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

BUT SD

BUT Science des données : Valorisation d'une production

Session 2026 • Code sujet : **26-BUT-SD-EP1** • Durée : **2 h 30** • Coefficient : **3**
(indicatif, valeur fournie pour cette demande)

Sommaire

Cliquez sur une partie pour y accéder directement

1. **Analyse fonctionnelle, qualité et cadre réglementaire** (3 questions) — **Page 3**
2. **Étude technique, calculs et choix d'intégration** (3 questions) — **Page 5**
3. **Préparation de la mise en production et transfert** (3 questions) — **Page 7**
4. **Prévention des risques et continuité de service** (2 questions) — **Page 9**
5. **Ce qu'il faut retenir & Questions fréquentes (FAQ)** (Synthèse & Q&A Examen) — **Page 10**

Sujet officiel (PDF) : https://butsd.fr/assets/annales/but-sd/C7 - Intégrer et valoriser sa production dans l'écosystème de l'entreprise/2026/BUT Science des Données - C7 - Intégrer et valoriser sa production dans l'écosystème de l'entreprise - Session 2026 - 1-Sujet.pdf

1. Analyse fonctionnelle, qualité et cadre réglementaire

Alpina Mobilité souhaite un tableau de bord quotidien présentant le chiffre d'affaires, le nombre de locations et le taux de clients récurrents. Il faut vérifier la qualité des données, encadrer leur diffusion et définir des exigences vérifiables de fraîcheur, de confidentialité et de traçabilité.

Question 1.1. Reformulez le besoin en trois exigences vérifiables : une exigence fonctionnelle, une exigence de fraîcheur des données et une exigence de confidentialité. Pour chacune, indiquez un critère permettant de vérifier qu'elle est satisfaite.

Fonctionnelle : afficher le chiffre d'affaires, le nombre de locations et le taux de clients récurrents, agrégés par jour et par agence. **Vérification :** contrôler la présence des trois indicateurs dans le tableau de bord.
Fraîcheur : actualiser quotidiennement, terminer avant 6 h 45 et publier avant 7 h 00. **Vérification :** consulter l'horodatage et confirmer que l'export a au plus 30 minutes au démarrage.
Confidentialité : ne publier aucun identifiant direct. **Vérification :** contrôler les champs transmis à l'outil, qui ne doit recevoir que les agrégats requis.

> ✓ **RÉPONSE À RETENIR**
>
> Trois indicateurs agrégés par jour et par agence, publiés avant 7 h 00 à partir d'un export âgé de 30 minutes au plus, sans identifiant direct.

Question 1.2. À partir de DT 2, calculez le nombre de lignes valides et le taux global de lignes en anomalie. Présentez le calcul du taux en pourcentage, arrondi à deux décimales, avec le nombre total de lignes comme dénominateur.

Les catégories d'anomalies sont disjointes : le total des lignes en anomalie est donc la somme des doublons et des dates invalides. Les lignes valides sont le total moins ces anomalies :
$120\,000 - (2\,400 + 1\,200) = 116\,400$ lignes.
Le taux global, avec le nombre total de lignes comme dénominateur, vaut :
$\frac{2\,400+1\,200}{120\,000} \times 100 = 3{,}00\,\%$.

> ✓ **RÉPONSE À RETENIR**
>
> 116 400 lignes valides ; taux global d'anomalie : 3,00 %.

Question 1.3. Complétez DR 1 en indiquant, pour chaque indicateur, sa définition opérationnelle et une vérification de qualité pertinente. Ajoutez deux mesures de protection des données et deux éléments de traçabilité à conserver, en tenant compte des règles internes de DT 3 ainsi que des principes de minimisation et de sécurité.

Compléter les huit cellules du DR 1 avec les définitions prévues par DT 2 et des contrôles vérifiables. Les identifiants directs ne sont pas nécessaires aux indicateurs agrégés : les exclure de la visualisation et réserver les exports bruts à un espace à accès restreint. Pour la traçabilité, conserver par exemple la date et la source, puis le volume reçu, les contrôles, le statut de publication et les erreurs. Les journaux ne doivent contenir aucun secret d'authentification.

DR 1 complété

Définition des indicateurs, contrôles et mesures de protection

Élément	Définition, contrôle ou mesure proposée
Chiffre d'affaires moyen par location : définition opérationnelle	CA valide / locations valides
Chiffre d'affaires moyen par location : contrôle de qualité	Contrôler total et dénominateur
Taux de clients récurrents : définition opérationnelle	Récurrents sur 90 j / éligibles
Taux de clients récurrents : contrôle de qualité	Récurrents ≤ éligibles
Protection des données : mesure 1	Aucun identifiant direct publié
Protection des données : mesure 2	Exports bruts : accès restreint
Traçabilité : élément à conserver 1	Date, source, volume, contrôles
Traçabilité : élément à conserver 2	Statut, erreurs ; aucun secret

✓ RÉPONSE À RETENIR

DR 1 complété : définitions, contrôles qualité, mesures de protection et traces de traitement.

2. Étude technique, calculs et choix d'intégration

L'étude porte sur le dimensionnement du flux quotidien de locations, le calcul des indicateurs métier et l'évaluation économique de l'automatisation. Les documents DT 1 à DT 3 précisent les données, les délais, les coûts et les architectures d'intégration possibles.

Question 2.1. À partir de DT 2, calculez le chiffre d'affaires moyen par location et le taux de clients récurrents. Donnez, pour chacun, la formule, les valeurs substituées et le résultat. Pour le taux de clients récurrents, utilisez comme dénominateur le nombre de clients éligibles indiqué dans DT 2.

Le chiffre d’affaires moyen par location se calcule en divisant le chiffre d’affaires des locations valides par le nombre de locations valides fourni pour cet indicateur : 8 400 000 / 120 000 = 70 € par location. Le taux de clients récurrents est le nombre de clients ayant effectué au moins deux locations en 90 jours divisé par le nombre de clients éligibles : (18 600 / 62 000) × 100 = 30 %. Le dénominateur utilisé pour ce taux est bien celui des clients éligibles indiqué dans DT 2.

> ✓ RÉPONSE À RETENIR
>
> Chiffre d’affaires moyen : 70 € par location. Taux de clients récurrents : 30 %.

Question 2.2. Calculez la taille théorique du fichier quotidien en MB et sa durée de transfert en secondes, avec les conversions imposées. Calculez le coût annuel d'hébergement et de supervision, le coût du travail initial et l'économie annuelle brute. Déduisez le bénéfice net de première année, sans actualisation ni autre coût, puis le ratio économie annuelle brute / (coût du travail initial + coût annuel d'hébergement) × 100.

Taille du fichier : 120 000 × 1,2 kB = 144 000 kB = 144 MB. Durée de transfert : 144 / 12 = 12 s. Coût annuel d’hébergement : 95 × 12 = 1 140 €. Coût du travail initial : 18 × 48 = 864 €. Économie annuelle brute : 3,5 × 46 × 32 = 5 152 €. Le bénéfice net de première année est 5 152 - 1 140 - 864 = 3 148 €. Le ratio demandé vaut 5 152 / (864 + 1 140) × 100 = 257,09 %.

> ✓ RÉPONSE À RETENIR
>
> Fichier : 144 MB ; transfert : 12 s ; hébergement annuel : 1 140 € ; travail initial : 864 € ; économie brute : 5 152 € ; bénéfice net : 3 148 € ; ratio : 257,09 %.

Question 2.3. À l'aide du schéma DT 1 et des caractéristiques de DT 3, choisissez une architecture parmi les options proposées. Justifiez votre choix par deux arguments techniques liés au flux, à la disponibilité ou à la sécurité. Précisez un contrôle qui empêche la publication si les données du jour sont incomplètes ou trop anciennes.

Choisir l’option B : transfert automatisé vers une zone sécurisée, contrôles, transformation planifiée et publication des agrégats. Elle reprend le flux de DT 1 et automatise l’actualisation, avec une fin de traitement exigée avant 6 h 45. Elle limite aussi l’exposition : les exports bruts restent en zone restreinte et seuls les agrégats nécessaires

sont transmis à l’outil de visualisation. Avant publication, vérifier la présence du fichier daté du jour, son schéma, son volume attendu de 120 000 lignes et sa fraîcheur, dont l’âge ne doit pas dépasser 30 minutes au démarrage. En cas d’échec, conserver la dernière publication valide.

✓ RÉPONSE À RETENIR

Option B : transfert sécurisé, contrôles automatisés et publication des seuls agrégats. Bloquer si le fichier du jour est absent, incomplet, non conforme ou âgé de plus de 30 minutes au démarrage. Ne pas remplacer la dernière publication valide en cas d’échec.

3. Préparation de la mise en production et transfert

Le candidat prépare la mise en service du flux quotidien de données de location et son transfert à l'équipe informatique. Il doit organiser les tests, les vérifications, la publication, le retour à la version précédente et la communication aux utilisateurs.

Question 3.1. Complétez DR 2 en ordonnant les opérations de mise en production et en indiquant, pour chacune, un contrôle observable et un résultat attendu. Le protocole doit inclure une exécution de test, la vérification des indicateurs, la publication et une procédure de retour à la version précédente en cas d'échec.

Compléter le DR 2 dans l'ordre : exécuter le test sans publier, valider les indicateurs, publier la version validée, puis restaurer la dernière version valide si une publication échoue. Le test vérifie le fichier du jour, le schéma attendu et un âge maximal de 30 minutes au démarrage. La publication n'est autorisée que si les contrôles passent et si le traitement se termine avant 6 h 45.

Ordre	Opération	Contrôle observable	Résultat attendu
1	Exécution de test	Fichier du jour, schéma conforme, fraîcheur ≤ 30 min	Test réussi, sans publication
2	Vérification des indicateurs	Valeurs et agrégats vérifiés, aucun identifiant direct	Indicateurs validés
3	Publication de la version validée	Contrôles réussis, fin avant 6 h 45	Agrégats publiés, heure de mise à jour visible
4	Retour à la version précédente en cas d'échec	Version antérieure restaurée et accessible	Dernière publication valide maintenue

DR 2 complété, protocole de mise en production

DR 1 : Document réponse à rendre avec la copie : questions 1.3 — **À RENDRE AVEC LA COPIE**

Compléter directement les cellules signalées par « ... ». Pour chaque ligne, proposer une définition ou une mesure concrète et vérifiable. Les réponses peuvent être formulées de manière concise.

Définition des indicateurs, contrôles et mesures de protection

Élément	Définition, contrôle ou mesure proposée
1 Chiffre d'affaires moyen par location : définition opérationnelle	Fichier du jour, schéma cor Test réussi, ≤ 30 min sans publier
2 Chiffre d'affaires moyen par location : contrôle de qualité	Valeurs contrôlées, aucun i Indicateurs direct validés
3 Taux de clients récurrents : définition opérationnelle	Contrôles OK, fin avant 6 h Agrégats publiés, heure visible
4 Taux de clients récurrents : contrôle de qualité	Version antérieure restauré Dernière version accessible valide maintenue

✓ RÉPONSE À RETENIR

Tester, vérifier les indicateurs, publier avant 6 h 45 si tous les contrôles passent, sinon conserver ou restaurer la dernière version valide.

Question 3.2. Rédigez une consigne de transmission de cinq à huit lignes destinée à l'équipe informatique. Elle doit préciser le nom du flux, sa fréquence, son horaire limite, les conditions de publication, le comportement en cas d'échec et les informations à consulter dans le journal d'exécution.

Consigne à transmettre à l'équipe informatique, en six lignes :

Flux : export quotidien des locations Alpina Mobilité vers le tableau de bord agrégé.
Fréquence : chaque jour ; export attendu à 5 h 30, traitement terminé avant 6 h 45.
Publier uniquement si le fichier du jour est présent, conforme et âgé de 30 minutes au plus au démarrage.
En cas d'échec, bloquer la publication, conserver la dernière version valide et déclencher une alerte.
Consulter le journal : date, source, nombre de lignes, contrôles, statut de publication et erreurs.
Ne jamais inscrire de mot de passe ni de jeton dans le journal.

✓ RÉPONSE À RETENIR

Flux quotidien automatisé : publication avant 6 h 45 après contrôles ; en cas d'échec, alerter et conserver la dernière version valide.

Question 3.3. Rédigez en anglais un message professionnel de trois à cinq phrases destiné aux responsables d'agence. Présentez les trois indicateurs livrés, la fréquence d'actualisation et la limite d'interprétation relative aux lignes écartées pour anomalie, dans un langage accessible à un public non spécialiste.

Professional message in four sentences:

> Dear Agency Managers, the dashboard shows daily revenue, the number of rentals, and the share of eligible customers who made at least two rentals in the last 90 days. It is updated every day, and the latest update time is displayed. Some records are excluded when they contain duplicate rentals or invalid dates. The figures therefore cover the valid records used in the calculations and may not represent every record received.

✓ RÉPONSE À RETENIR

The dashboard presents revenue, rental volume, and the repeat-customer rate, updated daily. Figures exclude records with anomalies.

4. Prévention des risques et continuité de service

Les situations concernent l'exposition d'identifiants directs, une demande d'accès excessif aux données détaillées et l'absence du fichier quotidien avant publication. Il faut associer les risques à des mesures vérifiables et éviter la diffusion d'indicateurs incomplets ou périmés.

Question 4.1. Pour les situations A et C, identifiez le principal risque et proposez une mesure de prévention adaptée. Pour chaque mesure, indiquez comment vérifier son application en vous appuyant sur les règles d'accès et de minimisation de DT 3.

Situation A : un export contenant des identifiants directs dans un espace partagé expose les clients à une divulgation non autorisée. Conserver les exports bruts dans un espace à accès restreint et exclure les identifiants du tableau de bord. Vérifier les droits de l'espace et contrôler les champs publiés.

Situation C : l'accès permanent à toutes les données détaillées dépasse le besoin professionnel et accroît le risque d'accès indu. Accorder uniquement les agrégats par jour et par agence utiles à la fonction, puis réexaminer les habilitations à chaque changement de fonction. Vérifier les droits attribués et leur adéquation au rôle.

> ✓ RÉPONSE À RETENIR
>
> A : export brut à accès restreint, identifiants exclus du tableau de bord. C : accès limité aux agrégats nécessaires, habilitations vérifiées et réexaminées.

Question 4.2. Pour la situation B, décrivez la conduite à tenir afin d'éviter la publication d'indicateurs incomplets ou périmés. Précisez la règle de blocage, l'information à transmettre aux utilisateurs, la trace à enregistrer et la condition de relance du traitement.

Si le fichier attendu à 5 h 30 est absent, bloquer le traitement et conserver la dernière publication valide. Informer les utilisateurs que les données du jour ne sont pas disponibles et leur communiquer l'heure de la dernière mise à jour.

- Journaliser la date, la source, le nombre de lignes reçues, les contrôles, le statut et l'erreur, sans secret d'authentification.
- Relancer après réception d'un fichier daté du jour, au schéma conforme et âgé de 30 minutes au plus au démarrage, uniquement si le traitement peut finir avant 6 h 45.

> ✓ RÉPONSE À RETENIR
>
> Bloquer la publication, conserver la dernière version valide, avertir les utilisateurs et journaliser l'échec. Relancer seulement si le fichier est conforme, assez récent et traitable avant 6 h 45.

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

Bilan de l'épreuve : Cette épreuve de BUT Science des données, session 2026, code 26-BUT-SD-EP1, dure 2 h 30 et indique un coefficient de 3, donné à titre indicatif pour cette demande. Le cas Alpina Mobilité porte sur un tableau de bord quotidien des locations de vélos. Les compétences évaluées couvrent la qualité des données, les indicateurs, le dimensionnement, l'intégration, la mise en production, la sécurité et la communication professionnelle.

Notions-clés, formules & règles à maîtriser

- Taux d'anomalie = nombre de lignes en anomalie / nombre total de lignes × 100; les catégories fournies sont disjointes.
- Chiffre d'affaires moyen par location = chiffre d'affaires total des locations valides / nombre de locations valides.
- Taux de clients récurrents = clients éligibles ayant réalisé au moins deux locations en 90 jours / nombre total de clients éligibles.
- Le flux bloque la publication si le fichier manque, si son schéma est incorrect ou si son âge dépasse 30 minutes; la dernière publication valide est conservée.

Points de vigilance & attentes du jury

- Pour calculer les lignes valides et les anomalies, additionnez les catégories disjointes sans compter de chevauchements.
- Respectez les dénominateurs indiqués dans DT 2 et présentez les unités ainsi que les étapes des calculs.
- Utilisez les conversions kB et MB imposées et distinguez coûts initiaux, coûts annuels et économie brute.
- Ne transmettez pas d'identifiants directs et ne remplacez pas la dernière publication valide après un échec.

Questions fréquentes sur cette épreuve (Q&A Examen)

Q : Comment calculer le taux de clients récurrents dans ce sujet ?

Divisez le nombre de clients éligibles ayant réalisé au moins deux locations pendant la période de référence de 90 jours par le nombre total de clients éligibles. Multipliez le résultat par 100 pour l'exprimer en pourcentage.

Q : Quand faut-il bloquer la publication du tableau de bord ?

Le traitement échoue si le fichier du jour manque, si son schéma est incorrect ou si son âge dépasse 30 minutes au démarrage. L'échec ne doit pas remplacer la dernière publication valide.

Q : Comment organiser son temps et ne pas oublier les documents réponses ?

Traitez les parties dans l'ordre et gardez du temps pour les calculs et les rédactions. Complétez DR 1 pour la question 1.3 et DR 2 pour la question 3.1, puis vérifiez unités, arrondis et longueurs demandées.

Q : Quels critères de qualité et de sécurité faut-il surtout respecter ?

Prévoyez des contrôles observables pour le schéma, les volumes, les anomalies et la fraîcheur. Limitez les accès aux besoins professionnels, ne publiez que les agrégats requis et consignez les événements sans enregistrer de secret d'authentification.

Sujet officiel associé (PDF) : https://butsd.fr/assets/annales/but-sd/C7 - Intégrer et valoriser sa production dans l'écosystème de l'entreprise/2026/BUT Science des Données - C7 - Intégrer et valoriser sa production dans l'écosystème de l'entreprise - Session 2026 - 1-Sujet.pdf

Ressources BUT Science des Données : 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.