



**Politecnico
di Torino**

Master's degree course in
Management Engineering

Redesign of Supplier Quality Performance Evaluation System at Ariston Group

Master's Thesis

Supervisor:
Prof. Luca Mastrogiacomo

Candidate:
Tommaso Taffi

Academic year 2024/2025

Acknowledgments

I would like to express my gratitude to Professor Luca Mastrogiacomo for his supervision and for the constructive feedback provided during the development of this thesis.

I am also thankful to my company supervisors, Roberto Mainero, Daniele Tomasetti and Marcello Zannotti, for giving me the opportunity to carry out this project within the Supplier Quality Assurance division at Ariston Group, and for their practical guidance when needed.

A special thanks goes to my colleagues at Ariston Group, whose collaboration made the internship both insightful and enjoyable, and to my university colleagues, for their support and companionship throughout the academic journey.

Finally, I am deeply grateful to my family and friends for their continuous encouragement and support.

Abstract

This thesis documents the redesign of supplier quality performance evaluation system at Ariston Group. It first clarifies the existing KPI and governance landscape, then examines the legacy Vendor Rating and Business Warning models, highlighting weaknesses related to evaluator dependence, coarse scaling, and volatility under low exposure. The redesign removes evaluator dependent artifacts by excluding the count of 8D reports, focuses on transactional indicators, namely Parts Per Million and Weighted Supplier Incidents, stabilizes skewed metrics through logarithmic transformation, and applies Bayesian shrinkage to temper noise for low-volume suppliers. The resulting Vendor Rating is more discriminative, interpretable, and statistically more stable, while the Business Warning logic becomes more proportionate and fair. Evidence from company data shows reductions in variance for low-exposure suppliers, improved rank stability, and clearer prioritization signals for supplier development and escalation. The thesis closes with actionable proposals for future evolution, including stronger governance of incident classification, the progressive inclusion of incident gravity within Weighted Supplier Incidents, and deeper integration between measurement and decision support. Overall, the contribution is a transparent and implementable redesign that increases the decision value of supplier performance information while remaining faithful to operational constraints.

Contents

Acknowledgements	ii
Abstract	iii
List of Tables	vii
List of Figures	viii
Glossary	ix
1 Introduction	1
2 Processes and Foundations	8
2.1 New Supplier Introduction and Qualification	9
2.2 Contracts with Suppliers: Key Documentation	12
2.3 Non-Conformity Management and Quality Monitoring	14
2.4 Supplier Audits: Types, Execution, and Follow-Up	18
2.5 Supplier Claim-Back Management	22
2.6 Concluding Remarks	24
3 Legacy Evaluation System	26
3.1 Vendor Rating Calculation Process	27
3.2 The Business Warning Logic	30
3.3 Limitations of the Vendor Rating	33
3.3.1 8D-Based Scoring	34

3.3.2	Fragmentation by Plant and Macrocategory	36
3.3.3	Low Granularity and Limited Discriminatory Power	37
3.4	Inconsistencies in the Business Warning Logic	38
3.5	Summary	40
4	New Evaluation System	42
4.1	Methodology	43
4.2	Architecture of the New Vendor Rating	45
4.3	Input KPIs and Preprocessing	46
4.3.1	Weighted Supplier Incidents (WSI)	47
4.3.2	Parts Per Million (PPM)	49
4.4	The Challenge of Instability in Low-Volume Suppliers	50
4.4.1	Bayesian Shrinkage	52
4.4.2	Logarithmic Normalization	54
4.4.3	Score Mapping and Aggregation	59
4.5	Recap of the New Vendor Rating Calculation Process	63
4.6	Evolution of the Business Warning Logic	66
4.7	Benefits of the New System	70
4.8	Transitioning to the New System	73
5	Future Improvements	75
5.1	Enhancing WSI with Incident Gravity	76
5.1.1	Limitations of Current WSI Indicator	76
5.1.2	Proposed Formula for the New WSI	78
5.1.3	Impact on Vendor Rating and Governance	79

5.2	Toward a Centralized Supplier Governance Platform	81
5.3	Expansion of the Vendor Rating Scope	83
6	Conclusion	86
	References	92

List of Tables

- 4.1 Example of instability for a low-volume supplier. 51
- 4.2 Quarterly average WSI without normalization. 56
- 4.3 Quarterly average log-normalized WSI. 57
- 4.4 Scoring bands for WSI. 60
- 4.5 Scoring bands for PPM. 61
- 4.6 Case study comparison: legacy vs. new Vendor Rating and BW. . . 72

List of Figures

- 1.1 Ariston Group Logo 3
- 2.1 Database extraction of notifications (Vendor names are censored). . 16
- 2.2 Excerpt of the 8D template (D4, D5, D6). 18
- 2.3 Example of evaluation criteria of supplier audits. 21
- 3.1 Legacy Vendor Rating calculation process. 30
- 3.2 Decision matrix of the legacy Business Warning logic. 32
- 4.1 New Vendor Rating calculation process. 64
- 4.2 Decision matrix of the new Business Warning logic. 68

Glossary

8D Structured problem solving method (Eight Disciplines) used to contain, analyse, correct, and prevent supplier non-conformity (incident). In Ariston it is typically opened by the plant Quality Control Manager or designated quality owner when a non-conformity (incident) requires formal corrective action. The process engages the supplier for problem description, root cause analysis, corrective and preventive actions, and validation.

Audit Systematic on-site assessment of a supplier's quality system and process capability.

Bayesian shrinkage Statistical technique that partially pools noisy individual estimates toward a prior or overall mean to reduce variance and improve estimation accuracy, especially with small samples. In this thesis each supplier's normalized KPI is shrunk toward the period overall mean with strength controlled by the smoothing factor (k) and dependent on exposure.

Business Warning (BW) Escalation framework built on the Vendor Rating (VR) that turns persistent low scores into timely, proportionate management actions. It highlights sustained supplier risk, focuses accountability, and triggers coordinated responses to protect operations and customers.

Claim-back Process to charge back the agreed cost of supplier caused non-conformities (incidents).

Corrective Action Plan (CAP) Structured plan to eliminate the cause of a non-conformity (incident) and prevent recurrence, with defined actions, owners, deadlines, and required evidence.

Critical to Function (CTF) Part whose failure compromises the intended function or performance of the product.

Critical to Safety (CTS) Part whose failure may create a safety risk for users.

Detection point Weight (DW) Weight factor used in the calculation of Weighted Supplier Incidents (WSI) indicator that reflects the impact and severity associated with the detection point of the non-conformity (incident); downstream detections (e.g., market or field) are more severe and therefore receive higher weights than upstream detections (e.g., incoming).

Exposure Amount of underlying activity supporting a KPI estimate in a given period. In this thesis it is the delivered volume, namely Goods Receipts (GR), used as the denominator for WSI and PPM. Low exposure increases statistical volatility and motivates the use of Bayesian shrinkage.

Goods Receipts (GR) Quantity of parts received (delivery volume) from the supplier in a given period that forms the exposure base (denominator) for WSI and PPM indicators.

Gravity weight (GW) Weight factor reflecting the technical severity of the non-conformity (incident).

Macrocategory High level classification grouping together different materials with similar characteristics or functions (e.g., *Steel components*, *Electronic components*).

Material (part) Individual item, component, or raw material supplied to Ariston and registered in the company's information system with a unique material code (i.e., part number). In this thesis the terms *material* and *part* are used interchangeably and refer to the same concept.

Non-conformity (incident) Technical event where one or more parts do not meet a requirement. In this thesis the terms *non-conformity* and *incident* are used interchangeably and refer to the same concept. Each event is recorded as a notification and feeds supplier KPIs (e.g., WSI, PPM) and governance decisions.

Notification Formal quality record opened in the company's information system to log a supplier-related non-conformity (incident). Typical fields include detection point, gravity, and scrap quantities. Notifications form the central data backbone for KPI computation and performance reporting.

Parts Per Million (PPM) Supplier quality KPI that counts Scrap Parts (SP) per one million received. Compared against targets defined in the Quality Assurance Agreement (QAA). Lower is better.

Production Part Approval Process (PPAP) Structured approval of supplier parts and processes before serial production.

Purchasing Terms and Conditions (PT&C) Contractual terms that govern commercial rights and obligations in purchasing.

Quality Assurance Agreement (QAA) Signed agreement that defines quality responsibilities, KPI targets, controls, and approval rules between Ariston and the supplier.

Scrap Parts (SP) Non recoverable defective parts that must be scrapped. They are the basis for the PPM calculation, providing the defective parts count used as the numerator.

Smoothing factor (k) Pooling constant used in Bayesian shrinkage to stabilize supplier scores at low exposure, higher k pulls scores more strongly toward the global mean.

Supplier Performance Evaluation (SPE) Structured process to measure and manage supplier performance using auditable data and defined KPIs. In this thesis SPE is mainly related to quality and implemented through Vendor Rating (VR) and governed by Business Warning (BW) to guide development and risk decisions.

Vendor Rating (VR) Composite supplier quality KPI that aggregates governed quality indicators into a single score for supplier evaluation and governance. It

supports supplier feedback, performance reviews, and the Business Warning (BW) escalation framework.

Weighted Supplier Incidents (WSI) Supplier quality KPI that measures the impact of non-conformities (incidents) by assigning each of them a detection point weight and, if used, a Gravity weight (GW). Lower is better.

1 Introduction

Quality has long been recognized as a paramount criterion in supplier evaluation across industries. Foundational studies show that when firms manage buyer-supplier relationships strategically, evaluation systems that emphasize quality and capability correlate with stronger performance outcomes for the buying firm (Carr & Pearson, 1999). Evidence from industrial marketing further links supplier quality and reliability to downstream distributor and customer performance, reinforcing quality as a dominant selection and monitoring criterion (Katsikeas et al., 2004).

In modern manufacturing, firms increasingly rely on extensive supplier networks, making supplier performance evaluation (SPE) a critical supply chain process to ensure product quality and reliability. An effective SPE focuses on tracking key performance indicators (e.g., defect rate, on-time delivery) and providing feedback, thus integrating suppliers into the purchasing firm's quality management system (Ma & Li, 2024). This is vital because a supplier's quality performance directly impacts the buyer's own product quality, customer satisfaction, and cost outcomes (Doshi, 2019). Indeed, research shows that when suppliers consistently fulfill quality requirements and other criteria, the performance of both the supplier and

the buying organization improves in tandem. SPE also serves a strategic role in risk mitigation, as maintaining records of supplier performance helps firms avoid the pitfalls of adverse supplier selection and supply disruptions (Hawkins et al., 2020).

Given these stakes, organizations invest heavily in monitoring and improving supplier quality performance as part of their overall quality management and continuous improvement efforts. Leading manufacturers employ structured supplier performance management programs (including regular evaluations, scorecards, audits, and development initiatives) to ensure that suppliers meet stringent quality standards (Salimian et al., 2021). Such supplier development activities (e.g., technical assistance, training, knowledge transfer) have been shown to elevate suppliers' capabilities and yield higher internal quality outcomes for the buying firm (Modi & Mabert, 2007; Salimian et al., 2021). The payoff from these efforts is significant. For example, one case study reported a 60% reduction in customer-reported defects after the introduction of a rigorous supplier performance evaluation and improvement process, along with a near-elimination of production line interruptions caused by supplier issues (Doshi, 2019). These findings justify why companies view supplier quality performance as a key driver of operational excellence and competitive advantage, warranting substantial investment in SPE systems to continuously monitor, rate, and improve supplier performance.

Against this backdrop, supplier quality performance evaluation emerges as a strategic capability for industrial firms, particularly those operating complex, multi-

tier supply networks. As sourcing footprints expand and product architectures become more modular, managers depend on a small number of indicators to steer supplier development, mitigate risk, and protect quality in the field. The central challenge is to transform heterogeneous operational events into signals that are comparable across sites and over time, sufficiently discriminative to separate truly different performance profiles, and statistically stable even when exposure is low. This thesis addresses that challenge in the concrete context of Ariston Group, a global manufacturer of high efficiency and renewable solutions for thermal comfort, with an international portfolio of brands and longstanding headquarters in Fabriano, Italy. The focus is explicitly applied and rooted in the author's internship experience, with the ultimate aim of improving managerial signaling, supplier feedback, and escalation choices.



Figure 1.1: Ariston Group Logo

The literature provides a clear set of principles that guide the work. First, credible indicators rest on governance rather than on formulas alone. Guidance on KPI development emphasizes unambiguous operational definitions, minimum data sets, documented data flows, and planned data quality checks. This ensures that the same measurement performed by different people in different locations yields the same result and triggers the same decision. In this view, KPIs are auditable signals that initiate investigation and corrective action, not ends in themselves

(Health Information and Quality Authority, 2013; Sreedharan et al., 2024). Second, subjectivity and evaluator behavior are known sources of distortion. Research in performance appraisal shows that discretion introduces leniency and compression effects that weaken the link between true outcomes and recorded scores, while studies in supply management document how procedural variation and rater effects reduce the credibility and risk mitigation value of supplier assessments (Hawkins et al., 2020; Moers, 2005). The practical implication is to engineer evaluator driven variance out of the measurement pipeline, to rely on transactional signals that follow uniform rules, and to avoid counting process compliance artifacts as if they were outcomes when their creation depends on local judgment.

A third theme concerns rating architecture. Psychometric evidence indicates that scales with too few categories lack discriminative power. Two, three, and four point scales tend to compress distinct performance levels into the same class and to dampen sensitivity to change over time. When the managerial task is to rank alternatives, set priorities, or track continuous improvement, the literature recommends scoring choices that preserve information and allow consistent differentiation across the observed distribution (Preston & Colman, 2000). A fourth theme addresses the statistical behavior of incident-based indicators, which are often heavy-tailed and sparse for low-volume suppliers. Two methodological ideas recur. Transformations, especially logarithmic transformations, compress extremes, reduce skewness, and preserve order, which makes downstream analysis more robust when working with long tailed operational data (Chuang & Oliva, 2015). Bayesian shrinkage stabilizes entity level estimates by combining limited, high

variance observations with population information, thereby reducing overreaction to isolated events when exposure is low. The logic has been demonstrated in operations and performance measurement and is well illustrated by empirical work in sports analytics, where shrinking the ratings of players with few matches toward the population mean protects rankings from single outlier results (Baker & McHale, 2017). This stabilization also speaks to the balance between statistical accuracy and incentives in performance measurement, since more accurate estimates can require careful communication to preserve motivational effects (Schwartz, 2021).

These strands of evidence shape how the thesis proceeds. The work privileges indicators with clear operational definitions and auditable data flows. It minimizes evaluator induced subjectivity in the inputs that feed any score. It avoids unnecessary coarseness in the mapping from indicators to ratings so that meaningful differences are not lost. It treats skewed and sparse data with transformations and principled pooling so that comparisons remain fair for both small and large suppliers. The thesis applies these ideas inside Ariston's operating environment in order to improve managerial signaling while keeping the system simple to operate, explain, and audit.

The document follows a practical arc. It begins by explaining in general terms the supplier performance evaluation system at Ariston. The description covers the indicator landscape used for monitoring and governance, the roles involved in producing and consuming those indicators, and the data and reporting flows that connect operational events to measurement. It then presents the legacy Vendor

Rating (VR) and Business Warning (BW) models and highlights their weaknesses and inconsistencies. The analysis considers issues such as evaluator dependent inputs, coarse scaling that compresses distinct performance levels, and the volatility that arises from heavy tailed incident distributions and heterogeneous exposure. The discussion is framed by the literature on subjectivity in appraisal, scale design, and statistical stabilization, which helps clarify why certain design choices reduce credibility and decision usefulness in practice.

Building on this critique, the thesis introduces redesigned Vendor Rating and Business Warning models. The redesign aligns measurement discipline with statistically sound treatment of the underlying data. Inputs are limited to transactional indicators that are created under uniform rules, avoiding evaluator dependent counts. Skewed distributions are stabilized through appropriate transformation. Low exposure volatility is addressed with partial pooling so that adjusted scores reflect underlying capability rather than sampling noise. The intent is to make the evaluation more reliable, more discriminative, and more stable, while also more transparent for internal stakeholders and for suppliers who receive feedback and targets.

The closing part of the work looks forward. It outlines future improvements that could be made to strengthen the governance of incident recording and classification, to deepen integration between measurement and decision support, and to extend the methodological toolkit where appropriate. These proposals are incremental by

design. They aim to evolve a working system in ways that preserve simplicity and auditability while raising the decision value of the information it produces.

In sum, this thesis operationalizes ideas often treated in isolation in the literature and consolidates them into a redesigned supplier quality performance evaluation system. It adopts governed transactional KPIs to remove subjectivity, applies logarithmic normalization and Bayesian shrinkage to sparse heavy tailed indicators, and introduces an interpretable 0 to 100 score mapping with stability conditions for proportionate escalation. Managerially, the redesign produces clearer and more comparable signals across plants, fairer accountability, earlier and more calibrated escalations, and stronger evidence for prioritization, risk assessment, and sourcing choices, thereby turning measurement into a reliable and valuable decision support.

2 Processes and Foundations

Before diving into the analytical core of this thesis - the redesign of the Vendor Rating (VR) and Business Warning (BW) systems - it is essential to contextualize the broader ecosystem in which these tools operate. Ariston Group manages an extensive and diversified network of suppliers, each playing a critical role in maintaining product quality, operational efficiency, and customer satisfaction. To ensure high performance across this supply base, the company has developed a structured and rigorous Supplier Quality Management System, formalized through a series of standardized procedures and aligned with international quality norms. This system governs the entire supplier lifecycle, starting from the early phases of qualification and onboarding, through continuous quality monitoring, issue resolution, and performance evaluation.

The present chapter explores the key components of this system, not only to provide the reader with the necessary background for understanding the subsequent methodological developments, but also to highlight the interdependencies between operational practices, quality control mechanisms, and strategic evaluation models. Each section is based on official internal documents used at Ariston, and aims

to offer a comprehensive and realistic overview of the company's approach to supplier quality management.

2.1 New Supplier Introduction and Qualification

The introduction of new suppliers within Ariston Group follows a clearly structured and detailed procedure aimed at ensuring that potential partners fully align with the company's stringent quality, safety, environmental, and compliance standards. This initial stage is critical, as it forms the foundation upon which all subsequent supplier quality management activities rely.

The process begins with the identification and preliminary screening of potential suppliers through Ariston's *SupplierNet* platform, an advanced portal managed via the Jagger system. SupplierNet enables a centralized evaluation that efficiently assesses multiple dimensions of a supplier's profile before formal engagement. The initial assessment typically comprises several essential evaluation criteria:

- **Health, Safety, and Environment (HSE) Compliance Questionnaire:** Suppliers must demonstrate full compliance with essential HSE standards to ensure responsible and safe operations aligned with Ariston's sustainability goals.
- **Product Conformity and Technical Competence:** Detailed verification of technical and operational capabilities, examining whether the supplier can consistently deliver components meeting Ariston's exact technical specifications and quality requirements.

- **Supplier Quality Assurance (SQA) Questionnaire:** Designed to evaluate the supplier's internal quality management systems, production processes, quality assurance practices, and capacity to meet continuous improvement objectives.
- **Buyer General Data and Trade Compliance Documentation:** Verification of fundamental legal and financial compliance requirements, including adherence to trade regulations, ethical standards, and financial stability.

Suppliers passing this pre-qualification screening phase enter a more comprehensive evaluation phase known as Supplier Qualification. This in-depth evaluation further validates suppliers' capability to consistently deliver high-quality components and services and evaluates their operational reliability, financial stability, and risk profile. The supplier qualification process includes:

- **Risk Analysis and Supplier Criticality Assessment:** Suppliers are assessed for their risk profile and classified according to product criticality as either "Critical to Safety" (CTS), "Critical to Function" (CTF), or non-critical. Suppliers providing safety-critical or critical to function components undergo more rigorous scrutiny and stricter qualification criteria to mitigate potential operational and customer safety risks.
- **Technical and Operational Verification (Potential Audit):** A detailed audit assesses production processes, safety protocols, quality control mechanisms, and management practices. The audit includes comprehensive verification

of documented procedures, operational capabilities, employee training, manufacturing equipment, and capacity to adhere to quality management systems. Suppliers scoring below the Ariston standard (minimum required score of 75) must submit a Corrective Action Plan (CAP) to address gaps and deficiencies identified during the audit.

- **Production Part Approval Process (PPAP):** Depending on supplier classification and the type of components supplied, suppliers may be required to undergo a PPAP. The PPAP is crucial in verifying that suppliers can consistently manufacture components according to Ariston's design and quality specifications before initiating large-scale production. For new and safety-critical parts, a detailed PPAP (level 3) is typically required, involving extensive documentation including process flow diagrams, control plans, product samples, and testing results. For existing or less critical components, a simplified PPAP (level 1) may be sufficient.

Upon successful qualification, suppliers are formally onboarded into Ariston's internal systems for ongoing monitoring and governance. The robust and methodical introduction and qualification process adopted by Ariston Group ensures that all new suppliers fully comply with the company's demanding standards from the outset. This structured approach mitigates risks associated with new supplier integration, fostering sustainable, reliable, and high-quality relationships that ultimately underpin Ariston's global market success.

The next step in the process, described in the following section, entails the formalization of the relationship through the signing of mandatory contractual documents that define legal, operational, and quality-related obligations.

2.2 Contracts with Suppliers: Key Documentation

Once a supplier successfully completes the qualification phase and is approved for collaboration, the relationship is formalized through the signing of mandatory contractual documents that define the legal, operational, and quality framework of the partnership. This step is essential not only for compliance purposes but also as a core component of Ariston Group's Supplier Quality Management System, ensuring that all expectations regarding product conformity, delivery performance, and continuous improvement are explicitly communicated and contractually enforced.

The **Purchasing Terms and Conditions (PT&C)** constitute the general commercial framework of the supplier relationship. This document establishes pricing and payment terms, delivery obligations, liability clauses, intellectual property protection, and dispute resolution mechanisms, as well as mandatory compliance requirements related to environmental, ethical, and legal standards. By signing the PT&C, suppliers formally acknowledge and accept Ariston's baseline expectations, which serve as a uniform contractual standard across the global supply base.

Equally critical is the **Quality Assurance Agreement (QAA)**, which forms the cornerstone of Ariston's supplier quality governance. The QAA translates the

company's quality philosophy into specific, measurable obligations for suppliers, defining indicators targets, epidemic thresholds, traceability rules, and non-conformity (incident) management procedures. It obliges suppliers to actively participate in root cause analysis and corrective actions for any non-conformity, typically applying the 8D problem-solving methodology, and to cooperate fully during process, product, or system audits. Additionally, the QAA promotes a continuous improvement approach, requiring suppliers to maintain and enhance quality performance over time. Importantly, one of the QAA appendices presents and details Ariston's Vendor Rating and Business Warning systems, defining the performance indicators, thresholds, and escalation logic that govern the monitoring of supplier performance. This aspect is central to the present thesis, as the following chapters focus on the redesign of the Vendor Rating and Business Warning frameworks to enhance objectivity, stability, and strategic value.

In addition to these two core contracts, Ariston often implements **Logistic Agreements** and **Spare Parts Agreements**, which provide further operational clarity. The Logistic Agreement specifies delivery schedules, packaging and labeling requirements, and inventory management rules, ensuring the efficiency and continuity of the supply chain. The Spare Parts Agreement guarantees long-term component availability to support after-sales service and product lifecycle management, which is crucial in the home appliance sector where product reliability and customer satisfaction are tightly linked to spare parts supply.

All contractual documents are then integrated into Ariston's Vendor Master Data Management system, linking legal and quality obligations directly to the supplier's digital profile. This integration enables centralized monitoring of compliance, facilitates coordination among Quality, Procurement, and Logistics teams, and provides the contractual foundation for supplier performance evaluation, including the Vendor Rating and Business Warning processes.

In summary, the formalization of supplier contracts in Ariston is a strategic quality management step: it translates supplier qualification outcomes into binding obligations, ensures that suppliers are aligned with the company's operational and quality standards, and establishes the framework for systematic performance monitoring and continuous improvement, forming the contractual backbone of the processes analyzed in the core chapters of this thesis.

2.3 Non-Conformity Management and Quality Monitoring

Once a supplier has successfully completed the qualification phase, signed all contractual agreements, and been formally onboarded, the real operational collaboration begins. At this stage, the supplier transitions from prospective partner to an active contributor to Ariston's production and supply chain, and the continuous monitoring of its quality performance becomes central to the relationship. From this moment, every delivery, component, and interaction generates performance data, which feeds into the company's supplier evaluation framework and ultimately influences Vendor Rating and Business Warning assessments.

Supplier quality monitoring in Ariston relies on a structured process of inspection controls and non-conformity management. All materials delivered by suppliers are subjected to multiple levels of quality control inspections, which can take place at different points in the value chain. Typical detection points include incoming inspection at the plant warehouse, in-process checks along the production line, statistical process control (SPC) checks, and field or market detections, such as customer returns or warranty claims. Each identified non-conformity is formally recorded in the internal database as a quality alert, known internally as a *notification*, which captures all the essential details of the event, such as:

- **Detection Point (W1, W2, W3, W4):** representing the stage where the non-conformity was identified, with W1 corresponding to field or market detections, W2 to incoming inspection, W3 to production line discovery, and W4 to statistical or sample-based checks.
- **Gravity (A0, A1, B, C):** classifies the severity of the non-conformity, ranging from critical issues that may trigger immediate containment actions to minor deviations that require monitoring but not immediate escalation.
- **Scrap Quantity:** specifies the exact number of scrap (defective) parts linked to the incident, later used to calculate the Parts Per Million (PPM) indicator.

Figure 2.1 provides an extract of the internal database where all supplier incidents are recorded as notifications.

Notification	Part	Material	Material Description	Vendor	Vendor name	Object start code text	Tap	Notification Status	Damage Code	Code group	Problem code text	DC	Notif. date	Qty (est.)
200849211	VN	000303022803	EIWH MED. HIGH HMI(VN)	7018679	*****	Vendor Cause	W3	OSNO NOTE	AA50	REI31-AA	PCB defect	C	03/06/25	1,000
200849076	WP	0003030504002	GAL-EVO2 HE MED	40005879	*****	Vendor Cause	W2	NOCO NOTE NOPT				A1	03/06/25	16,000
200849174	WU	035106003102	6L SAFETY VALVE 5.5BAR	40003546	*****	Vendor Cause	W2	NOPR NOTE NOPT				A1	03/06/25	250,000
200849176	WU	420050036101	NEUTRAL ADHESIVE LABEL THREE LABELS PAPE	40006191	*****	Vendor Cause	W2	OSNO NOTE NOPT				C	03/06/25	1,000,000
200849090	VN	460130022800	PCBA EWH BLUE-READY AN2 RS	7018621	*****	Vendor Cause	W3	OSNO NOTE	AA50	REI31-AA	PCB defect	C	03/06/25	1,000
200849175	WU	620000026800	SELF TAP. SCREW ST6.5*16	40002009	*****	Vendor Cause	W2	NOPR NOTE NOPT				A1	03/06/25	375,000
200849251	01	7200110244500	BRACKET REINFORCEMENT EASYHANDLING	1990	*****	Vendor Cause	W3	OSNO NOTE NOPT				B	03/06/25	100,000
200849276	WN	000303022803	EIWH MED. HIGH HMI(VN)	7018679	*****	Vendor Cause	W3	OSNO NOTE	AA50	REI31-AA	PCB defect	C	04/06/25	1,000
200849275	VN	000303022900	EIWH TEMP DISPLAY HMI	7018679	*****	Vendor Cause	W3	OSNO NOTE	AA50	REI31-AA	PCB defect	C	04/06/25	1,000
200849295	WP	440010044000	UPPER CARTON BOX 12-2Z ENTRY	40002330	*****	Vendor Cause	W2	OSNO NOTE NOPT				B	04/06/25	42,000
200849285	AK	580000056900	STEEL PIPE FLOW CONN. BOILER GENUS 45	4658	*****	Vendor Cause	W3	OSNO NOTE NOPT	PI02	REIAK-PI	PIPE	C	04/06/25	2,000
200849284	AK	580000057000	STEEL PIPE FLOW CONN. BOILER GENUS 65	4658	*****	Vendor Cause	W3	OSNO NOTE NOPT	PI02	REIAK-PI	PIPE	C	04/06/25	2,000
200849283	AL	580020381500	COPPER PIPE SYSTEM RETURN	4658	*****	Vendor Cause	W3	OSNO NOTE NOPT	PI04	REIAK-PI	PIPE	C	04/06/25	1,000
200849298	23	720010153801	SHELTER VASCHETTA PROTEZIONE ACQUA GALEV	1990	*****	Vendor Cause	W3	OSNO NOTE	Z02	REI23-Z	BROKEN/SCRATCHED	D	04/06/25	1,000
200849296	WH	740040020800	Overheat(thermostat)110°C (35°C) FFAST	40002112	*****	Vendor Cause	W3	NOCO NOTE NOPT				A1	04/06/25	2,000
200849299	23	7401300653402	FAN FAN 118 325V GALEVO2 HE FLANGE	6027841	*****	Vendor Cause	W3	OSNO NOTE	A04	REI23-A	DENTED	D	04/06/25	2,000
200849421	VN	000303022803	EIWH MED. HIGH HMI(VN)	7018679	*****	Vendor Cause	W3	OSNO NOTE	AA50	REI31-AA	PCB defect	C	05/06/25	1,000
200849482	05	440050158700	BOX BOARD ANDRIS RS 15	720	*****	Vendor Cause	W3	OSNO NOTE	IM09	REI05-IM		C	05/06/25	1,000
200849482	05	440050158700	BOX BOARD ANDRIS RS 15	720	*****	Vendor Cause	W3	OSNO NOTE	IM09	REI05-IM		C	05/06/25	2,000
200849448	AL	460090048203	GROUP CONNECTION CONTROLLER RV561.843/1	29006631	*****	Vendor Cause	W3	OSNO NOTE NOPT	003	REIAL-ME		C	05/06/25	1,000
200849440	23	680080047800	PUMP MODULATING SM 2ND STEP UPD	7527	*****	Vendor Cause	W3	OSNO NOTE	C02	REI23-C	DOESN'T WORK	D	05/06/25	2,000
200849449	AL	540070430601	GRILL FRONT GRILL ASS. AESTHETIC TOP COV	6608	*****	Vendor Cause	W3	OSNO NOTE NOPT	LV17	REIAL-LV		C	05/06/25	2,000
200849486	05	660030060000	ASSEMBLED HE2 VLS DRY-P-TECH 1.5KW IN	20339	*****	Vendor Cause	W3	OSNO NOTE	RE25	REI05-RE		A0	05/06/25	5,000
200849487	05	660030060100	ASSEMBLED HE1 VLS DRY-P-TECH 1.5KW OUT	20339	*****	Vendor Cause	W3	OSNO NOTE	RE25	REI05-RE		A0	05/06/25	4,000
200849334	WU	680010025000	SOUND INSUL SPONGE 70X15X10	40007965	*****	Vendor Cause	W2	OSNO NOTE NOPT				C	05/06/25	1,900,000
200849333	WU	720010275100	COVER II EVO UPD R290 BASEMENT-PP	40006192	*****	Vendor Cause	W3	NOPR NOTE NOPT	A802	REI31-AB	Look bad	B	05/06/25	1,000
200849450	AL	740030003100	IGNITION UNIT MODEL 271 W	6023606	*****	Vendor Cause	W3	OSNO NOTE NOPT	RE07	REIAL-RE		C	05/06/25	1,000
200849471	23	7401802330901	CABLE ASS. CABLAGGIO ASKOLL PUMP ENTRY H	2687	*****	Vendor Cause	W3	OSNO NOTE	Q06	REI23-Q	DEFECTIVE	D	05/06/25	1,000
200849625	AK	0ARP295	GASKET FOR GAS/AIR BUTTERFLY	6003733	*****	Vendor Cause	W3	OSNO NOTE NOPT	IS01	REIAK-IS	INSULATION / GASKETS	C	06/06/25	0,000
200849638	23	4000800080400	MIXER AIRGASZ CARTRIDGE 25	2663	*****	Vendor Cause	W3	OSNO NOTE	Z03	REI23-Z	DEFORMED/INCOMPLETE	D	06/06/25	1,000
200849634	23	420050046103	NEUTRAL ADHESIVE LABEL ENERGY COMBI	12417	*****	Vendor Cause	W3	OSNO NOTE	W01	REI23-W	WRONG/MISSING PRINT	D	06/06/25	1,800,000
200849635	23	540070317501	SIDE CASING PAN. PANN LAT FONQASS R/L W	6003374	*****	Vendor Cause	W3	OSNO NOTE	L3	REI23-L	SCRATCHED	D	06/06/25	6,000
200849578	23	540070409901	CASE BOTTOM EVO2 HE NEW STD PNEU FILL ID	2351	*****	Vendor Cause	W3	OSNO NOTE	Z05	REI23-Z	WRONG CODE	D	06/06/25	72,000
200849637	23	640060010600	MAIN DOOR CIRCOND 2 ELECTRODES XTRATECH	17036	*****	Vendor Cause	W3	OSNO NOTE	A03	REI23-AI		D	06/06/25	1,000
200849616	WU	680010024700	SOUND INSUL COMPRESSOR BODY INSULATION	40007965	*****	Vendor Cause	W2	OSNO NOTE NOPT				B	06/06/25	200,000
200849615	WU	680010024700	SOUND INSUL COMPRESSOR BODY INSULATION	40007965	*****	Vendor Cause	W2	OSNO NOTE NOPT				B	06/06/25	800,000
200849619	23	740180035101	CABLAGGIO LWI DISPLAY	2687	*****	Vendor Cause	W3	OSNO NOTE	Q06	REI23-Q	DEFECTIVE	D	06/06/25	1,000
200849620	23	740180271300	CABLE. NTC CONNECTION/H/E	2687	*****	Vendor Cause	W3	OSNO NOTE	Q06	REI23-Q	DEFECTIVE	D	06/06/25	1,000
200849631	23	740180271601	WIRING HV GAS VALVE GALEVO 2 HE WIRING	20344	*****	Vendor Cause	W3	OSNO NOTE	Q06	REI23-Q	DEFECTIVE	D	06/06/25	1,000
200849632	23	740180282500	CABLE LWI COMB. COND ENTRY HE UPDATE	2687	*****	Vendor Cause	W3	OSNO NOTE	Q06	REI23-Q	DEFECTIVE	D	06/06/25	1,000
200849630	05	440050158600	CARTON_BOX BOARD ANDRIS RS 10	720	*****	Vendor Cause	W3	OSNO NOTE	IM13	REI05-IM		B	09/06/25	1,000
200849831	05	440050221000	BOX BOARD ANDRIS RS 80 UL (BUFFER)	720	*****	Vendor Cause	W3	OSNO NOTE	IM03	REI05-IM		C	09/06/25	1,000
200849857	05	740170024700	POWER SUPPLY CABLE 3X1 L-1700 T125 NO PL	4945	*****	Vendor Cause	W3	OSNO NOTE	CV03	REI05-CV		A0	09/06/25	2,000
200849806	VN	000303015203	EIWH ENTRY(EP) G2	40005625	*****	Vendor Cause	W3	OSNO NOTE	AA50	REI31-AA	PCB defect	C	10/06/25	1,000
200849858	WP	099051012000	screw/cross recessed/M3.5x16	40008266	*****	Vendor Cause	W2	NOCO NOTE NOPT DFLF				A1	10/06/25	1,000,000
200849805	VN	440010183200	BOX EASY PH	7008663	*****	Vendor Cause	W3	OSNO NOTE	C802	REI31-CB	BROKEN	C	10/06/25	1,000
200849807	VN	460070029102	WATER FLOW SWITCH ENTRY SEP	7013626	*****	Vendor Cause	W3	OSNO NOTE	AA75	REI31-AA	Switch defect	C	10/06/25	1,000
200849759	AK	540070325700	Back Plate Low 1307 TXXL	3487	*****	Vendor Cause	W3	OSNO NOTE NOPT	005	REIAK-LA		C	10/06/25	1,000
200849961	VN	192860	H.E.T.P SS - S.B. BAD 18KW/240v	40007911	*****	Vendor Cause	W3	OSNO NOTE	AA13	REI31-AA	HE defect	C	11/06/25	13,000
200849955	VN	460000008700	CONTACT STAT THERMOSTAT 90°	40005455	*****	Vendor Cause	W3	OSNO NOTE	AA59	REI31-AA	Thermostat defect	C	11/06/25	1,000
200849956	VN	460070029102	WATER FLOW SWITCH ENTRY SEP	7013626	*****	Vendor Cause	W3	OSNO NOTE	AA75	REI31-AA	Switch defect	C	11/06/25	2,000

Figure 2.1: Database extraction of notifications (Vendor names are censored).

These systematically recorded incidents are not only used for immediate containment and resolution but also serve as input for supplier performance indicators.

The procedure defines clear rules for supplier involvement depending on the nature and severity of the non-conformity. For significant issues, the supplier is promptly notified and is required to analyze the root cause and propose corrective and preventive actions, often using the structured 8D problem-solving methodology.

An 8D report is issued when the non-conformity is considered impactful to quality, safety, or operational continuity, or when it indicates systemic weaknesses in the supplier's processes. The method was first documented in the Team Oriented Problem Solving manual by Ford Motor Company (1987), then evolved into the Global 8D format that is widely adopted in industry (AESQ Strategy Group, 2021; Barsalou & Perkin, 2025).

The "disciplines" are named, practical steps in a standard problem-solving sequence and the ones in current practice are (eight named disciplines plus the preliminary D0, hence nine steps overall):

- D0** Planning and immediate containment while preparing the team.
- D1** Form a cross-functional team with appropriate product and process knowledge.
- D2** Describe the problem precisely, including scope, timing, and evidence.
- D3** Install interim containment to protect customers while analysis proceeds.
- D4** Determine and verify root causes and identify any escape points in detection and control.
- D5** Select and verify permanent corrective actions that address the verified root cause.
- D6** Implement corrective actions and validate their effectiveness in operation.
- D7** Prevent recurrence by updating specifications, controls, training, and systems.
- D8** Congratulate to the team, close the case and capture lessons learned for reuse.

Figure 2.2 shows the section of the 8D template dedicated to D4 to D6, namely root cause analysis, selection and verification of permanent corrective actions, and implementation with effectiveness validation, ensuring that suppliers systematically investigate nonconformities and define and verify appropriate countermeasures.

ARISTON GROUP														
Supplier 8D Report - D4-D5-D6 OCCURENCE														
8D Number		First Issue Date		Update Date		Ariston Plant		Supplier						
Root causes analysis and Action Plan - Occurrence: Why did the problem occurred?														
Factors identification		D4 - Root cause						D5 - Corrective Action				D6 - Implementation		
		5 Why's analysis						Action Plan						
Shikane's Log	Influent Factors	Why 1	Why 2	Why 3	Why 4	Why 5	Reproduced?	Corrective Action	Photo Number	Owner	Planned	Done	Verified? Yes/No	Closure Date
Machine														
Method														
Material														
Man														
Design														
Environment														

Figure 2.2: Excerpt of the 8D template (D4, D5, D6).

Managerially, 8D provides a traceable and auditable path from notification to verified cause and action. Beyond industrial practice, peer reviewed evidence shows that structured 8D deployment with training can reduce the time needed to find root causes over multi year horizons (Barsalou et al., 2023).

In essence, non-conformity management is the operational backbone of supplier quality monitoring in Ariston, transforming raw notifications into actionable data, ensuring that suppliers actively participate in corrective and preventive actions, and laying the foundation for a data-driven and accountable supplier governance system.

2.4 Supplier Audits: Types, Execution, and Follow-Up

Complementing the continuous monitoring and non-conformity management activities, supplier audits represent a fundamental tool in Ariston Group's supplier

quality governance framework, ensuring that suppliers not only deliver conforming products but also maintain robust internal processes and quality systems capable of supporting long-term compliance. Audits serve both as preventive and corrective instruments, aimed at validating process maturity, identifying risks, and enforcing continuous improvement.

Ariston employs several types of supplier audits, each with a distinct purpose in the lifecycle of the supplier relationship:

- **Potential Audits:** performed during the qualification phase of new suppliers to assess whether they meet Ariston's baseline requirements before entering the approved vendor list.
- **Process Audits:** conducted when new products, new processes, or significant changes in supplier operations are introduced. They focus on the production flow, process controls, equipment, and operator competence to verify that the supplier can consistently meet Ariston's specifications.
- **System Audits:** conducted for suppliers that do not hold recognized ISO certifications, or when Ariston needs to verify the effectiveness and completeness of the supplier's quality management system against internal standards.

Each audit type is risk-based and linked to the criticality of the supplied product, with Critical to Safety (CTS) and Critical to Function (CTF) components subjected to stricter auditing requirements.

The audit execution process follows a structured and well-documented methodology. Audits are planned in advance and typically carried out by cross-functional teams, including Supplier Quality Assurance (SQA) representatives, procurement personnel, and, when necessary, product engineering specialists. Each audit includes preparation of an agenda, review of relevant documentation such as control plans, work instructions, and quality records, and on-site verification of production processes and inspection points. Special attention is paid to the supplier's ability to identify and react to non-conformities, traceability systems, calibration and maintenance of equipment, operator training, and implementation of corrective and preventive actions. More in detail, the audit agenda typically includes the following topics:

1. **Process Flow and FMEA:** verification of updated process flow diagrams and FMEA (Failure Mode and Effects Analysis) to ensure potential failure modes are identified and mitigated.
2. **Control Plans and Operator Instructions:** assessment of control plans, work instructions, and operator training to confirm consistent and correct execution of tasks.
3. **Capability Studies:** examination of process capability indices (C_p , C_{pk}) to ensure processes operate within specified tolerances.
4. **Specification Management:** verification that the supplier keeps up-to-date customer specifications and correctly applies them.

- 5. Warehouse and Reject Segregation:** review of warehouse layouts, FIFO (First In, First Out) methodology, and segregation of non-conforming components.
- 6. Traceability:** verification of systems linking raw materials, semi-finished goods, and finished components to ensure traceability back to suppliers and production batches.
- 7. Maintenance and Calibration:** assessment of maintenance schedules and calibration records for equipment and measuring devices.

Figure 2.3 provides an example of the evaluation criteria used during an audit, visually illustrating the type of questions posed to suppliers and the scoring structure adopted to measure compliance.

N°	Question	Score		
		100	50	0
1	Are the resources available and can the serial production be ensured?	The supplier has capacity for our products, even in the event of a bottleneck in the process	The supplier has only enough capacity with action (new shift, new machines)	Not enough capacity
2	Is a development and the process development available?	The supplier has a development and a process Development procedure for all products	The supplier has only a process development (depends on the product)	The supplier has no development and no process development
3	Supplier audits are planned and performed?	Internal process audits are regularly conducted	Internal process audits are conducted but not regularly	Internal process audits are not conducted
4	Are the in- / outgoing parts stored and checked in a appropriate way?	Parts / materials / semi finished and finished products are located in appropriate areas, with a physical layout aligned to production flow	Parts / materials / semi finished and finished products are located in appropriate areas, without a physical layout aligned to production flow	Parts / materials / semi finished and finished products are not located in appropriate areas
5	Is measurement room / laboratory available?	Is there a measuring room with the required equipment according to our requirements from the part	A measuring room is available but with poor equipment	No measuring room available
6	Is in the production quality gates / inspection area available?	The supplier has Q - Gates and make inspection, with documentation during the production	The supplier has Q - Gates and make inspection, without documentation	The supplier does not have Q-Gates and make no inspection
7	Control and monitoring of measuring equipment available?	Measuring equipment monitoring is available and the measuring equipment is on time	Measuring equipment monitoring is available but not all measuring equipment is on time	No Measuring equipment monitoring available
8	Are bad parts blocked in an appropriate way?	A lockable storage facility with an overview is available	The storage is not lockable without an overview is available	No blocking storage available
9	Are the employees trained and qualified in an appropriate way?	A skill matrix is available and the employees are trained	A skill matrix is available but the training is not on time	A matrix is not available
10	Does Supplier have any experience in PPAP execution and preparation?	The supplier knows the PPAP process, tracks in a database and the documentation is complete	The supplier knows the PPAP process but it is no database is available and/or PPAP documentation is not complete	The supplier does not know PPAP
11	Is the production (machines) able to ensure the product specification and characteristics?	Production can meet the required tolerances	Production can meet the required tolerances but only with rework or defined actions	No the Production can meet the required tolerances
12	Is there a maintenance plan for the production (machines / tools) available?	The machines are in a clean condition and have been maintained according to the maintenance schedule	The machines are maintained according to the maintenance schedule but in bad condition	No maintenance plan available
13	The handling of special / important characteristics is known?	The supplier has experience with special / important characteristic and a documentation is available	The supplier has no experience with special / important characteristic and a documentation is not available	No experience
14	Traceability of material is ensured?	The traceability is from the raw material to the finished product is ensured	Traceability is only possible to a limited extent	No traceability ensured
15	Are there methods in place to investigate on failure root causes in case of an product and process deviation?	Supplier has structured procedures, adequate competences and equipment in order to guarantee that all NCs are managed with 8D methodology	Supplier has structured procedures, adequate competences and equipment but no evidence that all NCs are managed with 8D methodology	Supplier does not have structured procedures, adequate competences and equipment in order to guarantee that all internal are managed with 8D methodology
16	Are the products stored in an appropriate way and is the quality of the packaging and the transportation ensured?	The condition from the warehouse is good and FIFO is ensured	The condition from the warehouse is not so good but FIFO is ensured	Bad condition without FIFO

Figure 2.3: Example of evaluation criteria of supplier audits.

Audit results are quantified through a scoring system, ensuring objective assessment and comparability across the supplier base. A minimum score of 75 points is required to achieve a "pass", confirming that the supplier meets Ariston's quality and process requirements. Suppliers scoring between 60 and 75 points are conditionally accepted but must submit and implement a Corrective Action Plan (CAP) within the defined timeframe, with follow-up verification by Ariston's SQA team. Scores below 60 points generally lead to rejection for new suppliers or potential escalation measures for existing ones, which may include temporary suspension of new orders. In exceptional cases where the supplier is considered strategically critical, derogation requests may be submitted, provided that a clear and enforceable improvement plan is in place.

This structured audit methodology not only enforces compliance with contractual and quality requirements but also serves as a preventive tool, enabling Ariston to identify risks before they result in major incidents. In combination with non-conformity management and continuous performance evaluation, supplier audits form a key pillar of Ariston's proactive supplier quality strategy, fostering both accountability and sustainable improvement across the supply base.

2.5 Supplier Claim-Back Management

As part of Ariston Group's supplier quality governance, claim-back management serves as a critical mechanism to ensure that suppliers remain financially accountable for the costs and operational disruptions generated by their non-conforming deliveries. Once a non-conformity is detected, registered in the database as a

notification, and eventually managed according to the 8D process, the potential economic implications are assessed following clear rules. This process is designed to systematically recover the internal and external costs that Ariston incurs as a direct result of supplier-related quality issues, promoting both accountability and continuous improvement.

A claim-back is typically initiated when an incident triggers tangible costs for the company, which may include:

- **Scrap costs** – calculated from the quantity of defective parts that cannot be reworked or used in production.
- **Sorting and rework costs** – including both internal activities and any external sorting interventions at the supplier's expense, either at the supplier site or at Ariston's plant.
- **Logistics costs** – in cases where defective batches must be returned, segregated, or require additional handling.
- **Field or warranty costs** – in the event that supplier defects lead to customer returns or service interventions, representing the most critical and expensive scenario.

If a cost recovery is warranted, a formal claim-back is opened in the system, detailing the type of non-conformity and the associated costs. Suppliers are required to

acknowledge the claim-back, and in cases of dispute, provide supporting evidence to challenge the assigned responsibilities.

Claim-backs are not only a financial recovery tool but also a behavioral lever in Ariston's supplier governance. By attributing concrete economic consequences to non-conformities, the process reinforces the importance of robust preventive quality practices on the supplier side. Suppliers with recurring claim-backs are closely monitored, as repeated financial impacts are a clear indicator of insufficient corrective action and systemic weaknesses. In severe cases, unresolved or excessive claim-backs may trigger a requalification process or influence sourcing decisions for future projects, linking operational quality directly with long-term business continuity.

By formalizing the claim-back process, Ariston ensures that quality deviations are not only corrected technically through the 8D methodology but also addressed economically, aligning supplier incentives with the Group's objective of zero-defect deliveries. This structured approach closes the loop between detection, corrective action, and accountability, creating a transparent and data-driven framework for supplier performance management.

2.6 Concluding Remarks

The procedures and mechanisms presented throughout this chapter offer a comprehensive overview of Ariston Group's supplier performance evaluation framework, which begins long before the first delivery and continues throughout the entire

lifecycle of the supplier relationship. From rigorous qualification and contractual alignment to systematic control of non-conformities, audits, and financial accountability, each element contributes to building a structured, repeatable, and transparent supplier management system.

However, while these processes are indispensable for ensuring baseline compliance and for promptly reacting to quality deviations, they remain largely transactional and reactive in nature. What transforms this wealth of operational data into strategic supplier oversight is the implementation of structured performance evaluation models. In this context, Vendor Rating and Business Warning represent the cornerstone of Ariston's transition from quality assurance to quality governance, enabling the company to track performance trends, identify systemic weaknesses, and proactively escalate or reward supplier behavior based on objective, data-driven metrics.

These models not only synthesize inputs from all the processes described in this chapter but also give them strategic relevance, offering a framework for classification, escalation, and decision-making that is consistent, traceable, and aligned with Ariston's long-term quality objectives. The following chapters will delve into the design and evolution of these two central tools, highlighting both the limitations of the previous system and the methodological innovations introduced to enhance their fairness, stability, and operational effectiveness.

3 Legacy Evaluation System

Vendor Rating (VR) and Business Warning (BW) systems are fundamental tools in Ariston Group's supplier quality governance framework, serving as the bridge between operational quality data and strategic supplier management decisions. The Vendor Rating is a structured, quantitative assessment that synthesizes key performance indicators, such as Parts Per Million (PPM), into a standardized score that reflects the overall quality performance of each supplier. Its purpose is to ensure objective and transparent evaluations across the global supplier base, enabling consistent benchmarking, early detection of performance deterioration, and data-driven decision-making in supplier development and sourcing.

Complementing this, the Business Warning system translates the Vendor Rating into a clear governance status by applying predefined thresholds and escalation logic. This logic assigns suppliers to categories, such as *Full Compliance*, *Admonition*, or *Business on Hold*, each of that triggers specific corrective actions, management involvement, or even commercial restrictions. Together, these systems operationalize Ariston's commitment to quality by transforming raw performance metrics into actionable governance signals. They support proactive risk management, foster

continuous improvement, and provide a formal mechanism for accountability across the supplier lifecycle.

The following chapter critically analyzes the structure and limitations of the legacy Vendor Rating and BW systems, establishing the rationale for the substantial methodological improvements presented in Chapter 4.

3.1 Vendor Rating Calculation Process

The legacy Vendor Rating for suppliers was calculated individually for each specific combination of plant and macrocategory. A *macrocategory* represents a classification of different supplied products sharing similar characteristics. For example, a front casing panel and a steel pipe both belong to the *Steel Components* macrocategory, whereas electronic components like transistors and transformers belong to the *Electronic Components* macrocategory. Consequently, if a supplier provided both steel and electronic components to plants *A* and *B*, it would receive four distinct Vendor Ratings, one for each unique plant-macrocategory combination.

The calculation of the Vendor Rating was structured, yet rigid, and involved the assignment of two partial scores for each supplier-plant-macrocategory combination:

- **PPM score**, reflecting the supplier's defect rate relative to contractual targets.
- **8D score**, indicating the frequency of corrective action reports issued.

Both scores ranged from 1 to 4 and were subsequently combined using a weighted average to produce the final Vendor Rating.

PPM Score

The PPM score was computed by first calculating the actual PPM:

$$PPM = \frac{SP}{GR} \times 1,000,000 \quad (3.1)$$

where:

- *SP*: number of defective parts detected within the given period for a specific supplier-plant-macrocategory combination (Scrap Parts).
- *GR*: total number of parts received from the supplier for the same combination within the given period (Goods Receipts).

This calculated PPM was then compared to the contractual PPM target stipulated in the Quality Assurance Agreement (QAA) by calculating their ratio. This ratio was then used to assign scores as follows:

- **Score 4:** ratio < 0.7 (excellent performance)
- **Score 3:** $0.7 \leq \text{ratio} < 1.0$
- **Score 2:** $1.0 \leq \text{ratio} < 1.3$
- **Score 1:** ratio ≥ 1.3

8D Score

The 8D score was determined by counting the total number of 8D reports issued for each supplier-plant-macrocategory combination during the quarter. Scores were assigned according to the following criteria:

- **Score 4:** 0 issued 8Ds
- **Score 3:** 1-2 issued 8Ds
- **Score 2:** 3-4 issued 8Ds
- **Score 1:** 5 or more issued 8Ds

Final Vendor Rating

The final Vendor Rating was then calculated by taking a weighted average of the two partial scores, applying a weight of 70% to the PPM score and 30% to the 8D score. The resulting value was rounded through predefined bands, generating a final Vendor Rating score ranging from 1 to 4. Figure 3.1 illustrates the calculation process, including the scoring criteria for PPM and 8D and the final aggregation bands used to classify suppliers.

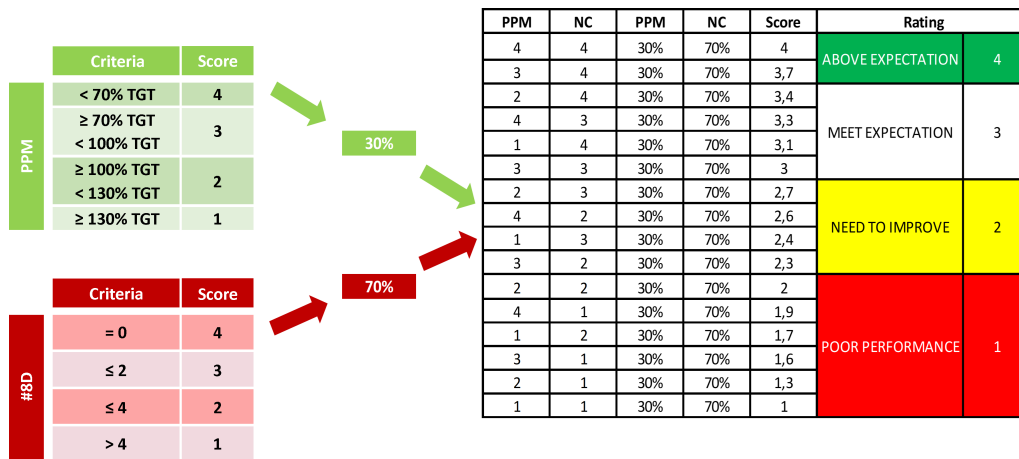


Figure 3.1: Legacy Vendor Rating calculation process.

This final score was then utilized to establish the supplier's Business Warning status and inform subsequent governance and corrective actions, as explained in the next section.

3.2 The Business Warning Logic

The Business Warning (BW) system function as a governance mechanism to monitor and manage supplier performance, enforcing progressive corrective actions based on the quarterly recalculated Vendor Rating scores. In the legacy BW, as in the Vendor Rating, each supplier was evaluated separately for each plant and macrocategory combination, ensuring that the assessments reflected localized performance conditions. Based on these evaluations, suppliers were classified into one of several predefined statuses:

- **Full Compliance:** Suppliers meeting or exceeding all quality expectations.

- **Admonition:** A cautionary status signaling a decline in Vendor Rating and the need for corrective action. Suppliers were required to submit a Corrective Action Plan (CAP) within three weeks of BW status notification and could continue supplying parts, remaining eligible for new PPAP requests.
- **Business on Hold:** A critical status indicating severely unacceptable performance. Suppliers were temporarily restricted from new PPAP requests, had to implement 100% control on all delivered batches, and were subjected to process audits. Supply continued only under strict monitoring and containment conditions.
- **Phase-Out:** The most severe status, indicating the supplier was to be progressively removed from sourcing. Existing orders could be fulfilled only with full certified controls and the supplier was required to cooperate for a smooth disengagement. Any future business was conditional upon requalification at Ariston's discretion.

Transitions between these statuses followed a structured decision matrix that incorporated both the current Vendor Rating and the previous evaluation status. This staged logic, illustrated in Figure 3.2, was designed to prevent abrupt changes and to encourage sustained improvement.

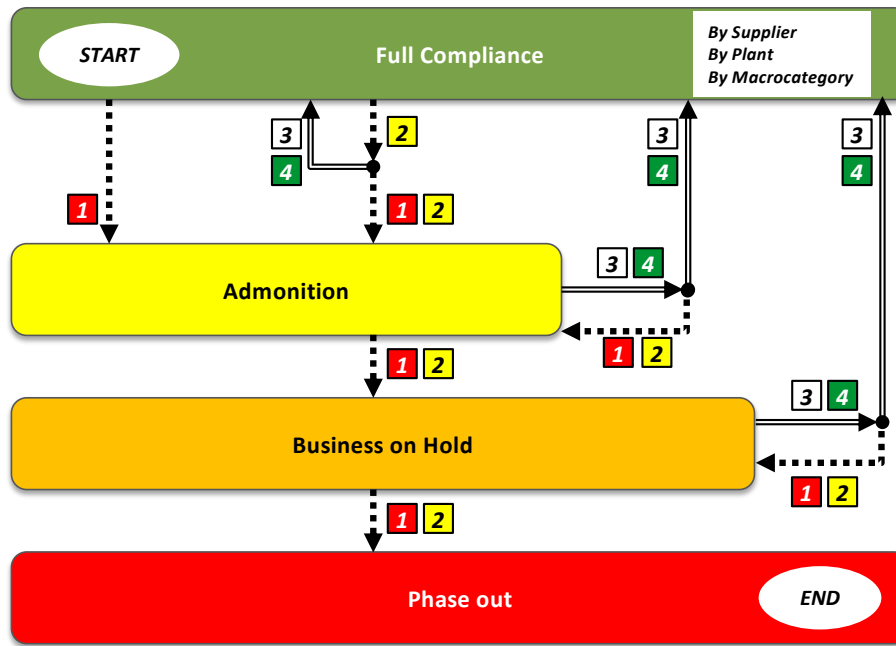


Figure 3.2: Decision matrix of the legacy Business Warning logic.

New suppliers always began in *Full Compliance*. After each quarterly evaluation, suppliers receiving a Vendor Rating of 1 or 2 were downgraded, while those with a rating of 3 or 4 were upgraded. To avoid abrupt transitions, intermediate statuses were implemented. For example, a downgrade from *Full Compliance* typically passed through an intermediate state termed *Still No Business Warning* before moving into *Admonition*. However, in cases of significantly poor performance, specifically when a Vendor Rating of 1 was assigned, the intermediate step was bypassed, and the supplier was directly moved to *Admonition*.

Conversely, suppliers experiencing performance improvements moved through intermediate statuses such as *Still Admonition* or *Still Business on Hold* before achieving a full upgrade to the next better status. This staged progression meant that a supplier needed to demonstrate consistently poor or excellent performance

for two consecutive quarters to experience a definitive status change. A single quarter's deviation, whether positive or negative, only resulted in an intermediate status adjustment, emphasizing the need for sustained trends before final status reassignment. The only exception was degradation to *Business on Hold* or *Phase-Out*, which could occur directly in cases of severe underperformance.

After two consecutive quarters of positive performance, suppliers were first upgraded to an intermediate status before returning directly to *Full Compliance*. This mechanism allowed suppliers to swiftly regain compliance following sustained improvement, in contrast to the gradual step-by-step degradation process. It thereby incentivized rapid recovery while maintaining supplier accountability and engagement.

3.3 Limitations of the Vendor Rating

The previous Vendor Rating system at Ariston, while functional as a basic governance tool, suffered from several structural and methodological weaknesses that compromised its reliability and strategic usefulness. Its design reflected historical practices rather than data-driven evaluation principles, leading to assessments that were often inconsistent, fragmented, and insufficiently granular. Three main limitations emerged as particularly critical.

First, the reliance on 8D-based scoring exposed the system to high levels of subjectivity, as the decision to open and manage 8D reports varied significantly

across plants and evaluators, undermining the objectivity and reproducibility of supplier assessments.

Second, the system was fragmented by Plant and Macrocategory, producing artificial separations in performance evaluation and further amplifying inconsistencies due to frequent misclassifications in the ERP system.

Finally, the discrete 1-to-4 scoring scale offered low granularity, failing to capture nuanced performance differences or reward incremental improvements, and limiting the organization's ability to detect emerging issues early.

Together, these limitations, discussed more in detail in the next sections, weakened the credibility of the Vendor Rating as a strategic decision-making tool, highlighting the need for a more standardized, data-driven, and continuous approach to supplier evaluation.

3.3.1 8D-Based Scoring

The reliance on 8D reports as a component of the Vendor Rating system introduced substantial limitations related to subjectivity, inconsistency, and data reliability. These shortcomings undermined the effectiveness of the 8D metric as a fair and robust performance indicator.

The decision to open an 8D report is typically left to the discretion of local Quality Control (QC) teams, without a standardized set of rules or thresholds across plants. As a result, the issuance of 8Ds is influenced by both human judgment and operational workload. During quieter periods, more 8Ds were opened,

including for minor issues, while in high-pressure times, even significant incidents might not result in formal reporting. Additionally, different manufacturing sites applied divergent interpretations for what warranted an 8D. The same type of incident might trigger a formal corrective action request at one plant while being resolved informally at another. Consequently, the 8D count often reflected internal procedural culture rather than objective supplier performance. This lack of harmonization in the application of 8D criteria led to severe evaluation inconsistencies across plants.

These observations are consistent with findings in the broader performance evaluation literature. Moers (2005) shows that subjectivity in performance appraisal introduces systematic bias, distorting the representation of true outcomes and reducing the reliability of metrics. Hawkins et al. (2020) similarly demonstrate that variations in evaluator behavior within supplier performance frameworks reduce both the credibility and the usefulness of the resulting data. Also in the healthcare industry, strong emphasis is placed on the need for procedural consistency in KPI systems. Sreedharan et al. (2024) stress that for performance indicators to be valid, they must be measured consistently across all evaluators and contexts. Likewise, Health Information and Quality Authority (2013) notes that reliable KPIs must yield the same result regardless of who performs the evaluation, something clearly compromised in Ariston's 8D-based approach.

In essence, the 8D-based scoring was highly vulnerable to internal procedural variability, undermining the objectivity, comparability, and reproducibility of supplier

assessments. Instead of reflecting consistent, data-driven quality outcomes, scores could be significantly skewed by local behaviors, discretion, and undocumented practices. As emphasized across the academic literature, KPI systems that lack methodological rigor and uniform application cannot serve as reliable foundations for governance and performance-based decision-making.

3.3.2 Fragmentation by Plant and Macrocategory

Another structural limitation of the previous Vendor Rating system was its fragmented evaluation by plant and macrocategory. While this approach was initially adopted to reflect local operations and purchasing structures, it is conceptually incorrect and generates several practical issues in supplier performance assessment.

From a quality evaluation perspective, assessing the same supplier separately for different plants does not provide an accurate representation of its production performance. If two plants source similar types of materials from the same supplier, evaluating them independently creates an artificial separation that does not reflect the supplier's true overall quality level, which depends primarily on its production processes rather than on the receiving plant.

The use of macrocategories, while theoretically reasonable, also introduces critical inconsistencies in practice. In the company information system, materials are often misassigned to the wrong macrocategories, leading to incoherent and misleading Vendor Rating results. A supplier might appear to perform poorly in one macrocategory simply due to incorrect material assignments, even if its actual performance is consistent.

This dual fragmentation, therefore undermines the reliability and coherence of supplier evaluations, making the Vendor Rating less representative of the supplier's real performance and less useful as a strategic decision-making tool.

3.3.3 Low Granularity and Limited Discriminatory Power

The previous Vendor Rating system's reliance on a 1-to-4 scoring scale significantly constrained its ability to accurately capture and reflect nuanced differences in supplier performance. This coarse scoring framework resulted in limited differentiation, as suppliers with notably different performance levels could end up receiving identical ratings. Research supports that scales with fewer than five rating levels tend to have insufficient discriminative capability, thus failing to accurately represent true performance differences (Preston & Colman, 2000). For example, suppliers experiencing vastly different numbers of incidents, such as two incidents compared to ten, might both receive the same score, obscuring real performance differences.

Beyond its weak discriminatory power, the 1-to-4 scale was also limited in terms of interpretability and universality. The meaning of an assigned score was not self-evident: for instance, a rating of "4" offered no intuitive sense of whether this represented excellence or poor performance, creating ambiguity for managers and suppliers alike. Such lack of clarity undermined the communicative value of the metric and reduced its acceptance as a shared reference point across plants, functions, and supplier networks. More universal and widely understood scales (e.g., percentile distributions, normalized indices, or target-based scores) provide

a clearer frame of reference, enabling users to immediately understand whether a score represents good, average, or poor performance.

Additionally, the discrete scoring structure lacked the capacity to recognize incremental improvements. Suppliers making consistent, gradual progress remained unnoticed unless improvements were substantial enough to cross arbitrary rating thresholds, thereby missing critical opportunities to reinforce and motivate continuous improvement efforts. The rigid thresholds and binary nature of the system also discouraged nuanced assessment and proactive performance management. Suppliers were categorized simplistically as either acceptable or unacceptable, eliminating mechanisms for early warnings or monitoring subtle performance trends. Consequently, the system limited the organization's ability to effectively detect and address emerging issues.

Transitioning toward a more refined, granular scoring model is therefore crucial to addressing these limitations and enabling more precise, fair, and effective supplier evaluations and management.

3.4 Inconsistencies in the Business Warning Logic

Although conceptually well-structured, the Business Warning (BW) system exhibited several significant structural inconsistencies that hindered its effectiveness as a governance mechanism. One notable issue was the rapid escalation of suppliers within the system: due to the absence of intermediate degradation steps, even minor or temporary drops in performance could quickly move a supplier from

Full Compliance to more severe statuses such as *Business on Hold*, or in some cases, the *Phase-Out* level. This swift escalation often surpassed the suppliers' ability to react appropriately, leaving insufficient time to implement effective corrective actions and adjustments to improve performance. Additionally, the *Phase-Out* status, though explicitly defined as a critical and ultimate level in the escalation framework, was seldom applied practically. Suppliers identified as candidates for removal or replacement were frequently provided with corrective action plans and reintroduced into the regular monitoring cycle instead of facing actual removal. As a result, this pivotal governance tool was rendered ineffective, becoming essentially symbolic and undermining the credibility of the escalation process.

Another crucial limitation was the absence of a stability condition. The system required suppliers' status to change at each evaluation, either positively or negatively, with no option to maintain their current status. This led to unnecessary volatility, especially in borderline cases where supplier performance was stable or showed minor fluctuations that did not warrant significant status shifts. Consequently, the governance model became excessively reactive, detracting from its ability to reliably monitor and manage supplier performance.

Collectively, these structural shortcomings weakened the overall robustness and operational effectiveness of the Business Warning logic, underscoring the necessity of developing a more stable, realistic, and strategically nuanced escalation framework.

3.5 Summary

The analysis of the legacy Vendor Rating and Business Warning system highlights how a framework that initially served as a basic governance tool had become methodologically weak and operationally unreliable. From a managerial perspective, these limitations had consequences that went beyond methodological accuracy. A supplier evaluation framework that is fragmented, subjective, poorly interpretable and lacking in granularity does not only reduce analytical reliability, but also compromises decision-making. Ambiguous ratings (such as the 1-to-4 scale) risked confusing both managers and suppliers about what “good” or “bad” performance actually meant, weakening the motivational and contractual force of the scorecard. Inconsistent use of 8D-based scoring and misclassifications across plants undermined fairness, made comparisons unreliable, and fuelled disputes between functions and sites. The volatility and symbolic nature of the Business Warning logic further eroded trust, limiting its role as an escalation mechanism. In practice, these weaknesses reduced the ability of Ariston’s Supplier Quality Assurance to prioritize resources, to identify risks early, and to support sourcing or negotiation with robust evidence—ultimately constraining the contribution of the system to strategic supplier management.

These findings underline the necessity of developing a new, data-driven, and continuous evaluation model, capable of delivering greater fairness, stability, and transparency in supplier assessments. The next chapter presents the redesigned Vendor Rating and Business Warning framework, addressing the shortcomings

identified in this analysis and establishing a more reliable foundation for supplier governance.

4 New Evaluation System

As extensively discussed in Chapter 3, Ariston Group's previous Vendor Rating and Business Warning framework exhibited several methodological and operational weaknesses, including fragmentation by plant and macrocategory, subjective reliance on 8D reporting, insufficient granularity and poor interpretability of scoring, and unnecessary volatility and symbolic nature of BW. Collectively, these limitations undermined the credibility, fairness, and effectiveness of supplier governance, necessitating a comprehensive redesign of the evaluation system.

This chapter introduces and analyzes the new Vendor Rating and Business Warning, specifically developed to address these identified deficiencies. Through the adoption of a holistic supplier-level evaluation approach, a refined and more granular 0-100 scoring system, and advanced statistical techniques such as logarithmic normalization and Bayesian shrinkage, the redesigned framework seeks to provide more accurate, objective, and actionable performance evaluations. Furthermore, by restructuring the Business Warning escalation paths and governance thresholds, the new model has been designed to establish a more transparent, stable, and proactive approach to managing supplier relationships. This chapter details

the architectural changes, explains the rationale behind each modification, and demonstrates through concrete examples how the new system delivers substantial improvements in supplier evaluation and governance effectiveness.

4.1 Methodology

The redesign of the Vendor Rating and the Business Warning followed a design-oriented, evidence-based process grounded in longitudinal company data and iterative validation with stakeholders. A comprehensive dataset was assembled covering the entire supplier base of more than 3,000 suppliers. The dataset integrated goods receipts and notifications from January 2023 to the present, with systematic checks for master data alignment, duplicate removal, and time series completeness. This longitudinal view was essential to evaluate behavior over an extended horizon rather than on a short or atypical window.

On this foundation, a simulation environment was built to run the legacy models in parallel with candidate redesigns month by month and supplier by supplier. The simulations reproduced the full monitoring cadence so that trajectories, thresholds, and escalations could be observed as they would have occurred in real time. Alternative specifications were tested in sequence, including the use of transactional indicators only, the exclusion of evaluator-dependent artifacts, the application of logarithmic transformations to stabilize skewed measures, and the adoption of principled pooling to temper volatility for low exposure suppliers. Each iteration was assessed on reliability, discriminative power, and stability over time.

The work proceeded through repeated cycles of testing, diagnosis, and fixing. Discrepancies surfaced by the backtests were traced to their root causes, which led to refinements in data preparation rules and adjustments in parameter ranges. Particular attention was paid to edge cases such as suppliers with very low monthly exposure and plants with structural shifts in mix, to ensure that the redesigned indicators remained fair and interpretable across heterogeneous operating conditions.

Stakeholder engagement was integral to the methodology. Interim results and exception lists were reviewed with supplier quality managers, plant quality teams, procurement, and operations to validate face validity and operational fit. Feasibility and integration aspects were discussed with IT to confirm that the redesigned logic can be generated reliably within existing data pipelines and reported through standard analytics tools. Feedback from these sessions informed successive design refinements and ensured alignment with governance and communication requirements.

To confirm robustness, the final specification was subjected to a shadow run over the full historical window, with side-by-side comparison of escalations and supplier rankings under the legacy and redesigned models. The evaluation focused on variance reduction for low exposure suppliers, persistence of rank ordering for mid-range performers, and proportionality of escalation signals, while preserving explainability for users and suppliers.

4.2 Architecture of the New Vendor Rating

A fundamental improvement in the redesigned Vendor Rating model is the transition from evaluations previously segmented by individual plant and macrocategory combinations to a unified, supplier-level assessment. This change recognizes that quality issues observed at a single plant often indicate systemic supplier deficiencies potentially affecting multiple locations. By adopting this holistic supplier perspective, the new framework ensures greater consistency, fairness, and comprehensive visibility into supplier performance.

Complementing this holistic view, the scoring logic has evolved from the legacy 1-to-4 scale to a 0-100 scale. This enhancement significantly increases granularity and interpretability, providing precise differentiation among suppliers and enabling clearer communication with both internal stakeholders and suppliers. Moreover, the new 0-100 scoring method aligns seamlessly with Ariston's existing supplier audit evaluations, creating a universally consistent framework that facilitates direct comparison and eventual aggregation of performance outcomes.

Another critical enhancement in the redesigned architecture is the complete elimination of the subjective and workload-biased 8D reporting mechanism, previously a significant source of inconsistency. The Vendor Rating now integrates the Weighted Supplier Incidents (WSI) indicator in place of 8D reports, thereby introducing an objective, impact-sensitive measure that accurately captures supplier quality performance. The methodological details and operational benefits of the

WSI indicator will be thoroughly examined in the subsequent sections of this chapter.

4.3 Input KPIs and Preprocessing

The redesigned Vendor Rating system relies upon two carefully selected Key Performance Indicators (KPIs): Weighted Supplier Incidents (WSI) and Parts Per Million (PPM). These metrics provide complementary dimensions of supplier quality performance, ensuring a holistic and balanced evaluation. The WSI indicator replaces the previous reliance on the subjective and inconsistent 8D count, offering a more objective, structured, and impact-oriented measure of non-conformities (incidents). On the other hand, the PPM indicator has been retained due to its recognized effectiveness in quantifying the material severity and tangible scrap consequences of supplier incidents.

The following subsections will detail each indicator's calculation methodology, operational significance, and their complementary roles within the Vendor Rating framework.

4.3.1 Weighted Supplier Incidents (WSI)

The Weighted Supplier Incidents (WSI) indicator captures the quality impact of supplier-related issues, weighing each non-conformity detection based on where it is detected in the production or use process.

Each incident is assigned a weight based on its detection point, with downstream detections (e.g., at customer or production line) considered more severe than upstream ones (e.g., incoming inspection). This allows the indicator to differentiate between minor and disruptive events. Although a future development (see Section 5.1) could include incorporating incident gravity (i.e., severity classification), this factor is not yet applied in the current WSI logic. At present, all weighting is based solely on the detection point.

All supplier-related notifications are extracted from the database and processed. Each incident is multiplied by its assigned detection point weight. These weighted incidents are then aggregated quarterly for each supplier. To account for differences in volume, the sum is divided by the number of goods receipts to produce the final WSI value. The calculation of the WSI for a given supplier is as follows:

$$WSI = \frac{\sum_{i=1}^n DW_i}{GR} \times 1,000,000 \quad (4.1)$$

where:

- n : total number of incidents of the supplier in the considered period.

- DW_i : Detection point Weight of incident i , depending on where the non-conformity is detected.
- GR : total number of parts received from the supplier in the considered period (*Goods Receipts*).

As already anticipated in Section 2.3, there are four possible detection points to be assigned to quality incidents. These detection points reflect the phase of the supply and production process at which the non-conformity is discovered, and they are weighted differently in the WSI calculation due to their varying impact on the business:

- **W1 – Field / Market:** Incidents detected directly at the customer site or in the field. These are the most critical, as they directly affect the end-user and brand perception. **Weight: 15.**
- **W2 – Incoming Inspection:** Non-conformities intercepted during the material reception and quality control checks before the components enter production. While these issues do not affect final products, they disrupt the flow and require containment. **Weight: 1.**
- **W3 – Production Line:** Issues identified during assembly or internal manufacturing processes. These defects may delay production but are still contained within the factory. **Weight: 2.**

- **W4 – Statistical Process Control (SPC):** Defects found during periodic or sample-based inspections. **Weight: 4.**

These weights are used in the WSI formula to account for the criticality of the detection point, ensuring that incidents are evaluated not only by their frequency but also by their business impact. The heavier weight assigned to W1 events, for example, reflects the reputational and customer satisfaction risks associated with field issues, whereas W2 events are considered less severe since they are contained early in the process.

This approach provides an impact-sensitive measure of supplier quality, ensuring that the WSI, and so the Vendor Rating, reflect the true operational relevance of each supplier's non-conformities.

4.3.2 Parts Per Million (PPM)

The PPM indicator reflects the number of defective parts per million received and is a widely recognized quality KPI. It complements the WSI by providing a quantitative measure of scrap intensity, offering a different but equally important dimension of supplier performance.

As already explained in Section 2.2, each supplier is assigned a PPM target, defined in the Quality Assurance Agreement (QAA). While many suppliers have standard targets (e.g., 233 or 500 PPM), these can vary depending on the criticality of the supplied parts. The actual PPM performance is then benchmarked against this target, allowing for contextual evaluation: a supplier consistently below the target

demonstrates strong conformance, while one exceeding it may require corrective actions.

The PPM indicator calculation happens in the same way as before (see Section 3.1). It is retained in the Vendor Rating because it provides critical information that WSI alone cannot capture. While WSI reflects the number and severity of incidents, PPM quantifies their material impact in terms of scrap volume: for instance, an incident that generates one scrap part differs significantly from one that results in 1,000 scrapped components.

Including both KPIs ensures a more meaningful and interpretable evaluation, addressing both the severity of non-conformities detection and their tangible impact in terms of scrap volume.

4.4 The Challenge of Instability in Low-Volume Suppliers

In performance measurement, particularly when evaluating suppliers with highly variable delivery volumes, significant fluctuations in data can substantially distort performance evaluations. Metrics such as the Weighted Supplier Incidents (WSI) and Parts Per Million (PPM), become disproportionately sensitive to outliers and random events when applied to suppliers with limited data.

Specifically, in Ariston's supplier evaluation system, the issue manifests clearly, as we can see in Table 4.1: if a low-volume supplier experiences even one single incident, their WSI metric immediately skyrockets, since the calculation of WSI involves dividing the weighted number of incidents times 1,000,000 by a small

number of delivered batches or goods receipts. Similarly, for PPM, a single defective batch can dramatically inflate the ratio of defects, thus severely penalizing low-volume suppliers in a disproportionate and unfair manner.

Table 4.1: Example of instability for a low-volume supplier.

Year	2023				2024				2025	
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
Goods Receipts (<i>GR</i>)	239	149	208	80	136	244	429	268	165	585
Inc. Weights ($\sum_{i=1}^n DW_i$)	2	0	0	0	5	0	4	4	4	2
Scrap Parts (<i>SP</i>)	1	0	0	0	3	0	4	2	1	2
<i>WSI</i>	8368	0	0	0	36765	0	9324	14925	12121	6838
<i>PPM</i>	4184	0	0	0	22059	0	9324	7463	6061	3419

Consequently, suppliers delivering fewer items face dramatic volatility in performance scores, which reflect statistical anomalies rather than genuine underlying quality trends. This instability undermines trust in the evaluation system and can lead to governance decisions that are not truly indicative of supplier reliability.

To address this, a correction mechanism is required, one that preserves genuine performance signals while reducing the noise introduced by sample variability. The following section introduces a Bayesian Shrinkage approach specifically designed to stabilize the scores of suppliers with low exposure without masking meaningful differences in performance.

4.4.1 Bayesian Shrinkage

Bayesian shrinkage provides a statistically robust solution to the problem by moderating individual supplier scores, effectively reducing their sensitivity to small-sample fluctuations. This method statistically "pulls" each supplier's observed performance metric toward the global mean or average of all suppliers, with the degree of shrinkage inversely related to data availability (Schwartz, 2021). The way the shrinkage happens is with the application of the formula below, taking as example the WSI indicator for a supplier i :

$$WSI_{shrink_i} = \frac{GR_i}{GR_i + k} \cdot WSI_i + \frac{k}{GR_i + k} \cdot \overline{WSI} \quad (4.2)$$

where:

- GR_i = Goods Receipts for supplier i during the evaluation period
- WSI_i = observed WSI value for supplier i (see Equation 4.1)
- $\overline{WSI}$ = average WSI across all suppliers in the evaluation period
- k = smoothing factor, a constant controlling the degree of shrinkage toward the average

The smoothing factor determines how quickly a supplier's score converges toward the population average when the data volume is low. A larger k produces stronger shrinkage (more conservative evaluations for low-volume suppliers), while a smaller k makes the observed score weigh more heavily. With the application of

Equation 4.2, suppliers with substantial delivery histories maintain scores close to their observed values, as their data are statistically reliable. However, low-volume suppliers, whose observed scores are inherently unstable and prone to extreme values, are adjusted more significantly toward the mean, thus achieving more stable and credible evaluations. In practice:

- High-volume suppliers ($GR_i \gg k$):

$$\frac{GR_i}{GR_i + k} \approx 1 \quad \Rightarrow \quad WSI_{shrunk_i} \approx WSI_i$$

- Low-volume suppliers ($GR_i \ll k$):

$$\frac{GR_i}{GR_i + k} \approx 0 \quad \Rightarrow \quad WSI_{shrunk_i} \approx \overline{WSI}$$

The theoretical underpinning of Bayesian shrinkage lies in its effective balance of variance and bias. Schwartz (2021) demonstrates how shrinkage significantly reduces measurement error, enhancing accuracy, especially in contexts sensitive to noisy performance signals. Similarly, Chuang and Oliva (2015) reinforce this concept by applying Bayesian methods to inventory management, showcasing reduced distortion in performance measures by leveraging historical group-level information, thus advocating for shrinkage estimators in operational decision-making environments characterized by sparse or variable data.

Particularly analogous to Ariston's supplier evaluation challenge is the empirical Bayes method used by Baker and McHale (2017) in their analysis of women's tennis rankings. In tennis player evaluations, infrequent participation by some players resulted in highly volatile performance ratings, as single wins or losses drastically shifted player rankings. Baker and McHale (2017) addressed this by applying Bayesian shrinkage, pulling the performance ratings of players with fewer matches toward the overall average. This approach effectively stabilized ratings, ensuring that players were neither unjustly rewarded nor penalized due to isolated performances. The parallel with the Ariston supplier scenario is clear: Just as the ranking of a tennis player should reflect true skill rather than one-off results, supplier performance ratings should accurately represent consistent capability rather than isolated incidents.

In conclusion, applying Bayesian shrinkage within Ariston's supplier quality performance evaluation model represents a scientifically rigorous, transparent, and equitable solution to the inherent instability observed in low-volume suppliers' performance metrics. By incorporating Bayesian principles, Ariston's new Vendor Rating system mitigates the disproportionate impact of individual incidents, stabilizes supplier scores, and ensures that governance decisions reflect true and sustained supplier quality performance.

4.4.2 Logarithmic Normalization

Before applying Bayesian shrinkage to key performance indicators such as Weighted Supplier Incidents (WSI) and Parts Per Million (PPM), normalization of these

metrics is not just recommended, it is methodologically essential. The reason for this step lies in the inherent statistical characteristics of the raw data. These metrics exhibit pronounced skewness and extreme variability, largely driven by suppliers with very low exposure. As explained in the previous subsection, the performance indicators of these low-volume suppliers are inherently unstable: as soon as a single incident occurs, their WSI or PPM can spike to extremely high values, also exceeding 10,000. Without appropriate normalization, these extreme values dominate and distort the arithmetic mean, precisely the central reference used by Bayesian shrinkage. Consequently, shrinkage adjustments performed directly on raw values would become biased and unreliable, disproportionately penalizing suppliers.

To address this critical issue, a logarithmic transformation is utilized. The main advantage of the logarithmic normalization is its capability to compress extreme values effectively, thereby substantially reducing the skewness and variance within the data while preserving the intrinsic ordering and relative distinctions among suppliers. Specifically, logarithmic transformations stabilize the distribution of the indicators, aligning them closer to a normal distribution, which is a core assumption underlying Bayesian statistical methods. As highlighted by Chuang and Oliva (2015), applying log-transformed metrics substantially enhances the robustness of subsequent statistical analyzes, particularly those employing shrinkage methods to manage performance metrics characterized by long-tailed distributions and sparse but significant outliers.

The normalization uses a base 10 logarithm and happens by using the following formula:

$$WSI_{norm} = \log(WSI + 1) \quad (4.3)$$

The addition of 1 ensures that suppliers with zero incidents are correctly handled, since $\log(0)$ is undefined, and maps them to zero in the log scale.

Example 4.1. Applying the transformation to the supplier in Table 4.1 for Q2 2025, where $WSI = 6838$:

$$WSI_{norm} = \log(6838 + 1) = \log(6839) \approx 3.83$$

The beneficial impact of applying logarithmic normalization to WSI values is clearly illustrated in the example tables below. The first table reports the quarterly average WSI computed on raw values, which results in very high averages dominated by a few extreme observations. The second table shows the average of the log-normalized WSI values, which are compressed into a stable and interpretable range, making them suitable as a baseline for Bayesian shrinkage:

Table 4.2: Quarterly average WSI without normalization.

Year	2023				2024				2025	
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
$\overline{WSI}$	255	378	352	433	475	204	224	603	417	609

Table 4.3: Quarterly average log-normalized WSI.

Year	2023				2024				2025	
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
$\overline{WSI}_{norm}$	0.26	0.28	0.27	0.30	0.26	0.22	0.23	0.24	0.21	0.20
Equivalent $\overline{WSI}$	1.83	1.90	1.86	2.01	1.81	1.65	1.71	1.75	1.62	1.58

From a practical perspective, logarithmic normalization ensures that Bayesian shrinkage operates as intended by stabilizing highly skewed performance data and preventing the use of distorted averages that would otherwise misrepresent supplier performance. This aligns closely with the findings of Schwartz (2021), who underscored the critical role of normalization prior to shrinkage to ensure fair and stable performance evaluations, particularly in contexts such as healthcare quality assessments, which face similar challenges due to data sparsity and variability in observed events. By transforming raw metrics into a more symmetric and interpretable form, logarithmic normalization provides the robust statistical foundation required to accurately adjust scores toward a genuinely representative mean, ultimately enhancing the fairness and decision-making value of Ariston's Vendor Rating system.

Thus, we can redefine the Equation 4.2 for Bayesian shrinkage, which is now applied to the normalized values rather than the raw values. The updated equation for shrinkage is the following:

$$WSI_{shrunk_i} = \frac{GR_i}{GR_i + k} \cdot WSI_{norm_i} + \frac{k}{GR_i + k} \cdot \overline{WSI}_{norm} \quad (4.4)$$

where:

- GR_i = Goods Receipts for supplier i during the evaluation period
- WSI_{norm_i} = Normalized WSI value for supplier i
- $\overline{WSI}_{norm}$ = average WSI_{norm} across all suppliers in the evaluation period
- k = smoothing factor, a constant controlling the degree of shrinkage toward the average

Example 4.2 (continuing Example 4.1). always looking to the data in Table 4.1, in Q2 2025 we have $GR = 585$ and $WSI = 6838$, we already calculated the normalized WSI:

$$WSI_{norm} = \log(WSI + 1) = \log(6839) \approx 3.83$$

So, using the period average $\overline{WSI}_{norm} = 0.23$ and smoothing factor $k = 1000$:

$$WSI_{shrunk} = \frac{585}{585 + 1000} \cdot 3.83 + \frac{1000}{585 + 1000} \cdot 0.23 \approx 1.54$$

We see that with only 585 goods receipts in the quarter, the value is pulled toward the period average, going from 3.83 to 1.54. Shrinkage reduces sensitivity to single quarter spikes while preserving signal as volume grows.

4.4.3 Score Mapping and Aggregation

Following the logarithmic normalization and Bayesian shrinkage of WSI and PPM indicators, the subsequent step in the Vendor Rating calculation involves mapping these adjusted metrics onto a standardized 0-100 scoring scale. This step serves two primary purposes: it enhances interpretability by transforming complex statistical measures into an easily understandable format, and it ensures comparability across suppliers. Furthermore, the adoption of a universal 0-100 scale is consistent with other quality evaluation tools currently utilized within Ariston Group, such as the supplier audit scoring framework (explained in detail in Section 2.4).

The mapping of normalized and shrunk metrics to 0-100 scores is accomplished through the use of predefined scoring bands. These bands assign scores according to the relative performance levels of suppliers, with lower normalized indicator values (representing fewer incidents or lower defect rates) receiving higher scores, thus reflecting superior supplier quality performance. Conversely, higher normalized values, indicative of poorer performance, correspond to lower scores.

The specific scoring bands currently employed in the Vendor Rating model are detailed in Tables 4.4 and 4.5. The tables include several columns designed to facilitate understanding of the scoring thresholds. Specifically, the *Threshold*

columns indicate the logarithmic, shrunk values utilized internally within the Vendor Rating calculation. To support practical interpretation, the tables also include columns showing these thresholds re-transformed back into their original, non-logarithmic scale (*WSI* for the Weighted Supplier Incidents, and *PPM/TARGET* for the Parts Per Million indicator). The *PPM/TARGET* column in Table 4.5 explicitly denotes how many times the supplier's actual PPM exceeds the target PPM, thus directly relating the scoring bands to operational targets.

Table 4.4: Scoring bands for WSI.

WSI	Threshold (WSI_{shrunk})	Score
1.0	0.00	100
10.0	1.00	90
31.6	1.50	75
70.8	1.85	60
316.2	2.50	45
1000.0	3.00	30
3162.3	3.50	15

Table 4.5: Scoring bands for PPM.

PPM/TARGET	Threshold (PPM_{shrunk})	Score
0.01	0.00	100
0.10	1.00	90
1.00	2.00	75
2.00	2.30	60
31.62	3.50	45
100.00	4.00	30
1000.00	5.00	15

To provide practical clarity, a WSI_{shrunk} value close to zero (corresponding to almost no supplier incidents) is assigned a score of 100, signifying exceptional supplier performance. Conversely, a high WSI_{shrunk} value, such as 3 (corresponding to a raw WSI value of 1,000), would drastically reduce the score to around 30 points, indicating significant quality issues. Similarly, for the PPM indicator, a value exactly equal to the target PPM ($PPM/TARGET = 1.0$) yields a Vendor Rating score of 75, whereas a PPM value twice the target ($PPM/TARGET = 2.0$) corresponds to a score of 60. Thus, the scoring bands ensure that minor deviations from optimal performance result in proportionate score reductions, while suppliers experiencing severe quality issues are quickly and appropriately penalized.

Example 4.3 (continuing Example 4.2). From supplier in Table 4.1 we obtained $WSI_{shrunk} \approx 1.54$ for Q2 2025. Using the scoring bands in Table 4.4 (thresholds

expressed on the WSI_{norm} scale), the value 1.54 lies in the band $[1.50, 1.85)$, therefore the mapped score is

$$Score_{WSI} = 75$$

Following the individual calculation of WSI and PPM scores, these metrics are combined into a single, final Vendor Rating through a weighted aggregation. Specifically, the Vendor Rating calculation applies a weighting of 70% to the WSI score and 30% to the PPM score:

$$Vendor\ Rating = 70\% \cdot Score_{WSI} + 30\% \cdot Score_{PPM} \quad (4.5)$$

The choice of weights reflects the importance assigned to WSI as an indicator capable of capturing the impact on business of incidents. Meanwhile, the PPM metric provides complementary insights into the quantity of scrap parts, ensuring a balanced evaluation perspective.

4.5 Recap of the New Vendor Rating Calculation Process

After detailing the methodological innovations introduced by the redesigned Vendor Rating, including logarithmic normalization, Bayesian shrinkage, and the adoption of a 0-100 scale, it is beneficial to summarize the overall calculation process comprehensively. Given the complexity and the multi-step nature of the procedures described in previous sections, Figure 4.1 provides a clear, visual overview of the sequential workflow underpinning the new indicator.

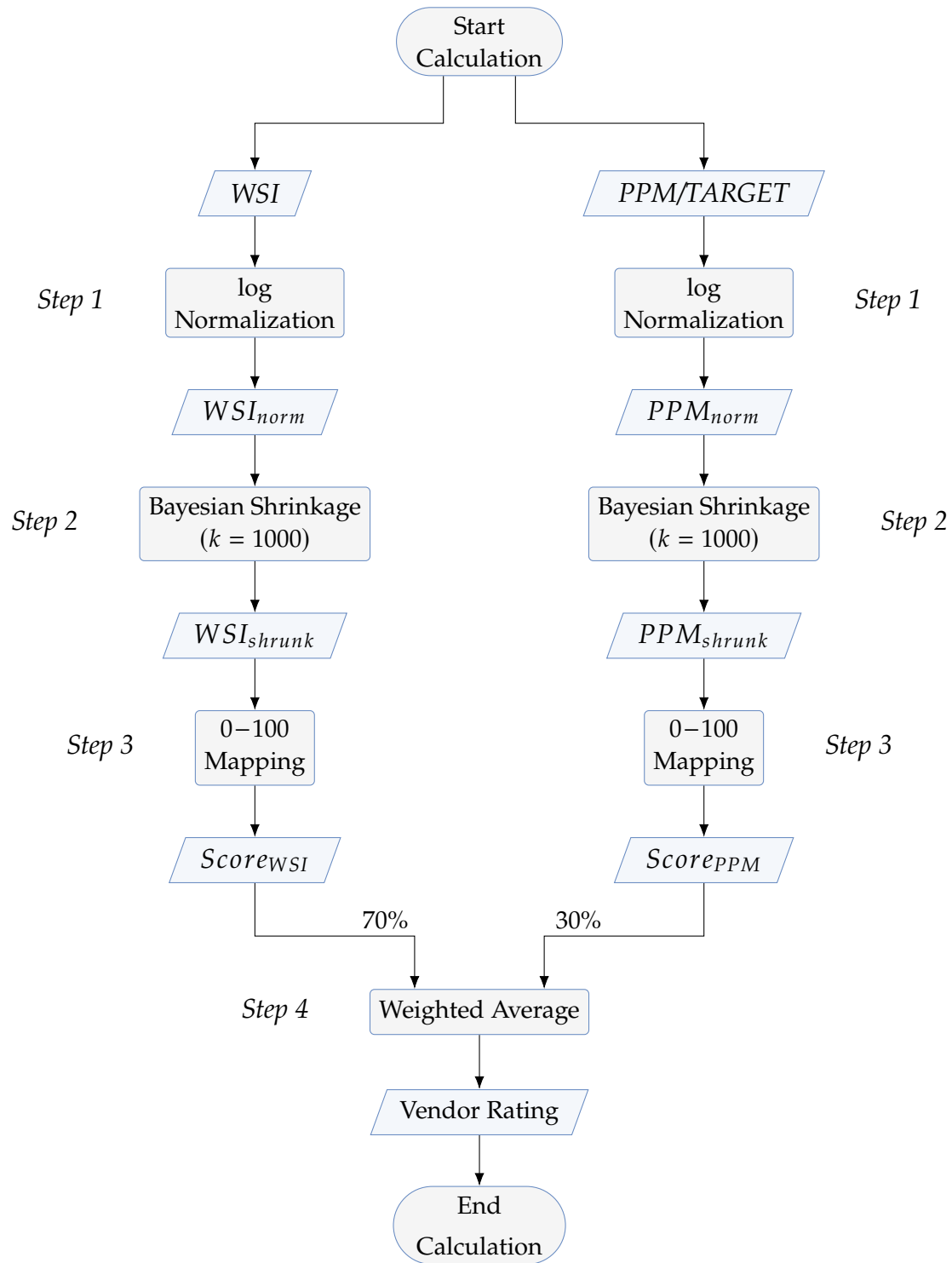


Figure 4.1: New Vendor Rating calculation process.

As illustrated, the calculation procedure follows four structured and clearly defined steps for both Weighted Supplier Incidents (WSI) and Parts Per Million (PPM). In the initial step, logarithmic normalization (Equation 4.3) is applied separately to each raw KPI, obtaining WSI_{norm} and PPM_{norm} (see Example 4.1). As previously explained, this transformation mitigates the pronounced skewness and the distorting impact of extreme values typically encountered in raw WSI and PPM data. By compressing these extreme values, the normalized data become more suitable and stable for subsequent analytical operations.

In the second step, Bayesian shrinkage (Equation 4.4) is introduced to address instability and high variability, especially among suppliers with limited or inconsistent delivery volumes (see Example 4.2). Through extensive testing and simulations, the smoothing factor k was optimally set to a value of 1,000, ensuring an effective balance between variance reduction for low-volume suppliers and the retention of genuine performance differentiation. The shrinkage step thus yields adjusted and stabilized KPIs, WSI_{shrunk} and PPM_{shrunk} , that are resilient to random fluctuations while preserving meaningful signals of performance variation across the supplier base.

Following this stabilization, the third step converts the shrunk values into standardized scores ranging from 0 to 100. This 0-100 mapping (Tables 4.4 & 4.5) facilitates clear interpretation and comparison, both internally and externally (see Example 4.3). A lower normalized and shrunk KPI corresponds to a higher mapped score, reflecting superior supplier quality performance, whereas higher

normalized values translate into lower scores, indicating potential quality concerns requiring governance intervention.

In the final, fourth step, these standardized scores, $Score_{WSI}$ and $Score_{PPM}$, are aggregated into a single composite Vendor Rating through a weighted average, with 70% weighting attributed to WSI and 30% to PPM (see Equation 4.5). The aggregated final score thus represents a comprehensive assessment of supplier performance, harmonizing operational severity and scrap intensity of incidents into a unified, easily communicable metric.

By delineating these four sequential and clearly defined steps - normalization, shrinkage, mapping, and aggregation - the redesigned Vendor Rating calculation ensures transparency, consistency, and statistical robustness.

Ultimately, this structured methodology provides Ariston Group with a sophisticated yet operationally practical framework, capable of effectively capturing nuanced supplier performance dynamics while solving the inconsistencies previously encountered in the legacy evaluation system.

4.6 Evolution of the Business Warning Logic

The introduction of the redesigned Vendor Rating system necessitated a comprehensive revision of the Business Warning logic, both to ensure alignment with the new quantitative performance scores and to correct the structural inconsistencies of the previous framework (see Section 3.4 for a detailed discussion). In the old model, suppliers frequently experienced rapid and sometimes disproportionate

escalations, moving from *Full Compliance* to *Business on Hold*, or even *Phase-Out*, after minor performance drops, often without sufficient time to implement corrective actions. *Phase-Out* itself was rarely applied in practice and had become largely symbolic, while the absence of a stability condition meant that suppliers were either promoted or downgraded at every quarterly evaluation, leading to unnecessary volatility and reactive governance.

The new BW logic, shown in Figure 4.2, directly integrates the new 0–100 Vendor Rating and introduces quantitative governance thresholds that define how supplier performance translates into actionable status decisions. Three bands now govern status transitions:

- **Score ≥ 75 :** The supplier is considered in full compliance. If previously in a lower status, this score triggers a promotion.
- **Score 60–75:** Performance is considered borderline, and the supplier retains the current BW status, introducing the long-missing stability condition that prevents unnecessary changes due to minor fluctuations.
- **Score < 60 :** Performance is deemed non-compliant, triggering a downgrade within the BW framework.

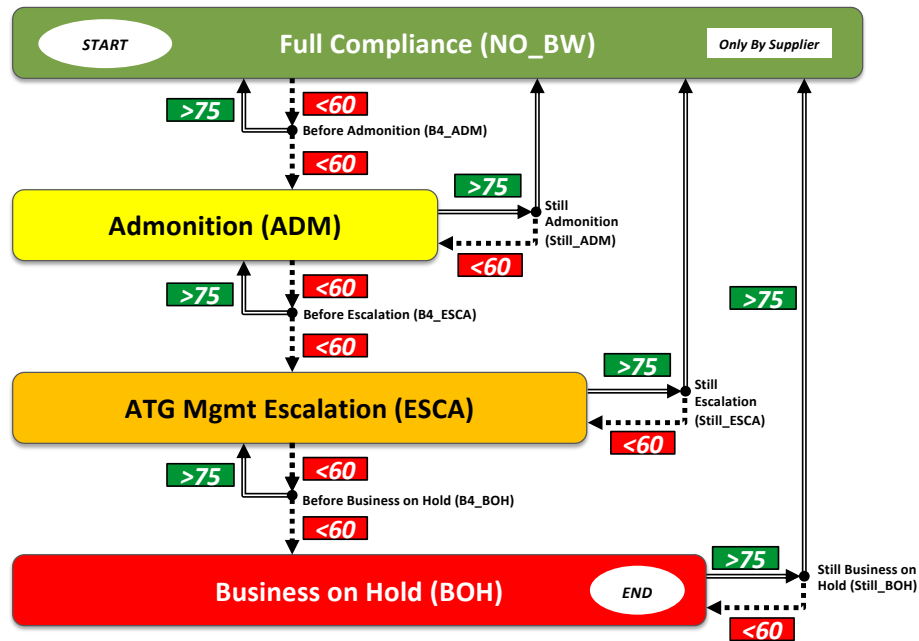


Figure 4.2: Decision matrix of the new Business Warning logic.

This new threshold system harmonizes with Ariston’s audit scoring methodology (detailed in Section 2.4), where 75/100 marks the benchmark for positive evaluation, thereby enhancing clarity and consistency for both internal teams and suppliers. The inclusion of a buffer zone (60–75) addresses one of the limitations of the old model, ensuring that suppliers with borderline performance are closely monitored but not arbitrarily escalated or promoted.

In addition to the adoption of quantitative thresholds, the structure of the Business Warning escalation path itself has been redesigned to resolve the operational weaknesses of the previous model:

- The *Phase-Out* status has been eliminated, as it was rarely applied and offered little added governance value.

- A new *Management Escalation (ESCA)* status has been introduced between *Admonition* and *Business on Hold*, creating a more gradual and realistic degradation path. Unlike *Admonition*, suppliers reaching the *ESCA* status are required to present their Corrective Action Plan (CAP) in person to Ariston's management team, demonstrating both commitment and the feasibility of their remediation plan. This in-person presentation differentiates *ESCA* from *Admonition*, where suppliers were only required to submit the plan remotely.
- Intermediate degradation steps now ensure that suppliers progress stepwise through *Admonition* and *Management Escalation* before reaching *Business on Hold*, giving them time to react and implement corrective actions before critical restrictions are applied.

The new BW logic achieves a balance between stability and responsiveness. It slows down negative escalations, providing suppliers with the opportunity to implement CAPs effectively, while accelerating positive transitions, allowing compliant suppliers to recover their status without unnecessary delays. Collectively, these modifications have made the BW system more understandable, fair, and operationally effective, aligning it closely with the data-driven nature of the new Vendor Rating and the strategic goal of proactive supplier development rather than punitive escalation.

4.7 Benefits of the New System

The implementation of the redesigned Vendor Rating and Business Warning system marks a fundamental advancement in Ariston Group's supplier governance framework, addressing the structural, methodological, and operational weaknesses identified in the previous evaluation system (detailed extensively in Chapter 3). By transitioning from a fragmented, subjective, and low-resolution rating process to a robust, data-driven approach, the new model delivers multiple strategic benefits.

Firstly, adopting a holistic, supplier-level perspective significantly mitigates the prior inconsistencies arising from plant and macrocategory segmentation. Suppliers are now evaluated comprehensively rather than in isolated, fragmented contexts. This change reflects operational reality more accurately, ensuring performance assessments represent true, overall supplier capabilities rather than isolated plant-based performance.

Secondly, the shift from a 1-to-4 scoring scale to a 0-100 scale introduces essential granularity and responsiveness to performance evaluation. The new scoring model addresses the old system's inability to recognize incremental improvements or nuanced performance fluctuations, thus enabling more precise and meaningful differentiation among suppliers. This change directly enhances the organization's capacity to detect emerging issues and trends proactively, moving beyond simplistic binary judgments toward a nuanced performance management paradigm.

Thirdly, the alignment of Vendor Rating scores with existing supplier audit evaluations has improved interpretability and integration across the entire quality assessment landscape. Suppliers and internal stakeholders now benefit from a unified scoring framework, eliminating confusion and streamlining communication. Additionally, the structured introduction of logarithmic normalization and Bayesian shrinkage methodologies has addressed the critical instability issue for low-volume suppliers, creating statistically reliable evaluations even for suppliers with limited data, and significantly reducing misleading volatility.

A crucial enhancement pertains to the newly restructured Business Warning logic, designed explicitly to resolve previously documented inconsistencies (see Section 3.4). The introduction of a clearly defined stability condition via the performance threshold band (scores 60–75) prevents unnecessary status volatility. Furthermore, by replacing the scarcely used *Phase-Out* status with the more practically relevant *Management Escalation (ESCA)* stage, the governance structure now provides suppliers sufficient opportunity to implement effective corrective actions. Notably, the *ESCA* status requires suppliers to present their Corrective Action Plans in person, strengthening accountability, management engagement, and ensuring corrective actions are credible and actionable.

The concrete advantages of these improvements are clearly demonstrated by the case study on Table 4.6 below. Under the old system, this supplier consistently maintained the highest rating (4) despite a substantial deterioration in quality performance, exemplified by rising Weighted Supplier Incidents (WSI) and Parts Per

Million (PPM) values. Due to the reliance on 8D reports, which were inconsistently issued, the previous model failed to capture or escalate the evident performance decline, allowing the supplier to avoid entering any corrective governance process. Conversely, the new scoring framework accurately captured the deterioration, reflecting it clearly in progressively decreasing Vendor Ratings (from 61 to 45 points) and triggering timely status escalations to *Admonition (ADM)* and subsequently to *Before Management Escalation (B4_ESCA)*, i.e., the intermediate degradation step between *ADM* and *ESCA*). This activation of the BW process ensures prompt governance intervention, corrective planning, and proactive supplier engagement, highlighting a direct operational advantage of the redesigned system.

Table 4.6: Case study comparison: legacy vs. new Vendor Rating and BW.

Year	2023				2024				2025	
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
Old Vendor Rating	4	4	4	4	4	4	4	4	4	4
Old BW Status	NO_BW	NO_BW	NO_BW	NO_BW	NO_BW	NO_BW	NO_BW	NO_BW	NO_BW	NO_BW
WSI	277	238	375	1,081	1,069	359	441	122	158	254
PPM (Target = 233)	258	357	375	1,689	2,475	382	1,259	61	7,853	190
New Vendor Rating	61	61	61	45	45	61	58	73	62	66
New BW Status	NO_BW	NO_BW	NO_BW	B4_ADM	ADM	ADM	B4_ESCA	B4_ESCA	B4_ESCA	B4_ESCA

In summary, the redesigned Vendor Rating and Business Warning system now offers a comprehensive, robust, and strategic approach to supplier governance. Its transparent, statistically sound methodology eliminates subjectivity, fragmentation, and granularity issues of the previous approach, enabling Ariston Group to more effectively manage supplier relationships and quality risks. By ensuring accurate, objective performance evaluation and timely, structured governance actions, the

new model positions supplier evaluation as a strategic driver for continuous improvement and long-term operational excellence, aligning closely with Ariston's broader quality management objectives.

4.8 Transitioning to the New System

Due to the substantial structural differences between the old and new Vendor Rating and Business Warning models, the transition will follow a carefully managed, stepwise process to ensure continuity and fairness. For the next few quarters, both logics will be applied in parallel, and supplier statuses will be cross-checked using a clear decision matrix: if a supplier is flagged as critical in the new logic only, it will temporarily receive an *Admonition* status and a formal *Admonition* letter, even if the old logic still shows *Full Compliance*; if it is critical in the old logic only, the old status will be maintained; and if it is critical in both logics, the status will either remain unchanged if consistent or be evaluated case by case if different. This staged approach avoids abrupt escalations, gives suppliers time to understand and familiarize with the new system before it fully replaces the old one.

To formalize the transition, a notification letter will be sent to all suppliers, informing them of the adoption of the new Vendor Rating and BW logic and the corresponding implications for their contractual obligations. Consequently, the Quality Assurance Agreement (QAA) will be updated to explicitly reflect the use of the 0-100 Vendor Rating with WSI and PPM indicators, as well as the new BW thresholds and escalation path. Internally, the IT infrastructure will be updated to automate the full calculation process, including WSI and PPM computation,

logarithmic normalization, Bayesian shrinkage and status assignment according to the new rules, while maintaining dual-logic evaluation during the transition period.

This comprehensive plan ensures that suppliers are properly informed and the organization can validate the new system's effectiveness while minimizing operational and relational risks.

5 Future Improvements

This chapter presents a roadmap that aligns the next evolution of Ariston Group's supplier evaluation framework with three enterprise priorities: digitalization, sustainability, and stronger integration between corporate governance and plant operations. Building on the advances introduced in the previous chapters, the proposals focus on three levers. First, enhancing WSI by incorporating incident gravity, in order to make escalation and supplier feedback more proportional to business impact. Second, implementing a centralized Supplier Governance Platform that unifies workflows and data, automates calculations and escalations, and provides explainable results. Third, expanding the Vendor Rating scope to include additional performance dimensions, including service and sustainability signals that are auditable and consistent with existing quality objectives.

The intention is to convert better analytics into faster and more credible decisions. Each section identifies the remaining limitations and introduces a method that connects to existing processes and data sources. The goal is to raise accuracy and stability while reinforcing traceability, reducing waste and field risk, and

supporting supplier development and risk management in a way that is coherent with Ariston's long term priorities.

5.1 Enhancing WSI with Incident Gravity

This section introduces a gravity-aware evolution of the Weighted Supplier Incidents indicator. The objective is to reflect not only where a non-conformity (incident) is detected along the value chain but also how severe it is. Recent governance changes have made the *gravity* field mandatory at alert creation, which provides a consistent basis to incorporate severity into the metric. The resulting model, referred to as WSI 2.0, preserves the existing detection point structure and augments it with a calibrated gravity weight for each incident. The subsections that follow set out the limitations of the current approach, present the proposed definition and weighting logic, and describe the expected impact on Vendor Rating and escalation governance.

5.1.1 Limitations of Current WSI Indicator

The current WSI indicator used at Ariston is based exclusively on the point of detection of each non-conformity. This means that incidents identified at different stages of the value chain, such as incoming inspection (W2), production line (W3), or in the market (W1), are assigned different weights depending on how late in the process the defect is discovered. While this structure encourages early detection and containment, it overlooks a fundamental dimension of quality: the actual severity of the non-conformity.

Until recently, incorporating gravity into the WSI formula was not feasible due to limitations in the way notifications were created and managed in the information system. Specifically, the *gravity* field, which classifies incidents into severity classes (A0, A1, B, C), was not a mandatory field for plants to fill when reporting a new alert. As a result, many incidents were recorded without a clearly defined severity classification, preventing a systematic and fair quantification of incident gravity across suppliers.

This gap significantly reduces the discriminative power of the WSI indicator. For example, an incident triggered by a minor cosmetic issue, such as a slight misalignment of the Ariston logo color, currently receives the same weight as a major technical failure, such as a dimensional error in a gas valve that could compromise product safety or functionality, if they are detected at the same stage. Such cases distort supplier performance assessments and weaken the ability of the Vendor Rating to prioritize the most impactful non-conformities.

Recognizing this limitation, Ariston recently implemented an important procedural update: the *gravity* field is now mandatory for all new quality alerts created in the system (third column from the right in Figure 2.1). This development opens the path to a significant evolution of the WSI indicator, allowing gravity to be incorporated alongside detection point as a second weighting factor. The resulting dual-weight model, referred to as WSI 2.0, will be discussed in the next subsection.

5.1.2 Proposed Formula for the New WSI

With the recent standardization of the *gravity* field as a mandatory input during the creation of each supplier-related quality alert, it is now possible to formally incorporate the gravity of non-conformities into the WSI calculation. The proposed formula builds upon the current framework (see Equation 4.1), maintaining the weighting for the detection point but adding a multiplicative gravity factor that allows a more granular and meaningful differentiation between minor and critical quality issues. The revised formula is as follows:

$$WSI = \frac{\sum_{i=1}^n (DW_i \times GW_i)}{GR} \times 1,000,000 \quad (5.1)$$

In this formula, n represents the total number of supplier-related non-conformities recorded within the analysis period. For a generic incident i , DW_i (Detection point Weight) reflects the location in the supply chain where the issue was identified, just like before. The new term GW_i (Gravity Weight) assigns a numerical value to the severity of the incident i , capturing the potential impact on product functionality, safety, and brand reputation. These weights must be carefully calibrated to reflect the true operational risk associated with each gravity class. As a preliminary proposal, the following mapping can be adopted:

Gravity Level	Weight
C	1
B	2
A0	4
A1	8

This weighting system ensures that severe issues, such as functional failures or safety-related defects, have a significantly larger impact on the final WSI score than cosmetic or minor deviations. For example, a defective gas valve that compromises product safety (A1) will influence the supplier's rating more heavily than a color mismatch in the logo (C), which would otherwise be treated similarly in the existing model. By multiplying the detection point weight by the gravity weight for each incident, the revised formula captures both the context and the criticality of quality issues, thereby improving the precision and strategic relevance of the WSI indicator.

5.1.3 Impact on Vendor Rating and Governance

The introduction of the updated WSI formula (WSI 2.0), which incorporates both detection point and gravity weights, represents a significant enhancement in the Vendor Rating system used at Ariston. Previously, the WSI metric used in the Vendor Rating calculation only considered the detection point, thus treating all quality incidents with the same weight regardless of their actual severity. As a result, minor issues detected late in the process could weigh more heavily than serious issues detected early, leading to distorted and sometimes counterintuitive

outcomes. By including gravity as a multiplicative factor, WSI 2.0 ensures that the severity of each incident is objectively represented in the supplier's performance score.

From a methodological perspective, the inclusion of WSI 2.0 in the Vendor Rating introduces greater discrimination in the evaluation process. Suppliers with frequent but low-severity issues will be differentiated from those with fewer but more critical problems. This change is expected to produce a more balanced and fair assessment, reducing the noise introduced by superficial deviations and emphasizing those non-conformities that truly impact product quality, safety, or customer satisfaction. Moreover, by amplifying the impact of serious incidents, the revised Vendor Rating encourages suppliers to proactively prevent and resolve critical issues, fostering a quality culture more aligned with Ariston's strategic priorities.

On a governance level, this proposal links directly to governance–supply chain integration. The improved sensitivity of the WSI 2.0-based Vendor Rating would allow better alignment between operational quality performance and strategic sourcing decisions. The new system provides a clearer and more defensible basis for triggering Business Warning statuses, escalating supplier interventions, or supporting requalification audits. It also enhances cross-functional transparency, as the refined WSI metric better communicates quality risks to procurement, engineering, and management stakeholders. Ultimately, the integration of WSI 2.0 into the Vendor Rating enables more effective supplier segmentation, risk

mitigation, and continuous improvement initiatives, reinforcing the credibility and utility of the supplier evaluation framework as a whole.

5.2 Toward a Centralized Supplier Governance Platform

One of the most promising directions for evolving Ariston's supplier evaluation framework is the creation of a centralized Supplier Governance Platform (SGP). The SGP is conceived as a dedicated operational environment that unifies supplier workflows and data across the lifecycle and serves as the single, auditable source of truth for supplier information and performance monitoring. It is not intended to replace the ERP, it remains the authoritative system of record for master data and transactions, while the SGP consolidates inputs from multiple sources, governs evaluation rules, and publishes consistent, explainable results.

Today, quality, procurement, and logistics often run parallel processes using fragmented tools, disconnected repositories, and offline spreadsheets. Activities such as calculating Vendor Rating, tracking Business Warning status, consolidating incident histories, and preparing cross-functional views require significant manual effort. This fragmentation generates version misalignments, slows decisions, and creates avoidable disputes about data. A centralized SGP addresses these pain points by automating calculations and escalations, enforcing common business rules, and aligning stakeholders on shared definitions and synchronized datasets.

The SGP would extend beyond static reporting. It would provide end-to-end workflow support and full traceability for supplier qualification, audits, non-

conformities, claim-back management, and the computation and publication of core quality indicators. Role-based access, explicit data lineage, and time-stamped versions of rules would strengthen governance and reproducibility. Key capabilities, kept at a high level, include:

- Data integration and validation across transactional and quality systems, with reconciliation routines that protect data quality.
- A governed rules engine for Vendor Rating and Business Warning, with version control, audit trails, and effective dating for any change in logic.
- Workflow orchestration for corrective actions and escalations, aligned with defined responsibilities and timelines.
- Diagnostics and explainability that allow users to drill from an overall score to the underlying incidents, deliveries, and applied rules.

Implementing an SGP also clarifies accountability. Definitions, ownership, and change management for indicators and rules are formalized. Monitoring and exception handling are standardized, and a small set of health checks ensures that live behavior remains aligned with governance expectations.

This platform advances digitalization and is a pragmatic and scalable target state. The ERP continues to provide the backbone for transactions and master data. The SGP sits above it to standardize rules, automate recurring activities, provide explainable analytics, and align all functions on a single reliable view of supplier

performance. On this foundation, the next section considers how the scope of the Vendor Rating can be broadened in a controlled and modular way.

5.3 Expansion of the Vendor Rating Scope

The redesign delivered in this thesis provides a robust and transparent quality-centric Vendor Rating. At the same time, supplier value creation extends beyond conformance to technical specifications. Building on the data foundation and rule governance enabled by the Supplier Governance Platform, Ariston would be able to progressively broaden the scope of the Vendor Rating so that it reflects a fuller view of performance while preserving objectivity, auditability, and comparability over time.

A quality-only scorecard, centered on Parts Per Million and Weighted Supplier Incidents, captures a critical but narrow dimension of behavior. In contemporary supply networks, superior performance also depends on the reliability of logistics execution, commercial discipline, operational maturity, responsiveness in problem-solving, and measurable progress on sustainability. Limiting the evaluation to quality risks overlooks material deficiencies outside conformance and can fail to recognize suppliers that create outsized value in adjacent domains. The signal that results is less informative for cross-functional decision-making and less aligned with strategic sourcing objectives.

A pragmatic path forward is to extend the Vendor Rating with a small number of well-governed blocks drawn from procurement, logistics, risk, and sustainability,

added in a modular way once data ownership, definitions, and controls are in place. Logistics performance can be represented through measures such as On Time In Full, lead time adherence, and compliance with agreed packaging and Incoterms, which stabilize production planning and reduce buffers. Procurement and commercial discipline can be reflected through price competitiveness relative to benchmarks, cost improvement trajectories, and contract compliance, so that sustained value delivery is visible without conflating price with quality. Audit and system maturity can be introduced through scores from process and system audits conducted with standardized checklists and rubrics, as signals of capability and risk mitigation. Collaboration and responsiveness can be captured by objectively measurable engagement in improvement projects, timeliness of containment and corrective actions, and participation in structured root cause analysis, all derived from workflow systems with time-stamped events rather than subjective scores. Sustainability can enter through verifiable environmental and social indicators such as certified management systems, energy efficiency programs, waste and recyclability metrics, and adherence to codes of conduct, sourced from audits, certifications, or validated disclosures mapped to recognized frameworks.

Any extension should meet the same standards that govern quality indicators. Operational definitions must be clear, data capture must be objective and repeatable across plants and months, lineage must be documented, and the rules that map indicators to scores must be versioned and effectively dated. Weighting across dimensions should be calibrated transparently and subjected to sensitivity checks

and backtesting before adoption, with phased rollout to safeguard continuity of comparisons.

A multidimensional scorecard further supports Ariston's strategic objectives, connecting governance–supply chain integration with sustainability and offering tangible benefits. It provides a more balanced view that improves prioritization and clarifies trade-offs across functions. It sharpens incentives by recognizing value creation beyond conformance and directs development resources to the areas with the highest impact. It strengthens internal alignment because quality, procurement, and logistics contribute to and consume a shared evaluation model. It advances sustainability objectives by integrating auditable environmental and social indicators into the score. Finally, it enables more nuanced segmentation of the supply base, supporting preferred programs, targeted development tracks, conditional onboarding, or exit planning based on a broader evidence set.

In sum, expanding the Vendor Rating beyond quality is a natural evolution once a consolidated data and rule environment is in place. By adding a limited number of governed, explainable dimensions in a staged manner, Ariston can increase the strategic relevance and depth of supplier evaluation while preserving the principles that make the current model credible and easy to operate.

6 Conclusion

The redesign of Ariston Group's supplier quality performance evaluation system tackled a complex, cross-functional problem: how to translate heterogeneous operational events into a single, comparable signal that supports objective decision-making across plants and over time. At the outset, the thesis established that stable supplier evaluation demands robust governance, minimal subjectivity, and statistically sound treatment of sparse, heavy-tailed data. Drawing on longitudinal datasets and stakeholder workshops, it devised a unified methodology in which transactional indicators - Weighted Supplier Incidents (WSI) and Parts Per Million (PPM) - are normalized, shrink-adjusted, and mapped to a 0–100 scale. The process integrates logarithmic transformation to mitigate skewness and Bayesian shrinkage to stabilize scores for suppliers with low exposure.

The resulting Vendor Rating (VR) is not simply a mathematical artifact; it codifies a cultural shift from siloed, plant-level ratings to a global supplier perspective. By separating detection from evaluation and removing the subjective 8D count, the approach encourages consistent quality reporting and empowers continuous improvement. The methodology balances granularity with interpretability, ensuring

that ratings remain sensitive to meaningful differences without overreacting to statistical noise. Moreover, by embedding a redesigned Business Warning (BW) logic, the system lays the groundwork for fairer supplier accountability and more predictable decision-making.

This thesis should also be read as a blueprint for future evolution that speaks directly to Ariston Group's strategic challenges. Its analysis exposed limitations of the current metrics, notably their inability to reflect the gravity of non-conformities and broader dimensions such as logistics, commercial behavior, sustainability, and collaboration. The proposed enhancements, including WSI 2.0 with gravity-weighted incidents and a centralized supplier governance platform to consolidate data and enforce rules, point toward a more holistic, real-time control tower for supplier performance. These improvements will require careful design to maintain objectivity, transparency, and cross-functional buy-in, but they hold the promise of turning supplier evaluation into a more powerful strategic lever that balances quality, cost, and sustainability, and that converts governed analytics into faster and more credible decisions across plants and functions.

Ultimately, the thesis demonstrates that rigorous statistical methods and thoughtful process design can transform quality management from a reactive reporting exercise into a proactive, strategic discipline. By combining empirical analysis with stakeholder engagement, it shows how data science can bridge functional silos and deliver a fairer, more stable evaluation of supplier performance. Its contributions extend beyond the immediate application at Ariston: the principles

of normalization, shrinkage, and evidence-based thresholds may inform vendor evaluation systems in diverse industries. Future work should focus on validating the proposed extensions, integrating additional performance dimensions, and measuring the impact of the new system on business outcomes and supplier behavior.

References

- AESQ Strategy Group. (2021, March 8). *RM13000: 8D problem solving method: An AESQ reference manual supporting SAE AS13100TM standard* (Reference Manual No. AESQRM000202103). SAE International: SAE Industry Technologies Consortia (SAE ITC). <https://aesq.sae-itc.com/binaries/content/assets/itc/content/aesq/download-tracking/aesqrm000202103.pdf>
- Baker, R. D., & McHale, I. G. (2017). An empirical bayes model for time-varying paired comparisons ratings: Who is the greatest women's tennis player? *European Journal of Operational Research*, 258(1), 328–333. <https://doi.org/10.1016/j.ejor.2016.08.043>
- Barsalou, M., Grabowska, M., & Perkin, R. (2023). Inquiry into the effectiveness of eight discipline-based problem-solving. *Quality Innovation Prosperity*, 27(2), 61–76. <https://doi.org/10.12776/qip.v27i2.1839>
- Barsalou, M., & Perkin, R. (2025). It's all in the approach: An 8d-based process for problem solving. *Quality Progress*, 58(6), 20–25.
- Carr, A. S., & Pearson, J. N. (1999). Strategically managed buyer–supplier relationships and performance outcomes. *Journal of Operations Management*, 17(5), 497–519. [https://doi.org/10.1016/S0272-6963\(99\)00007-8](https://doi.org/10.1016/S0272-6963(99)00007-8)

- Chuang, H. H.-C., & Oliva, R. (2015). Inventory record inaccuracy: Causes and labor effects. *Journal of Operations Management*, 39-40, 63–78. <https://doi.org/10.1016/j.jom.2015.07.006>
- Doshi, J. A. (2019). The significance of supplier performance management in quality improvement - a case of construction equipment manufacturing. *International Journal of Quality and Innovation*, 4(1-2), 88–98. <https://doi.org/10.1504/IJQI.2019.101409>
- Ford Motor Company. (1987). *Team oriented problem solving (tops) manual*. Dearborn, MI, United States.
- Hawkins, T. G., Gravier, M. J., & Muir, W. A. (2020). The role of supplier performance evaluations in mitigating risk: Assessing evaluation processes and behaviors. *Industrial Marketing Management*, 87, 2–17. <https://doi.org/10.1016/j.indmarman.2020.03.004>
- Health Information and Quality Authority. (2013, February). *Guidance on developing key performance indicators and minimum data sets to monitor healthcare quality*. Health Information and Quality Authority. <https://www.hiqa.ie/sites/default/files/2017-01/KPI-Guidance-Version1.1-2013.pdf>
- Katsikeas, C. S., Paparoidamis, N. G., & Katsikea, E. (2004). Supply source selection criteria: The impact of supplier performance on distributor performance. *Industrial Marketing Management*, 33(8), 755–764. <https://doi.org/10.1016/j.indmarman.2004.01.002>

- Ma, Q., & Li, H. (2024). A decision support system for supplier quality evaluation based on mcdm-aggregation and machine learning. *Expert Systems with Applications*, 242, 122746. <https://doi.org/10.1016/j.eswa.2023.122746>
- Modi, S. B., & Mabert, V. A. (2007). Supplier development: Improving supplier performance through knowledge transfer. *Journal of Operations Management*, 25(1), 42–64. <https://doi.org/10.1016/j.jom.2006.02.001>
- Moers, F. (2005). Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society*, 30(1), 67–80. <https://doi.org/10.1016/j.aos.2003.11.001>
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1–15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)
- Salimian, H., Rashidirad, M., & Soltani, E. (2021). Supplier quality management and performance: The effect of supply chain oriented culture. *Production Planning & Control*, 32(11), 942–958. <https://doi.org/10.1080/09537287.2020.1777478>
- Schwartz, A. L. (2021). Accuracy vs. incentives: A trade-off for performance measurement. *American Journal of Health Economics*, 7(3), 333–360. <https://doi.org/10.1086/714374>
- Sreedharan, J., Subbarayalu, A. V., Kamalasanan, A., Albalawi, I., Krishna, G. G., Alahmari, A. D., Alsalamah, J. A., Alkhathami, M. G., Alenezi, M., Alqahtani, A. S., Alahmari, M., Phillips, M. R., & Macdonald, J. (2024). Key performance indicators: A framework for allied healthcare educational institutions. *Clini-*

coEconomics and Outcomes Research, 16, 173–185. <https://doi.org/10.2147/CEOR.S446614>