

# Thème D Les données structurées et leur traitement

## Chapitre 8 Traitement des données structurées

Le thème D (*Les données structurées et leur traitement*) correspond aux deux chapitres suivants :

- Chapitre 7 : À la découverte des données structurées
- **Chapitre 8 : Traitement des données structurées**

Le chapitre 8 nécessite d'avoir traité le chapitre 7 au préalable.

### A. Le programme

Les capacités exigibles du BO pour ce chapitre sont données ci-dessous. Les autres contenus du thème *Les données structurées et leur traitement* ont été traités dans le chapitre 7.

Contenus	Capacités attendues du BO traitées dans le chapitre 8	Activités / Exercices
Traitement de données structurées	Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.	Activité 1 p. 120 Activité 2 p. 121  Exercices 1, 3, 4 et 5 p. 126-127
Données dans le nuage ( <i>cloud</i> )	Identifier les principales causes de la consommation énergétique des centres de données ainsi que leur ordre de grandeur.	Activité 3 p. 122-123 Exercice 2 p. 126

## B. QCM diagnostique p. 118

Ces questions vont instaurer le débat, ou la discussion.

Elles sont destinées à faire une évaluation diagnostique en début de chapitre et sont disponibles sur QCMCam et aux formats PDF, PPT et ODP sur le site web : [https://lienbordas.fr/740171\\_ch08\\_qcm](https://lienbordas.fr/740171_ch08_qcm).

1	<p><b>Avec quel logiciel peut-on trier et filtrer des données ?</b></p> <p>Réponses :</p> <ul style="list-style-type: none"><li>A. Un traitement de texte</li><li>B. Un tableur (bonne réponse)</li><li>C. Un navigateur</li></ul>
2	<p><b>Quel outil est généralement utilisé dans un tableur pour afficher uniquement certaines lignes en fonction de critères définis ?</b></p> <p>Réponses :</p> <ul style="list-style-type: none"><li>A. Mise en forme conditionnelle</li><li>B. Filtre (bonne réponse)</li><li>C. Tri alphabétique</li><li>D. Graphique croisé dynamique</li></ul>
3	<p><b>Qu'est-ce qu'une base de données ?</b></p> <p>Réponses :</p> <ul style="list-style-type: none"><li>A. Une base militaire secrète</li><li>B. Une application qui permet d'échanger des données entre internautes</li><li>C. Un service de stockage de données (bonne réponse)</li></ul>
4	<p><b>Quelle est l'une des principales causes de consommation énergétique dans un centre de données ?</b></p> <p>Réponses :</p> <ul style="list-style-type: none"><li>A. La consommation des écrans des utilisateurs</li><li>B. L'éclairage des locaux</li><li>C. Le refroidissement des serveurs (bonne réponse)</li><li>D. Le transport des données par câble</li></ul>
5	<p><b>Selon vous, quelle part de l'impact environnemental du numérique en France est liée à la consommation des centres de données ?</b></p> <p>Réponses :</p> <ul style="list-style-type: none"><li>A. 6 %</li><li>B. 26 %</li><li>C. 46 % (bonne réponse)</li></ul>

## C. Frise historique p. 119

### Réponses aux questions :

1. Le logiciel VisiCalc a facilité la percée des ordinateurs en entreprise en automatisant les calculs financiers sur tableur, ce qui les a rendus indispensables pour le traitement de grandes quantités de données.
2. Le langage SQL est devenu incontournable car il permet de manipuler, interroger et gérer efficacement les bases de données relationnelles, qui sont au cœur des systèmes d'information modernes.
3. Il est crucial de fixer des normes pour le développement du Web afin de garantir l'interopérabilité, la sécurité, et une expérience utilisateur cohérente sur différents appareils et navigateurs web. Cela permet également de promouvoir l'accessibilité, en veillant à ce que tous les utilisateurs, quels que soient leurs handicaps, puissent accéder aux contenus web. Enfin, les normes facilitent la maintenance et l'évolution des sites web, en assurant une compatibilité avec les technologies futures.

## D. Description des activités

### Activité 1 p. 120 Comment traiter des données structurées ?

#### Capacité travaillée :

- Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.

Cette activité propose de manipuler des ensembles de données structurées, présentées ici sous forme de tables, pour extraire, organiser et analyser l'information de manière efficace. Ces opérations permettent de :

- rechercher : déterminer des données spécifiques en fonction de critères prédéfinis ;
- filtrer : sélectionner des sous-ensembles de données qui répondent à certaines conditions (ou critères) ;
- trier : organiser les données selon un ordre spécifique (croissant, décroissant, alphabétique, etc.) ;
- calculer : effectuer des opérations mathématiques (somme, moyenne, etc.) sur des données numériques.

### Réponses aux questions :

1. Il est possible que certains défibrillateurs changent de place au fil du temps. Il est donc important d'impliquer la communauté des utilisateurs de l'application Staying Alive pour qu'ils vérifient et actualisent les emplacements des appareils.

2. Si la géolocalisation du smartphone est activée, l'application peut connaître la position du sauveteur et afficher à l'écran les défibrillateurs opérationnels les plus proches, idéalement dans un rayon de 500 mètres.

3. Le point bleu représente la position du sauveteur qui vient de lancer l'application Staying Alive, tandis que les points verts représentent les emplacements des défibrillateurs qui lui sont proches. N'apparaissent à l'écran que les défibrillateurs facilement accessibles, c'est-à-dire à proximité immédiate du sauveteur.

4. Il faut télécharger le fichier DAE . csv, qui est un fichier au format CSV d'après son extension de fichier.

5. Les données enregistrées dans le fichier sont au format CSV, ce qui signifie qu'elles sont séparées les unes des autres par des virgules. Afin que le logiciel tableur les interprète correctement, il est en effet important de préciser quel caractère séparateur a été utilisé (on aurait pu également utiliser des points-virgules ou des tabulations pour séparer les données, comme nous l'avons vu dans l'activité 3 p. 110).

**Options de séparateur**

Largeur fixe                       Séparé par  
 Tabulation     Virgule     Point-virgule     Espace     Autre

6. En examinant le tableau des données, on s'aperçoit qu'il y a 6 descripteurs : Identifiant, Code postal, Ville, longitude, latitude, pays. La première étape consiste à supprimer le descripteur inutile « pays », afin de ne garder que ceux demandés dans l'activité. Pour cela, il faut sélectionner toute la colonne à supprimer, effectuer un clic droit, choisir « supprimer des colonnes » et appuyer sur Entrée.

	A	B	C	D	E	F
1	Identifiant	Code postal	Ville	longitude	latitude	pays
2	100244	66100	Perpignan	2.903669	42.678202	FR
3	6079	66000	Perpignan	2.880616	42.669181	FR
4	58738	66000	Perpignan	2.886369	42.716957	FR
5	31530	66000	Perpignan	2.877569	42.668465	FR
6	100766	66000	Perpignan	2.906273	42.687328	FR
7	128353	66000	Perpignan	2.896104	42.712912	FR
8	10368	66000	Perpignan	2.890005	42.701104	FR
9	122577	66000	Perpignan	2.883159	42.69266	FR
10	29988	66000	Perpignan	2.947734	42.702564	FR
11	161082	66100	Perpignan	2.89003227	42.67869684	FR
12	100199	66000	Perpignan	2.845151	42.683832	FR
13	15789	66000	Perpignan	2.868494	42.680079	FR
14	81987	66000	Perpignan	2.853687	42.690049	FR
15	26777	66000	Perpignan	2.868356	42.73056	FR
16	167851	66000	Perpignan	2.89570883	42.73587054	FR
17	34947	66012	Perpignan	2.864805	42.695129	FR
18	9875	66000	Perpignan	2.889777	42.702061	FR
19	47442	66100	Perpignan	2.916321	42.669007	FR
20	104859	66000	Perpignan	2.91658	42.668958	FR
21	100253	66100	Perpignan	2.903455	42.678333	FR
22	9359	66000	Perpignan	2.89355	42.700059	FR
23	135049	66000	Perpignan	2.93186601	42.69902474	FR

7. Pour déterminer le nombre de défibrillateurs du tableau, il faut compter le nombre total de lignes en excluant la première, qui contient les descripteurs. Il y a 92 lignes, soit 92 DAE dans le fichier téléchargé.

**Remarque** : afin de faciliter le traitement de ces données, seul un petit nombre de défibrillateurs a été utilisé ici. Fin 2024, il existe 340 000 DAE enregistrés dans l'application en France métropolitaine.

8. Il s'agit ici de **trier** les villes par code postal croissant, c'est-à-dire du plus petit au plus grand. On effectue l'action « Données/Trier », puis on choisit la colonne « Code postal » comme critère de tri :

Identifiant	Code postal	Ville	longitude	latitude
6079	66000	Perpignan	2.880616	42.669181
58738	66000	Perpignan	2.886369	42.716957
31530	66000	Perpignan	2.877569	42.668465
100766	66000	Perpignan	2.906273	42.687328
128353	66000	Perpignan	2.896104	42.712912
10368	66000	Perpignan	2.890005	42.701104
122577	66000	Perpignan	2.883159	42.69266
29988	66000	Perpignan	2.947734	42.702564
100199	66000	Perpignan	2.845151	42.683832

**Remarque** : ne pas oublier d'« étendre la sélection » afin de trier l'ensemble des données.

9. On demande à présent d'afficher uniquement les défibrillateurs de la ville de Perpignan. Pour cela, il est indispensable de **filtrer** les données à l'aide du mot-clé « Perpignan ».

On effectue l'action « Données/Plus de filtres /Filtre Standard », puis on impose la condition : **Ville = Perpignan** comme critère de sélection :

The image shows a 'Filtre standard' dialog box. Under the 'Critères de filtre' section, there is a table with four columns: 'Opérateur', 'Nom de champ', 'Condition', and 'Valeur'. The 'Opérateur' column contains '=', the 'Nom de champ' column contains 'Ville', the 'Condition' column contains '=', and the 'Valeur' column contains 'Perpignan'. There is a red 'X' button to the right of the 'Valeur' field.

En effectuant cette action de sélection, toutes les lignes de données qui ne correspondaient pas à la ville « Perpignan » ont été automatiquement supprimées. Attention, car la numérotation des lignes est restée la même que dans le fichier initial : il se peut donc que la ligne numéro 80 ou 85 ne soit plus présente, ce qui est parfaitement normal !

10. On demande à présent d'appliquer la formule mathématique suivante à une case vide du tableau :  
**=NB.SI(C1:C93;"=Perpignan")**

Cette formule permet de compter automatiquement le nombre de DAE de la ville de Perpignan : on en dénombre 69.

Attention à bien respecter la syntaxe de la formule, notamment les « : », « ; », « . » et autres guillemets.

11. Grâce aux nombreuses fonctionnalités d'un logiciel tableur, il est possible de traiter un grand volume de données rapidement et sans risque d'erreurs.

## Activité 2 p. 121 Comment faire des requêtes dans une base de données ?

### Capacité travaillée :

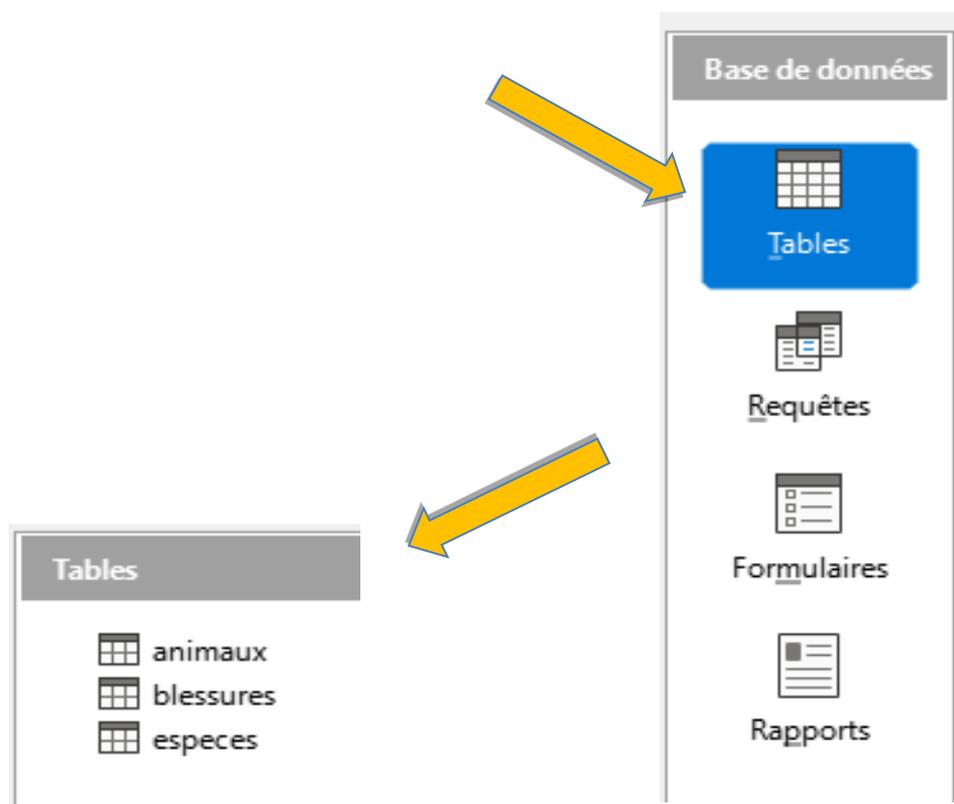
- Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.

Cette activité propose une initiation au langage SQL, qui est le langage le plus couramment utilisé pour gérer une base de données relationnelles. Il peut être intéressant de la compléter avec l'exercice 5 page 127.

Le langage SQL est au programme de terminale spécialité NSI. Précisons que ce n'est pas un langage de programmation.

### Réponses aux questions :

1. Dans une base de données relationnelles, les informations sont réparties entre différentes tables qui ont chacune un rôle spécifique. Dans cette activité, la base de données est composée de trois tables (que l'on appelle aussi des relations), nommées **animaux**, **blessures** et **especes**. Dans un tableau, les données sont organisées uniquement par les noms de colonnes.
2. Il faut télécharger le fichier **base\_zoo.odb**. Le format ODB correspond au format libre de la suite OpenOffice que nous allons utiliser ici avec le logiciel Base.
3. En cliquant sur **Tables**, on fait apparaître les trois tables de la base de données :



4. À l'aide d'un double-clic, on ouvre chacune des tables de manière à faire apparaître leurs descripteurs :

Table	Descripteurs
animaux	nom, espece, date_de_naissance, sexe
blessures	ID, nom, date, blessure
especes	espece, regime_alimentaire, aquatique_esperance_vie

5. a. C'est la table **animaux** qui contient les noms de tous les animaux du zoo.

b. C'est la table **especes** qui contient les espérances de vie des animaux, classés par race.

c. C'est la table **blessures** qui contient les dates des blessures des différents animaux.

6. En parcourant **visuellement** les tables, il est possible d'extraire l'information demandée dans cette question car chacune de ces tables a une taille limitée. En revanche, si elles contenaient plusieurs milliers d'animaux, il serait très fastidieux de déterminer combien le zoo en possède.

En parcourant visuellement le contenu de la table **animaux**, on constate que le zoo possède 20 animaux. La table contient en effet 20 enregistrements (une ligne par animal).

7. Pour répondre à la question A qui demande le nom des loups, il faut consulter la table **animaux** : les loups s'appellent Inca et Salma. Il s'agit ici, dans un premier temps, d'amener les élèves à parcourir les données de la table en recherchant les informations demandées. Dans un deuxième temps, ces informations seront extraites de la table grâce à une requête SQL.

8. Pour obtenir le nombre des loups en utilisant le langage SQL, il faut écrire la requête suivante :

```
SELECT nom
FROM animaux
WHERE espece = 'Loup';
```

Par convention, les mots-clés du langage SQL (**SELECT**, **FROM**, **WHERE**, etc.) sont écrits en majuscules, tandis que les noms des tables (**animaux**, **especes**, etc.) et des attributs (**nom**, **ID**, etc.) sont écrits en minuscules.

La requête se termine par un « ; ». Mais en toute rigueur, le langage SQL n'est pas sensible à la casse : on peut donc indifféremment tout écrire en majuscules ou en minuscules.

9. Deux valeurs ont été extraites de la table `animaux`, qui correspondent aux noms des deux loups : Inca et Salma. La requête SQL renvoie alors le résultat suivant :

nom
Inca
Salma

**Remarque** : si la requête renvoie une erreur ou ne renvoie rien, cela signifie qu'elle a été mal écrite. Les élèves oublient souvent la majuscule à 'Loup', qui est indispensable ici puisque 'Loup' est une chaîne de caractères.

10. Pour répondre à la question B, il faut avoir accès aux noms des animaux mais également à leur régime alimentaire. Il faut donc consulter deux tables : la table `animaux` et la table `especes`.

Comme ces informations sont contenues dans deux tables différentes, il faut effectuer une jointure entre les deux à l'aide du mot-clé `JOIN` :

```
SELECT nom, regime_alimentaire
FROM animaux JOIN especes
ON animaux.espece = especes.espece
WHERE nom = 'Boubou';
```

Le mot-clé `ON` associé à `JOIN` permet d'introduire la condition de jointure qui sert à relier les deux tables entre elles. Cette requête illustre parfaitement le terme « **relationnelle** » dans l'expression « base de données relationnelle ».

Le résultat de la requête est le suivant :

nom	regime_alimentaire
Boubou	Herbivore

Nous concluons que le régime alimentaire de Boubou est bien herbivore, ce qui permet de répondre à la question B.

**Remarques** : attention à ne pas confondre `especes` et `espece`. `especes` est le nom d'une table, alors que `espece` est un attribut. Remarquons également que `espece` est un attribut de la table `animaux` mais aussi de la table `especes` : il est donc essentiel de toujours préciser le nom de la table (`animaux` ou `especes`) lorsqu'on fait référence à un attribut, afin d'éviter toute ambiguïté dans l'exécution de la requête SQL. C'est pourquoi on écrit `animaux.espece` ou `especes.espece`. Ce souci d'exigence est pédagogiquement très intéressant, car il illustre l'importance des mots-clés et de la syntaxe dans les langages informatiques.

11. Pour répondre à la question C, il faut consulter la table `blessures` uniquement.

La requête qui permet d'y répondre est la suivante :

```
SELECT nom
FROM blessures
WHERE date_blessure = '12/05/17';
```

12. Si Robin avait géré le zoo avec une seule table, il aurait rencontré :

- des problèmes de redondance puisque certaines informations auraient été répétées plusieurs fois (espèces, noms des animaux, etc.) ;
- des difficultés de mise à jour : la modification d'une information aurait nécessité de changer plusieurs lignes ;
- des problèmes d'incohérence : des erreurs ou des contradictions entre les données auraient pu apparaître ;
- des problèmes de suppression : supprimer un animal aurait pu effacer des informations utiles au reste de la base.

### Activité 3 p. 122-123 Peut-on continuer à consommer du numérique sans modération ?

#### Capacité travaillée :

- Identifier les principales causes de la consommation énergétique des centres de données.

#### Capacité transversale travaillée :

- Développer une argumentation dans le cadre d'un débat et travailler l'oral.

Cette activité sociétale invite à réfléchir à l'impact environnemental de nos pratiques numériques et à encourager des usages moins énergivores.

### Réponses aux questions du parcours 1 :

1. Parmi les différents modes de stockage des données en local, on peut citer la mémoire de l'ordinateur ou du smartphone, les clés USB, ou encore les disques durs externes. Parmi les différents modes de stockage des données à distance, on peut citer les services de cloud, hébergés dans des *data centers* (cours du chapitre et doc. B). Tous les stockages de données en local sont moins énergivores que ceux à distance.

2. Voici quelques exemples parmi d'autres de gestes simples à adopter pour réduire sa consommation d'énergie grâce à un usage plus durable du numérique (cf. doc. A) :

- trier ses données numériques : supprimer systématiquement tous les fichiers inutiles ou non utilisés ;
- réduire la taille des fichiers lorsque c'est possible ;
- adopter un bon usage de l'e-mail en évitant les « répondre à tous » si ce n'est pas nécessaire ;
- utiliser des plateformes de transfert plus responsables ;
- utiliser des moteurs de recherche écoresponsables ;

- réduire l'utilisation de la caméra en visioconférence ;
- prendre soin des appareils électroniques, les faire réparer ou opter pour de la seconde main ;
- penser à la déconnexion.

3. En 2022, les data centers représentaient 46 % de la consommation énergétique du numérique en France. Le stockage des données numériques nécessite en effet de plus en plus de data centers qui sont très gourmands en électricité. Par ailleurs, le visionnage de vidéos en ligne et l'échange de photos requièrent des transferts de gros volumes de données qui sont eux aussi énergivores (doc. B et D).

4. L'empreinte carbone est un indicateur environnemental qui permet de mesurer les émissions de gaz à effet de serre d'une activité et donc son impact sur l'environnement. L'outil [impactco2.fr](https://impactco2.fr) nous permet de calculer les émissions de CO<sub>2</sub> émises par nos habitudes de consommation : dans l'exemple donné, 7 heures par semaine de visionnage en haute définition en streaming sur une télévision connectée en Wi-Fi équivaut à plus d'un aller-retour en TGV entre Paris et Berlin par an (doc. C).

## **Éléments de réponse aux questions du parcours 2 :**

Sont attendus les éléments suivants au cours du débat :

- lister les gestes utiles pour réduire sa consommation énergétique (doc. A) ;
- réfléchir à l'impact environnemental croissant du numérique et notamment de l'intelligence artificielle ;
- utiliser l'outil de calcul [impactco2.fr](https://impactco2.fr) pour avoir des exemples et des ordres de grandeur quant aux émissions de CO<sub>2</sub> des pratiques numériques quotidiennes des lycéens (doc. C) ;
- s'appuyer sur les statistiques proposées par les études de l'Ademe sur l'impact environnemental du numérique :  
<https://ecoresponsable.numerique.gouv.fr/actualites/actualisation-ademe-impact/>
- prolongement possible : se rapprocher des éco-délégués de l'établissement pour proposer de participer à un challenge autour du numérique durable. Exemple : élire la classe qui aura supprimé le plus grand nombre d'e-mails inutiles en une semaine.

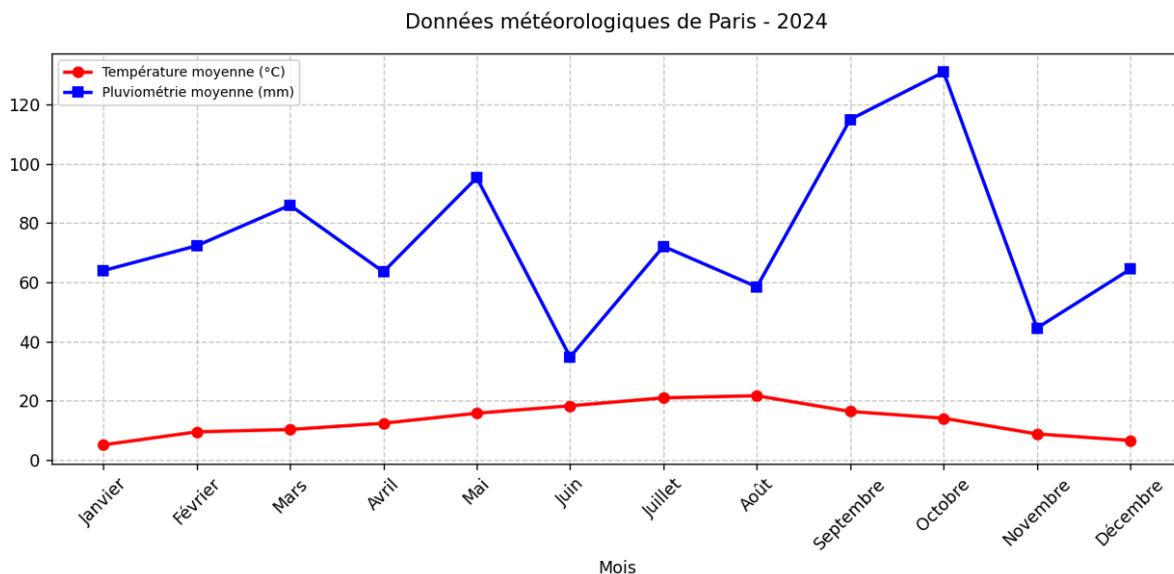
## E. Description des exercices

### Exercice 1 p. 126 Écologie

#### Capacité transversale travaillée :

- Écrire et développer des programmes pour répondre à des problèmes et modéliser des phénomènes physiques, économiques et sociaux.

1. On commence par télécharger le fichier `meteo.json`.
2. Après l'ouverture du fichier avec le Bloc-notes, on observe que les données sont structurées, à l'aide des caractères « { », « [ », « : » ou « , ». Il s'agit donc de données au format JSON, ce qui est conforme à l'extension JSON du fichier téléchargé.
3. Ces données structurées sont adaptées pour être traitées par un ordinateur. En revanche, l'être humain peut rencontrer quelques difficultés de lecture, notamment pour suivre l'évolution des valeurs numériques.
4. Il semble pertinent de représenter les données sous forme de graphiques comme un histogramme ou une courbe.
5. En allant à l'adresse proposée, on charge automatiquement le fichier Python ainsi que le fichier `meteo.json`, déjà téléchargé à la question 1, grâce au bouton .
6. Si on clique sur le bouton , le programme s'exécute et les courbes de température et de pluviométrie apparaissent à l'écran :



7. L'instruction Python en ligne 10 affiche les données météo pour l'année 2024 :

```
data_annee = data['donnees_meteorologiques'][-1]
```

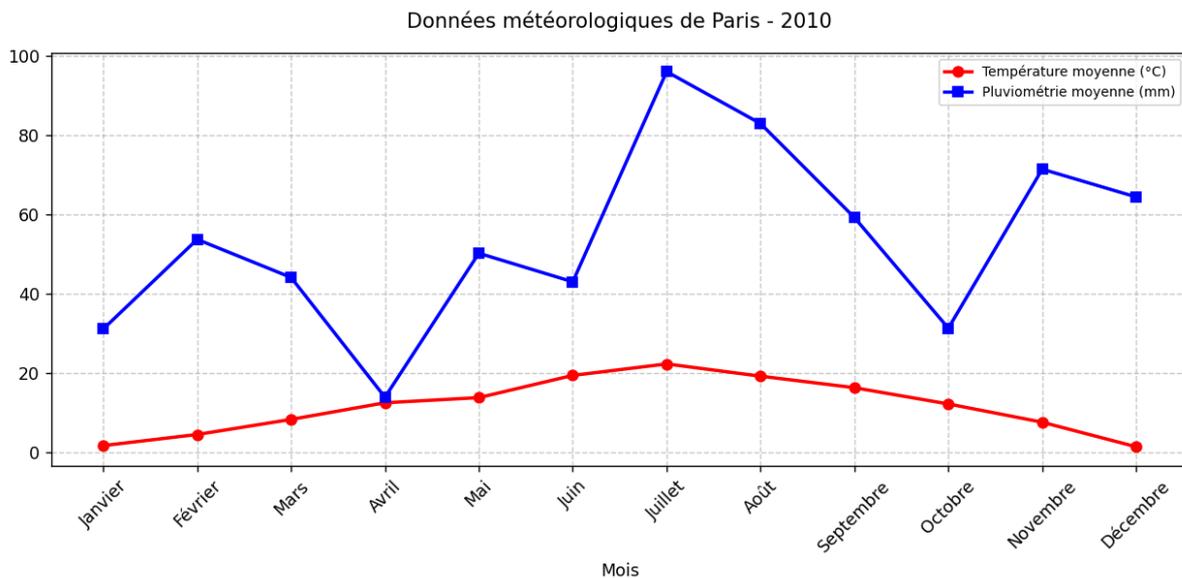
L'indice `-1` est ici utilisé pour extraire les données du dernier élément d'une liste, situé à l'indice `len(liste)-1` ou plutôt `-1` par écriture simplifiée.

Pour afficher les données de l'année 2010, il faut trouver l'indice de liste correspondant et modifier l'instruction Python avant de réinterpréter le programme.

En examinant la structure du fichier `meteo.json`, on s'aperçoit que l'année 2010 correspond au premier élément du fichier, qui sera donc à l'indice 0. L'instruction Python devient donc :

```
data_annee = data['donnees_meteorologiques'][0]
```

ce qui affiche :



8. Il est à présent plus aisé d'interpréter les courbes obtenues, car elles sont plus lisibles pour un humain. Le mois le moins pluvieux de l'année 2010 est repéré sur la courbe bleue, il s'agit du mois d'avril.

9. Collecter un historique des données météorologiques permet d'analyser l'évolution du climat sur le long terme afin de détecter, comprendre et anticiper les effets du changement climatique.

La correction dans l'éditeur WebPython est disponible à cette adresse :

[https://lienbordas.fr/740171\\_webpython16\\_OEX](https://lienbordas.fr/740171_webpython16_OEX).

## Exercice 2 p. 126 Impact énergétique du *cloud computing*

### Capacité travaillée :

- Identifier les principales causes de la consommation énergétique des centres de données ainsi que leur ordre de grandeur.

1. Il faut d'abord charger la page web du calculateur, qui va simuler la consommation d'énergie d'un centre de données (*data center*).

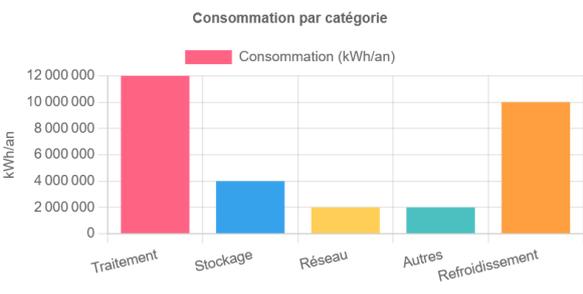
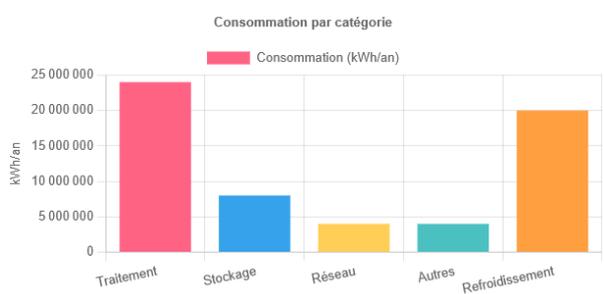
2. a. À l'aide de l'application web, il est possible de déterminer les trois causes principales de la consommation d'énergie du centre de données, en consultant le camembert de consommation :

1. traitement des données ;
2. refroidissement du data center ;
3. stockage des données.

b. Le système de refroidissement basique consomme 20,00 GWh/an, alors que le système très performant consomme 4,00 GWh/an, donc 5 fois moins.

3. Si le nombre de serveurs double (de 10 000 à 20 000) :

- a. la consommation des serveurs double ;
- b. la consommation du refroidissement double ;
- c. l'équivalent en nombre de foyers double également.

<b>Avec 10 000 serveurs :</b>	<b>Avec 20 000 serveurs :</b>
 <p>Consommation par catégorie</p> <p>Consommation (kWh/an)</p>	 <p>Consommation par catégorie</p> <p>Consommation (kWh/an)</p>
<p><b>Résultats :</b></p> <p>Consommation des serveurs seuls : 20,00 GWh/an</p> <p>Consommation du refroidissement : 10,00 GWh/an</p> <p>Consommation totale : 30,00 GWh/an</p> <p>Équivalent en maisons* : 6 000 maisons</p> <p>* Une maison moyenne consomme environ 5 000 kWh/an.</p>	<p><b>Résultats :</b></p> <p>Consommation des serveurs seuls : 40,00 GWh/an</p> <p>Consommation du refroidissement : 20,00 GWh/an</p> <p>Consommation totale : 60,00 GWh/an</p> <p>Équivalent en maisons* : 12 000 maisons</p> <p>* Une maison moyenne consomme environ 5 000 kWh/an.</p>

4. Pour faire le bon choix, il faut effectuer les deux simulations avec l'application. On obtient les résultats suivants :

<b>OPTION 1</b>	<b>OPTION 2</b>
10 000 serveurs avec un système de refroidissement basique	8 000 serveurs avec un système de refroidissement très performant
Consommation totale : 40,00 GWh/an	Consommation totale : 19,20 GWh/an

Bien que l'option 2 ait 2 000 serveurs de moins, sa consommation est plus de deux fois inférieure à celle de l'option 1. Cela montre que l'efficacité du refroidissement joue un rôle majeur dans la consommation totale d'un centre de données.

L'option 2 possède moins de serveurs, mais cela ne signifie pas automatiquement une baisse de performance :

- les serveurs peuvent être plus puissants ou plus récents ;
- la réduction du nombre de serveurs peut être compensée par une meilleure virtualisation ou une architecture optimisée.

### Exercice 3 p. 126 Exploiter les données d'une table

#### Capacité travaillée :

- Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.

1. Le produit ayant la meilleure note client est Ordinateur Z et son prix unitaire est de 1 299 €.

2. Les produits dont le stock est supérieur ou égal à 20 sont :

- Écouteurs Pro ;
- Casque Audio ;
- Tablette T10.

3. Si un client achète un produit de chaque type, le coût total sera la somme de la 3<sup>e</sup> colonne, soit :

$$899 + 199 + 1\,299 + 349 + 1\,499 + 499 = 4\,744 \text{ €}.$$

4. Si l'on s'intéresse aux produits ayant un stock inférieur ou égal à 10, celui qui a la meilleure note client est Ordinateur Z.

### Exercice 4 p. 127 Cybersécurité

#### Capacité travaillée :

- Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.

1. On commence par télécharger le fichier `connexions.csv`.

2. Il est possible d'ouvrir le fichier avec le logiciel Bloc-notes mais il contient un nombre important de données, ce qui rend leur traitement fastidieux avec un tel outil. Nous allons donc utiliser une application web destinée à traiter ce genre de données. Les experts en cybersécurité ont l'habitude de développer leurs propres outils orientés web ou programmation.

3. et 4. On charge le fichier `connexions.csv` à l'aide de l'application web. Nous pouvons voir que 1 000 connexions ont été enregistrées dans ce fichier journal de connexion.

Nombre d'enregistrements: 1000		
Adresse IP ▲ ▼	Heure de la connexion ▲ ▼	Statut de la connexion ▲ ▼
170.87.254.69	2025-01-16 16:23:58	Échouée
170.87.254.69	2025-01-16 16:29:25	Échouée
170.87.254.69	2025-01-16 16:28:55	Échouée
170.87.254.69	2025-01-16 16:25:31	Échouée

5. Le menu de l'application propose notamment de filtrer les données selon certains critères.

Voici les réponses attendues :

<p>Nombre d'échecs de connexion</p> <p><b>Filtres</b></p> <p><input type="radio"/> Tous <input type="radio"/> Réussites <input checked="" type="radio"/> Échecs</p> <p><input type="button" value="Réinitialiser"/></p>	<p>En cochant le mot « Échecs », le filtrage des données de connexion est automatique.</p> <p>Nombre d'enregistrements : 544</p> <p>Il y a donc eu 544 échecs de connexion et <math>1\ 000 - 544 = 456</math> réussites.</p>																								
<p>Nombre de connexions réussies depuis l'adresse 170.226.115.242</p>	<p>Pour effectuer le filtrage demandé, il faut cocher « Réussites » et fournir l'adresse IP que nous recherchons. On obtient un seul résultat :</p> <table border="1" data-bbox="587 1160 1455 1330"> <thead> <tr> <th colspan="3">Nombre d'enregistrements: 1</th> </tr> <tr> <th>Adresse IP ▲ ▼</th> <th>Heure de la connexion ▲ ▼</th> <th>Statut de la connexion ▲ ▼</th> </tr> </thead> <tbody> <tr> <td>170.226.115.242</td> <td>2024-12-21 22:56:23</td> <td>Réussie</td> </tr> </tbody> </table>	Nombre d'enregistrements: 1			Adresse IP ▲ ▼	Heure de la connexion ▲ ▼	Statut de la connexion ▲ ▼	170.226.115.242	2024-12-21 22:56:23	Réussie															
Nombre d'enregistrements: 1																									
Adresse IP ▲ ▼	Heure de la connexion ▲ ▼	Statut de la connexion ▲ ▼																							
170.226.115.242	2024-12-21 22:56:23	Réussie																							
<p>Nombre de connexions réussies entre le 01/01/2025 et le 15/01/2025 inclus</p>	<p>Pour effectuer le filtrage demandé, il faut cocher « Réussites » et renseigner les deux dates de début et de fin de filtrage. On obtient 220 enregistrements pour cette période.</p> <table border="1" data-bbox="587 1536 1417 2002"> <thead> <tr> <th colspan="3">Filtre par plage de dates</th> </tr> <tr> <td><input type="text" value="01 / 01 / 2025"/></td> <td><input type="text" value="15 / 01 / 2025"/></td> <td></td> </tr> <tr> <th colspan="3">Nombre d'enregistrements: 220</th> </tr> <tr> <th>Adresse IP ▲ ▼</th> <th>Heure de la connexion ▲ ▼</th> <th>Statut de la connexion ▲ ▼</th> </tr> </thead> <tbody> <tr> <td>47.98.69.205</td> <td>2025-01-05 19:23:07</td> <td>Réussie</td> </tr> <tr> <td>42.164.151.64</td> <td>2025-01-03 20:13:15</td> <td>Réussie</td> </tr> <tr> <td>6.81.48.243</td> <td>2025-01-15 13:52:09</td> <td>Réussie</td> </tr> <tr> <td>7.30.62.215</td> <td>2025-01-10 14:52:34</td> <td>Réussie</td> </tr> </tbody> </table>	Filtre par plage de dates			<input type="text" value="01 / 01 / 2025"/>	<input type="text" value="15 / 01 / 2025"/>		Nombre d'enregistrements: 220			Adresse IP ▲ ▼	Heure de la connexion ▲ ▼	Statut de la connexion ▲ ▼	47.98.69.205	2025-01-05 19:23:07	Réussie	42.164.151.64	2025-01-03 20:13:15	Réussie	6.81.48.243	2025-01-15 13:52:09	Réussie	7.30.62.215	2025-01-10 14:52:34	Réussie
Filtre par plage de dates																									
<input type="text" value="01 / 01 / 2025"/>	<input type="text" value="15 / 01 / 2025"/>																								
Nombre d'enregistrements: 220																									
Adresse IP ▲ ▼	Heure de la connexion ▲ ▼	Statut de la connexion ▲ ▼																							
47.98.69.205	2025-01-05 19:23:07	Réussie																							
42.164.151.64	2025-01-03 20:13:15	Réussie																							
6.81.48.243	2025-01-15 13:52:09	Réussie																							
7.30.62.215	2025-01-10 14:52:34	Réussie																							

6. Il est proposé ici d'utiliser le module de détection d'attaque par force brute, spécialement développé dans l'application web. Ce module va traiter les données numériques et rechercher quelles adresses IP présentent au moins 50 échecs de connexion sur une période de 10 minutes.

D'abord, il faut réinitialiser les données afin d'effacer tout filtrage antérieur des données et cocher l'option « Échecs ».

Nombre d'enregistrements: 50

Adresse IP ▲▼	Heure de la connexion ▲▼	Statut de la connexion ▲▼
170.87.254.69	2025-01-16 16:23:58	Échouée
170.87.254.69	2025-01-16 16:29:25	Échouée
170.87.254.69	2025-01-16 16:28:55	Échouée
170.87.254.69	2025-01-16 16:25:31	Échouée

Il apparaît que l'adresse IP 170.87.254.69 a effectué plus de 50 tentatives de connexion au serveur ayant mené à un échec. Cette adresse IP est donc suspecte.

## Exercice 5 p. 127 Question de goût

### Capacité travaillée :

- Réaliser des opérations de recherche, filtre, tri ou calcul sur une ou plusieurs tables.

Cet exercice propose d'effectuer quelques requêtes SQL en utilisant la base de données d'une pizzeria.

1. On télécharge d'abord le fichier `pizzeria.odt`, puis on l'ouvre à l'aide du logiciel Base.
2. Cette base de données possède 5 tables nommées `client`, `commande`, `compose`, `ingredient` et `pizza`.
3. En ouvrant la table `pizza` à l'aide d'un double-clic, on note que 5 pizzas sont proposées à la carte de la pizzeria puisque cette table possède 5 lignes (également appelées enregistrements).

4. La requête suivante affiche le nom et le prix des pizzas, par prix croissant, c'est-à-dire de la moins chère à la plus chère :

```
SELECT nom, prix
FROM pizza
ORDER BY prix;
```

Le résultat de la requête est le suivant :

nom	prix
Margherita	9,00
Pepperoni	10,00
Hawaienne	11,00
Végétarienne	12,00
Quatre Saisons	12,00

5. Cette requête affiche le nom et le prix des pizzas pour lesquelles le prix est strictement inférieur à 12 € :

```
SELECT nom, prix
FROM pizza
WHERE prix < 12;
```

Le résultat de la requête est le suivant :

nom	prix
Pepperoni	10,00
Hawaienne	11,00
Margherita	9,00

6. On souhaite à présent savoir quels ingrédients composent une pizza donnée. Pour répondre à cette question, il faut consulter simultanément les tables `pizza`, `ingredient` et `compose`.

En effet, la table `pizza` contient le nom des pizzas, la table `ingredient` contient le nom des ingrédients et la table `compose` est une table de liaison qui indique quelle pizza contient quels ingrédients.

7. Cette requête renvoie la liste des ingrédients qui composent la pizza dont le numéro unique d'identification est `id_p = 3`, c'est-à-dire les ingrédients de la pizza Hawaïenne.

Le résultat de la requête est le suivant :

nom
Sauce tomate
Mozzarella fraîche
Jambon cuit
Ananas
Pâte à pizza

8. Cette requête renvoie le montant total des dépenses effectuées par le client qui possède l'id\_client numéro 4.

```
SELECT SUM(montant)
FROM commande
WHERE id_client = 4;
```

En consultant la table `client`, nous voyons qu'il s'agit de Axel Aire, qui est l'un des meilleurs clients de la pizzeria.

Le résultat de la requête est le suivant : Axel Aire a dépensé un total de 36 € depuis l'ouverture de la pizzeria, à condition que la base de données soit à jour.

SUM("commande"."montant")
36

## F. Bilan du chapitre p. 128

Question	Réponse
1	<b>d. À organiser, analyser et manipuler des données sous forme de tableaux.</b>
2	<b>c. JSON.</b>
3	<b>a. CSV.</b>
4	<b>b. La virgule.</b>
5	<b>c. À localiser une ou plusieurs lignes d'une table en fonction d'un critère précis.</b>
6	<b>d. À organiser les données dans un ordre spécifique.</b>
7	<b>c. À effectuer des opérations mathématiques ou logiques sur des données.</b>
8	<b>b. À afficher uniquement les données qui répondent à certains critères.</b>
9	<b>b. Tri.</b>
10	<b>b. Un langage de requête structuré qui sert à interagir avec des bases de données.</b>
11	<b>c. Pour éviter la redondance des données et améliorer l'efficacité.</b>
12	<b>d. Du nombre de serveurs.</b>
13	<b>b. Les serveurs et leur refroidissement.</b>

Des QCM d'auto-évaluation sont disponibles pour un travail en autonomie de l'élève à l'adresse : [https://lienbordas.fr/740171\\_ch08\\_bilan](https://lienbordas.fr/740171_ch08_bilan).