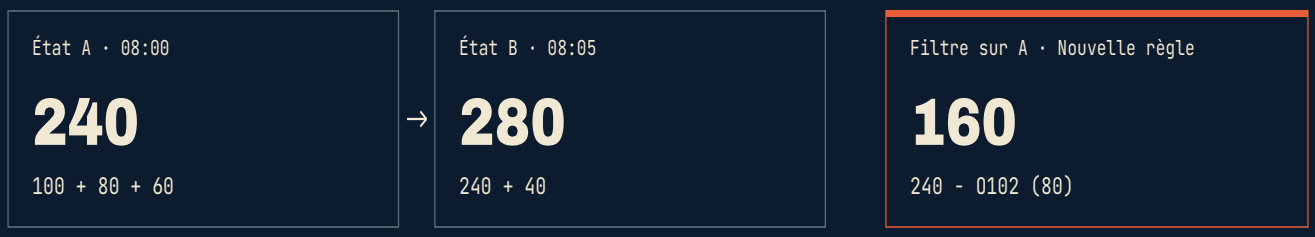
NOTES DE TERRAIN · 02

Conservez la démarche d’analyse, pas seulement le graphique

Une méthode pour transmettre une analyse avec sa référence de données, son calcul, les alternatives étudiées et les raisons des choix. Un exemple fictif passant de 240 à 280 distingue l’actualisation des données d’un changement de règle ; le parcours DUKE reste une direction de conception à évaluer.

Une référence, deux types de changement



Exemple fictif · Unités synthétiques

Direction de conception à évaluer

Préparé le 2 novembre 2026

URL canonique cible

https://duke-analytics.com/blog/conserver-analyse-contexte-et-raisonnement

Points à retenir

01 Documenter la référence exacte des données, la définition de la mesure, la requête et les filtres de chaque résultat.

02 Conserver le résultat de départ lors du test d’une nouvelle hypothèse et distinguer changement de règle et actualisation des données.

03 Garder les motifs d’un choix, les alternatives écartées et les hypothèses non résolues à côté du résultat.

04 Vérifier qu’un collègue peut reproduire, discuter et poursuivre le travail sans présentation orale.

05 Rendre visibles les droits d’accès et les limites de reproduction. L’exercice fictif ne prouve aucune fonctionnalité livrée.

L’ARTICLE INTÉGRAL

Un collègue ouvre le graphique que vous lui avez envoyé hier. Le chiffre a changé. Quelqu’un se demande si le modèle s’est trompé. Une autre personne propose de modifier le prompt. Avant de conclure, une question plus utile s’impose : consultent-ils le même état des données, avec la même définition et les mêmes filtres ?

Une analyse que l’on peut transmettre doit permettre de répondre à cette question. Le graphique rend le résultat visible. La démarche d’analyse explique ce qui a été compté, les autres pistes envisagées et ce que la personne suivante peut poursuivre en connaissance de cause. Cet article propose une méthode pour conserver ce travail. L’exemple chiffré est volontairement fictif. Les idées produit présentées sont des orientations de conception, pas l’affirmation que toutes ces capacités sont déjà disponibles dans DUKE.

Commencer par un résultat qu’un collègue peut reproduire

Prenons un petit jeu de démonstration contenant trois montants enregistrés, exprimés en unités synthétiques : 100, 80 et 60. À 08 h 00, l’instantané A contient ces trois lignes. Une requête qui additionne tous les montants enregistrés renvoie 240. À 08 h 05, l’instantané B contient une ligne supplémentaire de 40. Le même calcul renvoie 280.

Il n’y a aucune contradiction arithmétique. Les données en entrée ont changé. Désigner les deux ensembles par « les données de vente » masque précisément la différence qui explique le résultat. Ici, la mesure est simplement la somme des montants enregistrés. Elle ne décrit ni un chiffre d’affaires comptabilisé, ni un bénéfice, ni la performance d’un client. Ces mesures nécessiteraient leurs propres définitions métier.

La trace utile ne se limite donc pas à « le total vaut 240 ». Elle comprend la source choisie, l'instantané A, les trois lignes, le calcul et l'absence de filtre supplémentaire. Le total attendu peut être vérifié indépendamment : 100 + 80 + 60 = 240. L'instantané B possède son propre total attendu : 240 + 40 = 280. Conserver les deux permet d'expliquer le changement.

Distinguer la question de l'état des données

Trois éléments peuvent évoluer indépendamment : la question, le calcul et les données disponibles pour ce calcul. Une question portant sur tous les montants enregistrés peut devenir une question sur une région précise. Une définition métier peut commencer à exclure les commandes remboursées. De nouvelles lignes peuvent arriver alors que la formulation reste identique. Chaque changement doit être visible dans la démarche d'analyse.

Même une transaction de base de données ne garantit pas universellement que des requêtes successives verront des données identiques. PostgreSQL indique qu'au niveau d'isolation Read Committed, des commandes SELECT successives peuvent observer des données validées différentes. Il s'agit d'un comportement précis de cette base, pas d'une affirmation sur tous les entrepôts ou connecteurs. La leçon pratique consiste à documenter les conditions de cohérence réellement utilisées. Documentation PostgreSQL sur l'isolation des transactions

Un horodatage peut ne pas suffire à reproduire un ancien résultat. Si la source écrase les enregistrements sans conserver d'historique exploitable, connaître l'heure d'exécution d'une requête ne restitue pas les anciennes lignes. Conservez l'instantané, la version, l'export ou toute autre référence reproductible disponible. S'il n'en existe aucune, indiquez que la reproduction exacte n'est pas possible. Une limite explicite est plus utile qu'une étiquette « reproductible » trompeuse.

Conserver un résultat de référence avant de tester une nouvelle hypothèse

Imaginons maintenant la question suivante : que se passe-t-il si l'on exclut les commandes remboursées ? Traitez-la comme une nouvelle branche de l'analyse. Conservez le résultat de référence initial, puis notez le filtre modifié et la raison de ce changement. Gardez la même référence de données pendant le test de la nouvelle définition. Dans notre jeu fictif, la ligne O102 porte le montant 80 et le statut « refunded ». Son exclusion de l'instantané A donne 160. Le résultat initial de 240 reste conservé. Sans cette séparation, un nouveau résultat pourrait refléter à la fois le changement de filtre et l'arrivée de nouvelles lignes.

Cette méthode n'impose pas d'exclure les remboursements de toutes les mesures. Une équipe financière peut avoir besoin du chiffre d'affaires net, du montant brut des commandes, du volume des remboursements ou d'un autre indicateur. Définissez d'abord la mesure recherchée. Précisez comment le statut de remboursement, les dates, les devises et les corrections interviennent avant de qualifier un résultat de juste ou de faux.

Pour chaque branche, conservez une courte description de l'hypothèse, la version de la règle, la requête ou le calcul exact, les filtres actifs et le résultat. Indiquez ce qui a changé par rapport à la référence. Si une branche répond à une autre question métier, rendez cette différence

explicite. Deux chiffres plausibles peuvent constituer des réponses valides à deux questions différentes.

Des définitions centralisées des indicateurs peuvent aider à tenir cette discipline. Par exemple, dbt documente la définition de métriques dans une couche sémantique et leur utilisation par des outils en aval. C'est un élément de contexte utile, pas une capacité revendiquée comme exclusive à DUKE. Pour chaque résultat, l'analyse doit toujours préciser la définition et la référence de données effectivement utilisées. Documentation dbt Semantic Layer

Garder les raisons des choix

Reproduire un chiffre ne suffit pas à transmettre une analyse. La personne qui la reprend doit aussi comprendre pourquoi une piste a été retenue. Une comparaison portait peut-être sur des périodes incomplètes. Un segment était peut-être trop petit pour être interprété. Une hausse apparente a peut-être disparu lorsque la population a été maintenue constante. Ce sont des exemples de raisonnements à documenter, pas des événements présentés comme survenus dans un projet client.

Une note utile peut rester brève : « La branche qui exclut les remboursements modifie la règle de remboursement ; elle n'actualise pas les données. Nous conservons la référence A pour la question initiale. » Une autre peut indiquer : « Nous ne pouvons pas encore comparer ces deux résultats, car ils reposent sur des versions différentes de la source. » Placez cette explication à côté du résultat concerné, au lieu de la laisser enfouie dans une conversation séparée.

Conservez aussi la question encore ouverte. Une analyse doit rendre ses incertitudes assez visibles pour que quelqu'un puisse la poursuivre. Précisez quelle hypothèse reste à tester et quelles observations pourraient modifier la conclusion. Sinon, un graphique final soigné peut donner l'impression qu'un raisonnement inachevé est définitif.

Rendre le test de transmission concret

Confiez le dossier d'analyse à un collègue sans lui en faire une présentation orale. Demandez-lui de retrouver la question initiale, d'identifier la référence de données et de reproduire le résultat de départ. Demandez-lui ensuite de repérer le filtre modifié dans la branche alternative et d'expliquer pourquoi un résultat a été retenu, ou pourquoi aucune décision n'a pu être prise.

Ne considérez pas un désaccord comme un échec automatique de votre collègue. Il peut révéler une définition ambiguë, une preuve manquante ou un choix insuffisamment étayé dans le travail initial. Notez l'endroit où une clarification a été nécessaire. Vous obtenez une amélioration précise à apporter, plus instructive que la seule question « le graphique est-il clair ? ».

Laissez ensuite cette personne poursuivre avec une nouvelle question tout en conservant le résultat initial. Le test réussit lorsque l'ancienne conclusion reste consultable et que le nouveau travail montre ce qui a changé. Il doit aussi être possible de s'arrêter faute de preuves

suffisantes. Un système qui invente avec assurance le contexte manquant n’a pas réussi ce test.

Partager l’analyse dans le respect des droits d’accès

Conserver le contexte ne signifie pas copier toutes les sources dans un dossier accessible à tous. La personne qui reçoit l’analyse doit disposer des droits appropriés sur les éléments sous-jacents. Un résultat enregistré peut lui-même contenir des informations sensibles. Déterminez ce qui peut être partagé, quelles références nécessitent une nouvelle vérification des permissions et comment l’équipe traite une révocation ultérieure.

Distinguez l’accès à la trace d’un travail passé de l’autorisation d’exécuter une nouvelle requête. Une explication historique peut rester utile même si le destinataire n’a pas accès à toutes les lignes d’origine. Cette limite doit être visible. Le partage d’un graphique ne doit jamais laisser entendre qu’il autorise à récupérer ou à redistribuer les données sources.

La direction de conception envisagée pour DUKE

Nous souhaitons explorer un espace de travail où la question, la référence, les alternatives et l’explication restent réunies. Une nouvelle hypothèse créerait une branche visible. Une source actualisée rendrait le changement visible. Un collègue disposerait d’assez de contexte pour vérifier le résultat et décider de la prochaine question à examiner.

Ce sont des critères d’acceptation à démontrer, pas des fonctionnalités que la rédaction d’un article suffirait à déclarer achevées. Le jeu fictif fournit un premier test limité. Une évaluation utile devrait ensuite porter sur une question métier pertinente, une définition convenue et un résultat attendu connu, en consignant les échecs aussi soigneusement que les réussites.

Apportez une analyse qu’une autre personne doit pouvoir poursuivre : découvrir DUKE.

Questions & réponses

Comment 240 et 280 peuvent-ils être deux résultats justes ?

Dans le jeu fictif, l'instantané A contient 100, 80 et 60, soit 240. L'instantané B ajoute une ligne de 40. La même somme renvoie alors 280. Chaque résultat est correct pour l'état des données auquel il correspond.

Un horodatage suffit-il à reproduire une analyse ?

Pas nécessairement. Si les anciennes lignes ont été écrasées sans historique exploitable, l'heure de la requête ne permet pas de les retrouver. Il faut conserver un instantané, une version ou un export disponible, et signaler lorsque la reproduction exacte est impossible.

Que faut-il transmettre avec un résultat d'analyse ?

La question, la définition de la mesure, la référence de données, la requête ou le calcul, les filtres actifs, le résultat, les alternatives, les raisons du choix et les hypothèses encore ouvertes. Il faut aussi préciser les limites d'accès qui s'appliquent au destinataire.

Cet exemple montre-t-il un parcours déjà disponible dans DUKE ?

Non. Le calcul repose sur un exemple volontairement fictif et vérifiable. L'espace de travail et le comportement de transmission décrits pour DUKE sont des orientations de conception et des critères d'acceptation à évaluer, pas l'affirmation que toutes ces capacités sont livrées.

Sources

Documentation consultée le 2 novembre 2026.

1. Documentation PostgreSQL : isolation des transactions

2. Documentation dbt : Semantic Layer