

Séq. 1 – Les données structurées

Nom :

Prénom :

/6

	--	-	=	+	++
Études de documents faites					
Exercices faits & corrigés					
Activités faites					

=== Introduction ===

☐

<http://lienmini.fr/3389-101>

=== Études de documents ===

☐

1 Les données personnelles et métadonnées

1.1 Qu'est-ce qu'une Donnée personnelle ?

Lire P. 16-17 les doc. 1, doc. 2 et répondre à la question. Donnez des exemples

1.2 Que sont les métadonnées ?

☐

Lire P. 17 le doc. 3 et répondre à la question. Donnez des exemples

1.3 Qu'est-ce que le big data ?

☐

Lire P. 17 le doc. 3 et regarder la vidéo <https://lienmini.fr/3389-104> et répondre à la question.

2 Les données structurées

☐

2.1 Les formats de données

P. 18-19 DOC 3 - Nommez les 3 formes de données structurées - Quel est le principe de chacune ?

2.2 Quelles opérations peut-on réaliser sur ces données ? ☐

P. 20-21 DOC 1 - Quels sont les plats rapides à préparer ? Quelle recette pour faire un dessert sans œuf, sans gluten et rapide à préparer ?

3 Le cloud ☐

3.1 Principes

Après avoir visionné la vidéo (https://www.youtube.com/watch?v=5YawCCUxa_E), répondez aux questions suivantes :

1. Qu'est-ce que le cloud ?
2. D'où vient le terme cloud ?
3. Quelles sont les évolutions techniques qui ont permis le développement le cloud ?
4. Quels sont les avantages du cloud ? Et les désavantages ?

3.2 Consommation énergétique du cloud ☐

Après avoir visionné la vidéo (<https://ladigitale.dev/digiview/#/v/68b59e6bb86e6>), faites un inventaire des principales causes de la consommation énergétique des centres de données.

=== Cours === ☐

1 Les données et métadonnées numériques

1.1 Les nombres binaires

1.1.1 Système décimal vs. système binaire

Parce que les humains ont 10 doigts, les mathématiques que nous utilisons se basent sur 10 chiffres différents : 0, 1, 2, 3, 4, 5, 6, 7, 8 et 9.

Les machines n'utilisent que 2 chiffres différents : 0 et 1. Il s'agit d'un bit (Binary digIT).

Dans le monde numérique, toutes les données (.pdf, .jpg, .avi et même les .exe) sont codées en binaire.

1.2 Notion de taille de fichier (en octet) ☐

Un groupe de 8 bits s'appelle un octet (ou byte en anglais).

Exemples :

- 10101010 est un octet
- 01111011 est un octet
- 0101111011010010 : deux octets

Un groupe de 1000 octets (donc 8000 bits) est 1 kilo-octet ou 1 ko (en français) ou 1 kb (en anglais).

Un groupe de 1000000 octets (donc 8000000 bits) est 1 mega-octet ou 1 Mo (en français) ou 1 Mb (en anglais).

Un groupe de 1000000000 octets est 1 giga-octet ou 1 Go (en français) ou 1 Gb (en anglais)

Un fichier musical de 3-4 minutes fait 3 Mo.

Une photo numérique prise avec un Smartphone fait environ 10 Mo.

Une vidéo basse définition de 3-4 minutes fait 30 Mo.

Un film basse définition fait 700 Mo.

Un film HD peut faire 2-3 Go.

1.3 Les métadonnées ☐

Ces petits textes (en caractères ASCII) ne prennent vraiment pas de place et en mettre en début d'un fichier image, ou audio, ou vidéo n'alourdit pratiquement pas le fichier et passe inaperçu.

Donc on a l'habitude d'en mettre afin de donner des informations supplémentaires sur le fichier lui-même (son auteur, sa date de création, sa date de modification, le logiciel utilisé pour le créer, ...).

Toutes ces informations sont appelées métadonnées du fichier.

2 Les données personnelles ☐

Une **donnée personnelle** est toute information se rapportant à une personne physique identifiée ou identifiable, directement ou indirectement.

- **Identification directe** : L'information désigne explicitement la personne (ex. nom, prénom, photo, adresse email nominative).
- **Identification indirecte** : L'information permet d'identifier la personne par croisement d'un ensemble de données (ex. numéro de téléphone, numéro de sécurité sociale, adresse IP, plaque d'immatriculation, données de géolocalisation, ou combinaison de critères comme l'âge, la profession et la commune de résidence).

3 Les données structurées ☐

3.1 Introduction

<https://ladigitale.dev/digiview/#/v/68b5990319ec3>

3.2 Définition

Les **données structurées** sont des informations organisées de manière rigoureuse selon un schéma ou un modèle prédéfini, ce qui permet à un ordinateur ou un système informatique de les lire, les classer et les analyser automatiquement.

3.3 Caract. des données structurées – Vocabulaire

Lire P. 18 le doc. 1 et le doc. 2 et complétez le schéma ci-dessous :

Nom	Capitale	Hymne	Superficie (km²)
France	Paris	La Marseillaise	632 734
Chine	Pékin	La Marche des volontaires	9 596 961
États-Unis	Washington	The Star-Spangled Banner	9 833 517
Argentine	Buenos Aires	Himno Nacional Argentino	2 791 810

3.4 Les fichiers csv

Le format CSV est très courant sur internet. Voici ce que nous dit Wikipédia sur le format CSV :
« *Comma-separated values, connu sous le sigle CSV, est un format informatique ouvert représentant des données tabulaires sous forme de valeurs séparées par des virgules.* »
Chaque ligne du texte correspond à une ligne du tableau et les virgules correspondent aux séparations entre les colonnes. Les portions de texte séparées par une virgule correspondent ainsi aux contenus des cellules du tableau.

Voici un exemple du contenu d'un fichier CSV :

```
nom,prenom,date_naissance
Durand,Jean-Pierre,23/05/1985
Dupont,Christophe,15/12/1967
Terta,Henry,12/06/1978
```

Nous avons ici 3 personnes :

- Jean-Pierre Durand qui est né le 23/05/1985
- Christophe Dupont qui est né le 15/12/1967
- Henry Terta qui est né le 12/06/1978

nom, prenom et date_naissance sont appelés des descripteurs alors que, par exemple, Durand, Dupont et Terta sont les valeurs du descripteur nom.

3.5 L' open data

On trouve énormément de données sur internet. Une partie de ces données sont publiques, par exemple le site data.gouv.fr recense un grand nombre de données publiques. Ces données sont librement réutilisables.

4 Le cloud

Le **Cloud** (ou *cloud computing*, « informatique en nuage » en français) désigne l'accès à des services informatiques — comme le stockage, les serveurs, les bases de données ou les logiciels — directement via Internet, sans avoir à posséder ni à gérer d'infrastructures physiques chez soi ou dans son entreprise.

Quizz pages 12 et 13 (<http://lienmini.fr/3389-102>)



Historique en vidéo pages 14 et 15 : <http://lienmini.fr/3389-103>

=== Exercices ===



P. 29 ex 5

P. 29 ex 7

=== Activités ===



1 Notions de fichiers, dossiers, programmes



1.1 Les fichiers informatiques

Un fichier informatique a :

1. Un nom
2. Très souvent une extension

Exemple :

- Sans Nom.odt
- Photo soirée.jpg

Quand il y a beaucoup de fichiers, on ne s'y retrouve pas. IL Y A INTERET A DONNER DES NOMS SIGNIFICATIFS.

L'extension renseigne sur la nature du fichier. Il y a des fichiers bureautiques, des fichiers images, des fichiers vidéos, ...

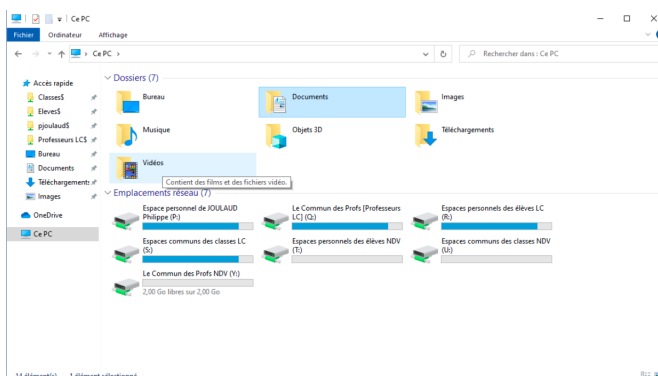
Nous pouvons manipuler les fichiers informatiques grâce à un « explorateur de fichiers ».

Il y a plusieurs façons de lancer un « explorateur de fichiers » :



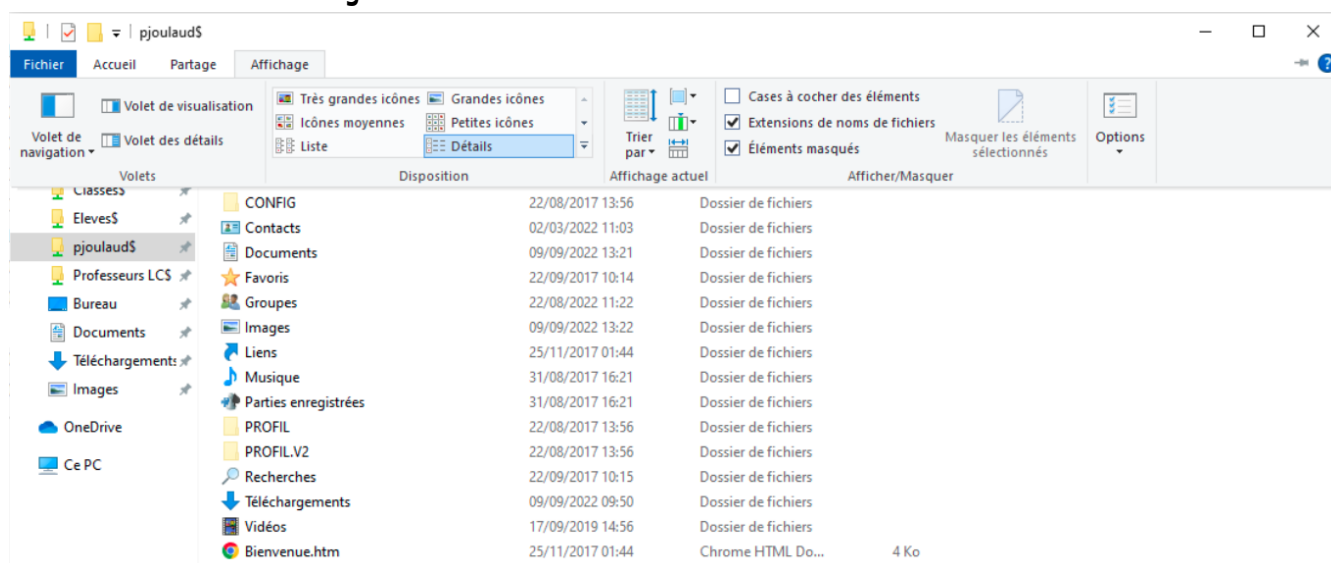
1. Clic droit sur le bouton démarrer (qui est tout à gauche sur la capture d'écran ci-dessus) et choisir Explorateur de Fichiers
2. Clic gauche sur l'icône Fichiers (qui est en jaune sur la capture d'écran ci-dessus)
3. Appuyer sur les touches Windows et E
4. Autres

Vous devez avoir une fenêtre comme ceci :



A faire vous même 1.

- Ouvrez un « explorateur de fichiers »
- Allez dans Affichage



- Afin de les afficher, cochez les cases Extensions de noms de fichiers et Eléments masqués
- Listez tous le fichier présent dans votre dossier personnel :
- Quelle extension a ce fichier ?
- Donnez 6 exemples d'extension. A quels types de fichiers correspondent-ils ?

1.2 Les dossiers informatiques (ou répertoires)

Un dossier informatique a :

1. Un nom (et c'est tout!)

Exemple :

- Mes Documents
- Téléchargements

Il s'agit d'éléments qui contiennent les fichiers.

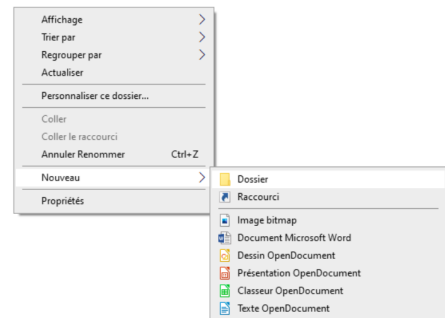
Si tous les fichiers étaient dans un même dossier, il arriverait un moment où on ne s'y retrouverait pas. D'OU L'INTERET DE CREER DES DOSSIERS AVEC DES NOMS SIGNIFICATIFS ET DE REFLECHIR A L'ENDROIT OU ON ENREGISTRE LES FICHIERS.

Nous pouvons voyager de dossier en dossier grâce à l'« explorateur de fichiers » vu plus haut.

A faire vous même 2.

Nous allons structurer notre espace numérique pour que nous puissions nous y retrouver tout au long de l'année.

- Ouvrez un Explorateur de fichiers
- Allez dans Documents
- Créez 4 dossiers: SNT, Math, Français, Histoire



- Allez dans SNT et créez les 8 dossiers suivants :
 - 0_programmation_python
 - 1_données_structurées
 - 2_internet
 - 3_le_web
 - 4_localisation_cartographie_mobilité
 - 5_images_numériques
 - 6_réseaux_sociaux
 - 7_informatique_embarquee_objets_connectés

Appelez le professeur

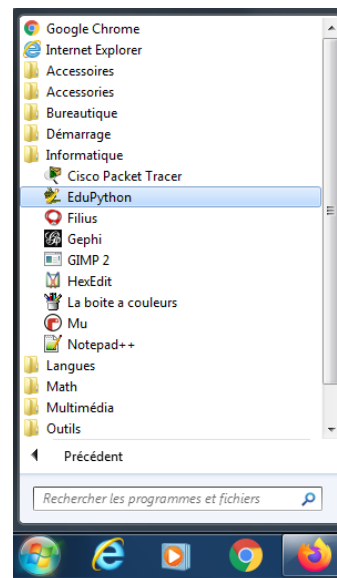


1.3 Des fichiers pas comme les autres : les programmes (ou logiciel)

Il existe des fichiers qui sont différents car ils peuvent se lancer ou s'exécuter. Vous double-cliquez dessus et il se passe quelque chose (souvent une nouvelle fenêtre s'ouvre).

La grande majorité de ces fichiers ont pour extension .exe

La plupart d'entre eux sont accessibles en cliquant sur le menu Démarrer



A faire vous même 3.

- Allez dans le menu Démarrer
- Cherchez et lancez les logiciels suivant :
 - Bureau => LibreOffice/OpenOffice
 - Multimédia => VLC
 - Informatique => Gimp
 - Informatique => Edupython
 - Informatique => Filius
- Trouvez à quoi servent tous ces logiciels :

- Fermez toutes ces fenêtres

A faire vous-même 4.



Sur le web, nous allons récupérer le fichier pdf qui correspond à votre support papier et l'enregistrer dans notre espace personnel.

- Cherchez et lancez le logiciel suivant : navigateur web Firefox :



- Dans Firefox, tout en haut à droite, cliquez sur le bouton : , cliquez sur Paramètres



- Descendez, trouvez Téléchargements et choisissez Toujours demander où télécharger le fichier. Fermez l'onglet
- Allez à l'adresse suivante : <http://ninoo.fr/LC>
- Choisissez 2nde_SNT puis seq1_les_donnees_structurees. Vous y trouverez 1_les_donnees_structurees_eleve.pdf .
- Enregistrez ce fichier pdf dans le dossier le plus adéquat parmi ceux que vous avez créés avant.

Ouvrir le lien dans un nouvel onglet
 Ouvrir le lien dans une nouvelle fenêtre
 Ouvrir le lien dans une fenêtre de navigation privée
 Marque-page sur le lien
 Enregistrer la cible du lien sous...
 Enregistrer le lien dans Pocket
 Copier le lien
 Recherche Google pour « coffre_fort.jpg »
 Inspecter

Appelez le professeur



1.4 Et les autres fichiers alors ?



Les autres fichiers ne font que stocker ou enregistrer des données. Il est possible de récupérer ces données, même si l'ordinateur a été éteint.

Voici quelques exemples de fichiers :

- fichiers en .pdf pour des fichiers non modifiables pdf
- fichiers en .odt pour des données LibreOffice Document
- fichiers en .mp3 pour des fichiers audio
- fichiers en .avi pour des fichiers vidéo
- fichiers en .jpg pour des images numériques

- fichiers en `.html` pour des fichiers web

Remarque ! Pour ne pas perturber les gens (ou pour les infantiliser) Ms-Windows ou autres ont décidé de masquer ces extensions par défaut. Mais elles sont bien là, nous les avons fait apparaître tout à l'heure.



Pourtant, quand je double-clique dessus, une fenêtre s'ouvre comme un programme ?

En fait, c'est l'ordinateur qui reconnaît quel est le type du fichier et qui, tout seul, ouvre le programme qui va bien avec le fichier dedans.

On peut s'amuser à ouvrir un fichier avec un autre logiciel.



A faire vous même 5.

☐

- Double-cliquez sur le fichier pdf que vous avez téléchargé. Vérifiez que celui-ci s'ouvre correctement (avec le logiciel Adobe Reader).
- A l'aide de votre navigateur web Firefox, cliquez sur le menu Fichier puis Ouvrir et allez chercher le pdf. Celui-ci devrait s'ouvrir correctement.
- Lancez le programme Notepad ++ et ouvrez le pdf. Celui-ci devrait s'ouvrir même si ce qui s'affiche est illisible.

A faire vous même 6.

☐

- Quelle est la taille de votre fichier pdf ?

A faire vous même 7.

☐

- Avec votre explorateur de fichiers, sélectionnez le fichier pdf
- Faites clic-droit et choisissez Comprimez le fichier
- Un nouveau fichier est créé, quel est son extension ?
- Quel est la taille de ce nouveau fichier ?
- Quelle conclusion en tirez-vous ?

Appelez le professeur



2 Données structurées – Les fichiers csv

☐

A faire vous même 8.

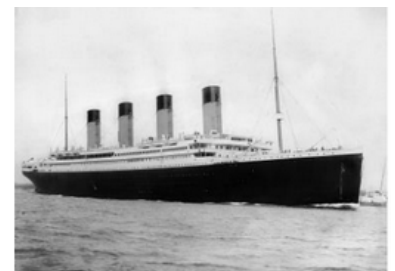
- Ouvrez le fichier http://ninoo.fr/LC/2nde_SNT/seq1_les_donnees_structurees/SPORT_BDD.csv avec un

éditeur de texte type notepad ou Bloc-notes et observer la façon dont sont présentées les données

- Que représente la première ligne ?
- Combien de données a-t-on sur chaque ligne ?
- Comment sont séparées les différentes données sur chaque ligne ?

Le RMS Titanic est un paquebot transatlantique britannique qui fait naufrage dans l'océan Atlantique Nord en 1912 à la suite d'une collision avec un iceberg, lors de son voyage inaugural de Southampton à New York. C'est l'une des plus grandes catastrophes maritimes survenues en temps de paix et la plus grande pour l'époque.

Ce paquebot de la White Star Line, construit à l'initiative de Joseph Bruce Ismay en 1907, a été conçu par les architectes Alexander Montgomery Carlisle et Thomas Andrews des chantiers navals Harland & Wolff. Sa construction débute en 1909 à Belfast et se termine en 1912. C'est à l'époque le plus luxueux et le plus grand paquebot jamais construit. Le Titanic est dirigé par le Capitaine Edward Smith, qui a également sombré avec le navire. Le paquebot portait certains des gens les plus riches de l'époque, de même que des centaines d'émigrants de Grande-Bretagne et d'Irlande et d'ailleurs en Europe qui cherchaient une nouvelle vie aux États-Unis.



Le 14 avril 1912, quatre jours après le commencement de son voyage inaugural, il heurte un iceberg à 23 h 40 (heure locale) et coule le 15 avril 1912 à 2 h 20 au large de Terre-Neuve. Le drame met en évidence l'insuffisance des règles de sécurité de l'époque, notamment le nombre insuffisant de canots de sauvetage et les carences dans les procédures d'évacuation d'urgence.

L'épave du Titanic est localisée le 1er septembre 1985 par le professeur Robert Ballard. Elle gît à 3 843 mètres de profondeur à 650 km au sud-est de Terre-Neuve. L'histoire du paquebot a marqué les mémoires, et suscité la publication de nombreux ouvrages (historiques ou de fiction) et la réalisation de longs métrages dont le film du même nom Titanic de James Cameron, sorti en 1997 et ayant entraîné un important regain d'intérêt pour le paquebot et son histoire.

Source : D'après <https://fr.wikipedia.org/wiki/Titanic>

A faire vous même 9. Exploration du fichier des passagers

Ouvrez le fichier https://ninoo.fr/LC/2nde_SNT/seq1_les_donnees_structurees/titanic.csv à l'aide du tableur de LibreOffice. Ce fichier contient de nombreuses données sur certains passagers du Titanic. Un formulaire de mise en forme est proposé, il est important de paramétrer « l'option de séparateur ». Faire plusieurs essais afin de comprendre le fonctionnement et jusqu'à obtenir un aperçu lisible et organisé. Validez enfin l'ouverture du fichier.

Le séparateur utilisé dans le fichier est le symbole et il sert à
.....

A faire vous même 10.

On retrouve les descripteurs du fichier dans la première ligne du fichier obtenu. Quels sont-ils ? Que permettent-ils ?

A faire vous même 11.

À part la première ligne, chaque autre ligne du fichier correspond à un objet.

Pour chaque objet, on a donc les valeurs pour chaque descripteur.

1. Pour le descripteur « sexe », à quoi correspond la valeur « 1 » ?
2. Pour le descripteur « tarif », à quoi correspond la valeur indiquée pour chaque objet ?
3. Pour le descripteur « survie », à quoi correspondent la valeur « 0 » et la valeur « 1 » ?
4. Choisissez un objet du fichier et rédiger un texte faisant référence aux descripteurs qui montre votre compréhension de l'organisation du fichier de données structurées.

Appelez le professeur



A faire vous même 12. Exploitation basique des données



À l'aide des fonctions usuelles du tableur que vous préciserez ci-dessous, déterminer l'âge moyen des passagers, le tarif moyen payé et le tarif le plus élevé qui a été payé.

A faire vous même 13.



L'outil Données/Autofiltre permet de réaliser des filtres sur les valeurs de chaque descripteur.

En activant cet outil, trier les tarifs par ordre croissant et compléter le tableau suivant :

Tarif (en £)	[0;50[	[50;100[	[100;150[	[150;200[	[200;250[	[250;300[	[300;350[	[350;400[	[400;450[	[450;500[	[500;550[
Effectif											

Recopier ce tableau dans une nouvelle feuille du tableur et représenter graphiquement ces données. Quel(s) commentaire(s) peut-on faire quant à la répartition des tarifs payés sur le Titanic ?

A faire vous même 14. Exploitation plus fine des données



À l'aide des fonctions usuelles du tableur, déterminer la moyenne des données du descripteur « survie ». Quelle interprétation peut-on en donner ?

Appelez le professeur



A faire vous même 15.



Il n'y avait pas suffisamment de places dans les canots de sauvetage du Titanic pour tous les passagers et les membres de l'équipage (et certains canots sont partis à peine remplis). On souhaite examiner l'influence de la classe sociale des passagers sur l'obtention d'une place sur un canot de sauvetage.

Trier les données du descripteur `classe` par ordre croissant à l'aide de l'autofiltre, copier et coller les données pour chaque classe dans des nouvelles feuille du tableur et déterminer la fréquence de survie pour chaque classe. Que peut-on en conclure ?

A faire vous même 16.



Dans le film de James Cameron, lors de l'évacuation du Titanic, on voit que les femmes embarquent davantage sur les canots que les hommes. On peut donc supposer que la fréquence de survie pour les femmes a été supérieure à celle des hommes...

Est-ce la réalité ? Répondre en adoptant une démarche similaire à celle utilisée à la question 15.

A faire vous même 17.



La fréquence de survie chez les femmes a-t-elle été indépendante de la classe dans laquelle voyageaient les passagères ? Pour répondre à cette question, extraire les données comme précédemment en n'oubliant pas de copier-coller la ligne des descripteurs pour pouvoir faire un deuxième tri.

Appelez le professeur



A faire vous même 18.



Afin de découvrir ce qu'est "l'open data", allez sur le site [data.gouv.fr](https://www.data.gouv.fr). Lisez la page suivante : https://www.data.gouv.fr/fr/pages/about/a-propos_data-gouv/

- A quoi servent les données ouvertes ?

A faire vous même 19. Pour les plus rapides



Explorez pendant quelques minutes le site [data.gouv.fr](https://www.data.gouv.fr). Dans Données, recherchez les données "Opérations coordonnées par les CROSS" à l'aide du moteur de recherche proposé par le site. Vous pouvez constater que ces données sont au format CSV