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

INFORMATIQUE

ÉVALUATION LOCALE C5 : EXPLOITATION DE DONNÉES À DES FINS D'ANALYSE

SESSION 2026

Durée : 3 heures

SPÉCIALITÉ / UNITÉ	COEFF
INFORMATIQUE	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 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 : ATELIER VÉLO MÉTROPOLE, SERVICE COMMERCE EN LIGNE, LYON

Le service commerce en ligne d'Atelier Vélo Métropole souhaite examiner les résultats d'une campagne de vente menée sur trois canaux : organique, courriel et publicité payante. Vous intervenez comme analyste de données. À partir des documents techniques fournis, vous devez contrôler les données, produire des indicateurs reproductibles, formuler une interprétation étayée et proposer des mesures de protection adaptées.

Les montants sont exprimés en euros. Les données du dossier sont fictives et les identifiants clients sont pseudonymisés. Pour les calculs, arrondir les taux à 0,01 point de pourcentage et les montants à 0,01 euro lorsque cela est nécessaire. Justifier les formules utilisées. Toute conclusion doit distinguer ce que les données permettent d'établir de ce qu'elles ne permettent pas d'affirmer.

Les modalités d'organisation, la durée et le coefficient sont ceux retenus pour cette évaluation locale de l'UE C5.

PARTIE 1 : Cadrage de l'analyse et qualité des données (5 points)

Définition du besoin, contrôle et sélection des données

Le responsable commercial demande : « Quel canal a généré les meilleures ventes et faut-il augmenter le budget publicitaire ? » Les documents DT 1 et DT 2 contiennent les données disponibles. Vous devez cadrer la question avant de produire une recommandation.

1.1 (1,5 point) Reformuler la demande sous la forme de deux questions d'analyse distinctes et définir, pour chacune, un indicateur calculable à partir des données fournies. Préciser le périmètre temporel et l'unité d'observation pertinente.

1.2 (2 points) À partir du DT 1, relever deux contrôles de qualité à effectuer avant agrégation. Pour chacun, indiquer le risque analytique associé et une règle de traitement reproductible. Distinguer explicitement les commandes payées, remboursées et annulées.

1.3 (1,5 point) Le responsable demande de relier les commandes aux clients pour expliquer les achats répétés. Indiquer une précaution à prendre concernant les identifiants pseudonymisés et une vérification nécessaire avant toute jointure. Préciser pourquoi un identifiant pseudonymisé ne constitue pas nécessairement une donnée anonyme.

PARTIE 2 : Calcul des indicateurs et interprétation (6 points)

Calculs professionnels, comparaison des canaux et limites d'interprétation

Utiliser les données du DT 1 et du DT 2. Dans cette épreuve, le chiffre d'affaires analysé correspond à la somme des montants des commandes au statut « payée ». Le taux de conversion conventionnel demandé est

le nombre de commandes payées divisé par le nombre de sessions du canal. Les commandes sont supposées attribuées au canal indiqué dans le fichier.

Données de fréquentation et dépenses par canal

Canal	Sessions	Dépense de campagne (€)
Organique	500	0
Courriel	200	40
Publicité payante	300	120

Source : DT 2. Une session est une visite enregistrée sur le canal pour la période étudiée.

2.1 (2 points) Calculer le chiffre d'affaires des commandes payées pour chacun des trois canaux et le chiffre d'affaires total payé. Présenter les étapes de calcul et exclure les commandes remboursées et annulées conformément à la définition fournie.

2.2 (2 points) Calculer le taux de conversion conventionnel de chaque canal à l'aide de la formule $taux\ de\ conversion = \frac{nombre\ de\ commandes\ payées}{nombre\ de\ sessions} \times 100$. Compléter le tableau du DR 1 et préciser pourquoi ces taux ne prouvent pas à eux seuls un effet causal du canal sur les ventes.

2.3 (1 point) Pour les canaux courriel et publicité payante, calculer le ratio $chiffre\ d'affaires\ payé / dépense\ de\ campagne$. Interpréter les deux résultats sans les qualifier de bénéfice ni de retour sur investissement net.

2.4 (1 point) Formuler une recommandation provisoire au responsable commercial, appuyée sur au moins deux résultats calculés. Citer une donnée supplémentaire nécessaire avant de décider d'augmenter le budget publicitaire.

PARTIE 3 : Requête et protocole de traitement reproductible (5 points)

Extraction, agrégation et vérification des résultats

Le DT 3 décrit le schéma relationnel utilisable dans un environnement SQLite. Les huit lignes du DT 1 sont chargées dans la table commandes. Les noms de tables et de colonnes respectent exactement les définitions du DT 3. Les fonctions d'agrégation SQL usuelles, notamment COUNT et SUM, sont disponibles.

3.1 (2 points) Écrire une requête SQL qui retourne, pour chaque canal, le nombre de commandes payées et la somme de leur montant. Les colonnes du résultat doivent être nommées canal, nb_commandes_payees et chiffre_affaires_paye. La requête doit exclure les autres statuts et regrouper correctement les lignes.

3.2 (1 point) Indiquer la sortie attendue de la requête de la question 3.1, en cohérence avec les données du DT 1. L'ordre des lignes n'est pas imposé.

3.3 (1 point) Décrire un protocole de traitement reproductible en quatre étapes, depuis le fichier source jusqu'au tableau d'indicateurs. Le protocole doit mentionner au moins un contrôle de qualité et une trace permettant de retrouver la version des données utilisée.

3.4 (1 point) Une extraction contient 8 lignes de commandes de 240 octets chacune, hors index. Calculer le volume de données correspondant en octets puis en kilooctets décimaux, avec $1\ kB = 1000\ octets$. Si une agrégation conservée ne représente que 2 lignes de 48 octets chacune, calculer le volume économisé. Indiquer une limite de cette estimation.

PARTIE 4 : Protection des données et communication des résultats (4 points)

Responsabilité, sécurité et restitution critique

Le DT 3 rappelle les éléments de contexte relatifs au traitement et à la sécurité. Les réponses doivent être adaptées au cas présenté et ne pas conclure à une conformité juridique complète sur la seule base des informations du dossier.

4.1 (1,5 point) Proposer deux mesures concrètes pour limiter l'exposition des données pendant l'analyse. Pour chacune, préciser le risque réduit. Appuyer la réponse sur les principes de minimisation et de sécurité mentionnés dans le DT 3.

4.2 (1 point) Le responsable souhaite recevoir le tableau détaillé contenant un identifiant client par commande. Indiquer une solution de restitution moins exposante qui réponde au besoin de comparaison des canaux, et justifier le niveau de détail retenu.

4.3 (1,5 point) Rédiger une synthèse de trois à cinq phrases destinée au responsable commercial. Elle doit présenter un résultat chiffré, une limite liée aux données disponibles et une prochaine vérification utile avant décision.

DOSSIER TECHNIQUE ET RESSOURCES

Documents ressources nécessaires à la résolution du sujet

DT 1 : EXTRACTION DE COMMANDES POUR LA PÉRIODE ÉTUDIÉE

Période : du 1er au 7 juin 2026 inclus. Chaque ligne représente une commande. Le champ montant est le montant de la commande en euros. Le statut décrit l'état enregistré au moment de l'extraction. Le champ client_id est un identifiant pseudonymisé. Le fichier ne contient ni coût des produits, ni marge, ni horodatage détaillé, ni information permettant de vérifier le parcours multi-canal.

Extrait complet de la table commandes

commande_id	date	canal	montant_eur	statut	client_id
O101	2026-06-01	Organique	120	payée	C01
O102	2026-06-02	Courriel	80	payée	C02
O103	2026-06-02	Publicité payante	150	payée	C03
O104	2026-06-03	Organique	60	remboursée	C04
O105	2026-06-04	Courriel	100	payée	C02
O106	2026-06-05	Publicité payante	90	payée	C05
O107	2026-06-06	Organique	110	payée	C06
O108	2026-06-07	Courriel	70	annulée	C07

Les huit lignes constituent l'intégralité du jeu de données utilisé pour l'épreuve.

DT 2 : FRÉQUENTATION ET DÉPENSES DE CAMPAGNE

Les valeurs ci-dessous concernent la même période que l'extraction du DT 1. Les sessions sont des visites enregistrées dans l'outil de mesure du site. Les dépenses correspondent aux montants déclarés par le service marketing pour la période. Le canal organique n'a pas de dépense de campagne directe enregistrée. Les conventions d'attribution et la déduplication des sessions n'ont pas été documentées dans l'extrait transmis.

Indicateurs de fréquentation et dépenses

Canal	Sessions	Dépense de campagne (€)
Organique	500	0
Courriel	200	40
Publicité payante	300	120

Source : extraction hebdomadaire du service marketing.

DT 3 : SCHÉMA DE TRAITEMENT, DICTIONNAIRE DE DONNÉES ET CONSIGNES DE PROTECTION

Schéma SQLite : commandes(commande_id TEXT, date TEXT, canal TEXT, montant_eur REAL, statut TEXT, client_id TEXT). La colonne commande_id identifie une commande dans l'extrait. Les valeurs possibles de statut observées sont payée, remboursée et annulée. Le champ client_id permet une récurrence d'identification au sein du jeu de données ; il n'est pas directement nominatif dans l'extrait, mais une réidentification peut rester possible si des informations supplémentaires sont accessibles.

Contexte de traitement fourni : l'accès à l'extraction est limité aux personnes habilitées ; les exports de travail doivent être conservés dans l'espace professionnel prévu à cet effet ; la politique interne indique une suppression des lignes détaillées après 90 jours et une conservation des indicateurs agrégés selon les besoins documentés. Ces éléments ne suffisent pas à eux seuls à établir la conformité complète du traitement.

Références de principes à mobiliser : règlement (UE) 2016/679, article 5, relatif notamment à la minimisation et à la limitation de la conservation ; article 25, relatif à la protection des données dès la conception et par défaut ; article 32, relatif à la sécurité du traitement. Le dossier ne fournit ni registre de traitement, ni analyse de risque, ni information sur la base juridique, l'information des personnes ou les éventuels sous-traitants.

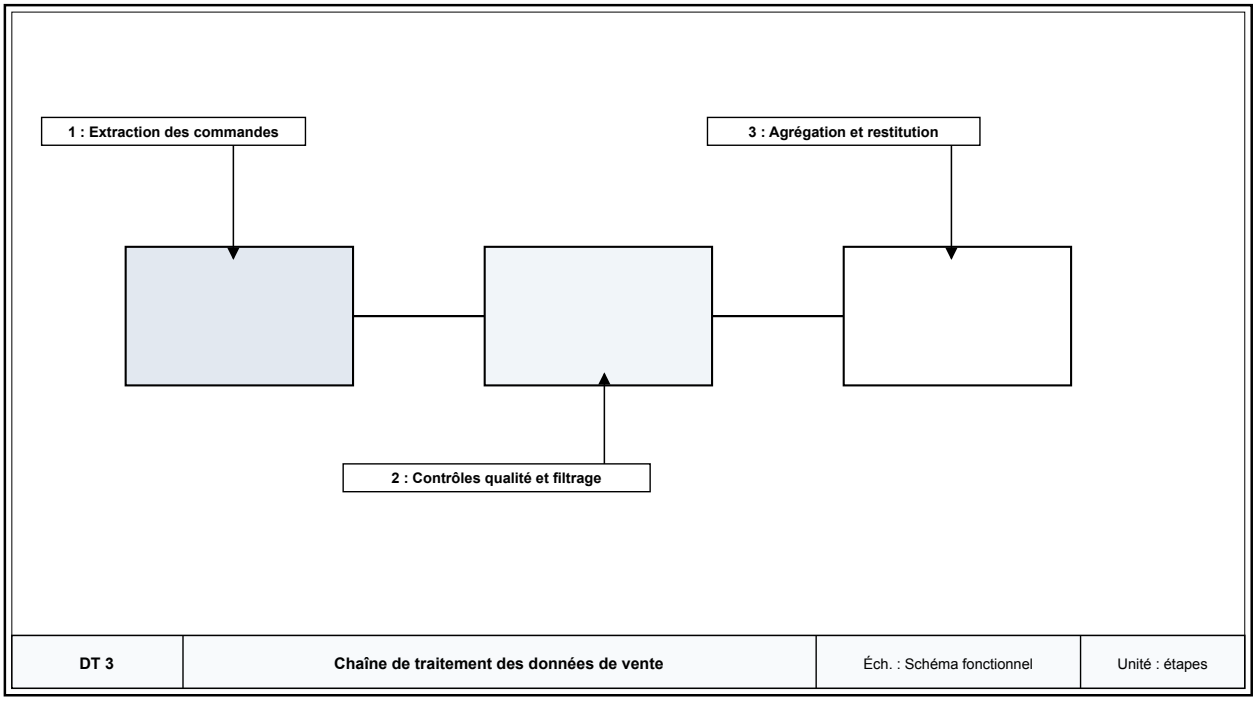


Schéma fonctionnel : contrôler les données avant agrégation et limiter les exports détaillés.

Dictionnaire des champs utiles

Champ	Type	Interprétation
commande_id	TEXT	Identifiant de commande
date	TEXT	Date au format AAAA-MM-JJ
canal	TEXT	Canal attribué à la commande
montant_eur	REAL	Montant de commande en euros
statut	TEXT	État enregistré de la commande
client_id	TEXT	Identifiant client pseudonymisé

Types et interprétations définis pour l'exercice.

NE RIEN ÉCRIRE DANS CETTE PARTIE

DR 1 : Document réponse : tableau des indicateurs par canal, questions 2.1 et 2.2

À RENDRE AVEC LA COPIE

Compléter les cases marquées « ... ». Indiquer les unités et les arrondis demandés dans l'énoncé.

Calcul des commandes payées, du chiffre d'affaires et du taux de conversion

Canal	Nombre de commandes payées	Chiffre d'affaires payé (€)	Sessions	Taux de conversion (%)
Organique	.....	.....	500	.....
Courriel	.....	.....	200	.....
Publicité payante	.....	.....	300	.....
Total ou valeur globale pertinente	.....	.....	1000	.....

À compléter par le candidat à partir des DT 1 et DT 2.

Cadre réservé aux justifications, calculs complémentaires ou schémas du candidat :

.....

.....

.....

.....

.....

.....

.....

NE RIEN ÉCRIRE DANS CETTE PARTIE

DR 2 : Document réponse : protocole de traitement reproductible, question 3.3

À RENDRE AVEC LA COPIE

Compléter chaque étape par une action concrète, un contrôle ou une trace adaptée au jeu de données étudié.

Étapes du traitement

Étape	Action et contrôle à réaliser	Trace ou résultat conservé
1 : Réception et identification de la source	.....	.....
2 : Contrôle et préparation	.....	.....
3 : Calcul et vérification des indicateurs	.....	.....
4 : Restitution et gestion de l'extraction	.....	.....

À compléter par le candidat. Une réponse structurée en quatre étapes est attendue.

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.