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 SCIENCE POLITIQUE • CORRIGÉ DÉTAILLÉ

BTS Science politique

Session 2026 • Code sujet : **26-LIC-SCIE-EP1** • Durée : **2 h 30** • Coefficient : **3,0**
indicatif

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Décrire et contrôler les données territoriales** (3 questions) — **Page 3**
2. **Mesurer une association et ajuster une droite de régression** (3 questions) — **Page 5**
3. **Estimer une proportion et tenir compte de la structure de l'échantillon** (3 questions) — **Page 7**
4. **Restituer les résultats et protéger les données** (3 questions) — **Page 9**
5. **Ce qu'il faut retenir & Questions fréquentes (FAQ)** (Synthèse & Q&A Examen) — **Page 11**

Sujet officiel (PDF) : https://licencesciencepo.fr/assets/annales/licence-science-po/C2 - Exploiter des données à des fins d'analyse/2026/Licence Science Politique - C2 - Exploiter des données à des fins d'analyse - Session 2026 - 1-Sujet.pdf

1. Décrire et contrôler les données territoriales

L'Observatoire métropolitain de la participation citoyenne souhaite comparer la participation électorale et la confiance institutionnelle dans quatre secteurs fictifs. Les données, agrégées par secteur, comprennent les effectifs d'inscrits, de votants et de répondants ainsi que les moyennes de confiance.

Calcul des taux de participation

Question 1.1. À partir du DT 1, calculez le taux de participation pour chacun des quatre secteurs en divisant le nombre de votants par le nombre d'inscrits. Présentez les résultats en pourcentage avec un chiffre après la virgule dans le DR 1.

Le taux de participation est le rapport entre les votants et les inscrits, multiplié par 100. Les calculs ci-dessous reprennent les effectifs du DT 1 et les résultats sont arrondis à un chiffre après la virgule. La comparaison porte sur la part des inscrits ayant voté dans chaque secteur. Centre et Est ont le taux le plus élevé, tandis que le Sud a le plus faible. Ces données décrivent la participation, sans en expliquer les raisons.

Secteur	Calcul	Taux
Nord	620 ÷ 1 000 × 100	62,0 %
Centre	520 ÷ 800 × 100	65,0 %
Sud	385 ÷ 700 × 100	55,0 %
Est	585 ÷ 900 × 100	65,0 %

DR 1 complété : taux de participation

Secteur	Calcul	Taux de participation
Nord	620 ÷ 1 000 × 100	62,0 %
Centre	520 ÷ 800 × 100	65,0 %
Sud	385 ÷ 700 × 100	55,0 %
Est	585 ÷ 900 × 100	65,0 %

✓ **RÉPONSE À RETENIR**

Nord : 62,0 % ; Centre : 65,0 % ; Sud : 55,0 % ; Est : 65,0 %.

Moyenne pondérée

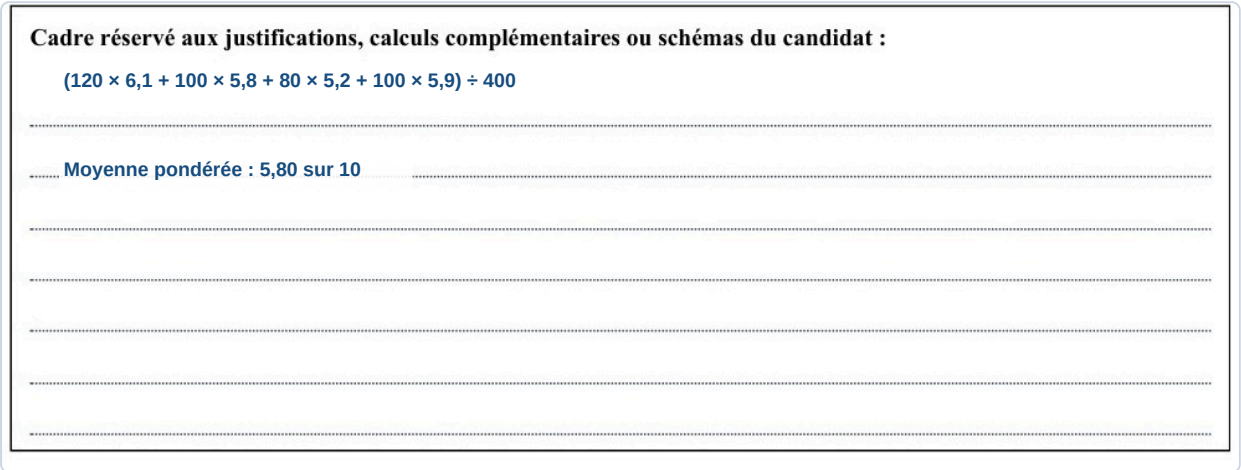
Question 1.2. Calculez la moyenne de confiance pondérée par le nombre de répondants. Écrivez le calcul et donnez le résultat sur 10, arrondi à deux décimales.

La moyenne pondérée donne à chaque secteur un poids correspondant à son nombre de répondants, plutôt que de traiter les quatre moyennes de façon égale. On additionne les produits de chaque effectif par sa moyenne, puis on divise par le total de 400 répondants :

$(120 \times 6,1 + 100 \times 5,8 + 80 \times 5,2 + 100 \times 5,9) \div 400 = 5,795$, soit 5,80 après arrondi à deux décimales.

Ce résultat décrit les répondants au questionnaire, pas nécessairement l'ensemble des habitants.

Calcul de la moyenne de confiance pondérée



✓ RÉPONSE À RETENIR

Moyenne de confiance pondérée : 5,80 sur 10.

Interprétation des données agrégées

Question 1.3. À partir des données disponibles, examinez l'affirmation selon laquelle les habitants du secteur Nord font davantage confiance aux institutions que ceux du secteur Sud. Indiquez ce que l'on peut effectivement comparer et formulez une limite liée au niveau d'agrégation des données.

Les données permettent de comparer les moyennes de confiance déclarées par les répondants des deux secteurs : 6,1 sur 10 au Nord contre 5,2 sur 10 au Sud, soit 0,9 point d'écart. Elles ne démontrent pas que chaque habitant du Nord fait davantage confiance que chaque habitant du Sud : les résultats sont agrégés et ne donnent ni les réponses individuelles ni leur distribution. La différence observée porte donc uniquement sur des moyennes de groupe.

✓ RÉPONSE À RETENIR

La moyenne des répondants est supérieure de 0,9 point au Nord. Les données agrégées ne permettent pas de conclure pour chaque habitant.

2. Mesurer une association et ajuster une droite de régression

Le DT 2 présente cinq observations communales fictives sur le nombre moyen de réunions publiques suivies par répondant et le score moyen de confiance institutionnelle. Les observations et les sommes statistiques servent à étudier une corrélation et une régression linéaire simple, sans constituer un échantillon représentatif de la population.

Corrélation de Pearson

Question 2.1. À l'aide de la formule fournie et des sommes du DT 2, calculez le coefficient de corrélation de Pearson entre x et y. Donnez le résultat à deux décimales et indiquez le sens ainsi que l'intensité de l'association observée.

Avec les sommes du DT 2, le numérateur vaut $5 \times 61 - 15 \times 18 = 35$. Les deux facteurs du dénominateur valent $5 \times 55 - 15^2 = 50$ et $5 \times 70 - 18^2 = 26$. Ainsi, $r = 35/\sqrt{50 \times 26} \approx 0,97$. L'association est positive et très forte parmi ces cinq communes fictives. Ce constat décrit une association linéaire entre les variables, mais ne démontre pas que la fréquentation des réunions cause une hausse de la confiance. Les communes ne sont pas représentatives de l'ensemble de la population.

> ✓ **RÉPONSE À RETENIR**
>
> $r = 0,97$: association linéaire positive très forte dans les cinq communes, sans preuve de causalité.

Régression linéaire

Question 2.2. Calculez la pente a et l'ordonnée à l'origine b de la droite de régression de y sur x à l'aide des formules fournies. Écrivez l'équation estimée ŷ = ax + b en arrondissant les coefficients à deux décimales.

La pente se calcule avec le numérateur et le premier facteur du dénominateur obtenus à la question 2.1. L'ordonnée à l'origine se déduit ensuite de la moyenne de y et de la pente.

Coefficient	Calcul	Résultat
Pente a	$(5 \times 61 - 15 \times 18)/(5 \times 55 - 15^2) = 35/50$	$0,70$
Ordonnée b	$(18 - 0,70 \times 15)/5$	$1,50$

La pente indique une hausse estimée de 0,70 point de confiance par réunion moyenne supplémentaire suivie. L'équation de la droite de régression, avec les coefficients arrondis à deux décimales, est $\hat{y} = 0,70x + 1,50$.

Régression linéaire complétée sur le DR 1

Régression linéaire de y sur x

Pente a	Ordonnée à l'origine b	Équation estimée
0,70	1,50	ŷ = 0,70x + 1,50

✓ RÉPONSE À RETENIR

$a = 0,70$, $b = 1,50$; $\hat{y} = 0,70x + 1,50$.

Estimation et interprétation

Question 2.3. À l'aide de l'équation calculée en 2.2, estimez le score de confiance pour x = 3. Expliquez en une phrase pourquoi ce résultat ne permet pas, à lui seul, d'affirmer que la participation aux réunions cause une hausse de la confiance.

Pour $x = 3$, on remplace x dans l'équation estimée : $\hat{y} = 0,70 \times 3 + 1,50 = 3,60$. Le score estimé est donc de 3,60 sur 10. Cette valeur est une estimation issue de la droite ajustée aux cinq communes observées. La corrélation et la régression décrivent une association, mais ne prouvent pas que suivre davantage de réunions augmente la confiance : d'autres facteurs peuvent intervenir, et les données ne permettent pas d'isoler un effet causal.

✓ RÉPONSE À RETENIR

Pour $x = 3$, score estimé : $3,60/10$; l'association n'établit pas de causalité.

3. Estimer une proportion et tenir compte de la structure de l’échantillon

Un questionnaire fictif sur le soutien à une proposition de consultation locale compte 400 répondants et présente une proportion brute favorable de 0,56. Des données par âge indiquent les parts de population, les parts dans l’échantillon et les proportions favorables nécessaires au calcul d’une estimation pondérée.

Erreur-type et intervalle de confiance

Question 3.1. Calculez l’erreur-type de la proportion observée, puis l’intervalle de confiance approximatif à 95 % avec la formule fournie. Donnez les bornes en proportion et en pourcentage, avec les pourcentages arrondis à un chiffre après la virgule, et précisez une hypothèse nécessaire à l’interprétation de cet intervalle.

Avec la proportion observée et les 400 répondants, l’erreur-type vaut $SE = \sqrt{0,56 \times 0,44/400} = 0,02482$. L’intervalle à 95 % se calcule par $0,56 \pm 1,96 \times 0,02482$, soit environ $[0,51135\ ;\ 0,60865]$. En pourcentage, les bornes arrondies à un chiffre après la virgule sont 51,1 % et 60,9 %. Son interprétation suppose un échantillonnage aléatoire simple et l’approximation normale indiquée dans le DT 3, sans correction de population finie.

Valeurs de l’intervalle de confiance dans le DR 2

Calcul ou résultat
0,02482
0,51135 (51,1 %)
0,60865 (60,9 %)

✓ RÉPONSE À RETENIR

SE : 0,02482 ; intervalle à 95 % : [0,51135 ; 0,60865], soit [51,1 % ; 60,9 %].

Estimation pondérée

Question 3.2. À partir du tableau, calculez la proportion favorable pondérée par les parts des groupes dans la population. Comparez-la à la proportion brute calculée avec les parts de l'échantillon et donnez l'écart en points de pourcentage.

Calcul	Résultat
$0,60 \times 0,60 + 0,40 \times 0,45$	0,54
$0,50 \times 0,60 + 0,50 \times 0,45$	0,525

La première ligne applique les parts de population, tandis que la seconde utilise les parts de l'échantillon brut. L'écart pondéré moins brut est $0,54 - 0,525 = 0,015$, soit 1,5 point de pourcentage. Dans ces données, la prise en compte de la structure par âge augmente donc l'estimation du soutien de 52,5 % à 54 %.

Estimations pondérée et brute dans le DR 2

Proportion favorable pondérée	0,54 (54 %)
Proportion favorable brute	0,525 (52,5 %)
Écart pondérée moins brute, en points de pourcentage	+1,5 point

✓ RÉPONSE À RETENIR

Proportion pondérée : 54 % ; proportion brute : 52,5 % ; écart : +1,5 point de pourcentage.

Limites de la pondération

Question 3.3. La pondération garantit-elle que l'estimation est exempte de tout biais ? Répondez en une ou deux phrases en mentionnant une limite possible.

Non. La pondération corrige ici le déséquilibre entre les deux groupes d'âge, mais elle ne garantit pas une estimation sans biais. Par exemple, au sein d'un même groupe, les personnes qui répondent peuvent différer des non-répondants dans leur soutien à la proposition. Une pondération fondée uniquement sur l'âge ne corrige pas cette différence, ni d'éventuels biais liés à d'autres caractéristiques non prises en compte.

✓ RÉPONSE À RETENIR

La pondération ne garantit pas l'absence de biais, notamment si répondants et non-répondants diffèrent au sein des groupes d'âge.

4. Restituer les résultats et protéger les données

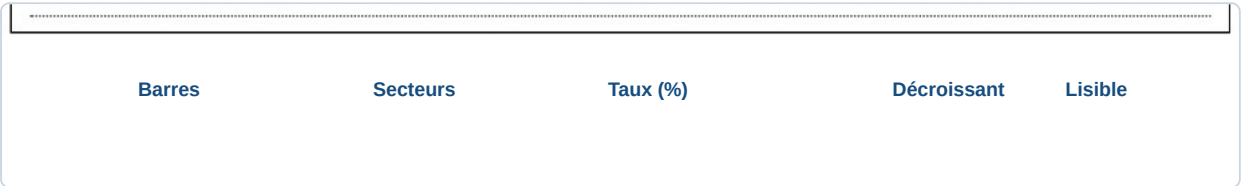
Les résultats doivent être présentés à des agents non spécialistes et aux élus sous une forme graphique et synthétique, sans transformer une association statistique en relation causale. Les règles de l'Observatoire encadrent la préparation, la diffusion et la conservation des données.

Représentation graphique

Question 4.1. Dans le DR 2, proposez une représentation graphique adaptée pour comparer les taux de participation calculés en 1.1. Indiquez le type de graphique, les axes, l'unité et un ordre de présentation pertinent, puis justifiez brièvement votre choix.

Choisir un diagramme en barres verticales : les secteurs figurent sur l'axe horizontal et le taux de participation, en pourcentage, sur l'axe vertical. Présenter les secteurs par taux décroissant, en conservant Centre et Est côte à côte ou en signalant leur égalité. Une échelle verticale de 0 à 100 % permet de comparer les taux sans exagérer les écarts. Chaque barre représente le taux calculé en 1.1 ; ce graphique rend les différences entre secteurs immédiatement lisibles.

DR 2 complété : projet de représentation graphique



✓ RÉPONSE À RETENIR

Diagramme en barres : Centre et Est, 65,0 % ; Nord, 62,0 % ; Sud, 55,0 %.

Synthèse destinée aux élus

Question 4.2. Rédigez une synthèse de deux phrases au maximum, destinée aux élus, qui rapporte un résultat descriptif issu de la partie 1 et le résultat de la partie 2 sans transformer une association en relation causale.

> Dans les quatre secteurs fictifs, le taux de participation observé varie de 55,0 % au Sud à 65,0 % au Centre et à l'Est ; ces taux décrivent les secteurs, sans renseigner à eux seuls sur les comportements individuels. Dans les cinq communes fictives, le nombre de réunions suivies et la confiance sont fortement et positivement associés (r = 0,97), mais ces données limitées n'établissent pas que les réunions causent une hausse de la confiance ni ne permettent de généraliser à toute la population.

✓ RÉPONSE À RETENIR

Participation observée : 55,0 % au Sud et 65,0 % au Centre et à l'Est. Association positive forte entre réunions suivies et confiance dans cinq communes (r = 0,97), sans preuve de causalité.

Protection des données

Question 4.3. Un tableau croisant un secteur et une catégorie d'âge contient une cellule avec trois personnes. Indiquez la mesure à prendre avant diffusion au regard des consignes du DT 3, puis citez une autre précaution concrète de protection des données applicable au fichier de travail.

La cellule de trois personnes ne peut pas être diffusée telle quelle, puisque l'effectif est inférieur au seuil de 5. Avant publication, il faut empêcher son identification directe ou par recoupement avec les totaux.

- Regrouper des catégories ou supprimer la cellule, avec suppression complémentaire si nécessaire.
- Dans le fichier de travail, retirer les identifiants directs, notamment les noms, adresses et coordonnées.
- Ne conserver que les variables nécessaires à l'analyse.

> **✓ RÉPONSE À RETENIR**
>
> Masquer ou regrouper la cellule de 3 personnes avant diffusion. Retirer les noms, adresses et coordonnées du fichier de travail.

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

Bilan de l'épreuve : Cette épreuve C2 de BTS Science politique, session 2026, porte le code 26-LIC-SCIE-EP1 et dure 2 h 30, avec un coefficient indicatif de 3,0. À partir de données fictives sur la participation citoyenne, la confiance institutionnelle et un questionnaire, elle évalue les indicateurs descriptifs, la corrélation, la régression, l'estimation pondérée, la restitution prudente des résultats et la protection des données.

Notions-clés, formules & règles à maîtriser

- Taux de participation = nombre de votants / nombre d'inscrits × 100.
- Moyenne pondérée = somme des produits effectif × moyenne / somme des effectifs.
- La corrélation de Pearson décrit le sens et l'intensité d'une association linéaire, sans établir une causalité.
- Proportion pondérée = somme des produits part de population du groupe × proportion favorable observée dans ce groupe.

Points de vigilance & attentes du jury

- Montrez les formules, les valeurs substituées et les arrondis demandés pour chaque calcul.
- Distinguez les résultats agrégés par territoire des caractéristiques ou comportements individuels.
- Présentez une association statistique sans la transformer en conclusion causale.
- Ne diffusez pas de cellule permettant d'isoler un effectif inférieur à 5 et retirez les identifiants directs du fichier d'analyse.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment calculer et présenter un taux de participation ?

Divisez le nombre de votants par le nombre d'inscrits, puis multipliez par 100. Pour cette épreuve, présentez les quatre résultats en pourcentage avec un chiffre après la virgule et indiquez clairement les valeurs utilisées.

Q : Une corrélation positive prouve-t-elle qu'une variable cause l'autre ?

Non. Une corrélation positive indique que les valeurs observées évoluent généralement dans le même sens. Elle ne démontre pas un lien de cause à effet, car d'autres facteurs peuvent intervenir et les données ne suffisent pas à établir une causalité.

Q : Comment organiser son temps pour cette épreuve ?

L'épreuve dure 2 h 30 et comporte quatre parties notées chacune sur 5 points. Répartissez le temps entre les calculs, les interprétations et la restitution, puis gardez une marge pour vérifier les résultats, les arrondis et les documents réponses.

Q : Quelles précautions de diffusion sont attendues pour les petits effectifs ?

Le DT 3 interdit de diffuser une cellule permettant d'isoler un effectif inférieur à 5. Supprimez ou regroupez les informations concernées avant publication, et appliquez aussi les règles de limitation des variables et de retrait des identifiants directs du fichier d'analyse.

Sujet officiel associé (PDF) : https://licencesciencepo.fr/assets/annales/licence-science-po/C2 - Exploiter des données à des fins d'analyse/2026/Licence Science Politique - C2 - Exploiter des données à des fins d'analyse - Session 2026 - 1-Sujet.pdf
Ressources Licence Science Politique : 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.