DUKE ANALYTICS

DE

PRAXISNOTIZEN · 02

Die Analyse bewahren, nicht nur das Diagramm

Eine Methode, um Analysen samt Datenreferenz, Berechnung, geprüften Alternativen und Begründungen zu übergeben. Ein fiktives Beispiel mit 240 und 280 trennt einen geänderten Datenstand von einer geänderten fachlichen Regel; der DUKE-Ablauf bleibt ein zu prüfendes Entwurfsziel.

Eine Referenz, zwei Arten der Änderung



Fiktives Beispiel · Synthetische Einheiten

Zu prüfendes Entwurfsziel

Vorbereitet am 2. November 2026

Vorgesehene kanonische URL

https://duke-analytics.com/blog/analyse-und-begruendung-nachvollziehbar-sichern

Das Wichtigste

01 Für jedes Ergebnis die genaue Datenreferenz, Kennzahlendefinition, Abfrage und aktiven Filter dokumentieren.

02 Die Ausgangsbasis beim Prüfen einer neuen Hypothese bewahren und eine Regeländerung von einer Datenaktualisierung unterscheiden.

03 Gründe für Entscheidungen, verworfene Alternativen und offene Annahmen beim Ergebnis festhalten.

04 Prüfen, ob ein Kollege die Arbeit ohne mündliche Einführung reproduzieren, hinterfragen und fortsetzen kann.

05 Zugriffsgrenzen und Einschränkungen der Reproduzierbarkeit sichtbar machen. Das fiktive Beispiel belegt keine ausgelieferte Produktfunktion.

DER VOLLSTÄNDIGE ARTIKEL

Ein Kollege öffnet das Diagramm, das Sie ihm gestern geschickt haben. Die Zahl ist eine andere. Jemand fragt, ob das Modell einen Fehler gemacht hat. Eine andere Person schlägt vor, den Prompt zu ändern. Bevor Sie eine dieser Erklärungen übernehmen, lohnt sich eine andere Frage: Betrachten alle denselben Datenstand, mit derselben Definition und denselben Filtern?

Eine Analyse, die sich sinnvoll übergeben lässt, muss diese Frage beantworten können. Das Diagramm macht das Ergebnis sichtbar. Die dokumentierte Analyse erklärt, was gezählt wurde, welche Alternativen geprüft wurden und woran die nächste Person fundiert weiterarbeiten kann. Dieser Artikel schlägt eine praktische Methode vor, um diese Arbeit zu bewahren. Das Zahlenbeispiel ist bewusst fiktiv. Die beschriebenen Produktideen sind Entwurfsziele. Sie besagen nicht, dass alle genannten Funktionen heute bereits in DUKE verfügbar sind.

Mit einem Ergebnis beginnen, das andere reproduzieren können

Betrachten wir einen kleinen Beispieldatensatz mit drei erfassten Beträgen in synthetischen Einheiten: 100, 80 und 60. Um 08:00 Uhr enthält Snapshot A diese drei Zeilen. Eine Abfrage, die alle erfassten Beträge summiert, ergibt 240. Um 08:05 Uhr enthält Snapshot B eine zusätzliche Zeile mit dem Betrag 40. Dieselbe Berechnung ergibt nun 280.

Die Arithmetik widerspricht sich nicht. Die Eingabedaten haben sich verändert. Beide Datenstände einfach als „die Verkaufsdaten“ zu bezeichnen, verdeckt genau den Unterschied, der das Ergebnis erklärt. Die Kennzahl in diesem Beispiel ist lediglich die Summe der erfassten Beträge. Sie ist keine Aussage über verbuchte Umsatzerlöse, Gewinn oder die Leistung eines Kunden. Solche Kennzahlen würden jeweils eigene fachliche Definitionen erfordern.

Eine brauchbare Dokumentation hält deshalb mehr fest als „Gesamtsumme gleich 240“. Sie umfasst die ausgewählte Quelle, Snapshot A, die drei Zeilen, die Berechnung und den Verzicht auf zusätzliche Filter. Die erwartete Summe lässt sich unabhängig prüfen: 100 + 80 + 60 = 240. Für Snapshot B gilt eine eigene erwartete Summe: 240 + 40 = 280. Wenn beide erhalten bleiben, ist die Veränderung erklärbar.

Die Fragestellung vom Datenstand unterscheiden

Drei Dinge können sich unabhängig voneinander ändern: die Frage, die Berechnung und die Daten, die dieser Berechnung zur Verfügung stehen. Aus einer Frage nach allen erfassten Beträgen kann eine Frage nach einer bestimmten Region werden. Eine fachliche Definition kann künftig erstattete Bestellungen ausschließen. Neue Zeilen können hinzukommen, obwohl die Formulierung unverändert bleibt. Jede Änderung sollte in der Analyse ausdrücklich erkennbar sein.

Auch eine Datenbanktransaktion ist keine allgemeine Zusage, dass aufeinanderfolgende Abfragen identische Daten sehen. PostgreSQL dokumentiert, dass auf der Isolationsstufe Read Committed aufeinanderfolgende SELECT-Anweisungen unterschiedliche bereits bestätigte Daten sehen können. Das ist ein konkretes Verhalten dieser Datenbank, keine Aussage über jedes Data Warehouse oder jeden Connector. Die übergeordnete praktische Konsequenz lautet: Halten Sie fest, welche Konsistenzbedingungen tatsächlich galten. PostgreSQL-Dokumentation zur Transaktionsisolation

Ein Zeitstempel allein reicht möglicherweise nicht aus, um ein früheres Ergebnis zu reproduzieren. Wenn die Quelle Datensätze überschreibt und keine nutzbare Historie vorhält, stellt die Kenntnis des Abfragezeitpunkts die früheren Zeilen nicht wieder her. Bewahren Sie den verfügbaren Snapshot, die Version, den Export oder eine andere reproduzierbare Referenz auf. Fehlt eine solche Grundlage, benennen Sie, dass eine exakte Wiederholung nicht möglich ist. Eine ehrliche Einschränkung ist hilfreicher als eine irreführende Kennzeichnung als „reproduzierbar“.

Eine Ausgangsbasis sichern, bevor Sie eine neue Hypothese prüfen

Stellen wir uns die nächste Frage vor: Was geschieht, wenn wir erstattete Bestellungen ausschließen? Behandeln Sie das als einen neuen Analysezweig. Bewahren Sie die ursprüngliche Ausgangsbasis auf und dokumentieren Sie anschließend den geänderten Filter und den Grund für die Änderung. Lassen Sie die Datenreferenz unverändert, während Sie die neue Definition prüfen. In unserem fiktiven Datensatz enthält die Zeile O102 den Betrag 80 und den Status „refunded“. Wird diese Zeile aus Snapshot A ausgeschlossen, ergibt sich 160. Das ursprüngliche Ergebnis von 240 bleibt erhalten. Andernfalls könnte ein verändertes Ergebnis sowohl auf einen neuen Filter als auch auf neu hinzugekommene Zeilen zurückgehen.

Dies ist ein Vorschlag für die Arbeitsweise, keine Anweisung, Erstattungen aus jeder Kennzahl auszuschließen. Ein Finanzteam benötigt vielleicht den Nettoumsatz, den Bruttowert der Bestellungen, das Erstattungsvolumen oder eine ganz andere Kennzahl. Definieren Sie zuerst,

was Sie messen wollen. Legen Sie fest, wie Erstattungsstatus, Datumsangaben, Währungen und Korrekturen berücksichtigt werden, bevor Sie ein Ergebnis als richtig oder falsch bewerten.

Bewahren Sie für jeden Analysezweig eine kurze Beschreibung der Hypothese, die Regelversion, die genaue Abfrage oder Berechnung, die aktiven Filter und das Ergebnis auf. Halten Sie fest, was sich gegenüber der Ausgangsbasis geändert hat. Wenn ein Zweig eine andere fachliche Frage beantwortet, machen Sie diesen Unterschied deutlich. Zwei plausible Zahlen können jeweils gültige Antworten auf unterschiedliche Fragen sein.

Zentral verwaltete Kennzahlendefinitionen können diese Vorgehensweise unterstützen. Beispielsweise beschreibt dbt, wie Kennzahlen in einem Semantic Layer definiert und von nachgelagerten Werkzeugen genutzt werden. Das ist hilfreicher Kontext, keine ausschließlich für DUKE beanspruchte Fähigkeit. Eine Analyse muss weiterhin festhalten, welche Definition und welche Datenreferenz einem konkreten Ergebnis zugrunde lagen. Dokumentation des dbt Semantic Layer

Die Gründe für Entscheidungen festhalten

Eine Zahl zu reproduzieren ist nur ein Teil der Übergabe. Die nächste Person muss auch verstehen, warum ein bestimmter Weg gewählt wurde. Vielleicht beruhte ein Vergleich auf unvollständigen Zeiträumen. Vielleicht war ein Segment zu klein, um es sinnvoll zu interpretieren. Vielleicht verschwand ein scheinbarer Anstieg, als die betrachtete Population konstant gehalten wurde. Das sind Beispiele für Überlegungen, die dokumentiert werden sollten, keine als tatsächlich geschehen dargestellten Ereignisse aus einem Kundenprojekt.

Eine hilfreiche Notiz darf kurz sein: „Der Zweig ohne Erstattungen ändert die Erstattungsregel; er aktualisiert nicht die Daten. Für die ursprüngliche Frage behalten wir die Ausgangsbasis A bei.“ Eine andere könnte lauten: „Wir können diese beiden Ergebnisse noch nicht vergleichen, weil sie auf unterschiedlichen Quellversionen beruhen.“ Speichern Sie diese Erklärung beim betreffenden Ergebnis, statt sie in einem getrennten Gespräch verschwinden zu lassen.

Halten Sie auch die noch offene Frage fest. Die Unsicherheit einer Analyse sollte so sichtbar sein, dass jemand daran weiterarbeiten kann. Benennen Sie, welche Annahme noch ungeprüft ist und welche Belege die Schlussfolgerung ändern würden. Sonst kann ein sauber gestaltetes Schlussdiagramm den Eindruck erwecken, ein noch unvollständiger Gedankengang sei bereits abgeschlossen.

Die Übergabe konkret prüfen

Geben Sie einem Kollegen die dokumentierte Analyse, ohne ihn mündlich hindurchzuführen. Bitten Sie ihn, die ursprüngliche Frage wiederzufinden, die Datenreferenz zu bestimmen und das Ausgangsergebnis zu reproduzieren. Anschließend soll er den geänderten Filter im alternativen Analysezweig erkennen und erläutern, warum ein Ergebnis ausgewählt wurde oder warum noch keine Entscheidung möglich war.

Betrachten Sie eine abweichende Einschätzung nicht automatisch als Versagen des Kollegen. Sie kann auf eine mehrdeutige Definition, fehlende Belege oder eine unzureichend begründete

Wahl in der ursprünglichen Arbeit hinweisen. Notieren Sie die Stelle, an der eine Erläuterung nötig wurde. Daraus ergibt sich eine konkrete Verbesserung. Das ist aufschlussreicher als die bloße Frage, ob das Diagramm „verständlich aussieht“.

Lassen Sie die Person anschließend eine neue Frage bearbeiten, während das ursprüngliche Ergebnis erhalten bleibt. Der Test ist bestanden, wenn die frühere Schlussfolgerung weiterhin überprüfbar ist und die neue Arbeit erkennen lässt, was sich geändert hat. Es muss auch möglich sein, wegen unzureichender Belege aufzuhören. Ein System, das fehlenden Kontext selbstsicher erfindet, hat diesen Test nicht bestanden.

Die Analyse innerhalb der richtigen Zugriffsgrenzen teilen

Kontext zu bewahren bedeutet nicht, jede Quelle in einen frei zugänglichen Ordner zu kopieren. Wer die Analyse erhält, benötigt die passenden Zugriffsrechte auf die zugrunde liegenden Informationen. Auch ein gespeichertes Ergebnis kann sensible Informationen enthalten. Legen Sie fest, welche Teile geteilt werden dürfen, bei welchen Referenzen die Berechtigungen erneut geprüft werden müssen und wie das Team mit einem späteren Entzug der Rechte umgeht.

Unterscheiden Sie den Zugriff auf die Dokumentation früherer Arbeit von der Berechtigung, eine neue Abfrage auszuführen. Eine historische Erklärung kann nützlich bleiben, auch wenn der Empfänger nicht auf jede ursprüngliche Zeile zugreifen darf. Diese Grenze muss sichtbar sein. Eine Übergabe darf niemals den Eindruck vermitteln, dass ein geteiltes Diagramm automatisch zum Abruf oder zur Weitergabe seiner Quelldaten berechtigt.

Die Entwurfsziele für DUKE

Wir möchten einen Arbeitsbereich untersuchen, in dem Fragestellung, Referenz, Alternativen und Begründung zusammenbleiben. Eine neue Hypothese würde einen sichtbaren Analysezweig erzeugen. Eine aktualisierte Quelle würde die Veränderung sichtbar machen. Ein Kollege hätte genügend Kontext, um das Ergebnis zu prüfen und zu entscheiden, welche Frage als Nächstes untersucht werden sollte.

Das sind Abnahmekriterien, die nachgewiesen werden müssen. Sie sind keine Funktionen, die allein durch ihre Beschreibung in einem Artikel als fertig gelten. Der fiktive Datensatz bietet einen kleinen ersten Test. Eine aussagekräftige Bewertung sollte anschließend eine relevante fachliche Frage, eine gemeinsam festgelegte Definition und ein bekanntes erwartetes Ergebnis verwenden. Fehlschläge müssen dabei ebenso sorgfältig festgehalten werden wie Erfolge.

Bringen Sie eine Analyse mit, an der eine andere Person weiterarbeiten soll: DUKE entdecken.

Fragen & Antworten

Wie können 240 und 280 beide richtig sein?

Im fiktiven Datensatz enthält Snapshot A die Beträge 100, 80 und 60, zusammen 240. Snapshot B ergänzt eine Zeile mit 40. Dieselbe Summenberechnung ergibt dann 280. Beide Ergebnisse stimmen für ihren jeweiligen Datenstand.

Reicht ein Zeitstempel aus, um eine Analyse zu reproduzieren?

Nicht unbedingt. Wurden frühere Zeilen ohne nutzbare Historie überschrieben, kann der Abfragezeitpunkt sie nicht wiederherstellen. Bewahren Sie einen verfügbaren Snapshot, eine Version oder einen Export auf und benennen Sie, wenn eine exakte Wiederholung unmöglich ist.

Was sollte zusammen mit einem Analyseergebnis übergeben werden?

Fragestellung, Kennzahlendefinition, Datenreferenz, Abfrage oder Berechnung, aktive Filter, Ergebnis, Alternativen, Gründe für die Auswahl und offene Annahmen. Hinzu kommen die Zugriffsgrenzen, die für den Empfänger gelten.

Zeigt dieses Beispiel einen bereits verfügbaren DUKE-Ablauf?

Nein. Die Berechnung verwendet ein bewusst fiktives und überprüfbares Beispiel. Der beschriebene Arbeitsbereich und die Übergabefunktionen für DUKE sind Entwurfsziele und zu prüfende Abnahmekriterien. Sie sind keine Aussage, dass alle Funktionen bereits ausgeliefert sind.

Quellen

Dokumentation geprüft am 2. November 2026.

1. PostgreSQL-Dokumentation: Transaktionsisolation

2. dbt-Dokumentation: Semantic Layer