Ce document a été mis en ligne par l'organisme FormaV®.

Toute reproduction, représentation ou diffusion, même partielle, sans autorisation préalable, est strictement interdite.

Pour en savoir plus sur nos formations disponibles, veuillez visiter :
formav.co/explorer

BACHELOR UNIVERSITAIRE DE TECHNOLOGIE

SCIENCE DES DONNÉES

ÉVALUATION DE LA COMPÉTENCE C1 : TRAITER DES DONNÉES À DES FINS DÉCISIONNELLES

SESSION 2026

Durée : 2 heures

SPÉCIALITÉ / UNITÉ	COEFF
SCIENCE DES DONNÉES	3,0 INDICATIF

Moyens de calculs et matériel autorisés :

L'usage de la calculatrice avec mode examen actif est autorisé. L'usage de la calculatrice sans mémoire, « type collège », est autorisé.

Aucun document ou matériel autorisé en dehors de ceux fournis avec le sujet.

Dès que le sujet vous est remis, assurez-vous qu'il est complet.

Ce sujet comporte 10 pages numérotées de 1/10 à 10/10.

Les documents réponses en fin de sujet sont à rendre avec la copie.

DOSSIER SUJET

MISE EN SITUATION PROFESSIONNELLE ET CAHIER DES CHARGES : DIRECTION DES BÂTIMENTS DE LA VILLE DE VALMONT-SUR-RIVE

La Direction des bâtiments de Valmont-sur-Rive souhaite fiabiliser le suivi des consommations électriques de trois bâtiments municipaux. Vous intervenez comme technicienne ou technicien en science des données. À partir des relevés du mois de mars 2026 fournis dans le dossier technique, vous devez proposer une organisation relationnelle, contrôler et traiter les données, produire des indicateurs et documenter le traitement.

Le besoin métier est de suivre les consommations relevées, de calculer des indicateurs comparables entre bâtiments et de repérer les relevés strictement supérieurs à 320 kWh. Un relevé est rattaché à un seul bâtiment. Les dates et valeurs présentées dans le dossier constituent l'intégralité des données nécessaires aux calculs demandés.

Le sujet porte sur la compétence C1 du BUT Science des Données. Le BUT ne comporte pas d'épreuve nationale unique de C1 avec durée, coefficient, règle de calculatrice ou code d'épreuve identiques pour tous les établissements. La durée de 2 heures et le coefficient indicatif de 3,0 sont les paramètres retenus pour cette proposition de sujet et ne valent pas modalités nationales de contrôle des connaissances.

PARTIE 1 : Analyse du besoin et modélisation des données (5 points)

Traduire le besoin métier en modèle de données cohérent

Vous vous appuierez sur le cahier des charges et le dictionnaire de données du DT 2. Les identifiants de bâtiments et de relevés sont uniques. Chaque relevé concerne un seul bâtiment. Les valeurs de consommation sont exprimées en kWh.

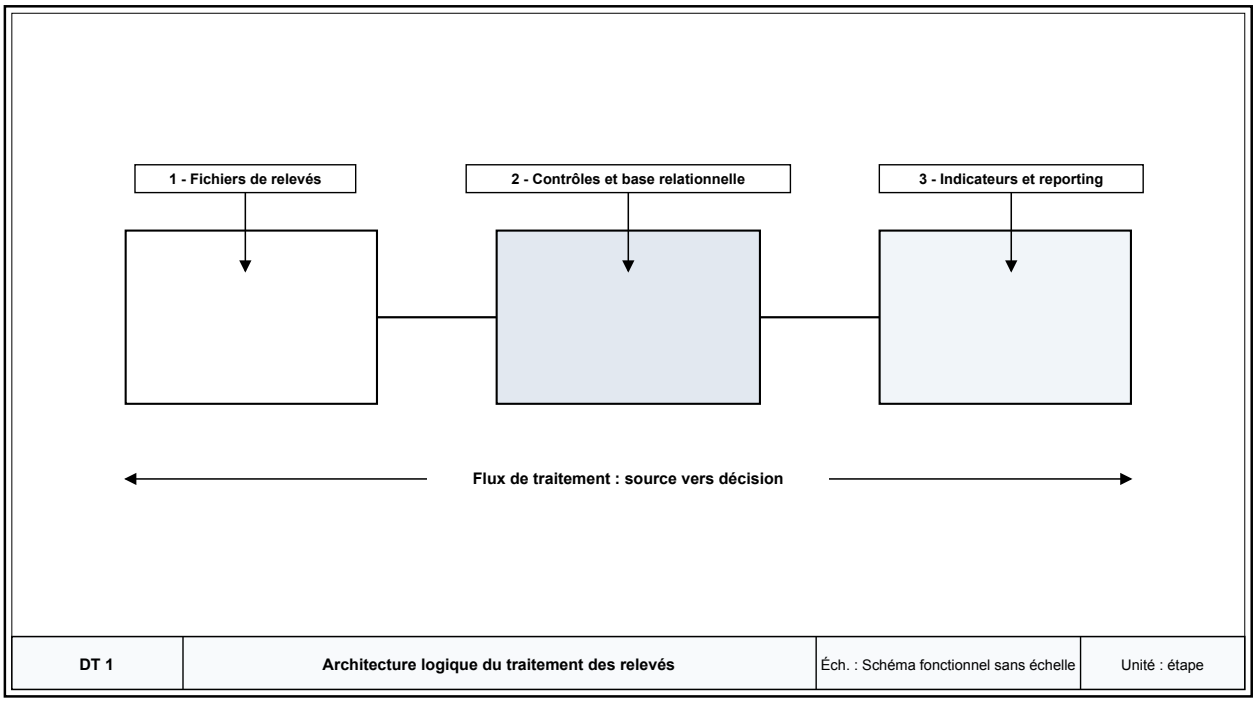


Schéma fonctionnel : les données sources sont contrôlées avant leur chargement et leur exploitation.

1.1 (2 points) Proposez un modèle relationnel minimal pour répondre au besoin. Indiquez les relations, leur clé primaire et la clé étrangère, ainsi que les types de données adaptés pour les champs principaux. Précisez la cardinalité entre bâtiment et relevé.

1.2 (1,5 point) Expliquez en quoi la séparation des bâtiments et des relevés évite une anomalie de mise à jour. Donnez une règle de validation pertinente pour la date et une règle pertinente pour la consommation.

1.3 (1,5 point) Complétez le DR 1 avec trois contrôles à effectuer avant le chargement en base. Pour chaque contrôle, indiquez la donnée concernée et l'action à mener si le contrôle échoue.

PARTIE 2 : Requêtes SQL et cycle de vie des données (5 points)

Corriger, insérer, supprimer et extraire des données

On utilise le schéma suivant : Batiment(id_batiment, nom, surface_m2) et Releve(id_releve, id_batiment, date_releve, conso_kwh, source). Les identifiants sont des chaînes de caractères. La clé étrangère Releve.id_batiment référence Batiment.id_batiment. Les relevés valides de la période sont ceux dont la date est comprise entre le 1er mars 2026 inclus et le 16 mars 2026 exclu. Un relevé est en alerte lorsque sa consommation est strictement supérieure à 320 kWh. Les opérations de nettoyage sont réalisées avant le calcul des indicateurs.

Résumé des données utiles aux requêtes

Bâtiment	Identifiant	Surface en m²	Règle d'alerte
Centre administratif	S01	1200	Consommation > 320 kWh
Gymnase municipal	S02	800	Consommation > 320 kWh
Médiathèque	S03	600	Consommation > 320 kWh

Les enregistrements détaillés et l'anomalie identifiée figurent au DT 2.

2.1 (2 points) Écrivez les instructions SQL permettant, dans cet ordre, de supprimer le doublon identifié au DT 2, de corriger la valeur du relevé R006 à 235 kWh, puis d'insérer le relevé R010 indiqué au DT 2. Utilisez des conditions suffisamment précises pour ne modifier que les enregistrements visés.

2.2 (1,5 point) Écrivez une requête SQL qui retourne, pour chaque bâtiment, le nom du bâtiment et la consommation moyenne des relevés valides entre le 1er et le 15 mars 2026 inclus. La moyenne doit être arrondie à deux décimales et le résultat trié par identifiant de bâtiment croissant. Le doublon doit être exclu après nettoyage.

2.3 (1,5 point) Écrivez une requête SQL qui compte le nombre de relevés en alerte sur la même période, après nettoyage. Indiquez le résultat attendu à partir des données du DT 2.

PARTIE 3 : Automatisation du traitement et production d'indicateurs (5 points)

Programmer, tester et documenter un traitement reproductible

La Direction souhaite un traitement automatisé à partir des relevés valides du DT 2, après correction et suppression du doublon. Pour les indicateurs de mars, les données sont limitées aux relevés du 1er au 15 mars inclus. Le relevé R010 du 22 mars est donc exclu. On calcule, par bâtiment, la consommation totale de la période rapportée à la surface, ainsi que l'évolution entre le premier et le dernier relevé de la période.

La formule de consommation surfacique est $I = \frac{\text{consommation totale en kWh}}{\text{surface en m}^2}$. L'évolution relative entre la première et la dernière valeur est $T = 100 \times \frac{C_{\text{dernier}} - C_{\text{premier}}}{C_{\text{premier}}}$.

Indicateurs à produire

Identifiant	Consommation totale en kWh	Surface en m²	Indice en kWh/m²	Évolution en %
S01	1215	1200	.....	.....
S02	965	800	.....	.....
S03	715	600	.....	.....

Valeurs calculées à partir des relevés valides du 1er au 15 mars inclus.

3.1 (2 points) Rédigez, en Python ou en pseudo-code structuré, une fonction qui reçoit les consommations d'un bâtiment et sa surface, puis renvoie la consommation totale rapportée à la surface. La fonction doit rejeter une liste vide et une surface inférieure ou égale à zéro. Commentez les contrôles réalisés.

3.2 (1,5 point) Calculez les trois indices en kWh/m² et les trois évolutions en pourcentage. Arrondissez les indices à quatre décimales et les évolutions à deux décimales. Complétez le DR 2. Pour chaque bâtiment, utilisez le premier et le dernier relevé valide de la période.

3.3 (1,5 point) Proposez deux tests unitaires vérifiables pour la fonction de la question 3.1, dont un test de cas limite. Indiquez pour chacun les données d'entrée et le résultat attendu.

PARTIE 4 : Prévention des risques et documentation (5 points)

Sécuriser et rendre le traitement compréhensible pour ses utilisateurs

Les indicateurs seront transmis à la Direction des bâtiments sous forme d'un tableau de bord mensuel. Les fichiers sources et les résultats sont conservés sur l'espace professionnel de la collectivité. Le dossier ne contient pas de données personnelles, mais les consommations constituent des données de gestion dont l'intégrité et la disponibilité doivent être préservées.

Situations à analyser

Situation	Risque principal
Un fichier source est remplacé sans conservation de version	Perte de traçabilité ou impossibilité de reproduire un résultat
Un utilisateur non habilité modifie un relevé	Altération des indicateurs
Un tableau de bord ne précise ni période ni unité	Interprétation erronée des résultats

Les mesures proposées doivent être proportionnées au contexte décrit.

4.1 (2 points) Pour les trois situations du tableau, proposez une mesure de prévention ou de réduction du risque. Votre réponse doit inclure une mesure portant sur les droits d'accès, une mesure portant sur la sauvegarde ou la traçabilité et une mesure portant sur la validation des données.

4.2 (1,5 point) Rédigez une courte documentation destinée à un responsable métier. Elle doit préciser la période étudiée, la définition de l'alerte, l'unité de l'indice surfacique et le sens d'une évolution négative.

4.3 (1,5 point) Expliquez pourquoi le traitement doit conserver séparément les données sources, les données corrigées et les résultats calculés. Indiquez un moyen concret de rendre le calcul reproductible par une autre personne de l'équipe.

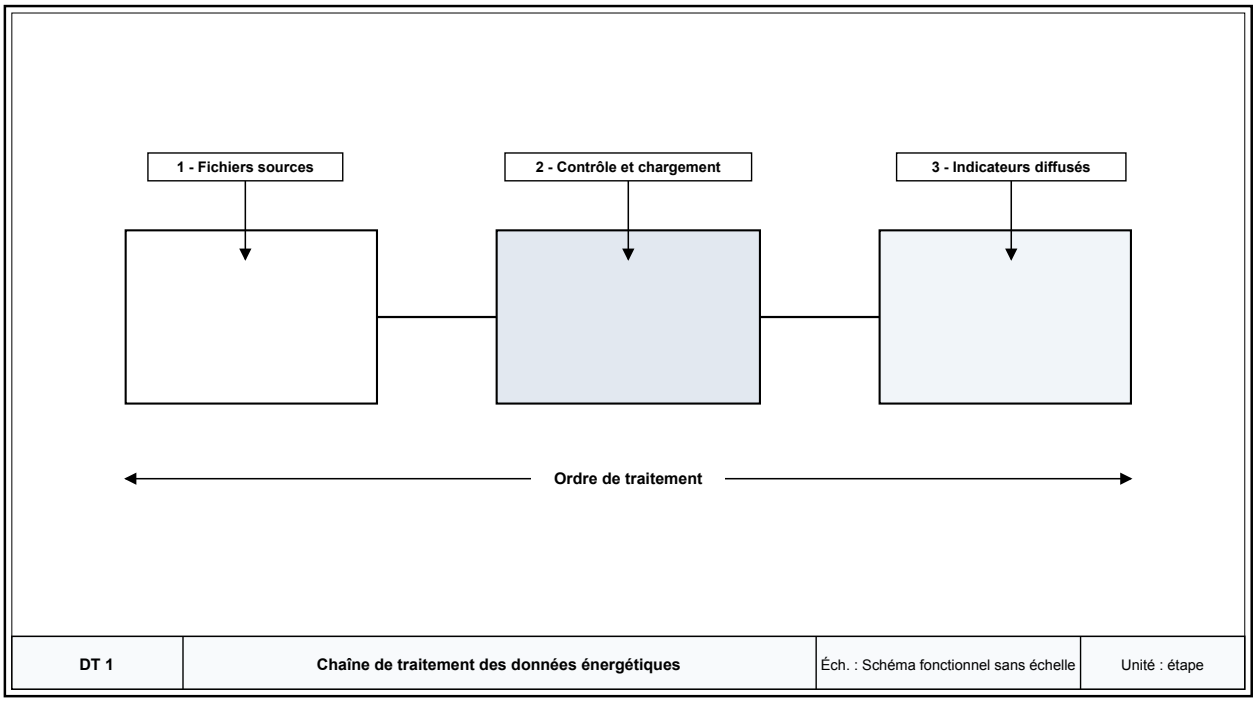
DOSSIER TECHNIQUE ET RESSOURCES

Documents ressources nécessaires à la résolution du sujet

DT 1 : ARCHITECTURE FONCTIONNELLE DU TRAITEMENT

Le flux de traitement attendu est : réception des fichiers, contrôle des valeurs et des identifiants, chargement dans les relations Batiment et Releve, calcul des indicateurs, puis diffusion d'un tableau de bord. Les données sources doivent être conservées sans modification. Toute correction est tracée avec l'identifiant du relevé, l'ancienne valeur, la nouvelle valeur, la date de correction et l'auteur de l'opération.

Le schéma fonctionnel représente un flux logique. Il ne prescrit ni logiciel ni architecture informatique particulière.



Les flèches indiquent l'ordre des étapes de traitement.

DT 2 : DICTIONNAIRE DE DONNÉES ET RELEVÉS DE MARS 2026

Les données suivantes sont fournies pour l'épreuve. Les surfaces sont exprimées en mètres carrés et les consommations en kWh. Les dates sont au format ISO AAAA-MM-JJ. La ligne R000 est un doublon exact de R001 et doit être supprimée. Après nettoyage, la consommation de R006 doit être corrigée à 235 kWh. Le relevé R010 est un nouvel enregistrement à insérer ; il est situé hors de la période d'indicateurs définie dans le sujet.

Types conseillés : identifiants en VARCHAR, noms et sources en VARCHAR, dates en DATE, surface et consommation en DECIMAL. Règles d'intégrité : id_batiment est une clé primaire dans Batiment ; id_releve est une clé primaire dans Releve ; Releve.id_batiment est une clé étrangère non nulle vers Batiment.id_batiment.

Relation Batiment

id_batiment	nom	surface_m2
S01	Centre administratif	1200
S02	Gymnase municipal	800
S03	Médiathèque	600

Données de référence des trois bâtiments.

Relation Releve avant nettoyage

id_releve	id_batiment	date_releve	conso_kwh	source
R000	S01	2026-03-01	420	compteur
R001	S01	2026-03-01	420	compteur
R002	S01	2026-03-08	390	compteur
R003	S02	2026-03-01	310	compteur
R004	S02	2026-03-08	330	compteur
R005	S03	2026-03-01	250	compteur
R006	S03	2026-03-08	230	compteur
R007	S01	2026-03-15	405	compteur
R008	S02	2026-03-15	325	compteur
R009	S03	2026-03-15	240	compteur

La valeur de R006 est à corriger ; R000 est le doublon exact de R001.

Nouvel enregistrement à insérer

id_releve	id_batiment	date_releve	conso_kwh	source
R010	S01	2026-03-22	400	compteur

Cet enregistrement est valide mais hors de la période du calcul de mars demandée dans le sujet.

DT 3 : RÈGLES DE QUALITÉ ET DE CALCUL DES INDICATEURS

Une date de relevé doit être une date valide et ne peut pas être nulle. Une consommation doit être numérique, non nulle et supérieure ou égale à zéro. La surface d'un bâtiment doit être strictement positive. Un identifiant de relevé ne peut apparaître qu'une seule fois après nettoyage. Un relevé doit faire référence à un bâtiment existant.

Pour la période du 1er au 15 mars inclus, la consommation totale d'un bâtiment est la somme de ses consommations valides. L'indice surfacique est calculé par $I = \frac{\sum C_i}{S}$, où S est la surface en mètres carrés. L'évolution entre le premier et le dernier relevé est calculée par $T = 100 \times \frac{C_{\text{dernier}} - C_{\text{premier}}}{C_{\text{premier}}}$. Une évolution négative signifie que la dernière consommation est inférieure à la première.

Les résultats du sujet sont arrondis à quatre décimales pour les indices surfaciques et à deux décimales pour les évolutions en pourcentage. Le seuil d'alerte est strictement supérieur à 320 kWh.

Données de référence pour les contrôles

Champ ou indicateur	Règle attendue
Identifiant de relevé	Unique et non nul
Identifiant de bâtiment	Présent dans la relation Batiment
Consommation	Valeur numérique supérieure ou égale à 0
Surface	Valeur strictement supérieure à 0
Alerte	Consommation strictement supérieure à 320 kWh
Période des indicateurs	Du 1er au 15 mars 2026 inclus

Règles définies par le cahier des charges de la Direction des bâtiments.

NE RIEN ÉCRIRE DANS CETTE PARTIE

DR 1 : Document réponse à rendre avec la copie : Question 1.3 **À RENDRE AVEC LA COPIE**

Complétez les cases. Pour chaque contrôle, précisez la donnée concernée et l'action à mener si le contrôle échoue.

Plan de contrôles avant chargement

Contrôle	Donnée concernée	Action en cas d'échec
.....	.....	.....
.....	.....	.....
.....	.....	.....

Cadre réservé aux justifications, calculs complémentaires ou schémas du candidat :

NE RIEN ÉCRIRE DANS CETTE PARTIE

DR 2 : Document réponse à rendre avec la copie : Question 3.2 **À RENDRE AVEC LA COPIE**

Complétez les deux dernières colonnes. Utilisez quatre décimales pour l'indice surfacique et deux décimales pour l'évolution en pourcentage.

Calcul des indicateurs par bâtiment

Identifiant	Consommation totale en kWh	Surface en m²	Indice en kWh/m²	Évolution en %
S01	1215	1200	.....	.....
S02	965	800	.....	.....
S03	715	600	.....	.....

Cadre réservé aux justifications, calculs complémentaires ou schémas du candidat :

Copyright © 2027 FormaV®. Tous droits réservés.

FormaV® est une marque déposée auprès de l'INPI sous le numéro 5227897.

Ce document a été élaboré par FormaV® avec le plus grand soin afin d'accompagner chaque apprenant vers la réussite de ses examens. Son contenu (textes, graphiques, méthodologies, tableaux, exercices, concepts, mises en forme) constitue une œuvre protégée par le droit d'auteur.

Toute copie, partage, reproduction, diffusion ou mise à disposition, même partielle, gratuite ou payante, est strictement interdite sans accord préalable et écrit de FormaV®, conformément aux articles L.111-1 et suivants du Code de la propriété intellectuelle. Dans une logique anti-plagiat, FormaV® se réserve le droit de vérifier toute utilisation illicite, y compris sur les plateformes en ligne ou sites tiers.

En utilisant ce document, vous vous engagez à respecter ces règles et à préserver l'intégrité du travail fourni. La consultation de ce document est strictement personnelle.

Merci de respecter le travail accompli afin de permettre la création continue de ressources pédagogiques fiables et accessibles.