© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

From Dashboard to Decision Evaluating KPI Governance and Corrective-Action Closure

ALLEN TEERAHUMBA[1], EMMANUEL HAGAN[2], FLORA PHIRI[3], TREVOR KAUYU[4], MUNASHE NAPHTALI MUPA[5]
[1]*University of Louisville, ORCiD: 0009-0001-7353-9205*
[2]*Park University, ORCiD: 0009-0000-9652-472X*
[3]*George Washington University, ORCiD: 0009-0007-6433-3428*
[4]*Arizona State University, ORCiD: 0009-0001-6019-2177*
[5]*Hult International Business School*

Abstract- ***Dashboards are often treated as neutral windows onto organizational performance, yet the decisions they prompt depend on definitions, denominators, aggregation rules, refresh cycles, thresholds and follow-through mechanisms. This study develops and tests a KPI Governance and Corrective-Action Closure Framework (KGCACF) for resource-constrained organizations. The empirical component uses the Brazilian E-Commerce Public Dataset by Olist, obtained from Kaggle, comprising 99,441 orders. Seven operational KPIs were calculated under paired definitions across 86 eligible weeks: delivery lateness, on-time delivery, cancellation, average order value, repeat-customer activity, delivery cycle and approval latency. Management signals were compared against stated thresholds. The paired definitions generated 47 disputed signals in 602 metric-week comparisons (7.81%). The largest decision instability occurred for on-time delivery (18 disputed weeks), delivery cycle (13) and average order value (11). A corrective-action scenario used 26,575 exception events derived from reproducible validation, service and anomaly rules. With identical modeled review capacity, FIFO processing produced a mean high-severity closure time of 482.3 days and 9.2% high-severity SLA compliance; severity/SLA sequencing produced 4.2 days and 100.0% compliance. These are queueing results under explicit assumptions, not observed organizational effects. An evidence-gate sensitivity scenario further shows that an assumed 12% retest-failure rate would reopen 3,191 administratively closed cases. The paper contributes a KPI contract, dispute log, dashboard heat maps, decision register, corrective-action schema and closure-verification protocol. The results show that dashboard value arises not from visualization alone but from governed meaning, accountable action and verified resolution.***

Keywords: key performance indicators; dashboard governance; decision latency; corrective action; action registers; performance measurement; data governance; SMEs; nonprofit organizations

I. INTRODUCTION

Dashboards promise visibility, speed and alignment. A manager can see delivery performance, revenue, customer activity and control exceptions without reading a long report. Yet the visual simplicity can conceal a complex measurement system. “On-time delivery” may divide on-time orders by delivered orders, all orders or orders due in the period. “Average order value” may be based on payment receipts, invoice lines or a line-weighted average. “Repeat customer” may refer to a persistent person identifier or a transaction-specific customer key. Each formulation can be technically executable and still answer a different question.

Performance measurement research has repeatedly shown that metrics influence behavior as well as describe it. Kaplan and Norton (1992) argued for a balanced set of measures rather than isolated financial indicators. Neely, Gregory and Platts (1995) framed performance measurement as the design, implementation and use of interconnected measures. Franco-Santos et al. (2007) later demonstrated that contemporary systems differ in their features, roles and processes, making definitional clarity essential. A dashboard therefore becomes a governance instrument when it defines what counts, what triggers action and who is accountable for the response.

Visualization alone does not complete the management cycle. Van Elten et al. (2022), drawing on dashboard development across 12 health-care innovation pilots in eight European countries, describe dashboarding as collective sensemaking in which stakeholders negotiate what is feasible, acceptable and informative. This insight travels beyond health care: a score must be understood and accepted before it can coordinate action. At the other end of the cycle, Chingezi et al. (2026), including Lucy Ganyani and Munashe Naphtali Mupa, argue that findings are not durably resolved by reporting them; ownership, root-cause correction, monitoring and closure evidence are required.

This study connects those two ends. It asks whether governed KPI definitions reduce decision inconsistency and whether a governed action register can accelerate attention to high-severity exceptions under fixed capacity. Because no authorized multi-organization pilot data were supplied, the paper does not invent interviews, users or pre/post organizational outcomes. It combines observed public-data variation with transparent workflow scenarios, creating a reproducible proof of concept and a protocol for later field validation.

1.1 Research questions

RQ1: How materially do alternative, plausible KPI definitions change reported values and threshold-based management signals?

RQ2: Which KPIs generate the most definition-driven decision disputes across operating weeks?

RQ3: Under identical modeled review capacity, how does severity/SLA sequencing change high-risk corrective-action closure compared with FIFO?

RQ4: What governance artifacts are required to connect a dashboard signal to a verified corrective-action closure?

II. LITERATURE REVIEW AND CONCEPTUAL FOUNDATIONS

2.1 KPI design and measurement validity

A KPI is not merely a displayed number; it is a compressed theory of performance. Good measures link an objective to an observable construct, specify a formula and support action. Neely, Gregory and Platts (1995) distinguish individual measures from the performance-measurement system and its organizational environment. Bourne et al. (2000) further emphasize that systems must be designed, implemented, used and refreshed. Measures can become obsolete when strategy, data or process conditions change.

KPI implementation also requires operational participation. Nunes, Pereira and Alves (2024), in an automotive-components case, show that KPI design benefits from systematic definition and connection to operational processes. Micheli and Manzoni (2010) warn that performance systems can support strategic alignment but also produce unintended consequences when measures are narrow or used mechanically. A governed KPI contract should therefore define objective, business question, owner, population, numerator, denominator, exclusions, refresh time, threshold, drill-down, limitations and revision history.

2.2 Dashboards, cognition and decision use

Dashboards reduce search costs by concentrating information, but display design and information quality affect judgment. Few (2006) argues that dashboards should emphasize comparison, exception and context rather than decoration. Yigitbasioglu and Velcu (2012) review dashboard research and identify visualization, information quality and fit with user tasks as important design considerations. Pauwels et al. (2009) similarly frame dashboards as a way to organize metrics into decision-relevant structures, while cautioning that development requires selection and interpretation rather than indiscriminate display.

Van Elten et al. (2022) add an institutional perspective: dashboards do not simply capture performance; they shape priorities and work routines. This is especially important in SMEs and nonprofits, where a small number of managers may both define

and act upon metrics. If the definition is undocumented, a meeting can become a dispute about the number rather than a decision about the process. The present study operationalizes this risk by calculating parallel definitions and counting when they cross opposite sides of a management threshold.

2.3 Data governance and metric contracts

Data governance specifies decision rights and accountabilities for data-related processes (Khatri and Brown, 2010). Abraham, Schneider and vom Brocke (2019) synthesize governance literature around structural, procedural and relational mechanisms. Metric governance is a focused application: it determines who may define, approve, calculate, change and certify a KPI. Data quality is necessary but not sufficient. Two accurate datasets can produce conflicting KPIs if their populations or aggregation logic differ.

The first article in this publication series established a reproducible data-quality and exception-management framework. The present paper shifts from whether source records satisfy rules to whether summarized measures carry stable meaning into decision forums. This distinction also reflects Mupa-linked research on audit analytics and decision intelligence. Taanisa et al. (2026) connect anomaly detection to SME control strengthening, while Taanisa et al. (2026b) connect variance analysis to risk-aware decisions. These studies support the need for traceability from data to decision, although the empirical estimates reported here arise solely from the Olist analysis.

2.4 Corrective action, recurrence and verified closure

An action item is not a corrective action unless it addresses a cause, has an accountable owner and is verified. Internal-control research shows that material weaknesses can persist and that remediation quality affects reporting outcomes (Goh, 2009; Johnstone, Li and Rupley, 2011). Continuous-audit research similarly emphasizes timely exception detection and follow-up (Alles, Kogan and Vasarhelyi, 2002; Vasarhelyi, Alles and Kogan, 2004). Recent exception-prioritization research identifies alert volume and class imbalance as practical constraints on follow-through (Svanberg et al., 2025).

Chingezi et al. (2026) develop a five-stage architecture for recurrence-sensitive planning, root-cause analysis, accountable remediation, exception monitoring and governance validation. The current paper extends that logic into a KPI-to-action chain. Okebu and Mupa (2026) use a similar feedback principle in competency-based aviation training: early signals have value when they prompt correction and re-verification. Machekera et al. (2026) emphasize lifecycle controls in infrastructure delivery, and Mudzingwa et al. (2026), including Admore Mungwadzi, emphasize prioritized response in resilience systems. These cross-domain studies are cited for governance design, not as evidence of dashboard effect size.

III. METHODOLOGY

3.1 Design and data source

The study used retrospective secondary analysis and transparent scenario modeling. The Brazilian E-Commerce Public Dataset by Olist was downloaded from Kaggle version 7. It contains approximately 100,000 marketplace orders from 2016–2018 and linked order, customer, payment, item, product, seller and review records (Olist, 2018). The analytical spine joined 99,441 orders to persistent customer identifiers, order-item totals, payment totals and exception flags. Purchase dates were grouped into weeks beginning Monday. Eighty-six weeks with at least 100 orders were retained for KPI-definition comparisons to reduce instability from boundary and very low-volume periods.

Data source: Kaggle — Brazilian E-Commerce Public Dataset by Olist

3.2 Paired KPI definitions

Seven KPIs were calculated under a governed definition and a plausible alternative. The governed definition was selected for conceptual alignment with the stated business question, not because it necessarily produced a more favorable value. For each week, both definitions were compared with a pre-specified management threshold. A dispute occurred when the definitions placed the same week on opposite sides of the threshold.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

Table 1. KPI-definition experiment

KPI	Governed definition	Alternative	Trigger
Late delivery	Late delivered orders ÷ delivered orders	Late delivered orders ÷ all orders	>10%
On-time delivery	On-time delivered orders ÷ delivered orders	On-time delivered orders ÷ all orders	<90%
Cancellation	Cancelled orders ÷ all orders	Cancelled ÷ delivered-or-cancelled orders	>1%
Average order value	Mean order-level payment total	Line-weighted invoice value	>R$150
Repeat customer	Orders from repeating persistent customer IDs ÷ orders	Repeated transaction-specific customer IDs ÷ orders	<5%
Delivery cycle	Median purchase-to-delivery days	Mean purchase-to-delivery days	>15 days
Approval latency	Median purchase-to-approval hours	Mean purchase-to-approval hours	>24 hours

3.3 Corrective-action scenario

The action-register scenario used 26,575 exception events generated by 12 reproducible rules: missing relationships, missing mandatory timestamps, invalid event sequences, payment reconciliation differences, non-positive values, late delivery and robust value, freight-ratio and duration anomalies. Critical, high and medium severities received target closure periods of 14, 30 and 60 days. Assumed review times ranged from 10 to 30 minutes by family. Weekly review capacity was fixed at 3,000 minutes for both strategies.

FIFO sorted cases by detection time. The governed strategy sorted by descending severity, then due date and detection time. Cases were processed serially until weekly capacity was exhausted. The scenario deliberately creates capacity pressure to test prioritization. Closure time was measured from the source order date to the end of the modeled closure week. Because these are assumptions rather than observed work logs, results are presented as queue mechanics, not causal productivity estimates.

3.4 Verification sensitivity test

Administrative closure can overstate remediation if the action is marked complete without retesting. A sensitivity test therefore assumed that 12% of administratively closed cases would fail an independent re-performance test and be reopened. A fixed random seed of 42 was used. The 12% value is not estimated from Olist or a field organization; it is a clearly labeled scenario parameter illustrating why “closed” and “verified closed” should be distinct dashboard statuses.

3.5 Analytical safeguards

All formulas, thresholds, filters and scenario parameters were defined explicitly. KPI values were not interpreted as causal outcomes. Persistent-customer identity was separated from transaction-specific customer identifiers. Means and medians were compared without declaring either universally superior. Statistical anomalies remained review candidates rather than confirmed errors. No interviews, user-adoption observations or consenting-organization pilots were fabricated. The framework proposed below is intended for prospective validation.

IV. RESULTS

4.1 Definition-driven value gaps

Across 86 eligible weeks and seven KPIs, alternative definitions produced visibly different levels. Mean weekly on-time delivery was 93.28% among

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

delivered orders but 89.94% when divided by all orders. Payment-based average order value averaged R$161.61, compared with R$180.12 under a line-weighted invoice formulation. Median delivery cycle averaged 10.65 days, while the weekly mean averaged 12.44 days. Median approval latency averaged 0.77 hours, compared with 10.45 hours for the mean, revealing a highly skewed tail.

Table 2. KPI-definition differences and decision agreement

KPI	Governed mean	Alternative mean	Mean relative gap	Disputed weeks	Agreement
Late-delivery rate	6.71	6.49	3.7%	1/86	98.8%
On-time delivery rate	93.28	89.94	3.7%	18/86	79.1%
Cancellation rate	0.80	0.83	2.8%	1/86	98.8%
Average order value	161.61	180.12	10.6%	11/86	87.2%
Repeat-customer rate	2.23	0.00	200.0%	1/86	98.8%
Delivery cycle	10.65	12.44	15.5%	13/86	84.9%
Approval latency	0.77	10.45	182.5%	2/86	97.7%

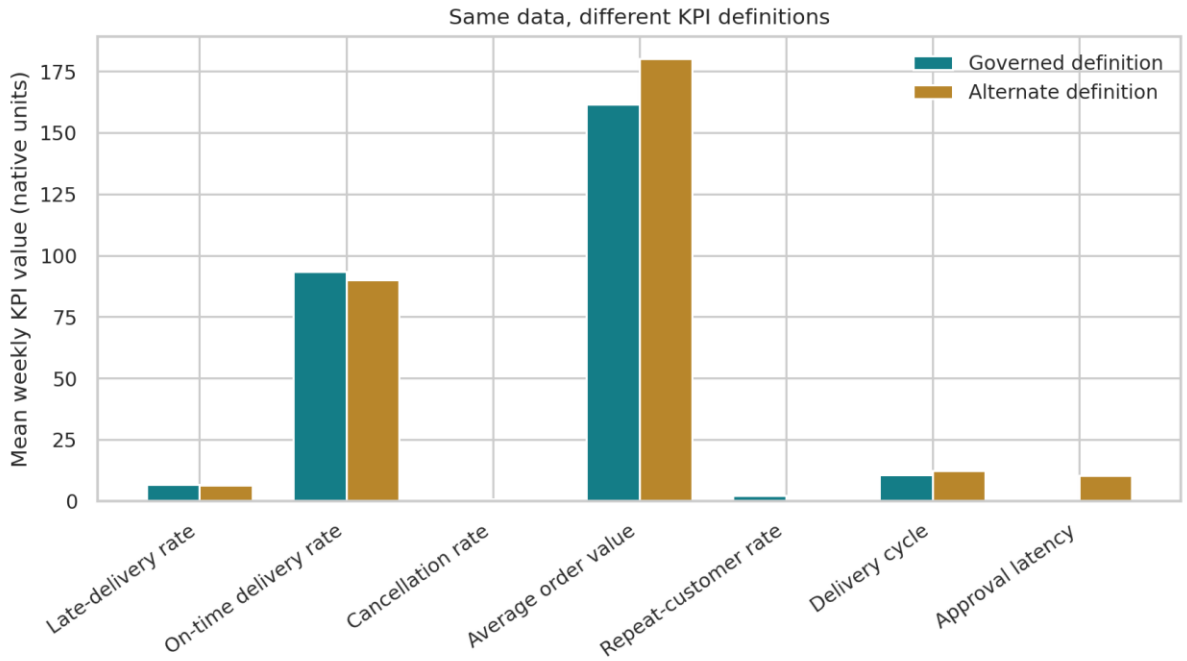


Figure 1. Same data, different KPI definitions

Source: Authors' analysis. Bars retain each KPI's native unit and are intended to show paired definition gaps, not compare magnitudes across KPIs.

4.2 Decision-signal disputes

The definition pairs generated 47 disputes in 602 metric-week comparisons (7.81%). On-time delivery created 18 disputed weeks because using all orders mechanically treats incomplete, cancelled or

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

otherwise non-delivered orders as not on time. Delivery cycle produced 13 disputes because the mean was pulled upward by long-tail cases. Average order value produced 11 disputes because line weighting gives multi-item orders disproportionate influence. The remaining metrics produced one or two disputes each under the selected thresholds.

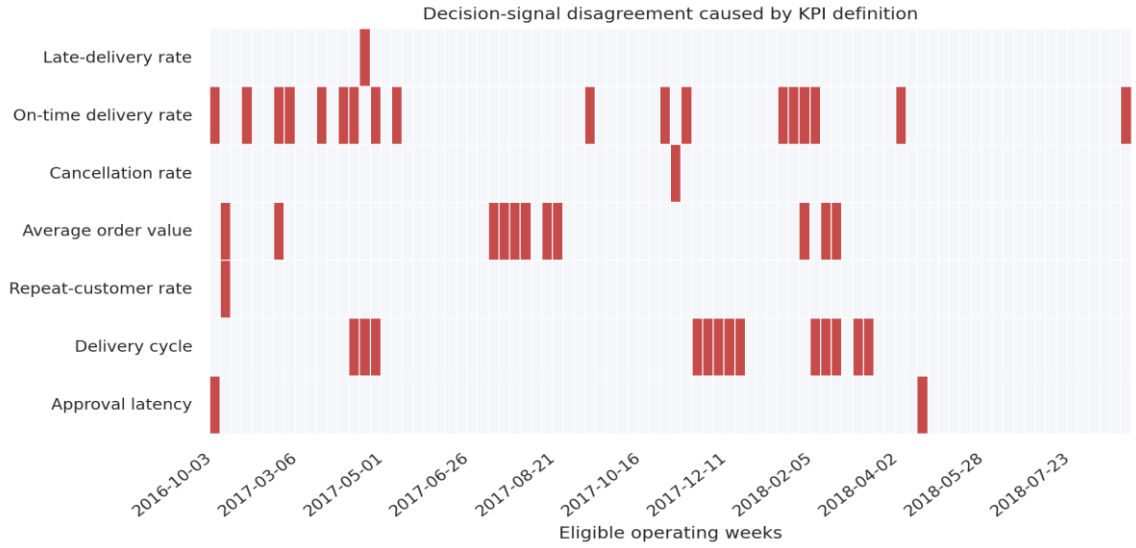
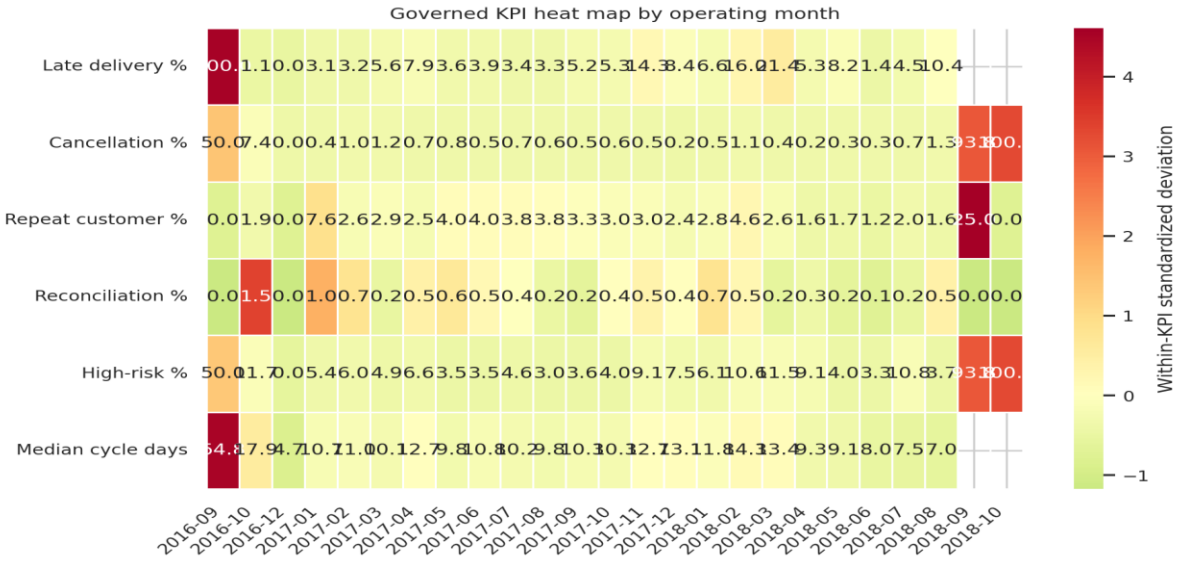


Figure 2. Decision-signal disagreement caused by KPI definition

Source: Authors' analysis. Red cells show weeks in which paired definitions fell on opposite sides of the stated trigger; pale cells show agreement.

The heat map shows temporal clustering rather than a constant error. This matters for governance: a definition may appear harmless in stable periods yet generate disagreement during operational stress or changes in order mix. A metric contract should therefore be tested against historical edge cases before publication, and definition changes should be back-cast to reveal whether prior trends or decisions would have changed.

4.3 Governed KPI operating heat map



© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

Figure 3. Governed KPI heat map by operating month

Source: Authors' analysis. Cell labels show native KPI values; color shows within-KPI standardized deviation, enabling directional comparison without conflating units.

The monthly heat map demonstrates a dashboard design principle: color should communicate deviation within a metric, while labels preserve the actual unit. It also shows why one red cell should not automatically create the same action as another. A cancellation spike, delivery delay and reconciliation exception involve different owners, evidence and response times. Governance connects each cell to a rule-specific playbook instead of treating dashboard color as a universal diagnosis.

4.4 Corrective-action prioritization

The modeled register contained 26,575 cases. Under FIFO, mean closure time for high-severity cases was 482.3 days and the 90th percentile was 814.0 days; only 9.2% met the applicable high-severity SLA. Under severity/SLA sequencing, mean high-severity closure fell to 4.2 days, the 90th percentile to 11.0 days and high-severity SLA compliance reached 100.0%. Overall mean closure remained very long because fixed capacity was insufficient for the high-volume medium-severity tail. Prioritization protected critical work; it did not remove workload.

Table 3. Modeled corrective-action closure performance

Strategy	Cases	Mean closure	Median closure	High mean	High P90	High SLA
FIFO	26,575	385.4 d	366.0 d	482.3 d	814.0 d	9.2%
Governed severity/SLA	26,575	378.1 d	377.0 d	4.2 d	11.0 d	100.0%

Source: Authors' scenario. Weekly capacity = 3,000 analyst minutes; service times and SLAs are specified assumptions. Results are not observed organizational performance.

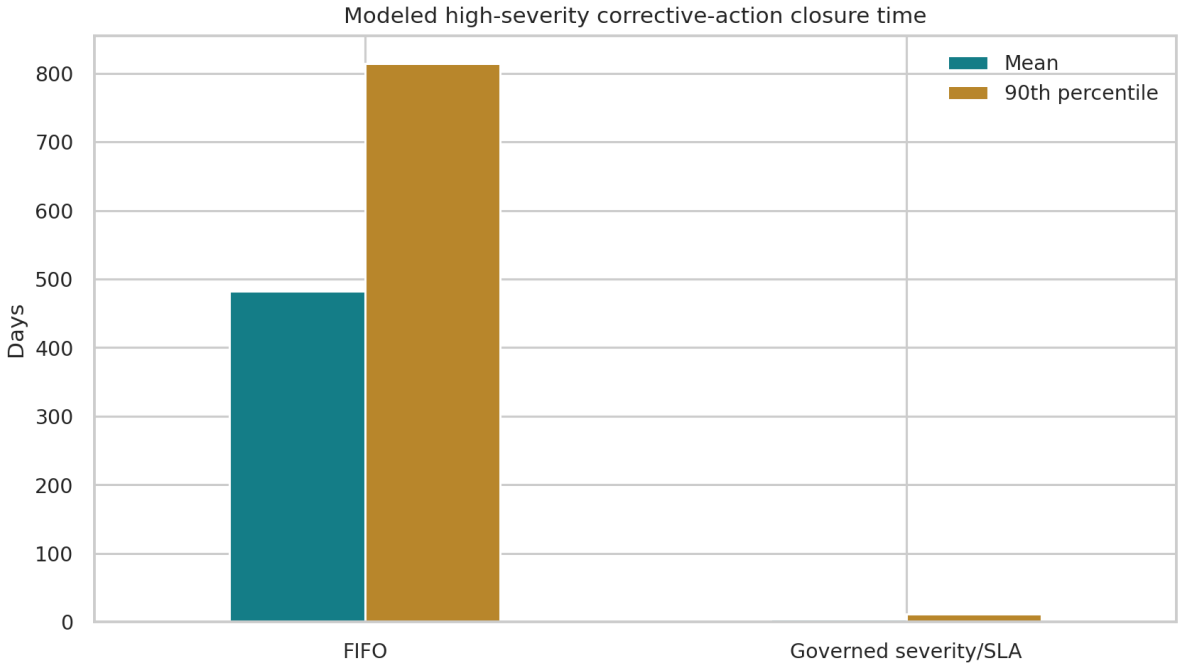


Figure 4. Modeled high-severity corrective-action closure time

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

Source: Authors' scenario analysis. The comparison isolates sequencing under equal capacity.

4.5 Backlog age and the limits of prioritization

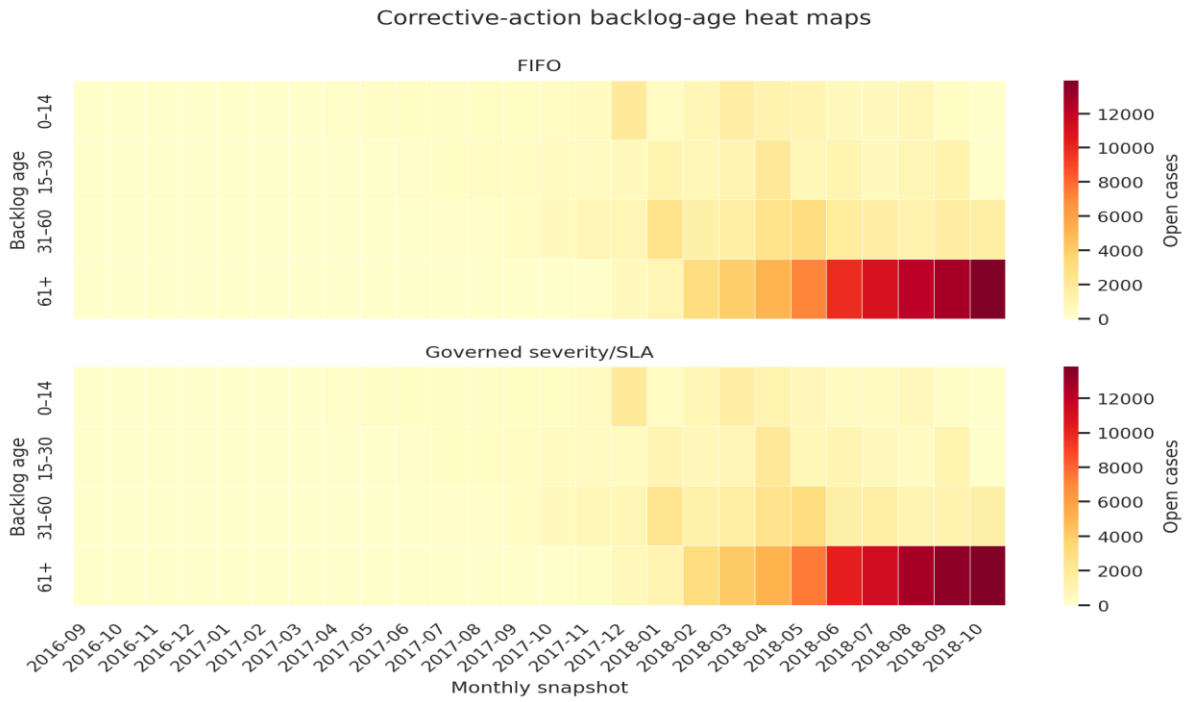


Figure 5. Corrective-action backlog-age heat maps

Source: Authors' scenario. Counts are monthly snapshots of modeled open cases by age band.

The backlog heat maps reveal a governance trade-off. Severity-first sequencing keeps critical cases young but can leave medium-severity service and anomaly candidates ageing when intake exceeds capacity. A credible dashboard must therefore report both risk-weighted SLA performance and the age distribution of the whole backlog. Management responses include increasing capacity, reducing duplicate alerts, batching homogeneous reviews, adjusting thresholds with evidence, automating deterministic checks and fixing recurring source causes.

4.6 Evidence-gated closure

In the sensitivity test, an assumed 12% retest-failure probability reopened 3,191 of 26,575 administratively closed cases. First-pass verified closure was 88%. The result is arithmetic, not an estimated field failure rate, but it illustrates a material reporting difference. A dashboard that displays administrative completion as closure would overstate the number of sustainably resolved cases by the reopened population.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

V. KPI GOVERNANCE AND CORRECTIVE-ACTION CLOSURE FRAMEWORK

5.1 KPI contract

Table 4. Minimum KPI contract

Field	Required content	Governance purpose
Identity	Stable KPI ID, name, version and effective date	Preserve lineage when definitions change.
Intent	Objective, business question and decision supported	Prevent measures without a use case.
Formula	Numerator, denominator, aggregation and unit	Eliminate semantic ambiguity.
Population	Inclusions, exclusions, time window and status logic	Make rates and averages reproducible.
Data	Source tables, owner, refresh time and quality checks	Connect metric trust to source controls.
Interpretation	Target, tolerance, trigger and known limitations	Separate signal from diagnosis.
Action	Decision owner, playbook, escalation and review cadence	Convert visibility into accountable response.
Change control	Approver, reason, test evidence and back-cast result	Avoid silent historical discontinuity.

5.2 Metric dispute and certification workflow

Before publication, the metric owner submits a contract and historical test. The data steward confirms source lineage and quality rules. Finance, operations or the relevant process owner validates the business interpretation. A governance forum approves the definition and threshold. Any later dispute is logged with the competing formula, affected period, decision impact and resolution. Certified metrics carry a version and timestamp; uncertified exploratory measures are visibly labelled and may not drive formal performance consequences.

5.3 Dashboard-to-decision register

Table 5. Decision-register fields

Stage	Required fields	Evidence of control
Signal	KPI ID/version, period, value, threshold, refresh timestamp	The displayed signal is reproducible.
Interpretation	Variance explanation, drill-down, assumptions, competing evidence	The team considered context rather than reacting to color.
Decision	Decision ID, option selected, decision owner, date, rationale	A named person accepted accountability.
Action	Action ID, task, owner, target, severity, dependencies	The decision became executable work.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723142

Stage	Required fields	Evidence of control
Closure	Completion claim, supporting evidence, retest result, verifier	Closure is separated from self-reported completion.
Learning	Outcome, recurrence, threshold or playbook change	The system updates from experience.

5.4 Corrective-action life cycle

The closure cycle contains seven gates. First, register the signal or finding. Second, triage severity and materiality. Third, diagnose the root cause rather than restating the symptom. Fourth, define a corrective action that changes process, system, competence or accountability. Fifth, assign an owner, resources and due date. Sixth, verify implementation and re-perform the relevant control. Seventh, monitor recurrence for a defined period. Administrative statuses—open, in progress, implemented, pending verification, verified closed, risk accepted and reopened—should not be collapsed into a single completion percentage.

5.5 Dashboard architecture

The executive layer should show no more than a small set of certified KPIs, high-risk open actions, SLA breaches, verified closure and recurrence. The management layer adds trends, heat maps and ageing. The analyst layer provides numerator, denominator, filters, rule outputs and record-level drill-through. Every displayed KPI should expose its version, last refresh and denominator. A tooltip or linked definition page is preferable to forcing users to remember formulas.

Table 6. Recommended dashboard measures

Measure	Formula	Why it matters
Metric dispute rate	Certified KPI-periods disputed ÷ KPI-periods reviewed	Tests semantic stability.
Reporting latency	Publication timestamp − data-period close	Measures information availability.
Decision turnaround	Decision timestamp − certified signal timestamp	Measures whether visibility leads to timely judgment.
Action conversion	Signals requiring action with registered actions ÷ actionable signals	Detects dashboard observation without follow-through.
High-risk SLA compliance	High-risk cases verified by due date ÷ high-risk cases due	Protects consequential exceptions.
Verified closure rate	Cases passing independent retest ÷ cases submitted for closure	Prevents administrative closure inflation.
Recurrence rate	Verified cases recurring within control window ÷ verified cases	Measures durability rather than paperwork.

5.6 Field-pilot protocol

A prospective pilot should recruit consenting SMEs or nonprofits and run for at least two baseline and two governed reporting cycles. Baseline measures would include definition disputes, preparation time, reporting latency, dashboard use, decision turnaround, action registration, closure age, SLA compliance, forecast error and recurrence. The intervention would implement KPI contracts, certification, a decision register, severity/SLA

prioritization and independent closure verification. Quantitative analysis could use interrupted time series or difference-in-differences where a comparable group exists. Interviews and meeting observation should examine interpretability, gaming, perceived burden and whether decisions actually change. Such a pilot not the current public-data demonstration would support causal claims.

VI. DISCUSSION

6.1 Governed meaning precedes dashboard design

The most important empirical result is that reasonable formulas can disagree at decision thresholds. In 7.81% of metric-week comparisons, paired definitions prescribed opposite signals. The dashboard did not resolve ambiguity; the metric contract did. On-time delivery was particularly sensitive to denominator choice, while cycle time and approval latency were sensitive to skew and aggregation. These findings reinforce the literature's treatment of performance systems as designed and socially interpreted rather than mechanically discovered (Neely, Gregory and Platts, 1995; Van Elten et al., 2022).

6.2 Priority protects risk but exposes capacity

The closure scenario demonstrates why raw closure counts are inadequate. Severity/SLA sequencing protected high-risk cases under overload, but the medium-severity backlog continued to age. An executive who viewed only 100% high-risk SLA compliance could miss growing service and anomaly debt. A balanced corrective-action dashboard therefore pairs consequence-based measures with backlog volume, age, false-positive rate and recurrence. Svanberg et al. (2025) similarly identify exception volume as a design constraint rather than an incidental inconvenience.

6.3 Closure must be independently verified

The sensitivity test separates activity from control effectiveness. A task can be completed while the failure mechanism remains. Goh (2009) and Johnstone, Li and Rupley (2011) show that remediation is consequential in financial-control settings; Chingezi et al. (2026) extend the principle to repeat operational findings. The framework operationalizes it through an evidence gate: the action owner may assert implementation, but a verifier re-performs the test before the case enters the verified-closed state.

6.4 Implications for resource-constrained organizations

The framework is intentionally low-technology. A KPI contract can be a controlled worksheet; a decision register can be a protected table; an action register can use unique IDs, owners and due dates; and heat maps can be generated in ordinary spreadsheet or business-intelligence software. The scarce resource is governance attention. Organizations should begin with five to ten consequential KPIs, one decision forum and one closure workflow. Adding visualizations before stabilizing definitions increases the speed at which ambiguity travels.

The cross-domain profile sources reinforce this design without substituting for direct evidence. Okebu and Mupa (2026) emphasize early signal, intervention and competence verification; Machekera et al. (2026) emphasize stage-gated lifecycle reliability; Mudzingwa et al. (2026) emphasize prioritized resilience responses; and Chingezi et al. (2026) emphasize closure quality and recurrence. The current analysis contributes a reproducible test of how those governance principles operate when definitions and action queues are made explicit.

VII. LIMITATIONS AND FUTURE RESEARCH

The Olist data represent a Brazilian marketplace from 2016–2018 and are not a representative sample of SMEs or nonprofits. The selected KPI thresholds are analytical triggers rather than contractual service targets. Paired definitions were chosen to demonstrate semantic risk and do not exhaust all valid formulations. Weekly observations are serially dependent, and dispute counts are descriptive. Dashboard use, user adoption, meeting quality and actual decision speed were not observed. The corrective-action scenario relies on assumed review times, capacity, SLAs and serial processing; its large backlog is designed to expose prioritization effects. The 12% retest-failure parameter is a sensitivity

assumption, not an empirical rate. Cross-domain profile publications support framework design but do not validate the Olist estimates.

Future studies should execute the field-pilot protocol, pre-register definitions and thresholds, and collect timestamped signal, meeting, decision, action and verification data. Researchers should test whether KPI contracts reduce meeting time and disputes; whether action conversion mediates the link between dashboard use and performance; whether evidence-gated closure reduces recurrence; and whether governance burden differs by organization size. Experimental work can vary display design independently from definition governance, clarifying whether better visualization adds value after semantic stability is achieved.

VIII. CONCLUSION

A dashboard becomes a decision system only when the measures have governed meaning and the resulting actions have governed closure. In this study, paired KPI definitions produced opposite management signals in 47 of 602 metric-week comparisons. Under modeled capacity pressure, severity/SLA sequencing protected high-risk cases far better than FIFO, but did not eliminate the broader backlog. An evidence-gate sensitivity test showed how administrative closure can overstate durable resolution.

The KPI Governance and Corrective-Action Closure Framework responds with a practical chain: contract the metric, certify the calculation, record the signal, document the decision, assign the action, verify the correction and monitor recurrence. Its contribution is a shift in emphasis—from attractive dashboards to traceable decisions, and from completed tasks to verified control recovery. That shift is especially valuable where staff and technology are limited, because it directs scarce attention toward the definitions and actions that most affect organizational reliability.

REFERENCES

[1] Abraham, R., Schneider, J. and vom Brocke, J. (2019) 'Data governance: A conceptual framework, structured review, and research agenda', *International Journal of Information Management, 49*, pp. 424–438. ScienceDirect

[2] Alles, M.G., Kogan, A. and Vasarhelyi, M.A. (2002) 'Feasibility and economics of continuous assurance', *Auditing: A Journal of Practice & Theory, 21*(1), pp. 125–138. Crossref

[3] Bourne, M., Mills, J., Wilcox, M., Neely, A. and Platts, K. (2000) 'Designing, implementing and updating performance measurement systems', *International Journal of Operations & Production Management, 20*(7), pp. 754–771. Crossref

[4] Chingezi, E., Chingezi, L., Ganyani, L., Yelduora, P.G. and Mupa, M.N. (2026) 'Reducing repeat findings and control failures in operationally intensive U.S. enterprises: A data-driven internal audit framework for remediation, loss prevention, and governance improvement', *World Journal of Advanced Research and Reviews, 30*(3), pp. 2188–2199. Crossref

[5] Few, S. (2006) *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, CA: O'Reilly Media.

[6] Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D. and Neely, A. (2007) 'Towards a definition of a business performance measurement system', *International Journal of Operations & Production Management, 27*(8), pp. 784–801. Crossref

[7] Goh, B.W. (2009) 'Audit committees, boards of directors, and remediation of material weaknesses in internal control', *Contemporary Accounting Research, 26*(2), pp. 549–579. Crossref

[8] Johnstone, K.M., Li, C. and Rupley, K.H. (2011) 'Changes in corporate governance associated with the revelation of internal control material weaknesses and their subsequent remediation', *Contemporary Accounting Research, 28*(1), pp. 331–383. Wiley

[9] Kaplan, R.S. and Norton, D.P. (1992) 'The balanced scorecard—Measures that drive performance', *Harvard Business Review, 70*(1), pp. 71–79.

[10] Khatri, V. and Brown, C.V. (2010) 'Designing data governance', *Communications of the ACM, 53*(1), pp. 148–152. ACM

[11] Machekera, G., Mtemeli, P.S., Dongo, M.M. et al. (2026) 'From trench safety to lifecycle reliability: A project-control framework for water and sewer reticulation delivery in U.S. communities', *World Journal of Advanced Research and Reviews*. ResearchGate (Accessed: 2 September 2026).

[12] Micheli, P. and Manzoni, J.F. (2010) 'Strategic performance measurement: Benefits, limitations and paradoxes', *Long Range Planning, 43*(4), pp. 465–476. ScienceDirect

[13] Mudzingwa, K., Nyamutswa, S.M., Mungwadzi, A.T., Dhegwa, P.Y. and Mupa, M.N. (2026) 'Cybersecurity of critical infrastructures: Challenges, empirical threat analytics and future resilience perspectives', *Iconic Research and Engineering Journals, 10*(2). ResearchGate (Accessed: 2 September 2026).

[14] Neely, A., Gregory, M. and Platts, K. (1995) 'Performance measurement system design: A literature review and research agenda', *International Journal of Operations & Production Management, 15*(4), pp. 80–116. Crossref

[15] Nunes, F., Pereira, R. and Alves, A.C. (2024) 'Implementing key performance indicators and designing a performance dashboard in an automotive components company', *Administrative Sciences, 14*(8), 175. MDPI

[16] Okebu, D.T. and Mupa, M.N. (2026) 'Predictive training analytics for competency-based pilot instruction: Early detection of checkride risk, procedural noncompliance and safety-relevant learning gaps', *World Journal of Advanced Research and Reviews*. ResearchGate (Accessed: 2 September 2026).

[17] Olist (2018) *Brazilian E-Commerce Public Dataset by Olist*, version 7. Kaggle. Kaggle (Accessed: 2 September 2026).

[18] Pauwels, K., Ambler, T., Clark, B.H., LaPointe, P., Reibstein, D., Skiera, B., Wierenga, B. and Wiesel, T. (2009) 'Dashboards as a service: Why, what, how, and what research is needed?', *Journal of Service Research, 12*(2), pp. 175–189. SAGE

[19] Svanberg, J., Öhman, P., Samsten, I., Neidermeyer, P.E., Rana, T. and Danielson, M. (2025) 'Addressing the exception prioritization problem in continuous auditing systems with thresholding', *Intelligent Systems in Accounting, Finance and Management, 32*(4), e70022. Wiley

[20] Taanisa, T., Mukwata, N.A., Sydney, J., Ganyani, L., Maturure, R.N., Chingezi, E., Yelduora, P.G. and Mupa, M.N. (2026a) 'AI-enabled audit analytics for SME financial reporting and anomaly detection: A risk-based framework for early irregularity identification and control strengthening', *World Journal of Advanced Research and Reviews, 30*(3), pp. 1113–1126. Crossref

[21] Taanisa, T., Mukwata, N.A., Sydney, J., Ganyani, L., Maturure, R.N., Chingezi, E., Yelduora, P.G. and Mupa, M.N. (2026b) 'From variance analysis to decision intelligence: Integrating finance audit and data science to improve SME budget governance and risk-aware growth', *World Journal of Advanced Research and Reviews, 30*(3), pp. 666–675. Crossref

[22] Van Elten, H.J., Sülz, S., Van Raaij, E.M., Wehrens, R. and De Bont, A. (2022) 'Big data health care innovations: Performance dashboarding as a process of collective sensemaking', *Journal of Medical Internet Research, 24*(2), e30201. JMIR

[23] Vasarhelyi, M.A., Alles, M.G. and Kogan, A. (2004) 'Principles of analytic monitoring for continuous assurance', *Journal of Emerging Technologies in Accounting, 1*(1), pp. 1–21. Crossref

[24] Yigitbasioglu, O.M. and Velcu, O. (2012) 'A review of dashboards in performance

management: Implications for design and research', *International Journal of Accounting Information Systems, 13*(1), pp. 41–59. ScienceDirect