

EXERCICES

2^{de}

Données — Structuration et traitement

Nom : _____

Classe : _____

Au contrôle, une réponse juste sans calcul ne rapporte presque rien : chaque problème se rédige en quatre temps — **justifier**, **poser** l'égalité, **mener le calcul** jusqu'au bout, **conclure** par une phrase avec l'unité. La place réservée est dimensionnée pour ces quatre temps.

1 Vocabulaire. Répondre en une phrase précise.

● ○ ○

- a. Qu'est-ce qu'une donnée structurée ?
- b. Différencier CSV et JSON.
- c. Qu'est-ce qu'une métadonnée ? Donner deux exemples.
- d. Qu'est-ce qu'une donnée personnelle au sens du RGPD ?
- e. Citer trois principes du RGPD.

2 Une table recense les emprunts d'une bibliothèque : identifiant, titre, catégorie, date d'emprunt, nombre de jours de retard.

● ● ○

- a. Quelle opération permet d'obtenir les seuls emprunts en retard ?
- b. Quelle suite d'opérations donne les cinq retards les plus longs ?
- c. La colonne « retard » comporte des cases vides. Peut-on les compter comme des zéros ?
- d. On veut le retard moyen par catégorie. Décrire la démarche.

3 Un établissement publie un fichier « anonymisé » : le nom est supprimé, mais il reste la date de naissance, le sexe, le code postal et la filière.

- Ce fichier est-il anonyme ? Justifier.
- Expliquer le mécanisme de réidentification par croisement.
- Quelles modifications rendraient la réidentification nettement plus difficile ?
- Quel principe du RGPD aurait dû s'appliquer dès la collecte ?
- Pourquoi la bonne intention de l'établissement ne suffit-elle pas ?

4 Big data et interprétation. Analyser.

- Qu'appelle-t-on big data ?
- En quoi la démarche diffère-t-elle d'une étude statistique classique ?
- Une corrélation forte entre deux colonnes prouve-t-elle un lien de cause à effet ?
- Pourquoi multiplier les croisements produit-il mécaniquement des corrélations sans signification ?

Corrigé détaillé sur progressschool.fr — chapitre « Données — Structuration et traitement ».