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

MATHÉMATIQUES ET INFORMATIQUE APPLIQUÉES AUX SCIENCES HUMAINES ET SOCIALES

ÉVALUATION DU BLOC RNCP40806BC02 : EXPLOITER DES DONNÉES À DES FINS D'ANALYSE

SESSION 2026

Durée : 2 heures

SPÉCIALITÉ / UNITÉ	COEFF
MATHÉMATIQUES ET INFORMATIQUE APPLIQUÉES AUX SCIENCES HUMAINES ET SOCIALES	3 (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 10 pages numérotées de 1/10 à 10/10.

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

DOSSIER SUJET

MISE EN SITUATION D'ANALYSE DE DONNÉES ET CONSIGNES GÉNÉRALES : OBSERVATOIRE DES SERVICES ÉTUDIANTS DE L'UNIVERSITÉ DES RIVES, SITE DE VALMONT

L'Observatoire des services étudiants étudie les demandes reçues par le service d'assistance numérique. L'équipe souhaite décrire les délais de traitement et les niveaux de satisfaction observés selon le canal de contact, puis formuler des recommandations prudentes pour améliorer le service.

Le jeu de données fourni dans le dossier est constitué pour cette étude et ne contient aucune donnée directement identifiante. Il comprend un extrait brut comportant des anomalies et un sous-ensemble nettoyé explicitement identifié. Les réponses doivent distinguer les constats calculés, les limites des données et les interprétations possibles. Les calculs peuvent être effectués à la main ou à l'aide d'une calculatrice. Les résultats numériques seront arrondis au dixième lorsque cela est nécessaire.

Le sujet évalue la sélection critique de ressources, l'analyse et la synthèse de données ainsi que la construction d'une argumentation raisonnée. Le barème total est de 20 points.

Partie 1 : Cadrage de l'analyse et qualité des ressources (4 points)

Définir une question d'analyse, sélectionner des ressources et diagnostiquer les données

Vous intervenez comme analyste auprès de l'Observatoire. La question de travail est la suivante : les demandes reçues par le canal « formulaire » présentent-elles des délais et des niveaux de satisfaction différents de ceux des demandes reçues par le canal « guichet » ? Le dossier technique fournit les informations disponibles sur la collecte, les variables et les ressources documentaires.

Ressources proposées pour documenter l'étude

Référence	Description	Limite à prendre en compte
R1	Dictionnaire des variables et protocole de collecte de l'Observatoire, version datée et identifiée	Décrit uniquement les données de l'étude
R2	Article de recherche évalué par les pairs sur la mesure de la satisfaction dans les services publics, avec méthode et échantillon explicités	Population étudiée différente de celle de l'Université des Rives
R3	Publication institutionnelle nationale présentant des données agrégées sur les usages des services étudiants	Indicateurs et période de collecte différents
R4	Message publié sur un forum sans auteur identifié, affirmant qu'un canal de contact est toujours plus efficace	Méthode, données et source non vérifiables

Les références et descriptions ci-dessous sont fournies pour les besoins du sujet.

1.1 (1,5 point) Reformulez la question de travail en précisant la population étudiée, les deux groupes à comparer et deux indicateurs pertinents. Indiquez en une phrase pourquoi cette comparaison ne permet pas, à elle seule, d'établir un effet causal du canal de contact.

1.2 (1,5 point) Sélectionnez deux ressources parmi R1 à R4 qui seraient les plus utiles pour conduire l'étude. Pour chacune, justifiez sa pertinence et indiquez une limite à signaler. Écartez explicitement une ressource dont la fiabilité ou la pertinence est insuffisante.

1.3 (1 point) À partir du DT 2, relevez les deux anomalies présentes dans l'extrait brut et indiquez le traitement à appliquer avant l'analyse. Justifiez brièvement chaque traitement.

Partie 2 : Exploitation statistique du jeu de données (7 points)

Nettoyer, décrire et comparer les données

Pour toutes les questions de cette partie, utilisez exclusivement le sous-ensemble nettoyé du DT 2. Les variables et leurs unités sont définies dans le DT 1. Présentez les formules ou les étapes de calcul utiles et donnez les résultats avec une précision adaptée.

Données nettoyées à analyser

ID	Canal	Délai de traitement (min)	Satisfaction (1 à 5)	Résolution au premier contact (1 = oui, 0 = non)
D01	Formulaire	12	4	1
D02	Formulaire	18	3	1
D03	Formulaire	25	4	1
D04	Formulaire	30	2	0
D05	Formulaire	15	5	1
D06	Guichet	20	4	1
D07	Guichet	28	3	1
D08	Guichet	35	2	0
D09	Guichet	22	4	1
D10	Guichet	50	1	0

Source : extrait préparé pour l'étude de cas. Chaque ligne correspond à une demande distincte.

2.1 (2 points) Calculez, séparément pour chaque canal, l'effectif, la moyenne et la médiane des délais de traitement. Pour une série ordonnée de cinq valeurs, la médiane est la troisième valeur. Détaillez les calculs des moyennes en utilisant $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

2.2 (2 points) Calculez, séparément pour chaque canal, la moyenne de satisfaction et le taux de résolution au premier contact. Le taux est défini par $T = \frac{\text{nombre de demandes résolues au premier contact}}{\text{nombre de demandes du groupe}} \times 100$. Inscrivez vos résultats dans le DR 1.

2.3 (1,5 point) Déterminez l'étendue des délais pour chaque canal, selon la formule $E = x_{\max} - x_{\min}$. Comparez les étendues et expliquez ce qu'elles indiquent, et ce qu'elles ne permettent pas de conclure, sur la dispersion des délais.

2.4 (1,5 point) Rédigez une synthèse de trois à cinq phrases comparant les deux canaux à partir d'au moins trois indicateurs calculés dans cette partie. Distinguez explicitement les résultats descriptifs d'une généralisation à l'ensemble des demandes du service.

Partie 3 : Restitution et argumentation à partir des résultats (5 points)

Choisir une représentation pertinente et argumenter avec esprit critique

L'équipe souhaite une restitution lisible pour une réunion de pilotage. Elle ne souhaite pas de conclusion qui dépasse les informations disponibles. Les résultats calculés dans la partie 2 et les limites précisées dans le dossier technique doivent guider votre proposition.

Repères de restitution

Élément à restituer	Valeur ou information disponible
Nombre de demandes	5 par canal
Unité du délai	Minute
Échelle de satisfaction	1 = très insatisfait ; 5 = très satisfait
Nature de la comparaison	Description d'un extrait de 10 demandes

Utilisez ces repères pour concevoir une comparaison compréhensible et fidèle aux données.

3.1 (2 points) Proposez une représentation graphique adaptée pour comparer les délais des deux canaux. Précisez le type de graphique, les variables portées sur les axes, les unités et au moins deux éléments nécessaires à une lecture correcte. Expliquez pourquoi un graphique circulaire serait peu adapté à cette comparaison.

3.2 (1,5 point) La responsable affirme : « Le formulaire provoque une baisse de satisfaction et doit être supprimé. » Évaluez cette affirmation à partir des résultats disponibles. Citez un résultat qui peut être décrit, deux limites de l'inférence et une donnée supplémentaire qui serait utile avant une décision.

3.3 (1,5 point) Rédigez une recommandation de deux à quatre phrases destinée à la réunion de pilotage. Elle doit s'appuyer sur les indicateurs calculés, formuler une réserve liée à la taille ou à la sélection de l'extrait et proposer une action de collecte ou d'analyse complémentaire.

Partie 4 : Protection des données et protocole d'analyse (4 points)

Définir des précautions de traitement et rendre l'analyse reproductible

Avant toute diffusion, l'Observatoire doit s'assurer que les données sont utilisées pour une finalité définie, que leur accès est maîtrisé et que les résultats ne permettent pas d'identifier une personne. Le DT 3 présente les consignes applicables à cette étude de cas.

Étapes minimales du protocole de traitement

Étape	Action à préciser
Avant le calcul	.....
Pendant le calcul	.....
Avant la diffusion	.....

Complétez les actions attendues pour produire une analyse vérifiable.

4.1 (1,5 point) Indiquez trois précautions concrètes à appliquer au traitement ou à la diffusion des données de demandes, en vous appuyant sur le DT 3. Pour chaque précaution, précisez le risque qu'elle vise à limiter.

4.2 (1,5 point) Complétez dans le DR 2 les actions à réaliser avant le calcul, pendant le calcul et avant la diffusion. Les actions doivent permettre de vérifier la qualité des données et de reproduire les résultats.

4.3 (1 point) Expliquez pourquoi les résultats de cet extrait de dix demandes ne peuvent pas être présentés sans réserve comme représentatifs de toutes les demandes de l'Université des Rives.

DOSSIER TECHNIQUE ET RESSOURCES

Documents ressources nécessaires à la résolution du sujet

DT 1 : DICTIONNAIRE DES VARIABLES ET PÉRIMÈTRE DE L'ÉTUDE

Unité d'observation : une demande adressée au service d'assistance numérique. Dans le tableau nettoyé, les dix demandes sont réparties en deux groupes de cinq selon le canal indiqué au moment de la prise en charge.

- **Canal :** variable qualitative nominale, avec les modalités « Formulaire » et « Guichet ».
- **Délai de traitement :** variable quantitative exprimée en minutes, mesurée entre l'enregistrement et la première réponse apportée.
- **Satisfaction :** variable ordinale déclarée sur une échelle de 1 à 5. Une valeur plus élevée correspond à une satisfaction déclarée plus importante.
- **Résolution au premier contact :** variable binaire codée 1 si la demande est résolue lors du premier échange, et 0 sinon.

Pour un groupe de valeurs $x_1, \ldots, x_n$, la moyenne arithmétique est $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. L'étendue est $E = x_{\max} - x_{\min}$. Pour une variable binaire codée 0 ou 1, la moyenne est égale à la proportion de valeurs égales à 1. Les comparaisons de cet extrait sont descriptives : le mode d'affectation au canal, la période et la représentativité de l'extrait ne sont pas documentés.

Typologie des variables et traitements adaptés

Variable	Type	Résumé ou traitement pertinent	Précaution d'interprétation
Canal	Qualitative nominale	Effectifs et comparaison par groupe	Ne constitue pas une mesure d'intensité
Délai de traitement	Quantitative	Moyenne, médiane, étendue et représentation de distribution	Vérifier les unités et les valeurs atypiques
Satisfaction	Ordinale	Distribution des réponses et moyenne à interpréter avec prudence	Les écarts entre niveaux ne sont pas nécessairement équivalents
Résolution au premier contact	Binaire	Effectif et proportion de réponses égales à 1	Indiquer le dénominateur et la taille du groupe

Repères méthodologiques fournis pour l'étude.

DT 2 : EXTRAIT BRUT, ANOMALIES REPÉRÉES ET SOUS-ENSEMBLE NETTOYÉ

L'extrait brut ci-dessous comporte deux anomalies intentionnelles : une ligne en doublon et une valeur de satisfaction hors de l'échelle définie dans le DT 1. Pour l'analyse statistique des questions 2.1 à 2.4, utiliser uniquement le tableau nettoyé de la partie 2, dans lequel ces deux anomalies ont été traitées. Aucune valeur manquante n'est présente dans le tableau nettoyé.

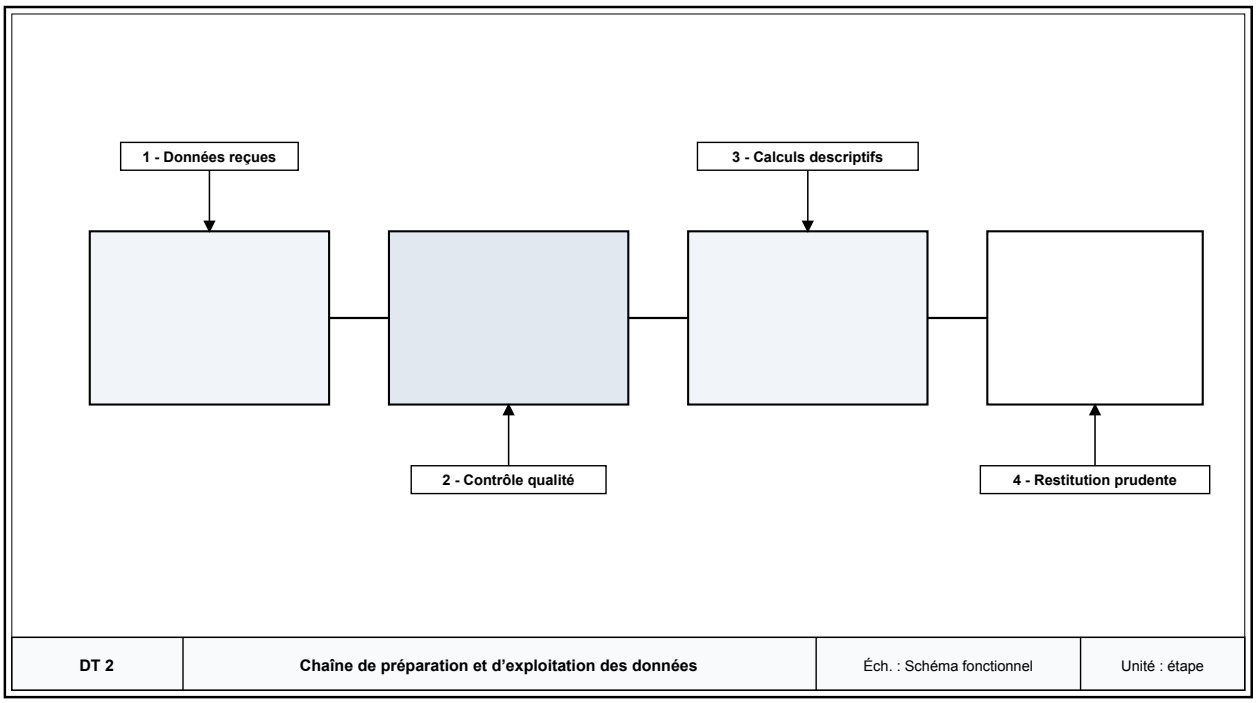
Règles de préparation documentées : vérifier l'unicité de l'identifiant de demande ; confronter les valeurs à la définition et aux bornes des variables ; ne pas remplacer une valeur invalide par une valeur supposée

sans justification ; conserver une trace des lignes écartées ou corrigées.

Extrait brut servant au diagnostic

ID	Canal	Délai (min)	Satisfaction	Résolution (1 ou 0)
D01	Formulaire	12	4	1
D01	Formulaire	12	4	1
D11	Guichet	24	6	1
D12	Guichet	31	3	0

Cet extrait de quatre lignes est présenté uniquement pour la question 1.3.



DT 3 : REPÈRES DE DOCUMENTATION, DE QUALITÉ ET DE PROTECTION DES DONNÉES

Les éléments suivants constituent les consignes documentaires de l'étude. Ils ne remplacent pas les procédures internes de l'établissement ni les obligations applicables en matière de protection des données.

- **Traçabilité :** consigner la source, la date d'extraction, le périmètre retenu, les règles de nettoyage et les éventuelles exclusions.
- **Qualité :** vérifier les identifiants, les types de variables, les unités, les bornes attendues et la cohérence des dénominateurs avant de calculer les indicateurs.
- **Accès et minimisation :** ne traiter que les variables nécessaires à la question d'analyse et limiter l'accès aux personnes autorisées selon les procédures de l'établissement.
- **Diffusion :** ne pas publier d'identifiants directs ; examiner le risque d'identification indirecte dans les petits groupes ; communiquer les effectifs et les limites de l'extrait.
- **Références :** apprécier l'autorité de l'auteur ou de l'organisme, la méthode, la date, la population étudiée et la proximité de la source avec la question. Une ressource externe ne remplace pas le dictionnaire ni le protocole de collecte des données étudiées.

Pour une proportion calculée sur un petit effectif, présenter le numérateur et le dénominateur avec le pourcentage. Une différence observée entre deux groupes ne démontre pas, à elle seule, que le canal est la cause de cette différence.

Éléments à conserver dans une trace d'analyse

Élément	Information à consigner
Jeu de données	Source, date d'extraction et périmètre
Préparation	Anomalies, exclusions et règles de traitement
Calculs	Formules, effectifs et dénominateurs utilisés
Restitution	Unités, limites et précautions de diffusion

Repères pour permettre la vérification des résultats.

NE RIEN ÉCRIRE DANS CETTE PARTIE

DR 1 : Document réponse à rendre avec la copie : questions 2.2 et 3.1 **À RENDRE AVEC LA COPIE**

Complétez le tableau avec les calculs demandés. Indiquez les étapes utiles et les unités. Pour les taux, mentionnez le numérateur et le dénominateur. La proposition de graphique peut être rédigée dans la dernière ligne.

Tableau de comparaison des deux canaux

Indicateur	Formulaire	Guichet
Effectif	.....	.....
Moyenne de satisfaction sur 5	.....	.....
Taux de résolution au premier contact : calcul	.....	.....
Taux de résolution au premier contact : résultat en %	.....	.....
Graphique proposé et axes	.....	.....

À compléter par le candidat à partir du tableau nettoyé.

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 : question 4.2 **À RENDRE AVEC LA COPIE**

Complétez chaque case par une action concrète adaptée à l'étude. Les actions doivent contribuer à la qualité, à la reproductibilité ou à la protection des données.

Protocole de traitement et de restitution à compléter

Étape	Action proposée par le candidat
Avant le calcul : contrôle du périmètre et des variables	.....
Pendant le calcul : traçabilité des méthodes et des résultats	.....
Avant la diffusion : confidentialité et interprétation	.....

Réponse attendue sous forme d'actions vérifiables.

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.