018	3939
6	44027	52	44	Casson	2331	2303	2217	2130
7	44039	52	44	Cheix-en-Retz	1086	1067	1047	1028
8	44073	52	44	Héric	6168	6049	5930	5811
9	44076	52	44	Jans	1378	1370	1342	1314
10	44096	52	44	Mésanger	4697	4687	4686	4686
11	44098	52	44	Missillac	5367	5342	5271	5199
12	44105	52	44	Mouais	378	385	384	384

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.

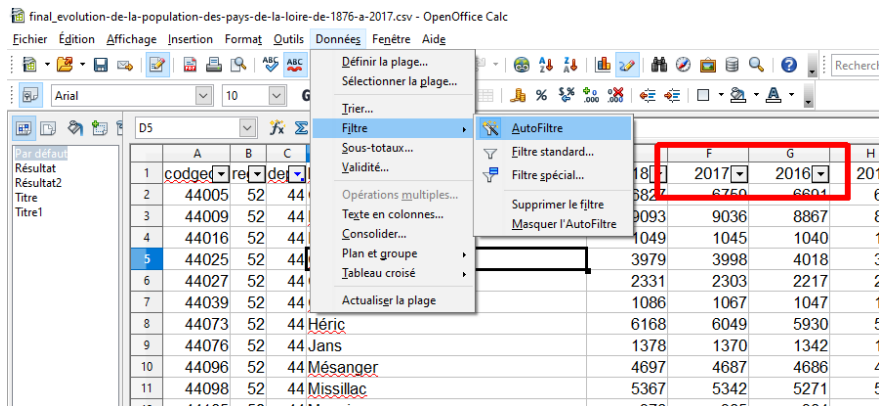
Ce filtre supprime-t-il les données ?

Votre réponse

Que se passe-t-il dans votre tableau en dessous de la ligne 24 ?

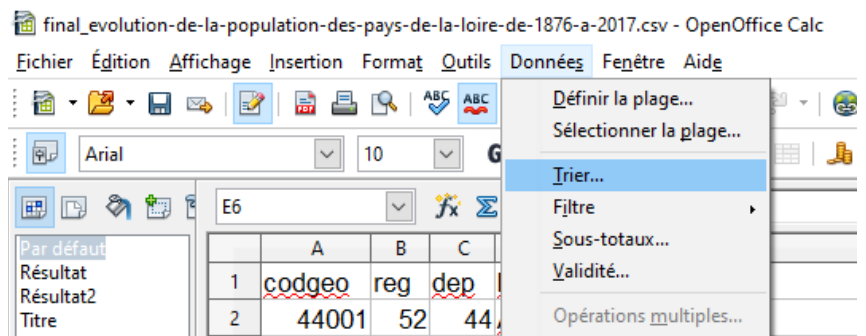
Votre réponse doit contenir le mot « masquée »

Supprimez l'autofiltre

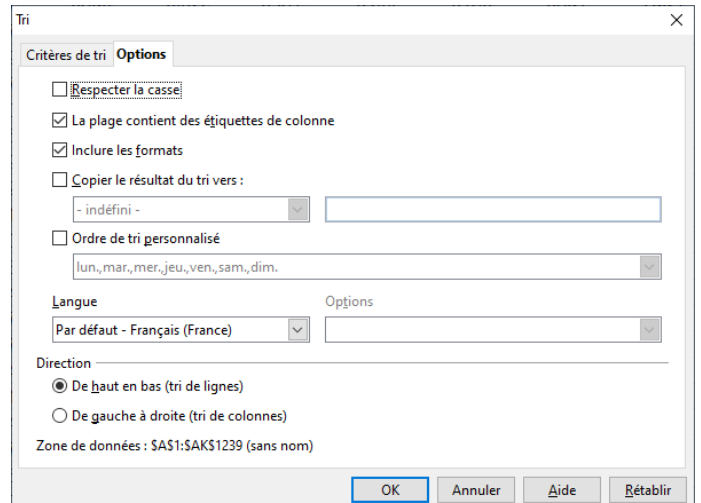


5.2- Le tri des données

Trier vos données par département PUIS par libgeo sans oublier de déclarer votre première ligne comme étiquette.



Thème n°5 – Données structurées



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 « département », vous pouvez les supprimer.

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

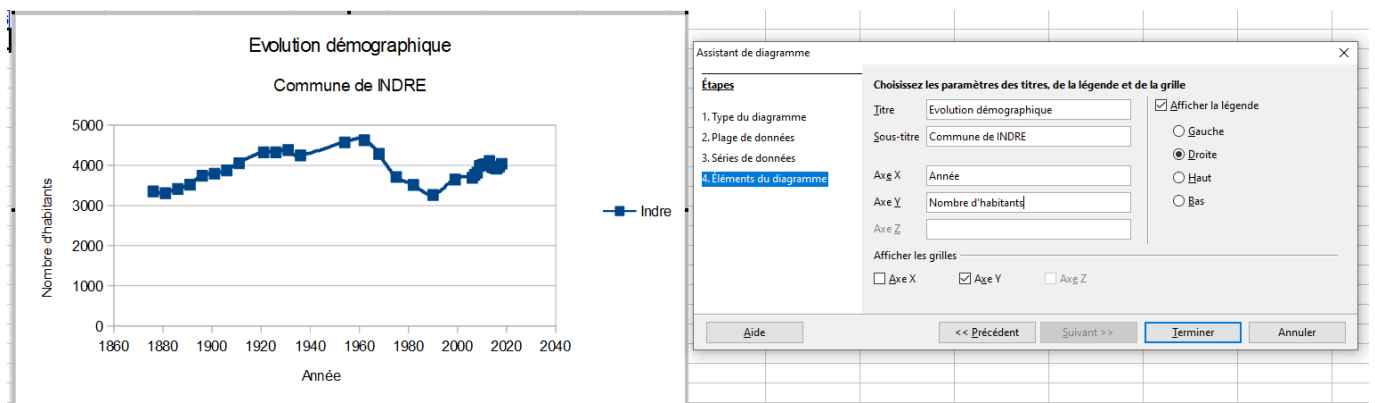
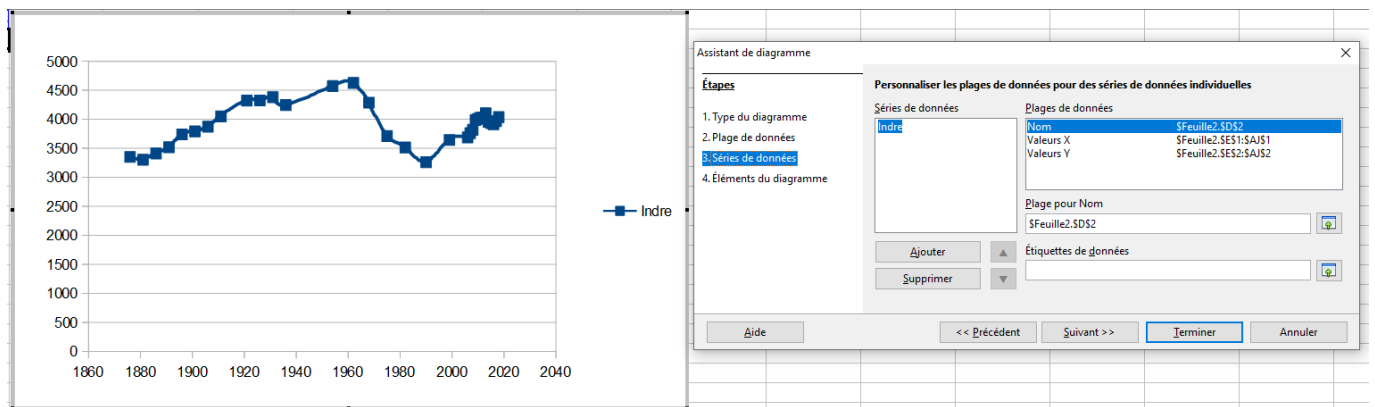
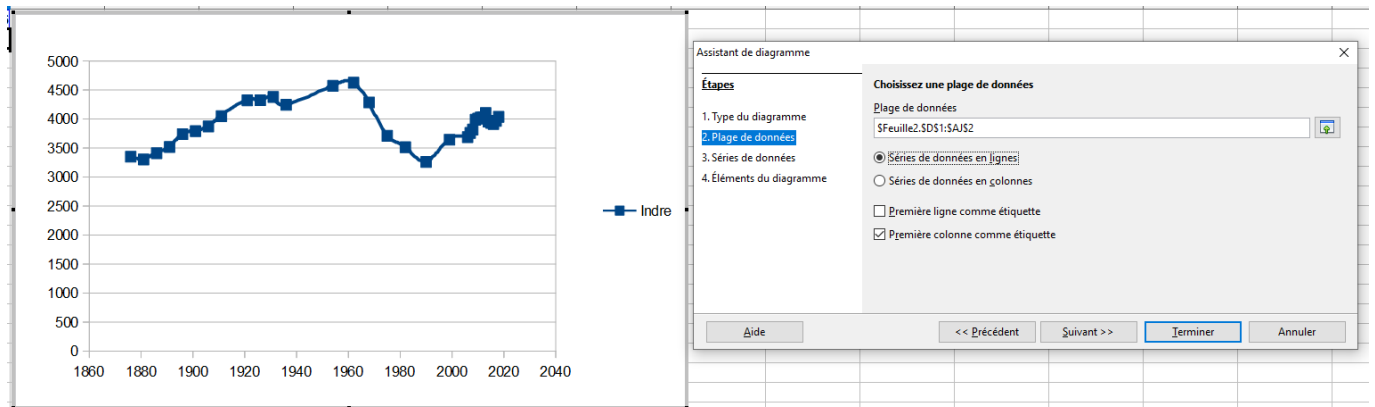
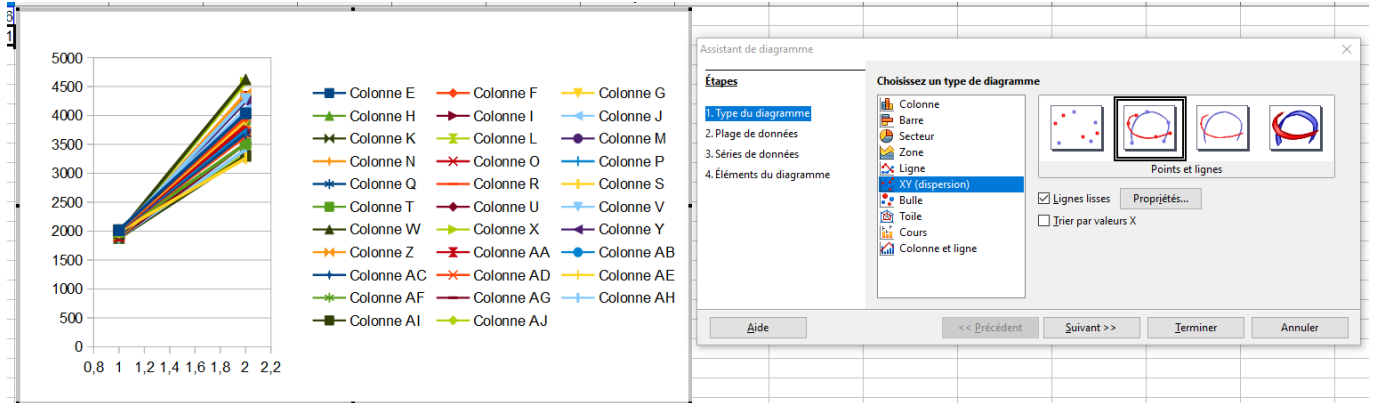
Mais qu'est ce donc que le clic droit contextuel ?

Votre réponse

Thème n°5 – Données structurées

6.2- Les 4 étapes de l'assistant graphique

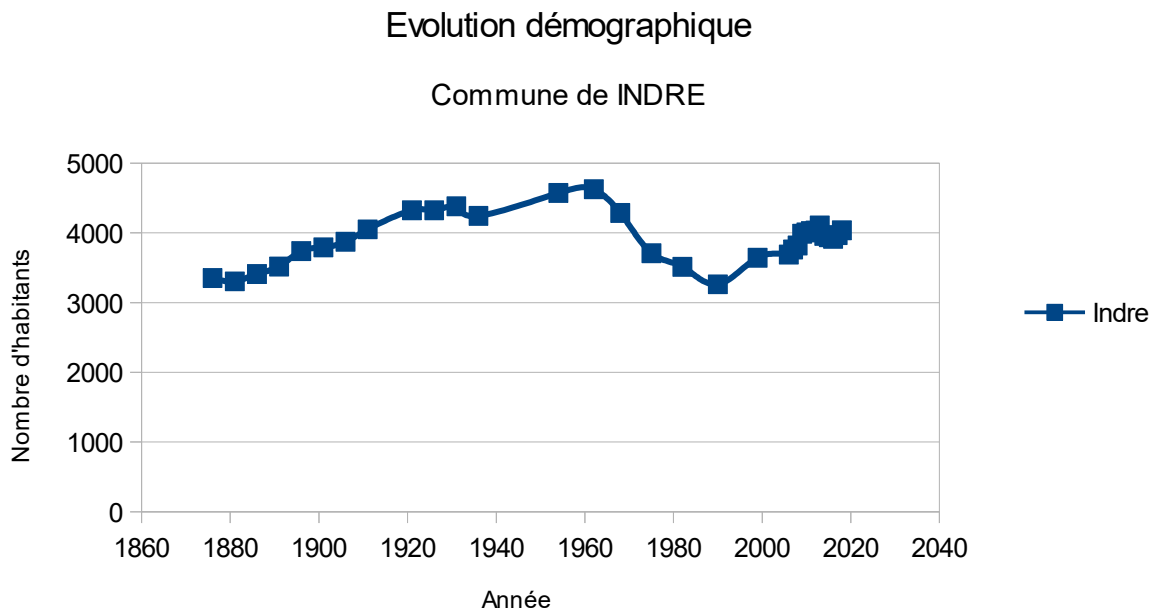
Après avoir sélectionner les données, lancer l'assistant graphique via le menu Insertion / Diagramme...



Thème n°5 – Données structurées

6.3- Bilan et remarque

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



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

Votre réponse

Ligne lissée, quel sens lui donnez-vous ?

Votre réponse

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