

Les données structurées.

Contenus	Capacités attendues
Données	Définir une donnée personnelle.
Données structurées	Identifier les différents descripteurs d'un objet. Distinguer la valeur d'une donnée de son descripteur.

Introduction

Les données constituent la matière première de toute activité numérique. Afin de permettre leur réutilisation, il est nécessaire de les conserver de manière persistante. Les structurer correctement garantit que l'on puisse les exploiter facilement pour produire de l'information. Cependant, les données non structurées peuvent aussi être exploitées, par exemple par les moteurs de recherche.

Repères historiques

- 1930 : utilisation des cartes perforées, premier support de stockage de données ;
- 1956 : invention du disque dur permettant de stocker de plus grandes quantités de données, avec un accès de plus en plus rapide ;
- 1970 : invention du modèle relationnel (E. L. Codd) pour la structuration et l'indexation des bases de données
- 1979 : création du premier tableur, VisiCalc ;
- 2013 : charte du G8 pour l'ouverture des données publiques.

Activité 1 : les cartes perforées.

Imaginons, vous avez une liste d'une trentaine de vos copains, et pour chacun vous avez noté sur une feuille la liste des plats parmi la vingtaine que vous aimez cuisiner. Ça ressemble à ça :

- Florent : pâtes carbonara, bouchées à la reine, ratatouille
- Claude : choucroute, bouchées à la reine, crumble ratatouille, pizza
- Liliane : lasagnes chèvre-épinards, gnocchi aux légumes, ratatouille, saucisse purée

Ça n'est pas particulièrement trié mais vous, au lieu de 3 amis, vous en avez 30. Maintenant choisissez une dizaine d'amis parmi les 30 et essayez de retrouver les plats aimés par tous ces amis. Pas facile ? Justement c'est là l'intérêt des bases de données.

- **Pour la cantine du lycée, essayons de voir le plat qui ferait plaisir à tout le monde afin d'avoir le moins de déchets possible.**
- **Par oral :**

Et puis quand tout le monde est d'accord sur le fait que c'est pénible et qu'il faut trouver une autre solution, on passe à la suite.

- **Les cartes perforées**

Pour cela il vous faut : une paire de ciseaux et des cartes (voir annexe) que l'on va vous distribuer.

Nom : <i>gross bouff</i>	Prénom : <i>Jolo</i>		
		Adore	Mange
Poulet frite			
Pizza			
Lasagnes épinard ricotta			
Cordon bleu petits pois			
Couscous			
Choucroute			
.....			

On réalise une fiche par élève, avec en première ligne son nom et sur chacune des lignes suivantes un des plats de la cantine. Ensuite pour chaque élève on fait un trou sur toutes les cartes en face de chacun des plats, s'il le mange mais n'en raffole (mange), ou s'il en raffole (adore).

➤ **Répondre aux questions suivantes :**

- Y a-t-il un plat que tout le monde adore ?
- Y a-t-il un plat que tout le monde mange ?
- Combien de personnes adorent le couscous ?

Activité 2 : notre premier fichier de données : vocabulaire.

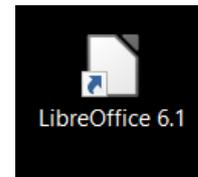
Les données et l'information

Une **donnée** est une valeur décrivant un objet, une personne, un événement digne d'intérêt pour celui qui choisit de la conserver. Par exemple, le numéro de téléphone d'un contact est une donnée. Plusieurs **descripteurs** peuvent être utiles pour décrire un même objet (par exemple des descripteurs permettant de caractériser un contact : nom, prénom, adresse et numéro de téléphone).

Une **collection** regroupe des objets partageant les mêmes descripteurs (par exemple, la collection des contacts d'un carnet d'adresses). La structure de table permet de présenter une collection : les objets en ligne, les descripteurs en colonne et les données à l'intersection. Les données sont alors dites structurées.

Pour assurer la persistance des données, ces dernières sont stockées dans des fichiers. Le format CSV (*Comma Separated Values*, les données avec des séparateurs) est un format de fichier simple permettant d'enregistrer une table.

➤ **Ouvrir un fichier calc dans libreoffice**



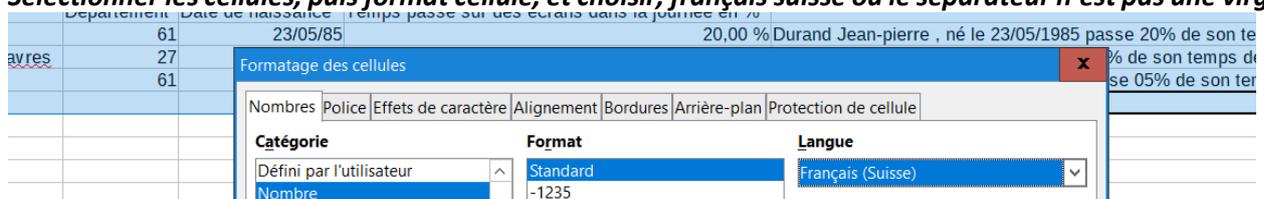
➤ **On crée les 6 Descripteurs suivants sur l'ordinateur du professeur :**

Fichier Édition Affichage Insertion Format Styles Feuille Données Outils Fenêtre Aide

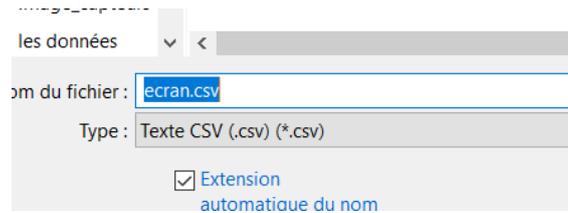
Liberation Sans 10 G I S A

	A	B	C	D	E	F
1	Nom	Prénom	Ville	Département	Date de naissance	Temps passé sur des écrans dans la journée en %
2						
3						
4						
5						

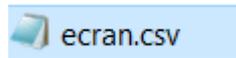
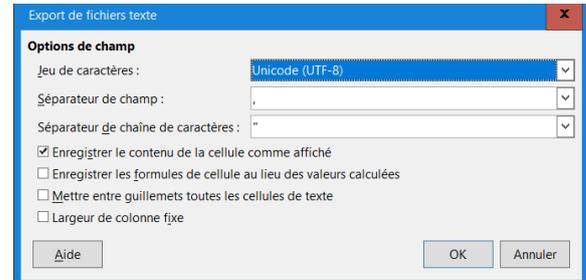
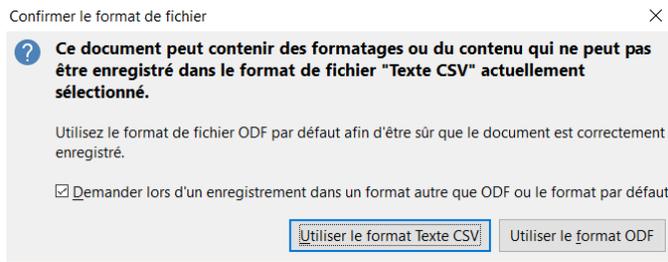
- **Chacun prépare par écrit ce qu'il va mettre dans chaque colonne, puis vient le rentrer dans le fichier.**
- **Quels problèmes apparaissent ?**
- **Comment y remédier ?**
- **Quelle est la collection des villes ?**
- **Sélectionner les cellules, puis format cellule, et choisir, français suisse où le séparateur n'est pas une virgule.**



- Une fois que tout le monde s'est mis d'accord enregistrez au format ODT.
- Puis enregistrez le fichier sous le format .csv ici ecran.csv

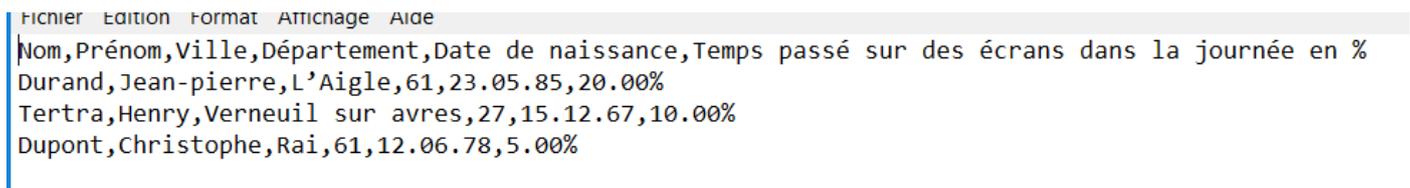


Apparaissent alors les fenêtres suivantes :



Retrouvez le fichier dans votre arborescence :

Puis, clic droit modifier. Cela aura pour effet d'ouvrir le fichier dans le bloc note. Vous devriez avoir un fichier de ce type :



- Que remarquez-vous ?

Activité 3 : premières manipulations de données : concaténation.

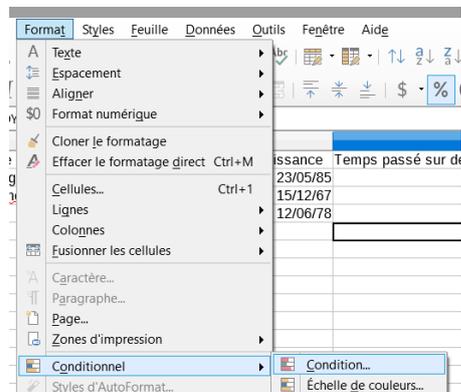
Dans un tableau, toute formule commence par le signe « = ». Dans une formule, on fait référence au nom de la cellule : on parle de référence relative. Si par exemple on a utilisé la cellule C4 alors par recopie verticale, le numéro de la ligne change et par recopie horizontale, la lettre donnant la colonne change. Pour empêcher le changement, on place un \$ devant la lettre ou le chiffre selon les besoins : on parle de référence absolue"

- Dans la case G2 taper le texte ci-dessous :
=CONCATENER(A2 ; " " ; B2 ; " ,né le " ; TEXTE(E2 ; "JJ/MM/AAAA"))

	A	B	C	D	E	F	G	H	I	J	K
1	Nom	Prénom	Ville	Département	Date de naissance	Temps passé sur des écrans dans la journée en %					
2	Durand	Jean-pierre	L'Aigle	61	23/05/85	20,00 %	=CONCATENER(A2;" ";B2;" ,né le ";TEXTE(E2;"JJ/MM/AAAA"))				
3	Tertra	Henry	Verneuill sur avres	27	15/12/67	10,00 %					
4	Dupont	Christophe	Rai	61	12/06/78	5,00 %					
5											
6											
7											

- Que remarquez-vous ?
- Répéter l'opération sur tout le fichier
- Quel descripteur pourrait-on donner à cette collection
- Modifier la formule de façon à faire apparaître : « Durand Jean-Pierre, né le 23/05/1985, passe 20% de son temps sur les écrans »
- Faire la moyenne du temps passer sur les écrans.
- En fonction du temps passé sur les écrans on veut affecter une couleur à la cellule. Pour cela on utilise la colorisation conditionnelle.

- Choisir, format, conditionnel, condition.



- Puis

Conditions

Condition1

La valeur de la cellule est entre et

Appliquer le style :

- Choisir : la valeur de la cellule est entre, mettre la valeur min et la valeur max.
- Au niveau du style, choisir nouveau style : la fenêtre suivante s'ouvre :

Style de cellule

Gestionnaire | Nombres | Police | Effets de caractère | Alignement | Bordures | Arrière-plan | Protection de cellule

Style

Nom :

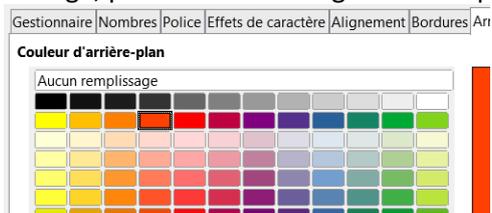
Hérité de :

Catégorie :

Contient

RGB(255, 64, 0), Non transparent

- Donner un nom au style, ici rouge, puis aller dans l'onglet Arrière-plan.



- Validez, revenir sur l'onglet gestionnaire, en bas choisir : plage de cellule

Plage de cellules

Plage :

- La sélectionner, ici de F2 à F4 :

Ville	Département	Date de naissance	Temps passé sur des écrans dans la journée en %
L'Aigle	61	23/05/85	20,00 %
Verneuil sur avres	27	15/12/67	10,00 %
Rai	61	12/06/78	5,00 %
			11,67 %

Formatage conditionnel pour \$F\$2:\$F\$4

- Et vous devriez avoir toutes les cellules comprises entre 0% et 8% en rouge

ce	Temps passé sur des écrans dans la journée en %	
i/85	20,00 %	Duranc
/67	10,00 %	Tertra
i/78	5,00 %	Dupon

- Faites de même avec deux autres couleurs pour 8% à 25% et 25% et plus.

