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 consulter tous nos sujets d'annales et corrigés de la Licence Maths, veuillez visiter :
licencemaths.fr/annales-sujets-et-corriges

LICENCE MATHÉMATIQUES • CORRIGÉ DÉTAILLÉ

EXPLOITATION DE DONNÉES

Licence Maths

Session 2026 • Code sujet : **26-LIC-MATH-EP1** • Durée : **2 heures** • Coefficient : **3, indicatif**

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Analyse fonctionnelle, qualité et réglementation des données** (4 questions) — **Page 3**
2. **Étude descriptive et calcul des erreurs** (4 questions) — **Page 5**
3. **Modélisation et exploitation de la relation entre référence et lecture** (4 questions) — **Page 8**
4. **Protocole d’analyse, décision et communication des résultats** (3 questions) — **Page 11**
5. **Ce qu’il faut retenir & Questions fréquentes (FAQ)** (Synthèse & Q&A Examen) — **Page 13**

Sujet officiel (PDF) : https://licencemaths.fr/assets/annales/licence-maths/C7 - Exploitation de données à des fins d'analyse/2026/Licence Mathématiques - C7 - Exploitation de données à des fins d'analyse - Session 2026 - 1-Sujet.pdf

1. Analyse fonctionnelle, qualité et réglementation des données

Le laboratoire ThermoMesure étudie des mesures comparant la température de référence à la lecture d'une sonde, avec des données d'identification et d'horodatage. Il faut préparer une analyse fiable, limiter les données conservées et circonscrire la portée des critères internes.

Question 1.1. Pour les variables temperature_reference, temperature_sonde et id_mesure, indiquez leur rôle dans l'étude et leur nature statistique. Précisez laquelle constitue la variable explicative dans une régression de la lecture de la sonde en fonction de la référence.

Variable	Rôle	Nature statistique
temperature_reference	Valeur de référence, variable explicative	Quantitative, en °C
temperature_sonde	Lecture observée, variable réponse	Quantitative, en °C
id_mesure	Repérage technique des observations	Identifiant entier, qualitatif nominal

Pour la régression demandée, la lecture de la sonde est modélisée en fonction de la référence : la référence est donc la variable explicative et la lecture la variable réponse. L'identifiant sert à distinguer les mesures, mais ses valeurs numériques ne représentent ni une quantité mesurée ni un ordre interprétable.

> **✓ RÉPONSE À RETENIR**
>
> Variable explicative : temperature_reference. Variable réponse : temperature_sonde. id_mesure est un identifiant technique, non une variable quantitative d'analyse.

Question 1.2. À partir du protocole du DT 1, identifiez deux contrôles de qualité à effectuer avant toute analyse statistique. Pour chacun, précisez le problème recherché et l'action adaptée en cas d'anomalie.

Deux contrôles adaptés au protocole du DT 1 sont nécessaires avant l'analyse :

- **Complétude et unicité :** rechercher les valeurs manquantes et les identifiants dupliqués, qui pourraient fausser l'effectif ou compter une mesure deux fois. Vérifier l'extraction source et corriger avec traçabilité, ou signaler toute exclusion.
- **Respect des conditions de mesure :** vérifier l'unité, l'association de chaque lecture à sa référence et la stabilisation préalable de l'affichage. En cas d'écart, consulter le relevé original et, si possible, refaire la mesure.

Le DT 1 précise qu'aucune valeur ne manque dans les six lignes fournies.

> **✓ RÉPONSE À RETENIR**
>
> Vérifier complétude, unicité, unités, correspondance des mesures et stabilisation; contrôler toute anomalie à la source et la documenter.

Question 1.3. **Puisque l'identité des opérateurs n'est pas nécessaire au fichier d'analyse, proposez une mesure de minimisation des données et une précaution de diffusion compatible avec la situation. Justifiez brièvement votre réponse.**

Le code opérateur peut être retiré du fichier d'analyse puisqu'il n'est pas nécessaire aux calculs. Si la traçabilité doit être préservée, conserver séparément toute table de correspondance et en restreindre l'accès. Pour la diffusion, ne transmettre que les données et résultats utiles aux destinataires autorisés. Les horodatages détaillés peuvent être supprimés ou réduits à la date, car leur précision, combinée au petit nombre de mesures, pourrait faciliter une identification indirecte.

> ✓ **RÉPONSE À RETENIR**
>
> Retirer le code opérateur de l'analyse et limiter la diffusion aux données nécessaires, avec accès restreint et horodatages réduits si leur précision est inutile.

Question 1.4. **Expliquez pourquoi le seuil interne du laboratoire ne permet pas, à lui seul, de conclure que les sondes sont conformes à toute exigence réglementaire ou normative. Donnez un exemple d'information complémentaire à vérifier.**

Les seuils de biais moyen et d'erreur quadratique sont des critères internes définis pour cette étude. Ils permettent de juger la série selon le cahier des charges du laboratoire, mais ne démontrent pas le respect de toutes les exigences réglementaires ou normatives liées au produit, à son usage ou à une opération réglementée.

- Vérifier les textes et normes applicables à l'usage prévu.
- Par exemple, rechercher la tolérance exigée pour la plage de température concernée et vérifier les exigences applicables à l'étalonnage de l'étalon.

> ✓ **RÉPONSE À RETENIR**
>
> Les critères internes ne valent pas conformité réglementaire. Vérifier les exigences applicables à l'usage prévu, notamment les tolérances et l'étalonnage exigés.

2. Étude descriptive et calcul des erreurs

Le DT 1 décrit six lectures d'une sonde confrontées à des températures de référence comprises entre 10 °C et 35 °C. Il faut calculer les erreurs signées, les indicateurs de biais et de dispersion, puis évaluer la série au regard des critères internes sans généraliser la conclusion.

Question 2.1. Complétez le DR 1 en calculant les six erreurs signées $e_i = y_i - x_i$, puis leur somme. Présentez les résultats en °C avec une précision de 0,1 °C.

Mesure	Erreur e_i (°C)
M1	0,4
M2	0,1
M3	0,6
M4	0,3
M5	0,7
M6	0,2
Somme	2,3

Pour chaque ligne, on soustrait la température de référence à la lecture de la sonde. Par exemple, M1 donne 10,4 − 10,0 = 0,4 °C. Les six erreurs étant positives, la sonde indique toujours une température supérieure à la référence sur cette série. La somme des erreurs signées est de 2,3 °C, à reporter dans la colonne et la ligne correspondantes du DR 1.

DR 1 complété : erreurs signées

Mesure	Référence x_i (°C)	Lecture y_i (°C)	Erreur e_i (°C)	e_i² (°C²)
M1	10,0	10,4	0,4	
M2	15,0	15,1	0,1	
M3	20,0	20,6	0,6	
M4	25,0	25,3	0,3	
M5	30,0	30,7	0,7	
M6	35,0	35,2	0,2	
Somme			2,3	

✓ RÉPONSE À RETENIR

Erreurs M1 à M6 : 0,4 ; 0,1 ; 0,6 ; 0,3 ; 0,7 ; 0,2 °C. Somme : 2,3 °C.

Question 2.2. **Calculez le biais moyen $\bar{e} = (1/n)\Sigma e_i$ et l'erreur absolue moyenne $Eabs = (1/n)\Sigma|e_i|$. Donnez les résultats en °C à 0,01 °C près et interprétez leur signe ou leur valeur.**

Le biais moyen est la moyenne des erreurs signées : (0,4 + 0,1 + 0,6 + 0,3 + 0,7 + 0,2) / 6 = 0,3833 °C, soit +0,38 °C à 0,01 °C près. L'erreur absolue moyenne est (|0,4| + |0,1| + |0,6| + |0,3| + |0,7| + |0,2|) / 6 = 0,3833 °C, soit 0,38 °C. Ces valeurs coïncident parce que toutes les erreurs sont positives. Le biais positif indique une surestimation moyenne de la température de référence par la sonde.

✓ **RÉPONSE À RETENIR**

Biais moyen : +0,38 °C. Erreur absolue moyenne : 0,38 °C.

Question 2.3. **Calculez l'erreur quadratique moyenne $EQ = \sqrt{((1/n)\Sigma e_i^2)}$, en °C à 0,01 °C près. Détaillez le calcul de la somme des carrés des erreurs.**

Les erreurs au carré valent respectivement 0,16 ; 0,01 ; 0,36 ; 0,09 ; 0,49 et 0,04 °C². Leur somme est 1,15 °C². On obtient alors EQ = √(1,15 / 6) = 0,4378 °C, soit 0,44 °C à 0,01 °C près. La racine carrée restitue une valeur dans la même unité que les erreurs initiales, le degré Celsius.

DR 1 complété : carrés des erreurs

Mesure	Référence x_i (°C)	Lecture y_i (°C)	Erreur e_i (°C)	e_i² (°C²)
M1	10,0	10,4	.....	0,16
M2	15,0	15,1	.....	0,01
M3	20,0	20,6	.....	0,36
M4	25,0	25,3	.....	0,09
M5	30,0	30,7	.....	0,49
M6	35,0	35,2	.....	0,04
Somme	.....	.....	.....	1,15

✓ **RÉPONSE À RETENIR**

Somme des carrés : 1,15 °C². Erreur quadratique moyenne : 0,44 °C.

Question 2.4. **Comparez les indicateurs obtenus aux deux critères internes du laboratoire. Concluez sur la série observée en distinguant clairement cette conclusion d'une conclusion générale sur toutes les sondes et toutes les conditions d'utilisation.**

Le biais moyen absolu est de 0,38 °C, inférieur au seuil interne de 0,50 °C, et l'erreur quadratique moyenne est de 0,44 °C, inférieure au seuil de 0,60 °C : les deux critères sont donc respectés pour la série étudiée. Cette conclusion porte sur une seule sonde, six mesures et une séance, entre 10,0 °C et 35,0 °C. Il n'y a qu'une mesure

par niveau. La série ne permet pas de conclure sur les autres sondes, les répétitions, les séances ou les conditions d’utilisation non testées.

✓ RÉPONSE À RETENIR

Les deux critères internes sont respectés pour la série observée seulement. Aucune conformité générale n’est établie.

3. Modélisation et exploitation de la relation entre référence et lecture

Les six couples de températures du DT 1 sont étudiés à l'aide des sommes récapitulatives et des formules de régression linéaire du DT 2. Il faut ajuster la lecture en fonction de la référence, interpréter l'association, réaliser une estimation et discuter les limites prédictives et celles de la corrélation.

Question 3.1. Calculez x̄, ȳ, Sxx et Sxy, puis déterminez la pente a = Sxy/Sxx et l'ordonnée à l'origine b = ȳ − ax̄ de la droite de régression ŷ = ax + b. Arrondissez les coefficients à 0,001 près, indiquez leurs unités et complétez le DR 2.

Les moyennes valent $\bar{x} = 135/6 = 22,5$ °C et $\bar{y} = 137,3/6 = 22,883$ °C. Les sommes centrées sont $S_{xx} = 3\,475 - 135^2/6 = 437,5$ °C² et $S_{xy} = 3\,528 - (135 \times 137,3)/6 = 438,75$ °C². Ainsi, $a = S_{xy}/S_{xx} = 1,003$ °C/°C et $b = \bar{y} - a\bar{x} = 0,319$ °C, en conservant les valeurs non arrondies pour le calcul de b. La droite ajustée est $\hat{y} = 1,003x + 0,319$, coefficients arrondis à 0,001 près. Les résultats sont reportés dans le DR 2.

DR 2 : résultats de la régression

Résultat
22,5 °C .....
22,883 °C .....
437,5 °C² .....
438,75 °C² .....
.....
1,003 °C/°C .....
0,319 °C .....
.....

> **✓ RÉPONSE À RETENIR**
>
> $\bar{x} = 22,5$ °C ; $\bar{y} = 22,883$ °C ; $S_{xx} = 437,5$ °C² ; $S_{xy} = 438,75$ °C² ; $a = 1,003$ °C/°C ; $b = 0,319$ °C.

Question 3.2. À l'aide du DT 2, calculez Syy = Σyᵢ² − (Σyᵢ)²/n puis le coefficient de corrélation linéaire r = Sxy/√(SxxSyy). Arrondissez r à 0,001 près et interprétez-le sans le confondre avec une preuve de causalité. Complétez le DR 2.

La dispersion des lectures autour de leur moyenne est $S_{yy} = 3\,582,15 - 137,3^2/6 = 440,268$ °C². Le coefficient vaut alors $r = 438,75/\sqrt{437,5 \times 440,268333} = 0,999699$, soit $1,000$ à 0,001 près. Il indique une association linéaire positive très forte dans cette série. Il ne prouve pas une causalité et ne mesure pas directement l'accord entre la référence et la sonde. La valeur de S_{yy} et le coefficient arrondi sont reportés dans le DR 2.

DR 2 : dispersion et corrélation

Résultat
.....
.....
.....
.....
440,268 °C² .....
.....
.....
1,000 .....

✓ RÉPONSE À RETENIR

$S_{yy} = 440,268$ °C² ; $r = 0,999699$, soit $1,000$ à 0,001 près. L'association linéaire observée est très forte et positive.

Question 3.3. À l'aide de la droite obtenue en 3.1, estimez la lecture de la sonde pour une température de référence de 27,0 °C. Calculez ensuite le résidu de la mesure M4, défini par $r_4 = y_4 - \hat{y}_4$. Donnez les résultats en °C à 0,01 °C près.

En utilisant les coefficients non arrondis, $\hat{y} = (438,75/437,5)x + 0,319048$. Pour une référence de 27,0 °C, la lecture estimée est $27,396$ °C, soit 27,40 °C. Pour M4, $x_4 = 25,0$ °C et $y_4 = 25,3$ °C ; la droite prévoit $\hat{y}_4 = 25,390$ °C. Le résidu vaut donc $r_4 = y_4 - \hat{y}_4 = -0,090$ °C, soit −0,09 °C. Son signe négatif indique que la lecture M4 est inférieure à la valeur ajustée.

✓ RÉPONSE À RETENIR

Lecture estimée à 27,0 °C : 27,40 °C. Résidu de M4 : −0,09 °C.

Question 3.4. Expliquez pourquoi une forte valeur du coefficient de corrélation, si elle est observée, ne suffit pas à démontrer l'accord entre la sonde et la référence. Appuyez votre réponse sur au moins un indicateur calculé dans la partie 2 et une limite liée à la taille ou à l'étendue de la série.

La corrélation décrit l'intensité d'une relation linéaire, non la proximité des mesures. Ici, malgré $r \approx 1$, le biais moyen est de $2,3/6 = 0,38$ °C et l'erreur quadratique moyenne vaut $\sqrt{1,15/6} = 0,44$ °C : ces indicateurs portent directement sur les erreurs. La conclusion reste limitée à six mesures d'une seule sonde, effectuées en une séance entre 10 et 35 °C. Une seule mesure par niveau ne renseigne pas sur la répétabilité ; les autres sondes, séances, opérateurs, régimes transitoires et températures hors de cet intervalle ne sont pas étudiés.

✓ RÉPONSE À RETENIR

Une corrélation presque parfaite ne garantit pas l’accord : biais moyen 0,38 °C et erreur quadratique moyenne 0,44 °C. Conclusion limitée à six mesures d’une sonde entre 10 et 35 °C.

4. Protocole d’analyse, décision et communication des résultats

Le DT 3 présente les conditions de mesure, les limites de la série et les exigences de restitution de l’équipe qualité. Il faut rendre l’analyse reproductible, envisager une collecte complémentaire et communiquer une décision proportionnée aux données et aux critères internes.

Question 4.1. Rédigez, sous forme de liste ordonnée, quatre étapes permettant à un autre analyste de reproduire l’analyse depuis le fichier brut jusqu’à la conclusion. Chaque étape doit préciser une opération vérifiable.

1. Importer le fichier brut et vérifier les six lignes, les champs attendus, les valeurs manquantes, les températures et les horodatages; consigner toute anomalie.
2. Retirer le code opérateur du fichier d’analyse, conserver les variables utiles et tracer toute transformation, correction ou exclusion.
3. Calculer les erreurs signées, le biais moyen et l’erreur quadratique moyenne, puis ajuster la régression de la lecture en fonction de la référence avec les formules du DT 2.
4. Comparer les indicateurs aux seuils internes, vérifier les arrondis, préciser les limites de la série et rédiger une conclusion proportionnée.

> **✓ RÉPONSE À RETENIR**
>
> Vérifier et tracer les données, calculer les indicateurs, appliquer les seuils internes et conclure en précisant les limites.

Question 4.2. Proposez une collecte complémentaire permettant de mieux étudier les performances des sondes. Précisez au moins deux paramètres expérimentaux à faire varier ou à contrôler et indiquez pourquoi.

Organiser une collecte répétée et documentée afin de distinguer la variabilité de mesure de celle liée aux sondes, aux séances et aux conditions d’utilisation. Maintenir une référence étalonnée et consigner les conditions de stabilisation à chaque essai.

Paramètre	Collecte complémentaire	Intérêt
Sondes	Tester plusieurs exemplaires	Estimer la variabilité entre sondes
Séances et opérateurs	Répéter sur plusieurs jours et avec plusieurs opérateurs	Évaluer la reproductibilité
Température et régime	Ajouter des niveaux et mesurer en régime transitoire et stabilisé	Étudier la plage de mesure et le retard de réponse

Répéter les mesures à chaque combinaison de conditions, conserver la traçabilité des essais et ne diffuser que les données opérateur strictement nécessaires.

✓ RÉPONSE À RETENIR

Répéter les essais sur plusieurs sondes, séances, opérateurs et températures, en régime stabilisé et transitoire.

Question 4.3. Rédigez une conclusion de 80 à 120 mots destinée à l'équipe qualité. Donnez les résultats essentiels, statuez au regard des critères internes, mentionnez au moins deux limites de l'étude et proposez une suite opérationnelle. Ne présentez pas cette série comme une certification générale du produit.

Sur les six mesures de cette sonde, le biais moyen est de +0,38 °C et l'erreur quadratique moyenne de 0,44 °C. Ces deux résultats respectent les critères internes, respectivement 0,50 °C et 0,60 °C. La droite ajustée est $\hat{y} = 1{,}003x + 0{,}319$, avec les températures en °C. Cette conclusion concerne uniquement une sonde, mesurée une fois à chacun de six niveaux entre 10 et 35 °C, au cours d'une seule séance stabilisée. La variabilité entre sondes, séances et opérateurs, la dérive temporelle et le régime transitoire restent inconnus. Nous recommandons des essais répétés sur plusieurs sondes, séances et niveaux avant toute décision générale.

✓ RÉPONSE À RETENIR

La série respecte les deux critères internes; des essais complémentaires sont nécessaires avant toute conclusion générale sur les sondes.

Ce qu'il faut retenir de ce sujet (Session 2026)

Bilan de l'épreuve : Cette épreuve de licence Mathématiques, session 2026, porte le code 26-LIC-MATH-EP1 et dure 2 heures, avec un coefficient indicatif de 3. À partir d'une série de six mesures comparant la lecture d'une sonde de température à une référence, elle évalue le contrôle des données, le calcul des erreurs, la régression linéaire, l'interprétation statistique et la communication d'une décision limitée aux critères internes et à la série étudiée.

Notions-clés, formules & règles à maîtriser

- L'erreur signée est $e_i = y_i - x_i$; une valeur positive signifie que la sonde indique une température supérieure à la référence.
- Le biais moyen est la moyenne des erreurs signées ; l'erreur absolue moyenne mesure la grandeur moyenne des écarts sans tenir compte de leur signe.
- L'erreur quadratique moyenne est $EQ = \sqrt{(\Sigma e_i^2/n)}$; elle s'exprime dans la même unité que les erreurs.
- La corrélation mesure l'intensité d'une association linéaire, pas l'accord entre deux méthodes ni une relation de causalité.

Points de vigilance & attentes du jury

- Vérifier valeurs, unités, identifiants, doublons et données manquantes avant les calculs, puis tracer toute correction ou exclusion.
- Respecter la définition $e_i = y_i - x_i$, les formules demandées et la précision d'arrondi indiquée pour chaque résultat.
- Comparer explicitement les indicateurs aux seuils internes et limiter la décision à la série observée.
- Signaler l'effectif réduit, la mesure unique par niveau, les conditions non étudiées et toute extrapolation hors de 10 °C à 35 °C.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment calculer les erreurs et les indicateurs demandés ?

Pour chaque mesure, soustrayez la référence de la lecture afin d'obtenir $e_i = y_i - x_i$. Calculez ensuite la moyenne des erreurs, la moyenne de leurs valeurs absolues et la racine carrée de la moyenne de leurs carrés.

Q : Pourquoi une corrélation élevée ne prouve-t-elle pas que la sonde est juste ?

Une corrélation élevée indique que les lectures et les références varient ensemble de façon linéaire. Elle ne mesure pas directement l'écart entre elles : il faut aussi examiner le biais et les erreurs, ainsi que les limites de la série.

Q : Comment organiser les calculs et la rédaction dans les deux heures ?

Commencez par contrôler les données et effectuer les calculs des parties 2 et 3 en gardant les formules visibles. Réservez ensuite du temps aux questions d'analyse et à la conclusion, puis vérifiez les unités, les arrondis et les documents réponses à rendre.

Q : Quels critères de notation et quelles normes faut-il appliquer ?

Le sujet demande de comparer la valeur absolue du biais moyen au seuil de 0,50 °C et l'erreur quadratique moyenne au seuil de 0,60 °C. Ce sont des critères internes à l'étude, pas une certification ou une norme nationale ; aucune exigence externe n'est fournie.

Sujet officiel associé (PDF) : https://licencemaths.fr/assets/annales/licence-maths/C7 - Exploitation de données à des fins d'analyse/2026/Licence Mathématiques - C7 - Exploitation de données à des fins d'analyse - Session 2026 - 1-Sujet.pdf
Ressources Licence Mathématiques : Toutes les annales & corrigés • Fiches de révision complètes • Simulateur de notes d'examen

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.