

NUSANTARA HOME SUPPLIES

Small Retail Business · Bandung, West Java

AI RESPONSE EVALUATION REPORT

Assessing AI-Assisted Business Operations

Report Date	June 2025
Prepared By	AI Quality Assurance Division
Report Version	v1.0 — Initial Release
Classification	CONFIDENTIAL — Internal Use Only

This report evaluates AI-generated responses across four critical business scenarios to determine their accuracy, relevance, practicality, clarity, and overall business impact for Nusantara Home Supplies.

Executive Summary

Nusantara Home Supplies, a small retail business based in Bandung, West Java, currently manages over 100 customer inquiries per week through WhatsApp, while simultaneously facing persistent challenges in inventory planning and operational reporting. To address these operational gaps, the company has initiated a structured evaluation of AI-assisted business tools.

This report presents the findings of four AI evaluation scenarios conducted by the AI Quality Assurance Division. Each scenario was assessed against five standardised criteria: Accuracy, Relevance, Practicality, Clarity, and Business Impact, using a 1–5 scoring scale.

Overall findings indicate that the AI performed most strongly in structured, data-driven tasks such as inventory analysis and weekly business reporting, while demonstrating moderate performance in open-ended creative tasks such as competitor research. Customer service response generation yielded consistently high scores, affirming AI's potential as a front-line support tool for WhatsApp-based interactions.

Scenario	Composite Score	Performance Band
1. Customer Service Response Generation	4.4 / 5.0	Excellent
2. Inventory Restocking Recommendation	4.2 / 5.0	Very Good
3. Weekly Business Reporting	4.6 / 5.0	Excellent
4. Competitor Research Summary	3.6 / 5.0	Satisfactory

Company Background

Company Name: Nusantara Home Supplies

Business Type: Small Retail Business — Home Goods & Household Supplies

Location: Bandung, West Java, Indonesia

Weekly Inquiry Volume: 100+ customer inquiries via WhatsApp

The company faces two primary operational challenges that AI tools are being evaluated to address:

- High-volume WhatsApp customer inquiries requiring timely, accurate, and professional responses
- Inconsistent inventory planning leading to stockout events and lost revenue opportunities
- Manual operational reporting processes that consume significant management time each week
- Limited visibility into competitor pricing, product range, and market positioning

Scenario 1: Customer Service Response Generation

Business Objective: Automate professional and empathetic responses to high-frequency customer inquiries received via WhatsApp, reducing average response time and improving customer satisfaction scores.

Original Prompt Submitted

Prompt: A customer sent the following WhatsApp message: 'Halo kak, apakah produk sapu lantai ukuran besar masih ada stoknya? Berapa harganya? Saya mau beli 3 pcs.' Please generate a professional, friendly, and informative reply in Bahasa Indonesia, confirming stock availability, providing pricing, offering a bundle discount if applicable, and closing with a clear call to action.

AI-Generated Response Summary

The AI produced a courteous, well-structured WhatsApp reply that addressed all four components of the inquiry: stock confirmation, unit pricing, a suggested bundle discount for the three-piece order, and a clear call to action directing the customer to confirm their order. The tone was consistent with conversational retail communication norms in Indonesia, appropriately using terms of address ('Kak'). The response also proactively included an estimated delivery timeline, a detail not requested but contextually valuable.

Evaluation Criteria Assessment

Evaluation Criterion	Score (1–5)	Assessment Rationale
Accuracy	5 / 5	All factual components — stock status, pricing, discount structure — were logically consistent and free from contradictions given the input data provided.
Relevance	5 / 5	Response directly addressed all elements of the customer inquiry. Proactive inclusion of delivery timeline added value without introducing irrelevant content.
Practicality	4 / 5	Immediately deployable to WhatsApp. Minor adaptation may be required to integrate real-time inventory data from POS systems for live deployment.
Clarity	5 / 5	Language was clear, concise, and appropriately informal. Message structure followed WhatsApp communication norms effectively.
Business Impact	3 / 5	High potential to reduce response workload, but ROI is contingent on integration with live inventory data. Standalone use requires manual verification.
COMPOSITE SCORE	4.4 / 5	<i>Weighted average across all evaluation criteria</i>

Strengths

- Natural language quality is indistinguishable from a trained human customer service agent
- Tone calibration to Indonesian retail communication norms demonstrates strong contextual understanding
- Proactive value-added content (delivery ETA) enhances customer experience

Weaknesses

- Cannot independently verify real-time stock levels without API integration
- Discount thresholds must be manually configured to align with actual store policy

Scenario 2: Inventory Restocking Recommendation

Business Objective: Generate data-driven restocking recommendations to prevent stockout events, optimise inventory turnover, and reduce carrying costs for slow-moving product categories.

Original Prompt Submitted

Prompt: Based on the following weekly sales data for our top 5 SKUs [Sapu Lantai: 45 units sold, 8 remaining; Kain Pel: 60 units sold, 3 remaining; Ember Plastik: 22 units sold, 30 remaining; Sikat WC: 15 units sold, 25 remaining; Deterjen Cair 1L: 90 units sold, 5 remaining], provide a prioritised restocking recommendation with suggested reorder quantities and urgency levels.

AI-Generated Response Summary

The AI produced a structured priority matrix ranking all five SKUs by urgency. Deterjen Cair 1L and Kain Pel were classified as Critical Priority (reorder within 24 hours), Sapu Lantai as High Priority (reorder within 48 hours), and Ember Plastik and Sikat WC as Low Priority with no immediate action required. Recommended reorder quantities were calculated using a days-of-stock methodology, with safety buffer multipliers applied. The response included a brief explanatory note on calculation logic, which enhanced transparency and business usability.

Evaluation Criteria Assessment

Evaluation Criterion	Score (1–5)	Assessment Rationale
Accuracy	4 / 5	Prioritisation logic was arithmetically sound. Reorder quantities applied a reasonable safety stock formula, though supplier lead times were assumed rather than sourced from actual data.
Relevance	5 / 5	All five SKUs were addressed systematically. Output format was directly applicable to purchasing decision workflow without modification.
Practicality	4 / 5	The recommendation is highly actionable for manual purchasing teams. Integration with a live inventory management system would elevate this to a fully automated workflow.
Clarity	4 / 5	Urgency classification and reorder quantities were presented in a clean, readable format. The explanatory methodology note improved confidence in the output.
Business Impact	4 / 5	Directly addresses the company's primary inventory pain point. Implementation could reduce stockout incidents by an estimated 60–70% based on comparable SME deployments.
COMPOSITE SCORE	4.2 / 5	<i>Weighted average across all evaluation criteria</i>

Strengths

- Quantitative prioritisation framework is immediately actionable for non-technical purchasing staff
- Safety buffer logic prevents reactive restocking and supports forward planning
- Transparent methodology note supports managerial trust and adoption

Weaknesses

- Assumes static supplier lead times; dynamic lead times would require additional data input
- Does not account for seasonal demand fluctuations or promotional event spikes

Scenario 3: Weekly Business Reporting

Business Objective: Automate the generation of structured weekly business performance reports, reducing reporting time from an estimated 3–4 hours per week to under 30 minutes, while improving consistency and analytical depth.

Original Prompt Submitted

Prompt: Generate a weekly business performance report for Week 23 (3–9 June 2025) using the following data: Total Revenue: Rp 14,750,000; Transactions: 312; Average Transaction Value: Rp 47,275; Top SKU: Deterjen Cair 1L (90 units); Returns: 4 items; New WhatsApp Inquiries: 107; Conversion Rate: 68%. Format it as a professional management report with sections for Executive Overview, Sales Performance, Customer Activity, Inventory Highlights, and Recommendations.

AI-Generated Response Summary

The AI generated a complete five-section management report within the specified structure. The Executive Overview provided a concise narrative of the week's trading performance against the prior week benchmark. The Sales Performance section presented the revenue figures with a computed daily average and transaction value analysis. Customer Activity highlighted the WhatsApp inquiry conversion rate (68%) alongside a recommendation to analyse drop-off reasons for the 32% unconverted inquiries. The Inventory Highlights section cross-referenced the top SKU with restocking urgency. The Recommendations section produced three specific, data-grounded action items for the following week.

Evaluation Criteria Assessment

Evaluation Criterion	Score (1–5)	Assessment Rationale
Accuracy	5 / 5	All figures were correctly reproduced and derived calculations (daily averages, conversion analysis) were arithmetically accurate. No data entry errors detected.
Relevance	5 / 5	All five requested sections were fully populated. Recommendations were grounded in the specific data provided rather than generic retail advice.
Practicality	5 / 5	Report is immediately submittable to management without editing. Format is consistent with standard SME management reporting conventions.
Clarity	4 / 5	Professional language and logical section flow. Some technical retail terminology (e.g., 'ATV', 'SKU velocity') may require brief orientation for non-technical stakeholders.

Evaluation Criterion	Score (1–5)	Assessment Rationale
Business Impact	4 / 5	Significant time savings (estimated 2.5–3.5 hours per week) for management. Consistent reporting format enables trend analysis across multiple reporting periods.
COMPOSITE SCORE	4.6 / 5	<i>Weighted average across all evaluation criteria</i>

Strengths

- Near-complete elimination of manual reporting effort with zero compromise on analytical quality
- Data-grounded recommendations distinguish this output from generic reporting templates
- Consistent output format enables longitudinal performance tracking

Weaknesses

- Comparison to prior periods requires historical data to be manually included in the prompt
- Technical retail terminology may reduce accessibility for non-specialist readers

Scenario 4: Competitor Research Summary

Business Objective: Produce a concise competitive intelligence summary of local home supplies retailers in Bandung to support pricing strategy, product range decisions, and promotional planning.

Original Prompt Submitted

Prompt: Summarise the competitive landscape for small home supplies retailers in Bandung, West Java. Include information on typical pricing strategies, popular product categories, customer service approaches, and any notable market trends. Use this to provide three strategic recommendations for Nusantara Home Supplies to differentiate itself in the local market.

AI-Generated Response Summary

The AI produced a four-section competitive landscape summary covering pricing strategies, product categories, customer service approaches, and market trends for the Bandung home supplies retail segment. The response identified the prevalence of price-competitive wet market vendors and emerging online marketplaces (Tokopedia, Shopee) as primary competitive pressures. Three differentiation recommendations were offered: (1) building a WhatsApp-native

loyalty programme, (2) introducing a curated premium product tier targeting middle-income households, and (3) leveraging hyperlocal delivery as a competitive moat against marketplace sellers. The response was well-structured but relied on generalised regional market knowledge rather than Bandung-specific primary data.

Evaluation Criteria Assessment

Evaluation Criterion	Score (1–5)	Assessment Rationale
Accuracy	3 / 5	General market observations were broadly accurate for West Java retail dynamics. However, absence of primary data sources or specific competitor analysis limits verifiability and precision.
Relevance	4 / 5	All four requested dimensions were addressed. Recommendations were contextually appropriate for a small retailer in an urban Indonesian market.
Practicality	3 / 5	Recommendations are strategically sound but require local market validation before implementation. They serve as a starting framework rather than a fully actionable plan.
Clarity	4 / 5	Well-organised with clear section headings and accessible language. Strategic logic was easy to follow for non-specialist readers.
Business Impact	4 / 5	If validated with primary research, the three differentiation strategies could meaningfully contribute to customer retention and margin improvement.
COMPOSITE SCORE	3.6 / 5	<i>Weighted average across all evaluation criteria</i>

Strengths

- Structurally sound competitive framework that serves as a useful starting point for deeper research
- Digital marketplace threat identification (Tokopedia, Shopee) reflects genuine and current market dynamics
- WhatsApp-native loyalty recommendation is directly aligned with the company's existing operational infrastructure

Weaknesses

- Lack of Bandung-specific primary data reduces precision and actionability of findings
- Competitor recommendations cannot substitute for physical market surveys or mystery shopping exercises
- Knowledge cutoff limitations may result in outdated competitive intelligence on fast-moving market entrants

Risk Analysis

The following risk matrix identifies key risks associated with deploying AI-generated outputs in a live business context at Nusantara Home Supplies, along with recommended mitigation strategies.

Risk Category	Likelihood	Impact	Mitigation Strategy
Inventory data disconnection — AI responses reflect manual inputs rather than live stock levels	High	High	Integrate AI prompting with POS system export; establish a daily data synchronisation routine
Overreliance on AI for competitor intelligence without primary data validation	Medium	Medium	Supplement AI research with bi-monthly competitor visits and customer feedback surveys
Customer-facing errors — incorrect pricing or stock confirmations sent via WhatsApp	Low	High	Implement a human review checkpoint for all customer-facing AI-generated messages prior to dispatch
Staff adoption resistance — team reluctance to integrate AI tools into existing workflows	Medium	Low	Conduct structured onboarding workshops; demonstrate time-saving benefits with measurable KPIs
Data privacy — customer inquiry data processed through external AI platforms	Low	Medium	Review AI vendor data processing agreements; anonymise customer data in prompts where feasible

Final Recommendations

Based on the evaluation findings and risk assessment, the AI Quality Assurance Division presents the following five strategic recommendations for the Board and Operations Management of Nusantara Home Supplies:

Recommendation 1: Prioritise Weekly Reporting Automation (Immediate)

Scenario 3 achieved the highest composite score (4.6/5.0) and presents the lowest implementation risk. The company should immediately adopt AI-assisted weekly reporting as a standard operational practice. Estimated time savings of 2.5–3.5 hours per week will generate compounding productivity dividends as the business scales.

Recommendation 2: Deploy AI Customer Service With Human Review Layer (Short-Term: 30–60 Days)

Scenario 1 demonstrated excellent language quality (4.4/5.0 composite). A phased deployment is recommended: begin with AI drafting responses and a staff member reviewing before sending, progressing to direct deployment for routine inquiry categories once accuracy benchmarks are confirmed. This approach balances speed and quality assurance.

Recommendation 3: Integrate Inventory AI With POS Export Data (Short-Term: 30–90 Days)

Scenario 2's performance (4.2/5.0) is constrained primarily by data connectivity. Establishing a structured weekly export from the POS system as a standardised prompt input will elevate this scenario's performance to Excellent band and materially reduce stockout incidents. This is the highest-ROI technical integration available to the company.

Recommendation 4: Use Competitor Research as a Starting Framework, Not Final Intelligence (Ongoing)

Scenario 4 (3.6/5.0) is best understood as a structured hypothesis generator rather than a definitive competitive intelligence source. The AI output should be used to design the questions for bi-monthly physical competitor surveys, not to replace them. The three differentiation strategies identified (loyalty programme, premium tier, hyperlocal delivery) merit validation through a 30-day customer feedback study before investment commitment.

Recommendation 5: Establish a Quarterly AI Performance Review Cycle

This report should serve as the first instance of a recurring quarterly review. As AI tools are integrated into operations, evaluation criteria should be updated to incorporate actual performance data (e.g., customer satisfaction scores, stockout frequency, reporting accuracy rates). The scoring framework used in this report provides a consistent baseline for longitudinal comparison.

Analyst's Closing Statement

AI tools demonstrate clear and measurable value for Nusantara Home Supplies across all four evaluated scenarios. The primary determinant of success in each case is not the capability of the AI, but the quality of data inputs and the governance framework surrounding deployment. With structured implementation, human oversight, and quarterly review, AI assistance has the potential to reduce operational overhead by an estimated 30–40% within 12 months while simultaneously improving the consistency and quality of customer-facing communications. The Board is advised to proceed with phased implementation beginning with Recommendations 1 and 2.

This report was prepared by the AI Quality Assurance Division | Nusantara Home Supplies | June 2025 | All rights reserved.