MEMORANDUM

TO: Executive Management Committee; Analytics Division

FROM: Eyasin, BI Analyst

DATE: January 18, 2026

SUBJECT: Data Quality Audit and Readiness Assessment – Q4 Global Sales Dataset

1.0 Executive Summary

This report outlines the results of the data quality audit conducted on the GameZone Global Sales dataset. The objective of this audit was to prepare the raw sales logs for strategic analysis regarding product portfolio performance, logistics efficiency, and regional market penetration.

The dataset has been successfully cleaned, standardized, and validated against core business logic. However, the audit revealed two systemic operational anomalies—specifically within the logistics tracking and web-checkout systems—that require immediate attention from the Operations and Engineering teams to prevent future reporting errors.

2.0 Scope and Methodology

The audit focused on the structural integrity and logical consistency of the sales data to ensure it can reliably answer the Executive Board's strategic inquiries.

- **Data Source:** Global Sales Master Table (Transactional) and Regional Code Table.
- **Methodology:** The dataset was subjected to a three-step validation process: (1) Standardization of geographical codes, (2) Validation of logical timelines (Order vs. Shipping dates), and (3) Identification of duplicate or erroneous transaction identifiers.

3.0 Critical Findings and Operational Risks

During the validation process, two significant anomalies were isolated that impact data reliability.

3.1 Technical Anomaly: Concurrency Error in Web Transactions

Analysis identified a subset of transactions (0.7% of total volume) where unique Order IDs were duplicated across different User IDs.

- **Root Cause Analysis:** This anomaly is exclusively isolated to **"Nintendo Switch"** consoles purchased via the **Website** channel. This pattern suggests a backend "race condition" during high-traffic intervals, where the system fails to differentiate simultaneous user sessions.
- **Business Impact:** This error compromises the accuracy of Customer Lifetime Value (CLV) calculations and poses a risk of shipping duplication.

- **Recommendation:** The Web Engineering team should review the checkout session handling for high-volume SKUs immediately.

3.2 Operational Discrepancy: Chronological Logic Failures

Approximately 9.15% of logistics records contain a "Shipping Date" that precedes the "Order Date."

- **Root Cause Analysis:** The magnitude of this error suggests a systemic failure in data entry procedures at specific fulfillment centers or a timezone synchronization error in the database.
- **Business Impact:** While revenue figures remain accurate, "Time-to-Ship" and logistics efficiency KPIs cannot be reliably calculated for these specific records.

4.0 Analyst Guidelines for Downstream Analysis

For the analyst tasked with generating the Executive Dashboard, the following constraints must be observed to ensure accuracy:

1. **Revenue & Portfolio Analysis:** The full dataset is cleared for use. The duplicate ID anomaly represents a negligible margin of error for aggregate revenue totals.
2. **Logistics Performance:** Records flagged with *[Error_Code: Date_Logic]* in the accompanying dataset must be excluded from any delivery time averages to avoid skewing results.
3. **Regional Analysis:** All Country and Region codes have been standardized and are cleared for heat-mapping and territory planning.

5.0 Conclusion

The "Golden Dataset" accompanying this report is now structurally sound and ready for ingestion into business intelligence tools. By isolating the anomalies listed above, we have ensured that the upcoming strategic review will be based on verified, high-integrity data.

Attached:

- *Project brief : Find here*
- *Cleaned file : Find here*