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

SCIENCE POLITIQUE

ÉPREUVE C2 : EXPLOITER DES DONNÉES À DES FINS D'ANALYSE

SESSION 2026

Durée : 2 h 30

SPÉCIALITÉ / UNITÉ	COEFF
SCIENCE POLITIQUE	3,0 INDICATIF

Moyens de calculs et matériel autorisés :

L'usage de calculatrice avec mode examen actif est autorisé. L'usage de 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 9 pages numérotées de 1/9 à 9/9.

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

DOSSIER SUJET

MISE EN SITUATION PROFESSIONNELLE ET CONSIGNES GÉNÉRALES : OBSERVATOIRE MÉTROPOLITAIN DE LA PARTICIPATION CITOYENNE, LYON

L'Observatoire métropolitain de la participation citoyenne prépare une note destinée aux élus et aux services de la Métropole. Il vous est demandé d'exploiter un jeu de données agrégé sur la participation électorale et la confiance institutionnelle, d'examiner une association statistique, puis de formuler des résultats prudents et compréhensibles.

Les données des DT 1 et DT 2 sont fictives et ont été constituées pour cette épreuve. Elles ne décrivent aucun territoire ni scrutin réel. Les résultats doivent être interprétés à partir des seules informations fournies. Toute conclusion causale doit être distinguée d'un constat de corrélation.

Répondez dans les espaces des documents réponses ou sur votre copie en indiquant clairement le numéro de chaque question. Les calculs doivent faire apparaître la formule utilisée, les valeurs substituées et le résultat arrondi au niveau demandé. Le sujet comporte quatre parties notées chacune sur 5 points, soit 20 points au total.

Pour les calculs, utiliser les formules suivantes si nécessaire : taux de participation = nombre de votants / nombre d'inscrits × 100 ; moyenne pondérée = somme des produits effectif × moyenne / somme des effectifs ; coefficient de corrélation de Pearson $r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$; erreur-type d'une proportion $SE = \sqrt{\frac{p(1-p)}{n}}$. Sauf indication contraire, arrondir les pourcentages à un chiffre après la virgule et les autres résultats à deux chiffres après la virgule.

PARTIE 1 : Décrire et contrôler les données (5 points)

Calcul d'indicateurs descriptifs et interprétation des unités d'analyse

À partir du DT 1, calculez les indicateurs demandés et explicitez ce qu'ils permettent, ou ne permettent pas, d'affirmer.

Indicateurs territoriaux fournis dans le DT 1

Secteur	Inscrits	Votants	Répondants au questionnaire	Confiance moyenne sur 10
Nord	1 000	620	120	6,1
Centre	800	520	100	5,8
Sud	700	385	80	5,2
Est	900	585	100	5,9

Données agrégées par secteur. Les inscrits et les votants concernent le scrutin étudié. La confiance est une moyenne sur une échelle de 0 à 10, calculée auprès des répondants au questionnaire.

1.1 (2 points) Calculez le taux de participation pour chacun des quatre secteurs. Présentez les résultats en pourcentage avec un chiffre après la virgule dans le DR 1.

1.2 (1,5 point) Calculez la moyenne de confiance pondérée par le nombre de répondants. Écrivez le calcul et donnez le résultat sur 10, arrondi à deux décimales.

1.3 (1,5 point) Un analyste affirme : « Les habitants du secteur Nord font davantage confiance aux institutions que ceux du secteur Sud. » À partir des données disponibles, indiquez ce que l'on peut effectivement comparer et formulez une limite liée au niveau d'agrégation des données.

PARTIE 2 : Mesurer une association entre deux variables (5 points)

Calcul et interprétation d'une corrélation linéaire

Le DT 2 fournit les observations de cinq communes ainsi que les sommes nécessaires aux calculs. La variable x mesure le nombre moyen de réunions publiques suivies par répondant, et la variable y est le score moyen de confiance institutionnelle sur 10.

Observations communales et statistiques récapitulatives

Commune	Réunions suivies par répondant, x	Confiance moyenne sur 10, y
A	1	2
B	2	3
C	3	4
D	4	4
E	5	5

Données fictives du DT 2. Chaque ligne correspond à une commune.

Sommes utiles au calcul

n	Σx	Σy	Σx²	Σy²	Σxy
5	15	18	55	70	61

Valeurs calculées à partir des cinq observations du tableau précédent.

2.1 (2 points) Calculez le coefficient de corrélation de Pearson entre x et y, en utilisant la formule fournie dans le contexte. Donnez le résultat à deux décimales et indiquez le sens ainsi que l'intensité de l'association observée.

2.2 (1,5 point) Calculez la pente a et l'ordonnée à l'origine b de la droite de régression de y sur x, avec $a = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$ et $b = \frac{\sum y - a\sum x}{n}$. Écrivez l'équation estimée $\hat{y} = ax + b$, avec des coefficients arrondis à deux décimales.

2.3 (1,5 point) 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.

PARTIE 3 : Estimer une proportion et tenir compte de la structure de l'échantillon (5 points)

Intervalle de confiance approximatif et pondération

Le DT 3 présente les données d'un questionnaire portant sur le soutien à une proposition de consultation locale. Pour le calcul de l'intervalle, l'échantillon comprend 400 répondants et la proportion observée de réponses favorables est $p = 0,56$. Utilisez l'approximation à 95 % : $[p - 1,96 \times SE\ ;\ p + 1,96 \times SE]$, où $SE = \sqrt{p(1-p)/n}$.

Pondération selon deux groupes de population

Groupe	Part dans la population	Part dans l'échantillon brut	Proportion favorable observée
18 à 39 ans	0,60	0,50	0,60
40 ans et plus	0,40	0,50	0,45

Parts de la population connues par la Métropole et proportions favorables observées dans chaque groupe.

3.1 (2,5 points) Calculez l'erreur-type de la proportion observée, puis l'intervalle de confiance approximatif à 95 %. Donnez les bornes en proportion et en pourcentage, arrondies à un chiffre après la virgule pour les pourcentages. Précisez une hypothèse nécessaire à l'interprétation de cet intervalle.

3.2 (1,5 point) 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. Donnez l'écart en points de pourcentage.

3.3 (1 point) La pondération garantit-elle que l'estimation est exempte de tout biais ? Répondez en une ou deux phrases en mentionnant une limite possible.

PARTIE 4 : Restituer les résultats et protéger les données (5 points)

Choix d'une représentation, rédaction analytique et précautions de traitement

Vous préparez une synthèse courte à destination d'agents qui ne sont pas spécialistes des méthodes quantitatives. Le DT 3 rappelle les règles de diffusion applicables au jeu de données.

Règles de gestion et de diffusion du jeu de données

Situation	Consigne
Fichier de travail	Ne conserver que les variables nécessaires à l'analyse.
Identifiants directs	Ne pas intégrer les noms, adresses ou coordonnées dans le fichier d'analyse.
Diffusion des résultats	Ne pas diffuser de tableau permettant d'isoler un effectif inférieur à 5.
Fin du traitement	Supprimer les fichiers de travail individuels après validation de la note, selon les règles de conservation de l'Observatoire.

Consignes internes de l'Observatoire pour ce dossier, à appliquer conjointement aux règles applicables en matière de protection des données personnelles.

4.1 (2 points) 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. Justifiez brièvement votre choix.

4.2 (1,5 point) 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.

4.3 (1,5 point) 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.

DOSSIER TECHNIQUE ET RESSOURCES

Documents ressources nécessaires à la résolution du sujet

DT 1 : DONNÉES AGRÉGÉES DE PARTICIPATION ET DE CONFIANCE PAR SECTEUR

Les quatre secteurs sont des unités territoriales fictives. Le taux de participation se calcule en divisant le nombre de votants par le nombre d'inscrits. La confiance institutionnelle est mesurée par une question notée de 0, confiance nulle, à 10, confiance très élevée. Les moyennes de confiance sont calculées sur les répondants au questionnaire indiqués dans le tableau. Ces observations agrégées ne renseignent pas directement sur les caractéristiques ou les comportements de chaque personne.

Base territoriale agrégée

Secteur	Inscrits	Votants	Répondants	Confiance moyenne sur 10
Nord	1 000	620	120	6,1
Centre	800	520	100	5,8
Sud	700	385	80	5,2
Est	900	585	100	5,9
Total	3 400	2 110	400	À calculer

Source : jeu de données fictif conçu pour l'épreuve.

DT 2 : DONNÉES COMMUNALES POUR L'ÉTUDE D'UNE ASSOCIATION

Les cinq communes ont été retenues pour illustrer un calcul de corrélation et une régression linéaire simple. La variable x est le nombre moyen de réunions publiques suivies par répondant. La variable y est le score moyen de confiance institutionnelle sur 10. Les sommes du tableau récapitulatif sont exactes au regard des cinq observations fournies. Les données ne constituent pas un échantillon représentatif de l'ensemble de la population.

Observations et sommes statistiques

Commune	x	y
A	1	2
B	2	3
C	3	4
D	4	4
E	5	5

Source : jeu de données fictif conçu pour l'épreuve.

Sommes calculées sur les cinq communes

n	Σx	Σy	Σx²	Σy²	Σxy
5	15	18	55	70	61

Ces valeurs peuvent être utilisées directement dans les formules indiquées dans le sujet.

DT 3 : ÉCHANTILLONNAGE, PONDÉRATION ET RÈGLES DE DIFFUSION

Le questionnaire sur la proposition de consultation locale compte 400 répondants. La proportion brute de réponses favorables est de 0,56. Pour l'intervalle demandé, on adopte l'approximation normale à 95 % avec le quantile 1,96, sous l'hypothèse d'un échantillonnage aléatoire simple et en l'absence de correction de population finie.

Pour l'estimation pondérée, les parts de population par groupe sont connues. La proportion pondérée est la somme des produits de la part de population et de la proportion favorable observée dans le groupe. Les consignes de l'Observatoire demandent de ne pas diffuser de cellule contenant un effectif inférieur à 5. Les données individuelles doivent être limitées aux besoins du traitement, protégées contre les accès non autorisés et conservées uniquement pendant la durée nécessaire, conformément aux règles applicables de protection des données personnelles.

Données disponibles pour la pondération

Groupe	Part de population	Part de l'échantillon	Proportion favorable
18 à 39 ans	0,60	0,50	0,60
40 ans et plus	0,40	0,50	0,45

Les parts de population totalisent 1,00. Les parts de l'échantillon brut totalisent également 1,00.

Consignes de traitement et de diffusion

Étape	Règle
Préparation	Limiter les variables aux besoins de l'analyse.
Fichier d'analyse	Retirer les identifiants directs tels que noms, adresses et coordonnées.
Publication	Ne pas diffuser de cellule permettant d'isoler un effectif inférieur à 5.
Clôture	Supprimer les fichiers de travail individuels après validation de la note, selon les règles de conservation de l'Observatoire.

Consignes internes applicables au dossier de l'Observatoire.

NE RIEN ÉCRIRE DANS CETTE PARTIE

DR 1 : Document réponse à rendre avec la copie : questions 1.1, 1.2 et 2.2 — **À RENDRE AVEC LA COPIE**

Complétez les cases demandées. Faites apparaître les calculs sur votre copie si l'espace est insuffisant. Pour la question 1.1, arrondissez les taux à un chiffre après la virgule.

Calcul des taux de participation

Secteur	Calcul	Taux de participation
Nord	.....	.....
Centre	.....	.....
Sud	.....	.....
Est	.....	.....

Régression linéaire de y sur x

Pente a	Ordonnée à l'origine b	Équation estimée
.....	.....	.....

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 : questions 3.1, 3.2 et 4.1 — **À RENDRE AVEC LA COPIE**

Complétez le tableau de calcul. Pour la question 4.1, décrivez précisément la représentation graphique prévue ; aucun tracé à l'échelle n'est exigé.

Intervalle et estimation pondérée

Élément	Calcul ou résultat
Erreur-type SE	.....
Borne inférieure de l'intervalle à 95 %	.....
Borne supérieure de l'intervalle à 95 %	.....
Proportion favorable pondérée	.....
Proportion favorable brute	.....
Écart pondérée moins brute, en points de pourcentage	.....

Projet de représentation graphique des taux de participation

Type de graphique	Axe horizontal	Axe vertical et unité	Ordre de présentation	Justification
.....	.....	.....	.....	.....

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.