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

LICENCE LETTRES • CORRIGÉ DÉTAILLÉ

BTS Lettres - C5 : Exploitation de données

Session 2026 • Code sujet : **26-LIC-LETT-EP1** • Durée : **2 heures** • Coefficient : **3,0** **indicatif**

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Comprendre et contrôler les données** (3 questions) — Page 3
2. **Résumer une série quantitative** (4 questions) — Page 5
3. **Analyser les réponses au questionnaire** (3 questions) — Page 8
4. **Présenter une conclusion et un protocole de suivi** (2 questions) — Page 10
5. **Ce qu'il faut retenir & Questions fréquentes (FAQ)** (Synthèse & Q&A Examen) — Page 12

Sujet officiel (PDF) : https://licencelettres.fr/assets/annales/licence-lettres/C5 - Exploitation de données à des fins d'analyse/2026/Licence Lettres - C5 - Exploitation de données à des fins d'analyse - Session 2026 - 1-Sujet.pdf

1. Comprendre et contrôler les données

La bibliothèque universitaire des Lettres souhaite mieux comprendre les usages de son espace de lecture et préparer une action de médiation autour des collections contemporaines. Cette partie examine dix observations afin de classifier les variables, contrôler la qualité des données et discuter les limites de représentativité et de causalité.

Nature des variables

Question 1.1. Pour les variables « créneau », « durée de lecture », « support principal » et « participation à l'atelier », indiquez si chacune est qualitative ou quantitative. Pour les variables qualitatives, précisez si leurs modalités sont nominales ou ordonnées.

Le type dépend de ce que la variable mesure : une catégorie ou une quantité numérique. Le créneau, le support et la participation sont des variables qualitatives nominales : leurs catégories ne suivent aucun ordre naturel. La durée est quantitative, car elle est mesurée en minutes. Les valeurs admises pour chaque variable sont celles du dictionnaire DT 3.

Variable	Type	Modalités ou unité
Créneau	Qualitative nominale	Matin ; Après-midi
Durée de lecture	Quantitative	Minutes
Support principal	Qualitative nominale	Imprimé ; Numérique
Participation à l'atelier	Qualitative nominale	Oui ; Non

DR 1 complété : types des variables

Variable ou indicateur	Réponse du candidat
Type de « créneau »	Qualitative nominale
Type de « durée de lecture »	Quantitative
Type de « support principal »	Qualitative nominale
Type de « participation à l'atelier »	Qualitative nominale
Somme des durées observées (min)	
Moyenne des durées (min)	
Rangs retenus pour la médiane	
Médiane (min)	
Premier quartile Q1 (min)	
Troisième quartile Q3 (min)	
Écart interquartile (min)	

✓ RÉPONSE À RETENIR

Créneau, support principal et participation : variables qualitatives nominales. Durée de lecture : variable quantitative, exprimée en minutes.

Contrôle de la qualité

Question 1.2. À partir des règles du dictionnaire de variables, proposez deux contrôles de qualité à effectuer avant tout calcul sur le fichier et indiquez, pour chacun, une anomalie qu'il permettrait de repérer.

Avant tout calcul, vérifier la cohérence des valeurs avec les règles de DT 3. Deux contrôles permettent de repérer des erreurs fréquentes :

- Contrôler les durées : chacune doit être numérique et strictement positive. Une valeur nulle ou une saisie textuelle est une anomalie.
- Contrôler les modalités : créneau « Matin » ou « Après-midi », support « Imprimé » ou « Numérique », participation « Oui » ou « Non ». Une valeur comme « Soir » révèle une erreur de codage.

Les valeurs absentes doivent également être signalées, sans remplacement arbitraire.

> **✓ RÉPONSE À RETENIR**
>
> Vérifier que les durées sont numériques et positives, et que les modalités respectent les catégories définies dans DT 3.

Représentativité et causalité

Question 1.3. L'observation a été réalisée sur deux créneaux, dans une seule bibliothèque et sur une journée. Formulez deux limites de représentativité de cet échantillon et expliquez pourquoi ces données ne permettent pas, à elles seules, d'affirmer que la participation à l'atelier cause une durée de lecture plus longue.

L'échantillon représente peu de situations : dix usagers observés dans une seule bibliothèque, pendant une seule journée. Il ne couvre que deux créneaux et ne permet donc pas de représenter les autres périodes ou établissements. Enfin, constater une différence de durée entre participants et non-participants ne prouve pas que l'atelier en est la cause. Les groupes ne sont pas comparés dans une expérience contrôlée et d'autres facteurs, comme les habitudes de lecture ou le créneau, peuvent influer sur la durée.

> **✓ RÉPONSE À RETENIR**
>
> Les résultats sont limités à dix usagers, un lieu, une journée et deux créneaux. Une différence observée ne démontre pas un lien de causalité.

2. Résumer une série quantitative

La série étudiée regroupe les durées de lecture de dix usagers observés, déjà rangées par ordre croissant. Il faut calculer et interpréter des indicateurs descriptifs, puis examiner l'influence de la valeur de 50 minutes sur la moyenne et la médiane.

Moyenne

Question 2.1. Calculez la moyenne des dix durées de lecture. Présentez la somme des valeurs et le calcul utilisé, puis donnez le résultat en minutes, arrondi à une décimale.

On additionne les dix durées, puis on divise la somme par l'effectif : $12 + 18 + 20 + 20 + 25 + 30 + 30 + 35 + 40 + 50 = 280$, puis $280 \div 10 = 28,0$. La durée moyenne observée est donc de 28,0 minutes. Elle résume les temps de lecture des dix usagers observés, sans décrire nécessairement ceux de l'ensemble des usagers de la bibliothèque.

DR 1 complété : somme et moyenne

Classification et indicateurs descriptifs

Variable ou indicateur	Réponse du candidat
Type de « créneau »	.....
Type de « durée de lecture »	.....
Type de « support principal »	.....
Type de « participation à l'atelier »	.....
Somme des durées observées (min)	280 min
Moyenne des durées (min)	28,0 min
Rangs retenus pour la médiane	.....
Médiane (min)	.....
Premier quartile Q1 (min)	.....
Troisième quartile Q3 (min)	.....
Écart interquartile (min)	.....

Compléter les cellules contenant « ... ».

✓ RÉPONSE À RETENIR

Somme : 280 min. Moyenne : 28,0 min.

Médiane

Question 2.2. Déterminez la médiane de la série en indiquant les deux rangs utilisés. Interprétez cette médiane dans le contexte de l'observation.

L'effectif est pair : la médiane est la moyenne des deux valeurs centrales, aux rangs 5 et 6.

1. Au rang 5, la durée est de 25 min ; au rang 6, elle est de 30 min.
2. Calcul : $(25 + 30) \div 2 = 27,5$ min.

Dans cette série, cinq durées sont inférieures à 27,5 minutes et cinq lui sont supérieures. La médiane partage donc les dix durées observées en deux groupes de même effectif.

DR 1 complété : rangs et médiane

Classification et indicateurs descriptifs

Variable ou indicateur	Réponse du candidat
Type de « créneau »	
Type de « durée de lecture »	
Type de « support principal »	
Type de « participation à l'atelier »	
Somme des durées observées (min)	
Moyenne des durées (min)	
Rangs retenus pour la médiane	Rangs 5 et 6
Médiane (min)	27,5 min
Premier quartile Q1 (min)	
Troisième quartile Q3 (min)	
Écart interquartile (min)	

Compléter les cellules contenant « … ».

✓ RÉPONSE À RETENIR

Rangs 5 et 6 : médiane de 27,5 min. Cinq durées sont de chaque côté de cette valeur.

Quartiles et dispersion

Question 2.3. En appliquant la définition de DT 3, déterminez le premier quartile et le troisième quartile. Calculez ensuite l'écart interquartile et interprétez-le en une phrase.

Selon DT 3, pour dix valeurs, Q_1 est la valeur de rang 3 et Q_3 celle de rang 8. La série étant ordonnée, on lit directement les durées correspondantes :

Indicateur	Rang ou calcul	Résultat
Premier quartile Q_1	3	20 min
Troisième quartile Q_3	8	35 min
Écart interquartile	$35 - 20$	15 min

L'écart interquartile mesure l'amplitude des 50 % de valeurs centrales. Ici, ces durées s'étendent sur 15 minutes, de 20 à 35 minutes.

DR 1 complété : quartiles et écart interquartile

Classification et indicateurs descriptifs

Variable ou indicateur	Réponse du candidat
Type de « créneau »	.....
Type de « durée de lecture »	.....
Type de « support principal »	.....
Type de « participation à l'atelier »	.....
Somme des durées observées (min)	.....
Moyenne des durées (min)	.....
Rangs retenus pour la médiane	.....
Médiane (min)	.....
Premier quartile Q1 (min)	..... **20 min** .....
Troisième quartile Q3 (min)	..... **35 min** .....
Écart interquartile (min)	..... **15 min** .....

Compléter les cellules contenant « ... ».

✓ RÉPONSE À RETENIR

$Q_1 = 20$ min ; $Q_3 = 35$ min ; écart interquartile : 15 min.

Comparaison des indicateurs

Question 2.4. Comparez la moyenne et la médiane. À partir des valeurs de la série, indiquez si la durée de 50 minutes semble influencer davantage la moyenne que la médiane et justifiez brièvement.

La moyenne est de 28,0 minutes et la médiane de 27,5 minutes. La durée de 50 minutes, supérieure à toutes les autres, tire la moyenne vers le haut puisque celle-ci utilise chaque valeur. La médiane dépend des deux rangs centraux, occupés par 25 et 30 minutes, et n'est donc pas directement affectée par la valeur maximale. Dans cette série, les 50 minutes influencent davantage la moyenne que la médiane.

✓ RÉPONSE À RETENIR

Les 50 min influencent davantage la moyenne (28,0 min) que la médiane (27,5 min), déterminée par les rangs centraux.

3. Analyser les réponses au questionnaire

Un questionnaire anonyme recueille les réponses de vingt usagers sur leur support principal et leur souhait de participer à un atelier de médiation littéraire. Il faut comparer les proportions de réponses positives, interpréter prudemment la différence observée et choisir une représentation graphique adaptée.

Proportions conditionnelles

Question 3.1. Calculez la proportion de réponses « Oui » parmi les répondants déclarant l'imprimé comme support principal, puis parmi ceux déclarant le numérique. Montrez les calculs et donnez les résultats en pourcentage entier.

Pour chaque support, la proportion conditionnelle se calcule sur le total de sa ligne : effectif « Oui » divisé par l'effectif du groupe, puis multiplié par 100. Pour l'imprimé : 6 ÷ 10 × 100 = 60 %. Pour le numérique : 3 ÷ 10 × 100 = 30 %. Les deux groupes comptent dix répondants ; les résultats sont déjà des pourcentages entiers, aucun arrondi supplémentaire n'est nécessaire.

DR 2, calculs des proportions complétés

Support principal	Effectif « Oui »	Effectif total du groupe	Calcul de la proportion	Proportion en %
Imprimé	6	10	6/10 × 100	60 %
Numérique	3	10	3/10 × 100	30 %

✓ RÉPONSE À RETENIR

Imprimé : 60 % ; numérique : 30 %.

Interprétation

Question 3.2. Décrivez la différence observée entre les deux groupes. Précisez qu'il s'agit d'une association observée dans ce questionnaire et non d'une preuve que le support principal détermine la réponse.

Les réponses positives sont plus fréquentes parmi les répondants déclarant l'imprimé comme support principal. La comparaison porte ici sur les proportions calculées pour les deux groupes de dix répondants.

- Imprimé : 60 % de réponses « Oui » ; numérique : 30 %, soit un écart de 30 points de pourcentage.
- Il s'agit d'une association observée dans ce questionnaire. Elle ne prouve pas que le support principal détermine le souhait de participer.

✓ RÉPONSE À RETENIR

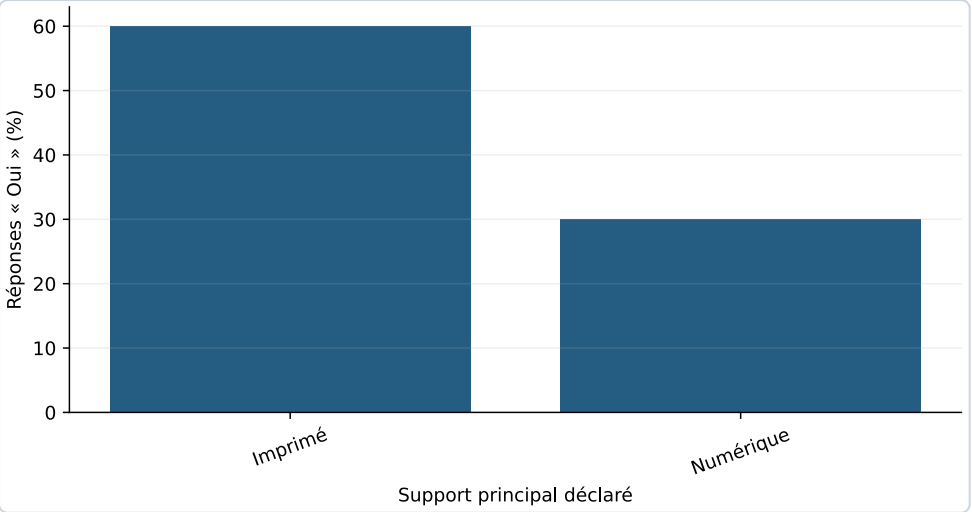
La proportion de « Oui » est supérieure de 30 points pour l'imprimé ; aucune causalité n'est établie.

Représentation graphique

Question 3.3. Proposez une représentation graphique adaptée pour comparer les proportions de réponses « Oui » entre les deux supports. Indiquez les catégories à placer sur chaque axe et les valeurs à représenter, puis justifiez votre choix en une phrase.

Élément du graphique	Contenu
Axe horizontal	Support principal déclaré : imprimé, numérique
Axe vertical	Proportion de réponses « Oui » (%)
Valeurs représentées	Imprimé : 60 % ; numérique : 30 %

Un diagramme en barres convient : chaque support forme une catégorie et la hauteur de sa barre indique la proportion correspondante. Les proportions, plutôt que les effectifs bruts, permettent de comparer directement les réponses positives dans les deux groupes.



✓ RÉPONSE À RETENIR

Diagramme en barres : imprimé, 60 % ; numérique, 30 %.

4. Présenter une conclusion et un protocole de suivi

L'équipe de la bibliothèque attend une note courte qui distingue les constats issus de l'observation et du questionnaire des limites de l'étude. Le candidat doit aussi définir un indicateur chiffré de suivi et proposer une amélioration concrète de la collecte.

Synthèse argumentée

Question 4.1. Rédigez une synthèse de trois à cinq phrases présentant un résultat descriptif issu de l'observation et un résultat issu du questionnaire. Mentionnez au moins une limite qui empêche de généraliser ces résultats à l'ensemble des usagers de bibliothèques universitaires.

Sur les dix usagers observés, les durées de lecture vont de 12 à 50 minutes et leur moyenne est de 28 minutes. Dans le questionnaire, 6 répondants sur 10 déclarant l'imprimé souhaitent l'atelier, soit 60 %, contre 3 sur 10 déclarant le numérique, soit 30 %. Cette différence constitue une association observée, et non la preuve que le support détermine la réponse. Les collectes portent sur une seule bibliothèque et une seule journée, avec seulement dix personnes observées et des réponses volontaires au questionnaire : elles ne permettent pas de généraliser à l'ensemble des usagers des bibliothèques universitaires.

> **✓ RÉPONSE À RETENIR**
>
> Observation : moyenne de 28 min. Questionnaire : intérêt déclaré de 60 % pour l'imprimé contre 30 % pour le numérique. Collectes limitées, sans généralisation possible.

Indicateur et amélioration de la collecte

Question 4.2. Proposez un indicateur chiffré à suivre lors d'une prochaine collecte pour évaluer l'intérêt pour l'atelier. Définissez son numérateur, son dénominateur et son unité, puis proposez une amélioration concrète de la collecte permettant de rendre la comparaison plus informative.

Élément	Proposition
Nom	Taux d'intérêt déclaré pour l'atelier
Numérateur	Nombre de réponses « Oui »
Dénominateur	Nombre total de réponses valides à la question
Unité	Pourcentage
Amélioration	Répéter la collecte sur plusieurs jours et créneaux, puis comparer les taux par support

Calculer le taux séparément pour chaque support permet de comparer les groupes. À chaque collecte, conserver la même question et le même mode de sollicitation, puis relever le support déclaré. Interroger des usagers sur plusieurs jours et créneaux réduira la dépendance à une seule période et rendra les comparaisons par support plus informatives. Les réponses incomplètes à la question sont exclues du dénominateur.

Document réponse DR 2 complété

Élément à définir	Proposition du candidat
Nom de l'indicateur	Taux d'intérêt déclaré
Numérateur	Réponses « Oui »
Dénominateur	Réponses valides
Unité	Pourcentage (%)
Amélioration de la collecte	Plusieurs jours et créneaux; comparaison par support

✓ RÉPONSE À RETENIR

Suivre le taux de réponses « Oui » parmi les réponses valides, en %. Répéter l'enquête sur plusieurs jours et créneaux, en comparant les résultats par support.

SYNTHÈSE DE L'ÉPREUVE & RÉFÉRENTIEL • LICENCE LETTRES

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

Bilan de l'épreuve : L'épreuve C5 du BTS Lettres, session 2026, porte le code 26-LIC-LETT-EP1 et dure 2 heures ; le coefficient est indiqué à titre indicatif comme 3,0. À partir de données constituées pour le sujet sur les usages d'une bibliothèque universitaire, elle évalue le contrôle des données, les statistiques descriptives, l'analyse de proportions, l'interprétation prudente et la proposition d'un suivi.

Notions-clés, formules & règles à maîtriser

- Une variable qualitative nominale décrit des catégories sans ordre ; une variable quantitative exprime une mesure numérique.
- La moyenne est la somme des valeurs divisée par l'effectif ; la médiane partage une série ordonnée en deux groupes.
- Pour dix valeurs, Q1 est la valeur de rang 3 et Q3 celle de rang 8 ; l'écart interquartile vaut Q3 - Q1.
- Une proportion conditionnelle est l'effectif recherché divisé par l'effectif du groupe, puis multiplié par 100.

Points de vigilance & attentes du jury

- Vérifier les identifiants uniques, les durées numériques strictement positives et les modalités autorisées avant tout calcul.
- Montrer les étapes de calcul, conserver les unités et respecter les arrondis demandés.
- Calculer les pourcentages sur l'effectif total du groupe concerné, et non sur l'effectif global.
- Distinguer constat et interprétation : une association observée ne démontre pas une causalité.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment déterminer la médiane et les quartiles de cette série de dix durées ?

La médiane est la moyenne des valeurs de rangs 5 et 6. Selon la convention fournie, Q1 est la valeur de rang 3 et Q3 celle de rang 8 ; l'écart interquartile est égal à Q3 moins Q1.

Q : Comment calculer une proportion de réponses positives pour un groupe de support ?

Divisez le nombre de réponses « Oui » dans le groupe par l'effectif total de ce groupe, puis multipliez par 100. Arrondissez le résultat à l'unité, comme demandé dans le document réponse.

Q : Comment organiser les deux heures et gérer le barème de l'épreuve ?

L'épreuve dure deux heures. Les quatre parties valent respectivement 5, 6, 5 et 4 points, soit 20 points au total. Répartissez le temps entre elles et gardez un moment pour vérifier calculs, unités et réponses.

Q : Quels critères de méthode faut-il respecter dans les réponses ?

Classez correctement les variables, appliquez les conventions statistiques du sujet, explicitez les calculs et leurs unités, puis formulez des conclusions proportionnées aux données. Ne présentez pas une association observée comme une preuve de causalité.

Sujet officiel associé (PDF) : https://licencelettres.fr/assets/annales/licence-lettres/C5 - Exploitation de données à des fins d'analyse/2026/Licence Lettres - C5 - Exploitation de données à des fins d'analyse - Session 2026 - 1-Sujet.pdf
Ressources Licence Lettres : 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.