



Data Management
DATA MANAGEMENT



ENSIAS
LECTURE ACADÉMIQUE

ENSIAS · UNIVERSITÉ MOHAMMED V DE RABAT · 1ÈRE ANNÉE · GÉNIE DE LA DATA

Initiation au management de la data

TD 1

TD1 — Auditer un dossier de données avant une décision

De la collecte à la décision : rendre les données fiables, gouvernées et utiles



Data Management
DATA MANAGEMENT

TP · SOMMAIRE

Table des matières

Organisation des 30 minutes	3
Mise en situation universitaire	3
Dossier fourni	4
Travail demandé — progression des questions	4
1. Lire le dossier — repérer les faits	4
2. Auditer la documentation — peut-on réutiliser le dossier ?	5
3. Évaluer rapidement la réutilisabilité — mini-grille FAIR	5
4. Construire un indicateur et formuler la question	6
5. Cartographier les responsabilités et sécuriser la décision	6
Critères de réussite — 10 points	7
Restitution	3
Pont vers les TP	8
Inspiration pédagogique	8





Cours : Initiation au management de la data **Durée :** 30 minutes de travail + 10 minutes de restitution et correction collective **Production :** un audit FAIR, une fiche d'indicateur, une cartographie annotée et une décision provisoire

Organisation des 30 minutes

Temps	Priorité	Production
0–7 min	lire les faits et l'anomalie	tableau observé/inconnu/à vérifier
7–15 min	auditer la documentation et FAIR	notes justifiées et deux améliorations
15–24 min	construire l'indicateur	formule, unité, couverture et calculs A/B
24–30 min	préparer la décision	cartographie courte, risques et vérification suivante

La restitution collective reprend ensuite les réponses enrichies et corrige les confusions entre ouverture, occupation, énergie observée et index de compteur.

Mise en situation universitaire

Dans le cadre de la démarche de sobriété énergétique du campus ENSIAS, le service patrimoine prépare une réunion avec la direction des études, la bibliothèque et les responsables des laboratoires. L'équipe doit décider quel bâtiment fera l'objet du premier diagnostic, sans dégrader les conditions d'enseignement, de recherche ou d'accueil des étudiants.

Vous intervenez comme une équipe d'analyse mandatée pour préparer ce diagnostic préliminaire. Votre rôle n'est pas de désigner un bâtiment « responsable » ni de conclure à un gaspillage. Vous devez d'abord vérifier si le dossier est compréhensible et réutilisable, puis construire une priorité de vérification, indiquer les preuves nécessaires et nommer les responsables de la suite.

Le dossier est pédagogique et anonymisé. Il ne contient aucune donnée individuelle sur les étudiants ou les personnels. Toute donnée supplémentaire doit répondre à une question explicite, une finalité et une durée de conservation.

Question directrice : peut-on défendre une décision à partir d'un dossier qui n'est pas encore complètement documenté ?

Dossier fourni

Le service patrimoine fournit un extrait de relevés énergétiques d'une journée commune. Les valeurs sont fictives, mais le raisonnement doit respecter les exigences d'un contexte universitaire réel : provenance, couverture, unités, documentation, responsabilités et limites d'interprétation.

Les consommations représentent des énergies par intervalle, et non des index cumulés de compteur.

Bâtiment	Usage	Surface	Énergie observée	Couverture horaire
A	Salles de cours	1 000 m ²	120 kWh	24/24
B	Laboratoires	500 m ²	90 kWh	24/24
C	Bibliothèque	800 m ²	70 kWh	22/24

Ressource	Producteur	Point à vérifier
Relevés horaires CSV	service patrimoine	C : deux intervalles absents
Surfaces des bâtiments	service patrimoine	date et méthode de mesure
Horaires d'ouverture	scolarité / bibliothèque	distinguer ouverture et occupation réelle
Température extérieure	station du campus	période commune aux relevés

Travail demandé — progression des questions

Répondez dans l'ordre. Chaque étape réutilise la réponse précédente. Une réponse doit indiquer ce qui est **observé**, ce qui est **inconnu** et ce qui doit être **vérifié**.

Noyau obligatoire dans les 30 minutes : faits et anomalie, audit FAIR, question et indicateur, puis décision provisoire. La cartographie complète des responsabilités et l'analyse des risques servent à enrichir la restitution et la correction collective si le temps manque.

1. Lire le dossier — repérer les faits

Répondez sans encore choisir de bâtiment :

1. Quels bâtiments ont une couverture complète ?
2. Quelle anomalie empêche une comparaison immédiate avec les autres bâtiments ?
3. Quelles sont les deux unités nécessaires pour comparer une consommation entre bâtiments de surfaces différentes ?

Notez dans votre fiche la différence entre un fait observé et une hypothèse. Les intensités observées à calculer sont $A = 120/1\,000 = 0,12 \text{ kWh/m}^2$ et $B = 90/500 = 0,18 \text{ kWh/m}^2$. Pour C, $70/800 = 0,0875 \text{ kWh/m}^2$ est seulement une valeur observée sur 22/24 heures ; elle ne doit

pas être classée avec A et B comme si les couvertures étaient équivalentes.

2. Auditer la documentation — peut-on réutiliser le dossier ?

Imaginez qu'une nouvelle équipe reçoive uniquement ce dossier. Pour chaque question, indiquez l'information disponible, l'information manquante et l'action à réaliser avant réutilisation.

Question de réutilisation	Information disponible	Information manquante	Action ou preuve à demander
Que mesure exactement chaque valeur ?			
Où, quand et avec quelle couverture les données ont-elles été produites ?			
Comment les relevés ont-ils été collectés et contrôlés ?			
Que signifie une valeur manquante ?			
Qui peut autoriser l'accès et l'usage ?			
Le format permet-il une réutilisation fiable ?			

Ne répondez pas « inconnu » sans proposer la trace, le document ou la personne qui permettrait de lever l'incertitude.

3. Évaluer rapidement la réutilisabilité — mini-grille FAIR

Attribuez à chaque dimension une note de 0 à 2 et justifiez-la :

- **0** : aucune preuve dans le dossier ;
- **1** : preuve partielle ou seulement locale ;
- **2** : preuve explicite et réutilisable.

Dimension	Question à poser	Note / 2	Justification et amélioration prioritaire
Findable — trouvable	Le fichier et sa description sont-ils identifiables ?		
Accessible — accessible	Sait-on où obtenir les données et sous quelles conditions ?		
Interoperable — interoperable	Les formats, unités et variables sont-ils compréhensibles par un autre outil ?		
Reusable — réutilisable	Provenance, qualité, limites et droits d'usage sont-ils documentés ?		

Choisissez ensuite les deux améliorations qui augmenteraient le plus la réutilisabilité du dossier.

4. Construire un indicateur et formuler la question

Transformez le besoin de la direction en une question d'analyse précise, puis complétez la fiche :

Champ	Votre réponse
Question d'analyse	
Décision provisoire	
Nom de l'indicateur	
Formule	
Unité	
Période et périmètre	
Couverture minimale	
Limite d'interprétation	

Calculez l'intensité observée kWh/m^2 pour A et B. Vous pouvez calculer la valeur observée pour C, mais vous ne devez pas la présenter comme un total journalier complet : sa couverture est de 22/24.

5. Cartographier les responsabilités et sécuriser la décision

Complétez le parcours suivant. Sous chaque étape, indiquez l'acteur et la trace qui reste :

CODE

```
[Question] → [Source / producteur] → [Brut daté] → [Contrôles]
                                     ↓
                               [Anomalie / action]
                                     ↓
                        [Indicateur] → [Décision]
```

Acteurs possibles : producteur, propriétaire métier, référent énergie, data steward, data engineer, data analyst, décideur.

Traces possibles : question et périmètre, fiche de source, fichier brut, rapport de couverture, fiche d'anomalie, formule, validation métier et règle d'accès.

Notez deux risques et une action de maîtrise pour chacun. Avant toute diffusion, vérifiez la pertinence, la qualité, la couverture, l'autorisation d'usage et la responsabilité de la décision.

Ne déduisez pas le comportement d'un étudiant ou d'un agent à partir de ces mesures. Une observation énergétique sert ici à préparer un diagnostic du bâtiment, pas à surveiller des personnes.

Risque	Action de maîtrise

Critères de réussite — 10 points

Critère	Points	Preuve attendue
Faits et anomalies	2	couverture, unités et inconnues distinguées
Documentation et FAIR	2	informations manquantes reliées à des preuves et améliorations
Indicateur et question	3	formule, calcul, unité, couverture et limite
Parcours et responsabilités	1	sources, contrôles, acteurs et traces représentés
Risques et décision	2	maîtrises, priorité provisoire et vérifications argumentées

Un résultat numériquement correct mais fondé sur un dossier mal documenté ne satisfait pas tous les critères.

Restitution

En deux minutes, présentez dans cet ordre :

1. un fait observé et l'anomalie principale ;

2. une lacune documentaire importante et la preuve à demander ;
3. l'indicateur et son calcul ;
4. la décision provisoire, ses risques et la vérification indispensable.

La restitution doit distinguer clairement un fait, une interprétation, une incertitude et une décision proposée.

Pont vers les TP

Conservez votre fiche de source, votre formule, la limite de couverture et la preuve à demander. Ces éléments seront réutilisés dans le TP1 pour documenter une acquisition et produire un jeu contrôlé, puis dans le TP2 pour relier qualité, indicateurs et décision.

Inspiration pédagogique

Les activités d'audit documentaire et de grille FAIR sont adaptées de ressources ouvertes du UK Data Service, notamment *Exercise: Data documentation* et *Exercise: FAIRness of a dataset* (CC BY 4.0). L'adaptation utilise ici un dossier énergétique universitaire fictif et ne reproduit pas leurs jeux de données.

