

Data Quality Score: Measure of Master Data Reliability in Procurement

The data quality score is a metric that assesses the quality of data stocks – in particular item master data, supplier data and transaction data – on a scale. For buyers and supply chain managers, it is a central steering instrument, because faulty master data leads to mis-ordering, incorrect delivery dates and unnecessary queries. A systematically collected score makes data problems visible, quantifiable and thus prioritisable. In this glossary article we explain the definition, calculation methods, KPIs and practical measures relating to the data quality score.

Key facts

- The data quality score quantifies data quality typically in per cent or on a scale of 0 to 100.
- Commonly used dimensions are completeness, correctness, consistency, currency, uniqueness and plausibility.
- In practice, values above 95 per cent are considered a good benchmark for critical master data fields such as purchase prices or material numbers.
- Faulty master data typically cause queries, mis-ordering, dunning runs and returns along the entire source-to-pay chain.
- An effective score requires clear field priorities, defined validation rules and one responsible party per data object.
- Data quality scores can be determined automatically and monitored continuously in dashboards.

Definition: Data quality score

The data quality score is an aggregated metric that assesses the quality of individual data records, data fields or entire data stocks. It combines several quality dimensions into a single value, enabling quick comparison across objects, teams or systems.

Distinction from data quality in general

Data quality describes the business condition of data based on criteria such as completeness or correctness. The data quality score turns this into a measurable, comparable number. While data quality is a target state, the score is a steering instrument: it shows where the actual state lies, how it develops over time and where measures need to be applied. Without a score, data quality often remains a diffusely perceived shortcoming; with a score, it becomes a controllable metric alongside delivery reliability or procurement costs.

Relevant data objects in procurement

For buyers, the focus is above all on item master data with material numbers, units of measure, technical characteristics and purchase prices, as well as supplier master data with addresses, bank details and certifications. Added to this are transaction data from purchase orders, delivery advices and invoices. Errors in these objects have an immediate impact: an incorrect unit of measure leads to

over-ordering, an outdated bank detail jeopardises payment, and an incomplete supplier status complicates risk assessment. Especially in supplier bank detail checks and creditor onboarding, high data quality is a prerequisite for smooth processes.

Methods and approach for score determination

The calculation of a data quality score usually follows a fixed process: definition of the quality dimensions, determination of the assessment rules per field, weighting according to business relevance and aggregation into an overall value. Depending on maturity, the assessment is carried out manually, rule-based or automatically with data profiling and matching tools.

Defining quality dimensions

Orientation towards six dimensions has proven effective: completeness (all mandatory fields filled), correctness (values match reality), consistency (no contradictions between systems), currency (timestamps and revisions are correct), uniqueness (no duplicates) and plausibility (values within sensible ranges). Uniqueness is central to duplicate management, while currency is decisive, for example, in supplier location risk assessments, as location data must be checked regularly. Each dimension is assessed individually per object type.

- Completeness: proportion of mandatory fields filled
- Uniqueness: result of the duplicate check
- Plausibility: value range and reference checks

Rule set and weighting

Each data field receives a validation rule (e.g. mandatory field, format, referential integrity, threshold value) and a priority. Critical fields such as purchase price, unit of measure or supplier number are given a higher weight than optional descriptive fields, because errors there have direct operational consequences. A weighting of roughly 70:20:10 between core, standard and additional fields is common. The weighting should be agreed jointly by the purchasing, controlling and IT departments and reviewed periodically, as business relevance can shift.

Automated profiling and audits

In practice, companies combine automated rule checks with sample-based business audits: software detects gaps, format errors and duplicates daily, while factual correctness (e.g. whether a technical characteristic matches the component) can only be verified through expert review. Analogous to the Gage R&R measurement system analysis from quality assurance, it is advisable to check occasionally how reliable the checking method itself is, i.e. whether different reviewers arrive at the same result. Only then is the score itself robust.

Key KPIs around the data quality score

The data quality score is itself a metric, but it only unfolds its steering effect in combination with further metrics. A small set of metrics that links score values, error rates and business impacts makes sense.

Score values by object type and dimension

Reporting is typically separated by item master data, supplier master data and transaction data, as well as by dimension. An overall score for the company is useful for communication, but the detailed values are what guide action: for example, a high completeness of 98 per cent may exist alongside weak

currency of 82 per cent – the measures differ considerably. In practice, benchmarks are values above 95 per cent for critical fields and above 85 to 90 per cent for secondary data as acceptable.

Linking with process metrics

The score only shows its effect in conjunction with process metrics: proportion of purchase orders with price or quantity deviations, rate of invoices rejected due to master data errors, proportion of duplicate suppliers or creditors, and queries resulting from incomplete enquiry data. Delivery date tracking also suffers from imprecise dates and quantities in order data. Those who evaluate these metrics together can demonstrate the economic benefit of data cleansing and justify investment in data maintenance.

- Error rate in purchase orders per 1,000 line items
- Proportion of invoices rejected due to master data errors
- Duplicate rate in supplier and creditor stocks

Trend and benchmark analysis

Individual values say little; trends are decisive. A monthly or quarterly reported trend shows whether cleansing projects are working or whether new sources of error – for example through mergers, system changes or rapidly growing catalogues – are depressing the score. Internal comparison between sites, product groups or teams creates healthy competition for data quality. An external supplier benchmark of data delivery quality, for example in delivery advice (ASN) processing, complements the picture with the view of business partners.

Risks, causes and controls

Poor data quality rarely arises maliciously, but rather through unclear responsibilities, lack of validation, interface losses and organisational changes. The data quality score makes these causes measurable – controlling must address them systematically.

Typical causes in the procurement environment

Common causes are manual maintenance without mandatory field logic, multiple source systems without clear leadership, merger- and migration-related duplicates, and inadequate requirements management for new materials and suppliers. Errors arise especially in creditor onboarding when incoming data is not validated in a structured way. Company growth also has an effect: new sites, additional supplier locations and expanding contract manufacturing grow the data stock faster than maintenance capacity.

- Lack of clear data ownership
- Interface losses between ERP, SRM and catalogue systems
- Migration and merger duplicates without ongoing duplicate checks

Controls and process measures

Effective controls start at the source: validation rules already during the creation of materials and suppliers, mandatory field concepts in requirements management, four-eyes principle for critical fields such as bank details, and periodic cleansing runs. Quality checks integrated into the source-to-pay (S2P) process – for example at order entry and before invoice approval – prevent errors from propagating further. Data quality scores belong on the agenda of regular procurement reviews, alongside supplier status and risk assessments.

- Validation rules directly in the creation processes

- Four-eyes review of critical payment and price data
- Quarterly cleansing and duplicate check runs

Organisation and responsibility

Sustained good data quality requires clear roles: a data owner per data object decides on definitions and priorities, data stewards maintain operationally, IT operates tools and interfaces. A procurement rapid response team can prioritise and rectify at short notice in the event of acute data problems – for example after system failures or faulty migrations. Training the staff involved, anchored for example in a procurement competency matrix, ensures that the rules are understood and applied.

Data quality score in industrial procurement practice

The score unfolds its benefit along concrete procurement processes – from the enquiry through supplier management to the centralisation of master data in the supplier network. A high score is a prerequisite for further digitalisation projects.

Enquiries, awards and supplier management

Complete and correct master data accelerate enquiry creation and award: drawing references, bills of materials and delivery terms are correct, and queries decrease. In the bidder register and supplier selection, well-maintained certificate and capacity data help to identify suitable providers more quickly. In make-or-buy decisions and toolmaking procurement, calculation models rely on correct bills of materials and process data – data errors here lead to distorted cost comparisons and false advantages or disadvantages.

Operational running and logistics

In day-to-day operations, data errors directly affect availability and delivery dates: incorrect replenishment lead times distort the safety stock calculation, imprecise call-off data leads to shortages in the call-off order, and inconsistent packaging and advice data complicate delivery advice (ASN) processing. The delivery schedule and delivery date tracking also presuppose that quantities, units and dates in the systems are correct. High data quality here reduces dunning runs, firefighting and expensive express deliveries.

- Correct replenishment lead times as the basis for safety stock calculation
- Clean advice data flow for stable ASN processing
- Consistent call-off data for reliable delivery schedules

Financial processes and payment transactions

In invoicing and payment, data quality is directly measurable: duplicate invoices, incorrect bank details or differing creditor names lead to process stoppages and compliance risks. In self-billing (ERS), where invoicing is done from goods receipt data without a classical invoice, correct price and quantity data are an essential prerequisite – every data error here translates directly into payment differences. Financial processes in areas such as forfaiting and balance management also rely on clean creditor and transaction data.

Current developments: automation, AI and data ecosystems

Data quality scores are increasingly collected automatically and monitored continuously. Machine learning detects anomalies in master data, matching processes identify duplicates across system

boundaries, and dashboards make scores usable for business departments without database knowledge.

AI-supported checking and enrichment

Language models and classification methods help to standardise free-text material descriptions, categorise items and extract missing attributes from documents. This significantly reduces manual maintenance effort and makes duplicates findable that can only be identified through semantic similarity. Control remains important: AI suggestions should serve as review templates and be approved by the business departments, so that no systematic false assumptions creep into the master data.

Data quality across company boundaries

In the supplier network and integrated source-to-pay (S2P) platforms, a shared understanding of data emerges between customer and supplier: master data is maintained once and used many times. This shifts responsibility for data quality partly into the value chain. Buyers should therefore anchor data quality requirements in supplier agreements – for example in service contracts with a development or logistics partner or in long-term framework agreements – and regularly check compliance using score-like metrics.

Regulatory requirements

Increasing reporting obligations, for example on supply chains and sustainability, raise the bar for supplier data: certificates, locations, proof of origin and metrics must be up to date, complete and verifiable. A low data quality score for these fields thus becomes a compliance risk. Legal requirements in detail are highly context-dependent and should be examined case by case – for example with the legal or compliance department. The score provides an early warning instrument that makes gaps visible in good time.

Typical metrics

Data quality score core fields: > 95%

Benchmark for critical master data fields such as prices, units, bank details

Data quality score secondary data: > 85–90%

Acceptable range for optional descriptive and additional fields

Duplicate rate: < 1%

Target value for supplier and creditor stocks after cleansing

Invoice rejections due to master data errors: < 2%

Indicator of the process benefit of high data quality

Score trend: rising monthly

Trend more important than individual values; setbacks indicate new sources of error

Practical example

A medium-sized mechanical engineering company with around 4,000 active suppliers noticed increasing process stoppages when introducing automated invoice matching: price deviations, duplicate invoices and incorrect bank details. An initial score measurement showed only 78 per cent quality for supplier master data, mainly due to outdated bank details and duplicates from an earlier ERP migration. The company established a structured cleansing programme with data ownership in procurement and

quarterly rule reviews.

- Score measurement per object type with weighted field priorities, starting with supplier and item master data
- Ongoing duplicate checks and supplier bank detail checks with the four-eyes principle for critical fields
- Validation rules and mandatory fields anchored directly in creditor onboarding
- Monthly score reporting to procurement and controlling, target value above 95 per cent for core fields

Conclusion

The data quality score translates an abstract aspiration – good data – into a controllable metric. It makes sources of error visible, enables prioritisation according to business relevance and provides evidence that cleansing measures work. In procurement this pays off directly: fewer queries, more reliable delivery dates, smoother invoice processes and sound decisions in make-or-buy and supplier selection. Clear responsibilities, validation at source and regular reporting are decisive. Those who make data quality measurable create the foundation for digitalisation, automation and reliable procurement metrics.

Frequently asked questions

What is a data quality score?

A data quality score is a metric that expresses the quality of data – such as item or supplier master data – as an aggregated value, typically in percent or on a scale of 0 to 100. It combines dimensions such as completeness, correctness, consistency, currency and uniqueness. This allows data stocks to be compared, trends to be tracked and cleansing measures to be prioritised.

How is a data quality score calculated?

First, data fields are defined with validation rules (mandatory field, format, value range, reference) and weighted according to business criticality. Each field is then checked for rule compliance, and the results are aggregated into field, record and stock scores. The weighting is agreed jointly between the business departments and IT; automated data profiling performs the ongoing checks.

Which data quality dimensions are there?

Six dimensions are common: completeness, correctness, consistency, currency, uniqueness and plausibility. Completeness and uniqueness can be checked automatically, for example through duplicate checks. Correctness in the business sense, however, usually requires manual review by the responsible departments.

From what value is a data quality score considered good?

In practice, values above 95 per cent are considered a good benchmark for critical master data fields such as purchase prices, material numbers or bank details. For secondary data, 85 to 90 per cent is often acceptable. The trend is more important than individual values: falling scores indicate new sources of errors, rising values demonstrate effective measures.

What are the consequences of poor master data in procurement?

Faulty master data leads to mis-ordering, incorrect units of measure, rejected invoices and unnecessary

queries. They also distort safety stock calculations and delivery date tracking, which can result in express deliveries and dunning runs. In addition, compliance risks arise, for example with outdated bank details or missing certificates in supplier management.

How can data quality in procurement be improved?

Proven measures include validation rules directly during the creation of materials and suppliers, clear data ownership, mandatory field concepts in requirements management and periodic cleansing runs including duplicate checks. In addition, automated profiling, four-eyes review of critical fields and monthly score reporting help. Training the staff involved ensures that the rules are applied in day-to-day work.