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 Info

LICENCE INFORMATIQUE : C5

Session 2026 • Code sujet : **2026-LICINFO-C5** • Durée : **3 heures** • Coefficient : **3, indicatif**

Sommaire

Cliquez sur une partie pour y accéder directement

1. Cadrage de l’analyse et qualité des données (3 questions) **Page 3**

2. Calcul des indicateurs et interprétation (4 questions) **Page 5**

3. Requête et protocole de traitement reproductible (4 questions) **Page 8**

4. Protection des données et communication des résultats (3 questions) **Page 10**

5. Ce qu’il faut retenir & Questions fréquentes (FAQ) (Synthèse & Q&A Examen) **Page 12**

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

1. Cadrage de l’analyse et qualité des données

Atelier Vélo Métropole veut comparer les ventes issues des canaux organique, courriel et publicité payante à partir d’une extraction de commandes et de données de fréquentation. Il faut cadrer les questions d’analyse, contrôler les données et encadrer l’usage des identifiants pseudonymisés.

Question 1.1. Reformuler la demande commerciale 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. Comparer les ventes : quel chiffre d’affaires payé chaque canal a-t-il généré ? Indicateur : somme des montants des commandes au statut « payée », par canal. L’unité d’observation est la commande, agrégée par canal.

2. Éclairer la décision budgétaire : quel chiffre d’affaires payé est associé à chaque euro dépensé en campagne ? Indicateur : chiffre d’affaires payé divisé par la dépense de campagne, pour les canaux payants. L’unité d’analyse est le canal, sur la période du 1er au 7 juin 2026 inclus. Les commandes sont attribuées au canal indiqué dans le fichier.

> ✓ **RÉPONSE À RETENIR**
>
> Comparer le chiffre d’affaires payé par canal et, pour les canaux payants, le chiffre d’affaires payé rapporté à la dépense, du 1er au 7 juin 2026.

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

Effectuer ces contrôles avant agrégation pour éviter des totaux erronés.

Contrôle	Risque analytique	Règle reproductible
Unicité de commande_id	Un doublon compterait deux fois une commande.	Repérer les doublons, les mettre en attente et les rapprocher de la source avant correction.
Validité des champs et des statuts	Une valeur absente ou invalide fausserait le classement ou le montant.	Vérifier les champs requis, les montants numériques et les statuts autorisés ; corriger à partir de la source ou exclure de l’agrégation en conservant une trace.

Inclure les commandes « payée » dans le nombre et le chiffre d’affaires payés. Exclure les commandes « remboursée » et « annulée » de ces indicateurs, sans les confondre avec les commandes payées.

> ✓ **RÉPONSE À RETENIR**
>
> Contrôler l’unicité des commandes et la validité des champs ; agréger uniquement les commandes payées, en excluant les remboursées et les annulées.

Question 1.3. Pour relier les commandes aux clients et expliquer les achats répétés, indiquer une précaution concernant les identifiants pseudonymisés et une vérification nécessaire avant toute jointure. Préciser pourquoi un identifiant pseudonymisé ne constitue pas nécessairement une donnée anonyme.

Limiter l'accès aux identifiants pseudonymisés aux personnes habilitées et ne les utiliser que si le suivi des achats répétés est nécessaire. Avant la jointure, vérifier la compatibilité du référentiel d'identifiants, la correspondance des clés et l'unicité de la clé côté table clients ; contrôler que la jointure ne multiplie pas les commandes.

Un identifiant pseudonymisé permet de relier plusieurs achats. Des informations supplémentaires accessibles peuvent permettre de retrouver l'identité du client : cet identifiant reste donc une donnée personnelle pseudonymisée, pas nécessairement anonyme.

✓ RÉPONSE À RETENIR

Vérifier la compatibilité et la cardinalité des identifiants avant jointure ; en limiter l'accès, car des recoupements peuvent réidentifier les clients.

2. Calcul des indicateurs et interprétation

L'analyse porte sur huit commandes enregistrées du 1er au 7 juin 2026 et sur les sessions et dépenses de campagne des trois canaux. Il faut calculer chiffre d'affaires, taux de conversion et ratios chiffre d'affaires sur dépenses, puis interpréter les résultats sans conclure à une causalité ou à un bénéfice net.

Question 2.1. Calculer le chiffre d'affaires des commandes payées de chacun des trois canaux ainsi que le chiffre d'affaires total payé. Présenter les étapes de calcul et exclure les commandes remboursées et annulées conformément à la définition fournie.

Le périmètre est la période du 1er au 7 juin 2026 inclus. Chaque ligne représente une commande. On additionne les montants uniquement lorsque le statut est « payée » : Organique, O101 et O107 ; Courriel, O102 et O105 ; Publicité payante, O103 et O106. La commande O104 est remboursée et O108 est annulée : toutes deux sont exclues du chiffre d'affaires payé.

Canal	Commandes payées	Calcul du chiffre d'affaires payé
Organique	2	120 + 110 = 230 €
Courriel	2	80 + 100 = 180 €
Publicité payante	2	150 + 90 = 240 €
Total	6	230 + 180 + 240 = 650 €

DR 1 complété : commandes payées et chiffres d'affaires

Canal	Nombre de commandes payées	Chiffre d'affaires payé (€)	Sessions	Taux de conversion (%)
Organique	2	230 €	500	
Courriel	2	180 €	200	
Publicité payante	2	240 €	300	
Total ou valeur globale pertinente	6	650 €	1000	

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

✓ **RÉPONSE À RETENIR**

Organique : 230 € ; courriel : 180 € ; publicité payante : 240 €. Total payé : 650 €.

Question 2.2. Calculer le taux de conversion conventionnel de chaque canal avec la formule : nombre de commandes payées divisé par nombre de sessions, puis multiplié par 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.

Pour chaque canal, le taux conventionnel est le nombre de commandes payées divisé par les sessions, puis multiplié par 100. Les données donnent deux commandes payées par canal. Les résultats sont arrondis à 0,01 point de pourcentage. Le taux global utilise les six commandes payées et les 1 000 sessions.

Canal	Commandes payées	Sessions	Taux de conversion
Organique	2	500	0,40 %
Courriel	2	200	1,00 %
Publicité payante	2	300	0,67 %
Total	6	1 000	0,60 %

Ces taux décrivent une association observée, pas un effet causal. L'attribution des commandes est supposée, tandis que la déduplication des sessions et les parcours multi-canaux ne sont pas vérifiés.

DR 1 complété : taux de conversion

Canal	Nombre de commandes payées	Chiffre d'affaires payé (€)	Sessions	Taux de conversion (%)
Organique	.....	.....	500	0,40 %
Courriel	.....	.....	200	1,00 %
Publicité payante	.....	.....	300	0,67 %
Total ou valeur globale pertinente	.....	.....	1000	0,60 %

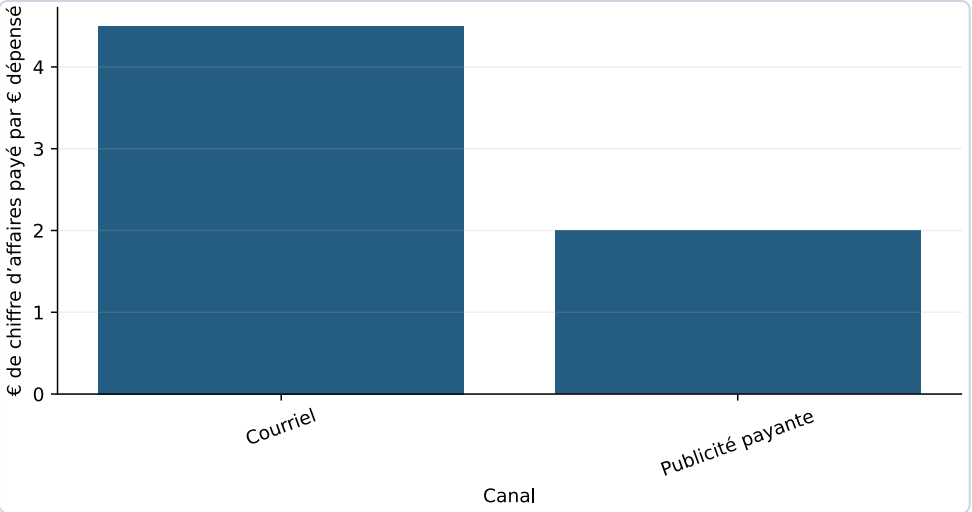
À compléter par le candidat à partir des DT 1 et DT 2.

✓ RÉPONSE À RETENIR

Organique : 0,40 % ; courriel : 1,00 % ; publicité payante : 0,67 % ; global : 0,60 %.

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

Le ratio demandé est le chiffre d'affaires payé divisé par la dépense déclarée pour le canal. Pour le courriel : 180 € / 40 € = 4,50. Pour la publicité payante : 240 € / 120 € = 2,00. Ainsi, chaque euro dépensé correspond à 4,50 € de chiffre d'affaires payé pour le courriel et à 2,00 € pour la publicité payante. Ces ratios ne tiennent pas compte du coût des produits ni des autres charges : ils ne mesurent donc ni un bénéfice ni un retour sur investissement net.



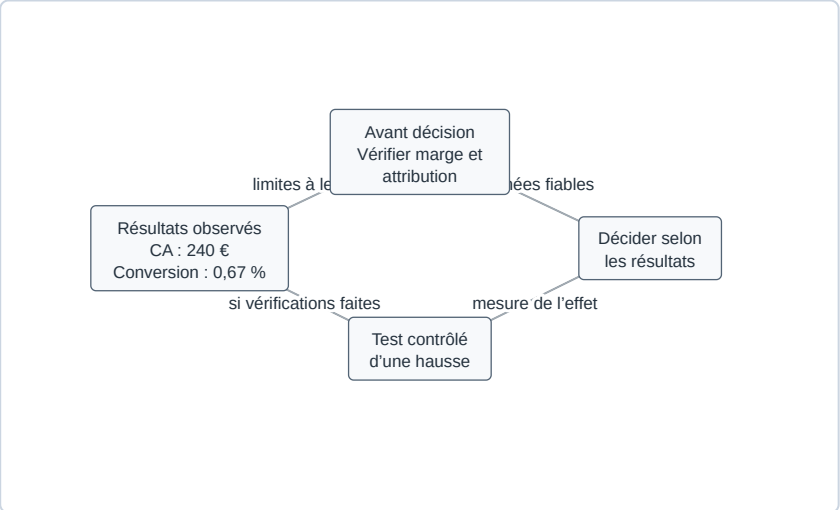
✓ RÉPONSE À RETENIR

Courriel : 4,50 € de chiffre d'affaires payé par euro dépensé ; publicité payante : 2,00 € par euro dépensé.

Question 2.4. 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.

La publicité payante a généré 240 € de chiffre d'affaires payé, avec une conversion de 0,67 % et un ratio de 2,00 € par euro dépensé. Le courriel affiche une conversion de 1,00 % et un ratio de 4,50. Ces résultats portent sur une seule semaine et ne justifient pas, à eux seuls, une hausse budgétaire. Il faut vérifier la marge ou le coût des produits ainsi que la fiabilité de l'attribution et des sessions. Une hausse pourra ensuite être évaluée par un test contrôlé.

Décision provisoire sur le budget publicitaire



Décider après vérification des données et mesure de l'effet d'une hausse.

✓ RÉPONSE À RETENIR

Ne pas augmenter immédiatement le budget publicitaire. Vérifier la marge et l'attribution, puis mesurer l'effet d'une hausse par un test contrôlé.

3. Requête et protocole de traitement reproductible

Les huit commandes du DT 1 sont chargées dans une table SQLite dont le schéma et les champs sont décrits dans le DT 3. Il faut produire une agrégation par canal, en vérifier le résultat, décrire un traitement traçable en quatre étapes et estimer un volume de données économisé.

Question 3.1. Écrire une requête SQL qui retourne, pour chaque canal, le nombre de commandes payées et la somme de leur montant. Nommer les colonnes canal, nb_commandes_payees et chiffre_affaires_paye, exclure les autres statuts et regrouper correctement les lignes.

La requête filtre les commandes avant l'agrégation : seules les lignes dont le statut est « payée » sont retenues. `COUNT(*)` compte les commandes conservées et `SUM(montant_eur)` additionne leurs montants. Le regroupement par `canal` produit une ligne par canal présent parmi ces commandes. Les alias donnent exactement les noms de colonnes demandés, conformément au schéma SQLite du DT 3.

✓ RÉPONSE À RETENIR

```
SELECT canal,
       COUNT(*) AS nb_commandes_payees,
       SUM(montant_eur) AS chiffre_affaires_paye
FROM commandes
WHERE statut = 'payée'
GROUP BY canal;
```

Question 3.2. 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.

On retient les six commandes payées du DT 1 et on exclut la commande remboursée O104 ainsi que la commande annulée O108. Les deux commandes payées de chaque canal sont additionnées : Organique, 120 et 110 ; Courriel, 80 et 100 ; Publicité payante, 150 et 90. Le tableau ci-dessous donne les effectifs et sommes attendus, sans imposer d'ordre aux lignes.

✓ RÉPONSE À RETENIR

Canal	Commandes payées	Chiffre d'affaires payé (€)
Organique	2	230
Courriel	2	180
Publicité payante	2	240

Question 3.3. Décrire un protocole de traitement reproductible en quatre étapes, depuis le fichier source jusqu'au tableau d'indicateurs. Mentionner au moins un contrôle de qualité et une trace permettant de retrouver la version des données utilisée, puis compléter le DR 2.

Le protocole suit les quatre étapes du DR 2, de la réception à la restitution. Il conserve les éléments nécessaires pour refaire les calculs et retrouver la version des données utilisée.

1. Réceptionner le fichier et relever sa source, sa période et le nombre de lignes.
2. Contrôler les champs requis, l'unicité de `commande_id` et les statuts ; préparer les données sans confondre commandes payées, remboursées et annulées.
3. Filtrer les commandes payées, agréger les montants et les effectifs par canal, puis vérifier les résultats.
4. Restituer le tableau agrégé dans l'espace professionnel prévu et supprimer les lignes détaillées après 90 jours.

Conserver le nom et la version du fichier, la date d'extraction, la requête SQL et les résultats des contrôles.

DR 2 complété : protocole en quatre étapes

Étape	Action et contrôle à réaliser	Trace ou résultat conservé
1 : Réception et identification de la source	Fichier, période, 8 lignes	Nom, date, version
2 : Contrôle et préparation	Champs, doublons, statuts	Contrôles consignés
3 : Calcul et vérification des indicateurs	Filtrer payées, agréger	Requête SQL, résultats vérifiés
4 : Restitution et gestion de l'extraction	Publier l'agrégat, supprimer le détail jours	Agrégat, date de suppression

✓ RÉPONSE À RETENIR

Source identifiée, données contrôlées, indicateurs vérifiés et tableau agrégé conservé avec les traces de traitement ; lignes détaillées supprimées après 90 jours.

Question 3.4. Calculer en octets puis en kilooctets décimaux le volume de huit lignes de commandes de 240 octets chacune, hors index, avec 1 kB = 1000 octets. Calculer le volume économisé si l'agrégation conservée comprend deux lignes de 48 octets chacune, puis indiquer une limite de cette estimation.

Le volume de l'extraction est le produit des huit lignes par 240 octets chacune. Pour l'agrégation, les deux lignes de 48 octets représentent le volume conservé. La différence entre ces volumes donne l'économie ; la conversion en kilooctets décimaux se fait en divisant les octets par 1 000. L'estimation exclut les index, les métadonnées et les coûts de stockage propres au système.

✓ RÉPONSE À RETENIR

Extraction : 1 920 octets, soit 1,92 kB. Agrégation : 96 octets. Volume économisé : 1 824 octets, soit 1,824 kB.

4. Protection des données et communication des résultats

Le DT 3 décrit les mesures internes de contrôle d'accès, de conservation et de protection des données, ainsi que les limites des informations disponibles pour conclure sur la conformité du traitement. Il faut proposer des mesures de réduction de l'exposition, une restitution adaptée et une synthèse commerciale nuancée.

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

Deux mesures réduisent l'exposition. **Minimisation :** pour les indicateurs, ne conserver que le canal, le montant et le statut, et écarter `client_id` ; cela réduit le risque de réidentification et d'accès inutile aux données. **Sécurité :** limiter l'accès aux personnes habilitées et stocker les exports dans l'espace professionnel prévu ; cela réduit le risque de divulgation ou d'accès non autorisé. Appliquer aussi la règle interne de suppression des lignes détaillées après 90 jours et conserver seulement les agrégats nécessaires. Ces mesures ne suffisent pas, à elles seules, à établir la conformité complète du traitement.

> ✓ **RÉPONSE À RETENIR**
>
> Limiter les champs aux données nécessaires, restreindre l'accès et supprimer les lignes détaillées après 90 jours.

Question 4.2. 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.

Transmettre un tableau agrégé par canal pour la période du 1er au 7 juin 2026, avec le nombre de commandes payées, le chiffre d'affaires payé, les sessions et le taux de conversion. Ce niveau de détail permet de comparer les canaux sans communiquer un identifiant client par commande. Ne pas joindre les lignes détaillées ni les `client_id` : leur pseudonymisation n'exclut pas une réidentification si d'autres informations sont accessibles. L'extraction détaillée reste réservée aux personnes habilitées qui en ont besoin.

> ✓ **RÉPONSE À RETENIR**
>
> Restituer les indicateurs agrégés par canal, sans identifiant client ni détail par commande.

Question 4.3. 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.

Du 1er au 7 juin, les commandes payées attribuées à la publicité représentent 240 € de chiffre d'affaires. Ce montant ne permet pas de conclure à la rentabilité, car les coûts des produits et les marges ne sont pas fournis. L'effet propre du canal n'est pas non plus établi : les règles d'attribution et la déduplication des sessions ne sont pas documentées. Avant d'augmenter le budget, vérifions ces règles et intégrons les coûts ou les marges pour évaluer la contribution réelle de la campagne.

✓ RÉPONSE À RETENIR

La publicité totalise 240 € de chiffre d’affaires payé, sans preuve de rentabilité ni d’effet causal. Vérifier l’attribution, la déduplication des sessions et les coûts ou marges avant toute hausse budgétaire.

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

Bilan de l'épreuve : L'évaluation locale C5 de la Licence informatique, session 2026, porte le code 2026-LICINFO-C5, dure trois heures et indique un coefficient 3 à titre indicatif. À partir d'une campagne fictive d'Atelier Vélo Métropole, elle évalue le contrôle et l'agrégation de commandes, le calcul d'indicateurs de vente, les requêtes SQL, la reproductibilité des traitements, l'interprétation prudente des résultats et la protection des données.

Notions-clés, formules & règles à maîtriser

- Le chiffre d'affaires étudié est la somme des montants des seules commandes au statut « payée ».
- Taux de conversion conventionnel = nombre de commandes payées / nombre de sessions du canal × 100.
- Le ratio chiffre d'affaires payé / dépense de campagne ne mesure ni le bénéfice ni le retour sur investissement net.
- Un identifiant pseudonymisé reste potentiellement réidentifiable si des informations supplémentaires sont accessibles.

Points de vigilance & attentes du jury

- Avant agrégation, contrôler les doublons, les champs manquants et la cohérence des montants et des statuts.
- Exclure explicitement les commandes remboursées et annulées du chiffre d'affaires payé.
- Distinguer association et causalité, car l'attribution des canaux et la déduplication des sessions ne sont pas documentées.
- Restituer des indicateurs agrégés utiles, limiter les données détaillées et conserver une trace de la version traitée.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment calculer le taux de conversion demandé dans ce sujet ?

Pour chaque canal, divisez le nombre de commandes au statut « payée » par le nombre de sessions correspondant, puis multipliez par 100. Présentez le résultat en pourcentage et arrondissez-le à 0,01 point de pourcentage.

Q : Pourquoi un identifiant pseudonymisé ne suffit-il pas à rendre les données anonymes ?

La pseudonymisation remplace les identifiants directs, mais ne supprime pas nécessairement la possibilité de relier les données à une personne. Des informations complémentaires accessibles peuvent permettre une réidentification, comme le rappelle le DT 3.

Q : Comment organiser le temps et les réponses dans cette épreuve ?

L'épreuve dure trois heures et totalise 20 points répartis entre quatre parties. Traitez les calculs et les documents réponses, puis gardez du temps pour vérifier les unités, les statuts retenus, les formules et la prudence des conclusions.

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

Le sujet demande des formules justifiées, des résultats cohérents avec les données, une distinction claire entre commandes payées, remboursées et annulées, et des conclusions qui reconnaissent les limites des données. Il ne permet pas de conclure à lui seul à une conformité juridique complète.

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