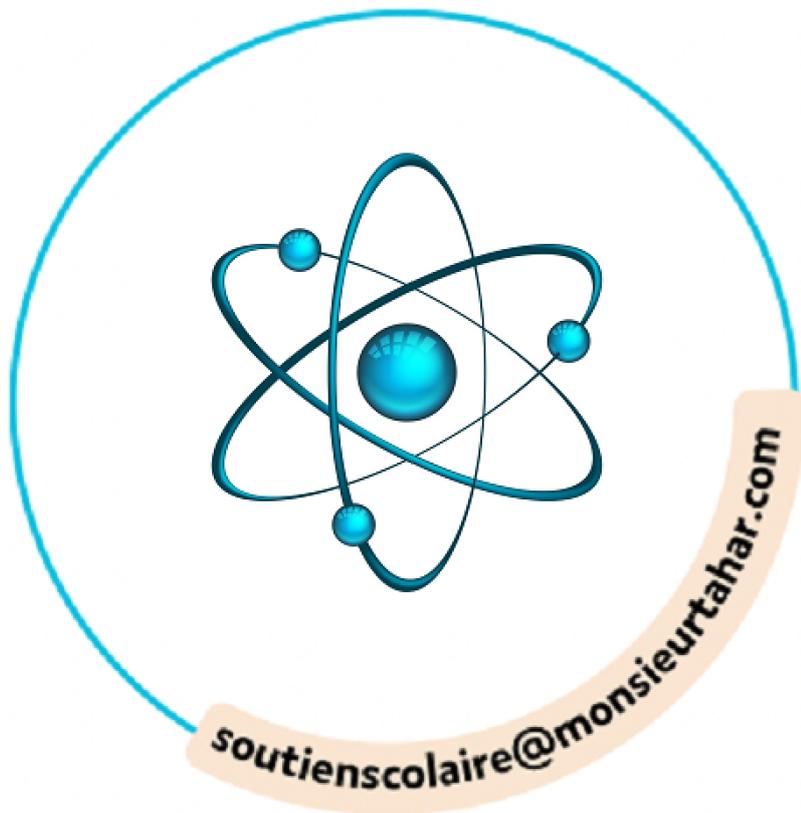


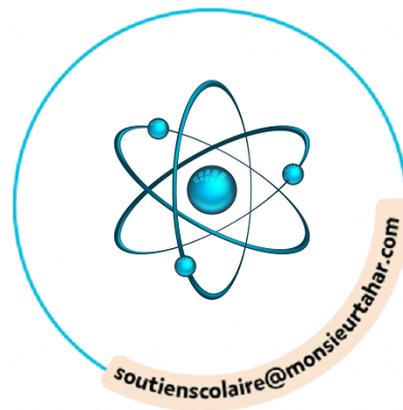
SNT



CHAPITRE 4

Exercices

Autoévaluation



1. CSV

Que signifie le sigle .csv ?

Le sigle .csv signifie *Comma Separated Values*, c'est-à-dire, valeurs séparées par des virgules.

2. Panne de *cloud*

Pour prévenir les conséquences de ce type de panne, qu'est-il judicieux de mettre en place sur son ordinateur personnel ?

Pour éviter les problèmes de connexion à un cloud, ou éventuellement à une panne qui risquerait de détruire une partie des données stockées dans les *data centers*, il est pertinent de mettre en place une synchronisation des fichiers sur son propre appareil.

3. Données publiques

Quel site Internet permet d'accéder aux données rendues publiques par l'État français ?

Le site d'accès aux données rendues publiques par l'État français est le site www.data.gouv.fr.

4. Logiciels

Parmi les logiciels suivants, lesquels sont des tableurs ?

Microsoft Excel et Framacalc sont des tableurs. Il convient de noter que Microsoft Excel est un logiciel propriétaire alors que Framacalc est libre. Inkscape est un logiciel de dessin vectoriel tandis que Google Slides est un logiciel de présentation.

5. Texte à trous

Compléter le texte à trous suivant :

Dans un fichier de données structurées, une ligne correspond à un même objet, une colonne correspond à un même descripteur. L'intersection entre une ligne et une colonne correspond à une donnée.

6. Vrai/Faux

1. L'état civil d'une personne est une donnée personnelle.

Vrai. Ces informations sont d'ailleurs sensibles car elles peuvent être l'objet d'usurpation d'identité.

2. Stocker ses données dans le cloud limite fortement les risques de perte de données.

Vrai. Cependant, il n'évite pas totalement ce risque comme l'illustre l'incendie dans un *data center* de la société d'hébergement Internet OVH les 9 et 10 mars 2021 qui a entraîné la perte des données de certains internautes.

7. Frise chronologique

Classer par ordre chronologique les éléments suivants.

Il existait des bibliothèques dans l'Antiquité. La carte perforée a été inventée au XVIII^e siècle, Microsoft Excel a été inventé à la fin du XX^e siècle et Google Sheets a été développé au XXI^e siècle.

2 : Invention de la carte perforée.

1 : Invention des bibliothèques.

3 : Lancement de Microsoft Excel.

4 : Lancement de Google Sheets.

8. Format de fichier

Pour chaque format de fichier, déterminer s'il est judicieux de s'en servir pour stocker des données.

Un fichier .txt est pertinent pour stocker des données car il occupe peu de place en mémoire, mais le fichier .csv l'est encore plus car il est facilement reconnu par les tableurs. Un fichier .html n'est pas pertinent, il est utilisé pour écrire une page Web. Les fichiers .pptx sont utilisés pour les présentations, ce qui n'a pas de rapport direct avec le stockage des données.

9. Qui suis-je ?

1. Je suis un type de logiciel permettant de stocker, trier, comparer et effectuer des opérations sur des données.

Je suis un tableur. Parmi les tableurs connus on peut citer Microsoft Excel, LibreOffice Calc, Google Sheets, et le premier tableur Visicalc.

2. Je suis une méthode de stockage dématérialisé permettant aux utilisateurs d'accéder à leurs données depuis n'importe quel appareil connecté.

Je suis un stockage dans le *cloud*. On peut notamment citer NextCloud, Onedrive, Google Drive etc..

Application

10. Données et métadonnées

1. Citer deux métadonnées qui pourraient être associées au fichier.

On pourrait utiliser les données présentes pour :

- calculer une moyenne générale, en utilisant éventuellement des coefficients ;
- classer les élèves selon leurs performances.

2. Les données contenues dans ce fichier sont-elles des données personnelles ?

Oui, ce sont des données personnelles car elles concernent des individus.

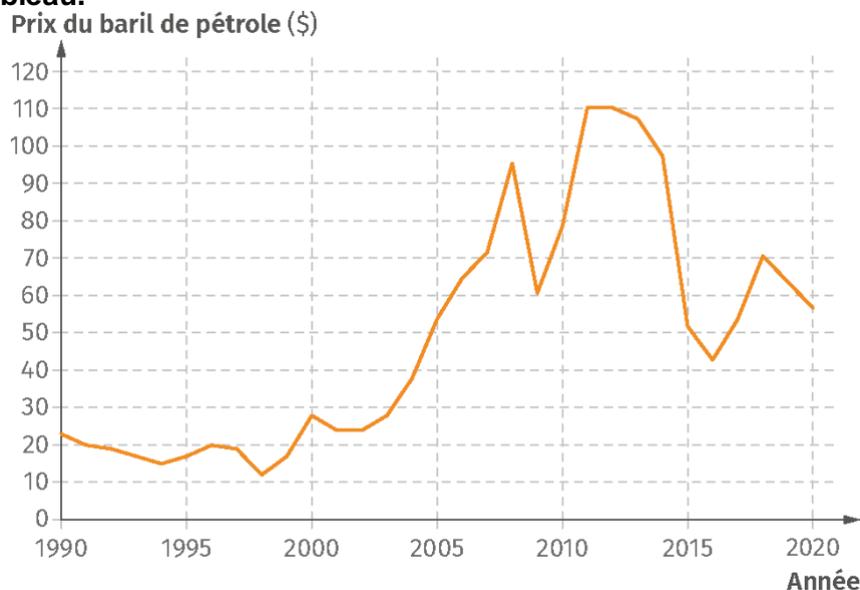
11. Stockage local

◆ Le stockage local a pour avantage :

Le seul avantage parmi les trois choix présentés est qu'il est moins polluant que le *cloud*. En effet, pour maintenir à disposition les données par connexion Internet, les *data centers* qui hébergent les données consomment de grandes quantités d'énergie électrique.

12. Représentation de données

1. Représenter l'évolution des prix en janvier de chaque année sur le graphe suivant ou sur un tableau.



2. Quel est le prix moyen du baril de pétrole depuis 1990 ?

Le prix moyen du baril de pétrole depuis 1990 s'élève à 49,196 \$.

3. Trier les années par ordre croissant de prix du baril de pétrole.

Cette opération, même s'il est possible de la faire à la main, est plus simple à réaliser à l'aide d'un tableur en utilisant ses fonctionnalités concernant le tri des données :

| | |
|------|-------|
| 1998 | 12,76 |
| 1988 | 14,91 |
| 1994 | 15,86 |
| 1993 | 17,1 |
| 1995 | 17,2 |
| 1999 | 17,9 |
| 1989 | 18,23 |
| 1987 | 18,53 |
| 1997 | 19,11 |
| 1992 | 19,32 |
| 1991 | 20,4 |
| 1996 | 20,64 |
| 1990 | 23,76 |
| 2001 | 24,46 |
| 2002 | 24,99 |
| 2000 | 28,66 |

| | |
|------|--------|
| 2003 | 28,85 |
| 2004 | 38,26 |
| 2016 | 43,64 |
| 2015 | 52,32 |
| 2017 | 54,13 |
| 2005 | 54,57 |
| 2009 | 61,74 |
| 2019 | 64,3 |
| 2006 | 65,16 |
| 2018 | 71,34 |
| 2007 | 72,44 |
| 2010 | 79,61 |
| 2008 | 96,94 |
| 2014 | 98,97 |
| 2013 | 108,56 |
| 2011 | 111,26 |
| 2012 | 111,63 |

13. Base de données

1. 🖱️ Avec quel type de logiciels pourrait-on les visualiser ?

Il est possible d'ouvrir une base de données à l'aide d'un tableur-grapheur et d'un traitement de texte. Bien que ce soit possible, il est peu pertinent de le faire à l'aide d'un navigateur Internet qui empêcherait la modification des données.

2. 🖱️ Identifier deux descripteurs de cette table de données.

Les descripteurs de cette table de données sont :

- le numéro ;
- l'âge ;
- le nom ;
- le prénom ;
- la cotisation annuelle payée.

14. Métadonnées

1. 🖱️ Rechercher comment accéder aux métadonnées d'un fichier sur votre *smartphone*.

L'accès aux métadonnées est différent selon les modèles de *smartphone*. Sur un iPhone, il est possible d'y accéder en sélectionnant un fichier et en accédant à ses Infos.

2. L'image ci-contre représente les métadonnées d'un fichier. Donner le nom, la taille et le nombre de pages du fichier.

Il apparaît sur les métadonnées les informations suivantes :

- nombre de pages : 17 ;
- taille : 3,2 Mo ;
- nom : Rapport_Stage.pdf.

3. Le fichier est stocké localement.

Vrai. Il est toutefois possible que ce fichier se trouve dans un dossier synchronisé avec un serveur distant. Dans ce cas, une copie est sauvegardée et peut être récupérée en cas de détérioration de l'appareil.

15. Vrai/Faux

1. Le stockage local consomme plus d'énergie que le stockage dans le *cloud*.

Faux. Le stockage local consomme peu d'énergie électrique, notamment lorsque la sauvegarde (par exemple un disque dur externe) est débranché. Le stockage par le *cloud* nécessite des serveurs toujours allumés.

2. Le RGPD permet de protéger les entreprises qui exploitent des données.

Le RGPD (Règlement Général sur la Protection des Données) a pour objectif de protéger les données personnelles des internautes européens et leur traitement.

16. Opérations sur un tableur

1. Préciser l'opération à écrire dans un tableur pour effectuer la moyenne des valeurs de la colonne A, des lignes 1 à 10.

L'instruction à utiliser est **=MOYENNE(A1:A10)**. Attention, les symboles sont importants. Ici, le symbole *deux points* signifie que ce sont toutes les cellules comprises entre A1 et A10 incluses qui sont utilisées ; si l'on met un autre symbole, comme une virgule ou un point-virgule, le tableur ne fera la moyenne qu'entre les valeurs des cellules A1 et A10, sans prendre en compte les cellules de A2 à A9.

2. Préciser l'opération à écrire dans un tableur pour effectuer la somme des valeurs de la colonne D, des lignes 15 à 30.

L'instruction est **=SOMME(D15:D30)**.

17. Base de données

◆ L'intitulé caractérisant une ou des données stockées est appelé :

L'intitulé caractérisant une ou des données stockées est appelé descripteur. De nombreux descripteurs permettent de classer les données relatives à des objets dans une table structurée.

18. Base de données géographiques

1. 🖱 En cherchant sur Internet, compléter le tableur suivant.

En s'appuyant sur les données collectées par Wikipédia, on a :

| | A | B | C |
|---|-----------|------------------------|------------|
| 1 | Ville | Superficie | Population |
| 2 | Lyon | 47,87 km ² | 518 635 |
| 3 | Rennes | 50,39 km ² | 217 728 |
| 4 | Marseille | 240,62 km ² | 868 277 |
| 5 | Nantes | 65,19 km ² | 314 138 |
| 6 | Paris | 105,40 km ² | 2 175 601 |
| 7 | Toulouse | 118,30 km ² | 486 828 |

2. Préciser la formule à taper dans le tableur pour calculer la superficie moyenne de ces six villes.

L'instruction à noter est **=moyenne(B2:B7)**.

19. Vrai/Faux

1. Un fichier .csv ne peut être ouvert que par un logiciel de type tableur-grapheur.

Faux. On peut tout à fait ouvrir un fichier .csv avec un traitement de texte.

2. Il n'existe qu'un seul moyen de trier les données d'une base de données.

Faux. Les données peuvent être triées selon tous les descripteurs, dans un ordre croissant ou décroissant (alphabétique ou inversé), qu'il s'agisse d'une table de données en chaînes de caractères ou en nombre.

20. Recherche de données

1. 🖱️ Le nombre de personnes inscrites au registre des Français vivant à l'étranger en 2020 qui habitent au Danemark.

Sur le site [data.gouv.fr](https://www.data.gouv.fr/fr/datasets/francais-de-letranger-inscriptions-au-registre-des-francais-etablis-hors-de-france/), on peut avoir accès à ces données depuis la page Web <https://www.data.gouv.fr/fr/datasets/francais-de-letranger-inscriptions-au-registre-des-francais-etablis-hors-de-france/>.

D'après ce document, 5321 français de l'étranger vivaient au Danemark en 2020.

2. 🖱️ Les trois films les plus diffusés à la télévision française depuis 1957.

On peut avoir accès à cette information depuis la page Web <https://www.data.gouv.fr/fr/datasets/classement-des-films-les-plus-diffuses-sur-les-chaines-nationales-gratuites-depuis-1957-1/>.

Les trois films les plus diffusés depuis 1957 sont « *Deux heures moins le quart avant Jésus-Christ* », « *Delphine 1, Yvan 0* » et « *Kirikou et les bêtes sauvages* ».

21. Pandas

1. Que représente la base de données créée dans le script ?

La base de données recense les notes de 5 élèves dans 3 disciplines.

2. Calculer la moyenne de l'élève 4 sachant que toutes les notes affichées ont le même coefficient.

Les notes de l'élève 4 s'élèvent à 12, 18 et 14. La moyenne est donc de 14,7 (arrondi au dixième).

Pratique

22. Consommation électrique d'un *data center*

Correction :

1. 🖱 Relever l'énergie (en kW·h) consommée dans le monde par les *data centers* en 2020.

D'après l'article, cette consommation annuelle s'élève à 650 TW·h à l'échelle mondiale, équivalent donc à $6,5 \times 10^{11}$ kW·h.

2. Sachant qu'une centrale nucléaire produit une puissance de 1 000 MW, soit une énergie de 9 TW·h sur une année, déterminer le nombre de centrales nécessaires pour alimenter en énergie tous les *data centers* du monde pendant un an.

Le nombre de centrales N correspond au rapport entre la consommation mondiale d'énergie électrique et l'énergie moyenne produite par une centrale nucléaire, soit :

$$N = \frac{E_{\text{consommée}}}{E_{\text{centrale}}}$$

AN : $N = \frac{650}{9} = 70$

La valeur est arrondie à un chiffre significatif, en cohérence avec l'énergie produite par une centrale nucléaire au cours d'une année. Il faut donc près de 70 centrales pour alimenter en énergie tous les *data centers* du monde pendant un an. À titre de comparaison, en 2017, 250 centrales nucléaires étaient recensées dans le monde, ce qui représentait environ 10 % de la production d'énergie électrique mondiale.

Parcours pix

Compétence principalement travaillée : Traiter des données

Correction :

Niveau 1. 🖱 Rechercher le nombre de filles nées à Rennes en 2020.

À Rennes, en 2020, 3 588 filles sont nées. Cette information est accessible depuis la page Web <https://www.data.gouv.fr/fr/datasets/naissances-a-rennes/>.

Niveau 2. 🖱 Rechercher le montant de la subvention 2019 de la ville de Rennes accordée à son club d'athlétisme.

Cette subvention s'élève à 34 950 €. Cette information se trouve sur la page Web <https://www.data.gouv.fr/fr/datasets/ca-2019-ville-de-rennes-subventions-aux-associations/>.

Niveau 3. 🖱 Rechercher quels légumes ont été servis dans les cantines des écoles de Rennes le 18 décembre 2019.

Des petits pois ont été servis le 18 décembre 2019, dans les cantines de Rennes. Cette information peut être trouvée sur les fichiers téléchargeables de la page Web <https://www.data.gouv.fr/fr/datasets/menus-des-cantines-des-ecoles-ville-de-rennes/>.