

Uncovering

The Hidden Cost of

Poor Product Data in Indian E-Commerce



Loading ... 76%



FOREWORD

India's digital commerce ecosystem is undergoing a profound transformation. Rapid growth in E-commerce and Q-commerce has reshaped how Indian consumers discover, evaluate, and purchase products, while creating unprecedented opportunities for brands, marketplaces, and logistics providers. This expansion, however, has also placed new demands on the completeness, accuracy, and consistency of product information flowing across increasingly complex digital and physical supply chains.

As digital commerce scales, high quality product data has become a core enabler of growth. From online discovery and conversion to fulfilment, returns, and regulatory compliance, every transaction today is mediated by data. Yet, much like physical supply chains once struggled with fragmentation and opacity, digital commerce now faces its own structural weakness: inconsistent and inaccurate product data. Gaps in standardisation across identifiers, attributes, and content continue to create friction across platforms, eroding value in ways that often remain hidden within day-to-day operations.

This challenge is no longer marginal. As competition intensifies and margins tighten across E-commerce and Q-commerce, even small inefficiencies in product data compound rapidly at scale – suppressing conversion rates, increasing returns, inflating operating costs, and constraining profitability for both platforms and sellers. Strengthening the foundations of digital commerce therefore requires the same discipline that transformed physical supply chains – common standards, shared data languages, and robust data governance frameworks.

It is in this context that GS1 India and Kanvic Consulting have jointly undertaken this study. GS1 India, as part of the global GS1 system, brings deep expertise in product identification, standards, and data interoperability. Kanvic brings a strategic and analytical perspective on India's e-commerce ecosystem, with experience translating operational challenges into quantified business and economic impact. Together, we aim to move the conversation beyond anecdotal observations and provide a fact-based assessment of how product data quality directly affects revenue, margins, costs, and consumer trust.

Through this report, we seek to quantify the economic consequences of poor product data quality and demonstrate how standards-led, collaborative data governance can become a strategic lever for improving marketplace efficiency, seller profitability, and regulatory compliance. By strengthening the digital foundations of commerce, India can ensure that the next phase of E-commerce growth is not only faster – but smarter, more resilient, and more sustainable.

S. Swaminathan
Chief Executive Officer
GS1 India

Deepak Sharma
Co-founder and Director
Kanic Consulting

TABLE OF CONTENTS

1	Executive Summary	
2	India E-commerce Market	2
2.1	India's E-commerce Growth Outlook	
2.2	Structural and Technology growth drivers	3
2.3	Category penetration is evolving at different speeds	4
2.4	Platform concentration across E-commerce and Q-commerce	5
2.5	Category - Platform specialisation across E-retail	6
2.6	Product data as a critical enabler	7
3	Product Data Quality	9
3.1	Product data quality - Outlook	
3.2	E-commerce data lifecycle and challenges	10
3.3	Data quality methodology	11
3.4	Status of data quality in numbers	13
3.5	Completeness and Accuracy - Parameters of data quality risk	
3.6	Consumer attributes impact on data quality failures	14
3.7	Data fragmentation across platforms	
4	Business and Economic Impact	18
4.1	Economic cost of poor data quality	
4.2	Financial impact - Methodology	
4.3	Financial exposure by marketplace and category	20

4.4	Financial impact - Marketplace	21
4.5	Financial impact - FMCG & Food	
4.6	Financial impact - Fashion Apparel & Accessories	
4.7	Financial impact - Beauty & Personal Care	
4.8	Financial impact - E-Pharma	
5	Global Standards	23
5.1	Why standards matter for E-commerce	24
5.2	Standards across E-commerce value chain	
5.3	Standards for the next-generation E-commerce	25
6	Way Forward	27
6.1	For Brands	28
6.2	For Marketplaces	
6.3	For Regulators	29
7	Conclusion	30
8	Methodology	31
9	Disclaimer	32
10	Abbreviations	33



EXECUTIVE SUMMARY

India's E-commerce and Q-commerce sectors are entering a decisive phase of growth. Expanding digital adoption, rapid category diversification, and the rise of fast delivery models are reshaping how consumers discover and purchase products. Yet as online commerce scales, a critical structural weakness is constraining performance across the ecosystem: poor product data quality.

This study finds that product information across Indian digital commerce platforms remains uneven and materially poor. A majority of SKUs fail to meet acceptable thresholds of completeness, accuracy, and conformity, undermining the efficiency of digital transactions at scale. These shortcomings do not remain confined to listing quality. Instead, they propagate across the entire value chain from discovery and conversion to fulfilment, returns, inventory management, and regulatory compliance.

Our analysis shows that poor product data quality translates directly into significant economic loss. Across India's digital commerce ecosystem, poor data quality is estimated to result in approximately ₹5,000 crore in annual revenue loss, in addition to ₹2,000 crore in lost gross margin and ₹1,900 crore in excess return costs. These impacts are structural rather than episodic – embedded within everyday operations and amplified by the scale of digital commerce.

Financial exposure varies sharply by channel and category. Marketplace and D2C e-retail accounts for the majority of losses, driven by data-intensive categories such as Beauty & Personal Care, FMCG & Food, and Fashion. Q-commerce, while smaller in absolute scale, also exhibits material exposure, particularly in grocery and non-grocery essentials where speed of fulfilment magnifies the cost of inaccurate product data. These losses arise not from a single failure point, but from repeated inefficiencies across multiple stages of the transaction lifecycle.

Crucially, the largest share of economic impact is concentrated in two areas: gross margin erosion as a result of lost revenue and return costs, which together account for the majority of total losses. Lower conversion from poor product content, increased return rates driven by misleading or inaccurate listings, and operational inefficiencies across fulfilment and post-purchase processes combine to systematically weaken profitability for both marketplaces and sellers.

At a time when competition is intensifying and margins are tightening, improving product data quality is no longer a technical hygiene issue - it is a strategic imperative. High-quality, standardised product data is foundational to automation, scalability, and trust in digital commerce.

GS1 standards provide the global language required to address this structural weakness. By enabling consistent product identification, harmonised attributes, and interoperable data exchange across platforms, GS1 helps create a single source of truth for product information. Strengthening adoption of global standards across India's digital commerce ecosystem can unlock measurable improvements in marketplace efficiency, seller profitability, regulatory compliance, and consumer confidence.

As India's digital economy continues to expand, the quality of product data will increasingly determine the quality of growth itself.

LINE SHOP

ME MOBILE ELECTRONICS TECHNOLOGY CLOTHING SHOES & BEAUTY FURNITURE GROCERIES FOOD

Search



Mobile Devices
Mobile Phone ~~\$849~~ \$699
★★★★★
[+ MORE INFO](#)



Home Electronics
Laptop Computer ~~\$1299~~ \$999
★★★★★
[+ MORE INFO](#)



Technology
Virtual Reality Goggles ~~\$299~~ \$199
★★★★★
[+ MORE INFO](#)



Clothing
Blue Polo Shirt ~~\$49~~ \$29
★★★★★
[+ MORE INFO](#)



Shoes & Beauty
Classic Oxfords ~~\$179~~ \$99
★★★★★
[+ MORE INFO](#)



Furniture & Home
Home Office Chair ~~\$199~~ \$119
★★★★★
[+ MORE INFO](#)



Groceries
Eco Groceries Bag ~~\$9~~ \$5
★★★★★
[+ MORE INFO](#)



Food Delivery
Sushi Set ~~\$9.99~~ \$7.99
★★★★★
[+ MORE INFO](#)

INDIA E-COMMERCE MARKET

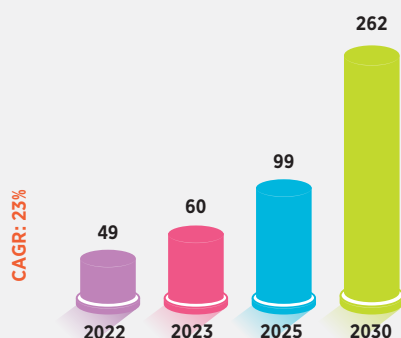
2.1 India is set to be the world's fastest growing E-commerce market through 2030

India's E-commerce market has expanded rapidly over recent years and is projected to grow further, increasing from \$99 billion in 2025¹ to \$262 billion by 2030 reflecting a 23% CAGR² when compared with 2022. Alongside this, the Q-commerce segment is forecast to grow from \$8 billion in 2025³ to \$46 billion by 2030⁴ representing a significantly higher 74% CAGR over the same period.

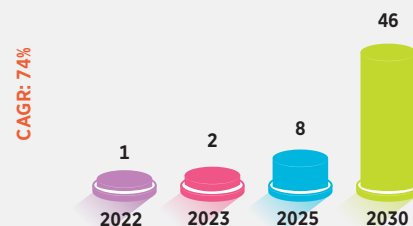
Exhibit 1

India's digital commerce sector growth

Market size of Indian E-Commerce (GMV) in \$ Bn



Market size of Indian Q-Commerce (GMV) in \$ Bn



¹ Study E-commerce in India, statista, ² Kanvic's research, ³ India's rapid delivery race, CareEdge, ⁴ Kanvic's research

In a global context, India’s growth rates outpace those of other major E-commerce markets such as China, the United States, Europe, and Asia Pacific⁵.

Indian E-Commerce is also ahead of the pack across key metrics globally

Table 1

Metric	India	China	United States	Europe	Asia Pacific*
GMV	99.4 Bn	3.2 Tn	1.8 Tn	UK: 141.8 Bn / DE: 105.3 Bn	1.12 Tn
Market Penetration	10-11%	47%	15.80%	UK: 30.6% / Germany: 66% (user-base)	South Korea: 30% / Indonesia: 31.9%
Mobile-Commerce Share	>80%	>85%	~50%	UK: 60% / Germany: 66%	Mobile-first; strong social commerce orientation
Growth (CAGR)	~19%	~10.4%	~10.7%	UK: ~3.95%	~7.8%
Average Delivery Time	2-4 days	2-3 days	2-5 days	3-5 days	NA
Average Order Value (AOV)	33	86	153	53-65	NA

*Excludes India and China,
All figures are in USD, based on the year 2025

2.2 Structural and technological shifts are jointly reshaping the next phase of Indian E-commerce growth

India’s E-commerce expansion is being driven not only by market scale, but by a combination of structural demand shifts and rapid technology adoption that are fundamentally changing how digital commerce operates. Growth is increasingly coming from beyond Tier-1 cities, with Tier-2 and Tier-3 markets now accounting for 60% of new online shoppers and 45% of total orders, supported by regional language platforms, improved logistics networks, lower warehousing costs, and wider internet access under Digital India.

At the same time, real-time digital payments are structurally reducing friction in online purchasing, with UPI surpassing 208 billion transactions in FY2024 and BNPL growing at 22.7% CAGR, expanding addressable demand by making E-commerce more accessible across income segments.

In parallel, technology is no longer just an efficiency layer but a core growth engine. Platforms are deploying AI-driven

personalization and hyper-local optimization to operate at pin code level dynamically tailoring assortment, pricing, and recommendations to local demand patterns. These capabilities are already delivering 30-40% improvements in conversion rates⁶, while reducing inventory waste by aligning supply more tightly with localized consumption. This marks a structural shift from scale-led to intelligence-led growth, where data quality and precision execution increasingly determine competitive advantage.

Together, these forces are accelerating a more granular, high-frequency commerce model – one that favours speed, relevance, and execution excellence over pure marketplace reach. This sets the foundation for the rapid emergence of Q-commerce, particularly in categories where immediacy and availability matter most.

⁵ The Biggest Ecommerce Markets, Netguru, ⁶ India Ecommerce Market Size & Share, Mordor Intelligence

2.3 E-commerce growth in India is uneven across categories, with penetration and scale evolving at different rates

India's e-commerce market has developed asymmetrically across categories. Electronics and smartphones emerged early as the largest online category, achieving scale first and driving initial adoption, but their relative category share is projected to decline as the market matures.

In contrast, categories such as fashion and apparel, which started from a smaller base, are expected to grow faster over time and gain a larger share of overall e-commerce GMV. More recently, food and FMCG categories have gained significant traction, driven by the rapid expansion of Q-commerce models⁷.

These categories, which historically lagged online sales due to immediacy needs and

low tolerance for delivery delays, are now gaining disproportionate momentum as ultra-fast fulfilment models remove the primary barrier to adoption.

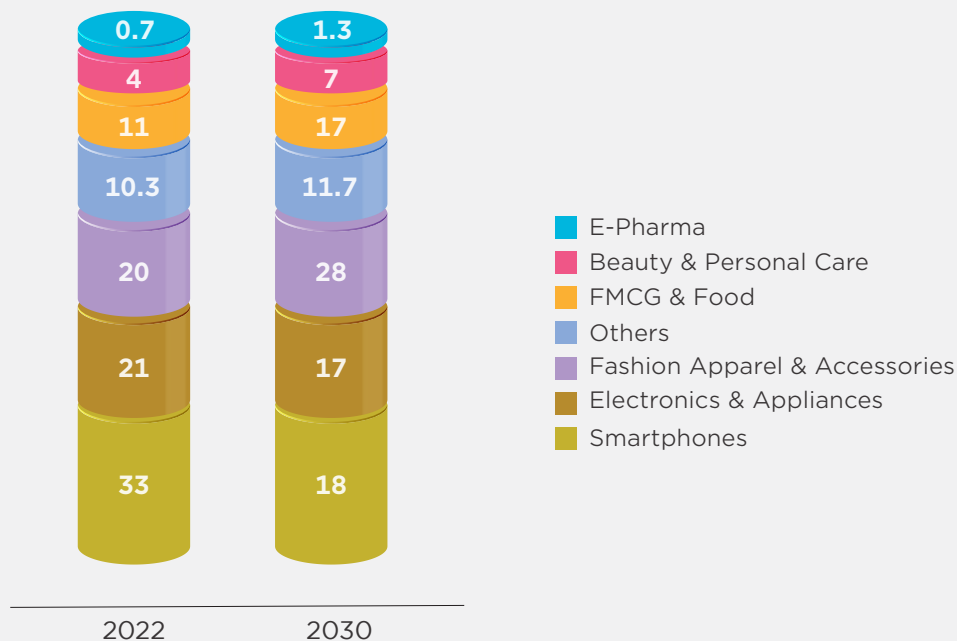
The expansion of dense dark-store networks and micro-fulfilment infrastructure enables delivery within minutes, fundamentally changing the role of e-commerce from planned purchasing to on-demand consumption.

As a result, categories such as grocery and daily essentials are no longer constrained by traditional e-commerce logistics economics. Instead, they are becoming structurally scalable digital categories, contributing meaningfully to incremental GMV growth.

Exhibit 2

Fashion, FMCG, and Food categories will increase their penetration, making almost half of the online purchases by 2030

Share of Product Category in the Indian E-Commerce in percentage



⁷ State of Indian Ecommerce H1 2025, INC 42, Ecommerce Market in India 2024, brickworks

2.4 India’s E-commerce and Q-commerce markets are highly concentrated, with leadership shared among a small number of platforms

India’s e-commerce market is characterized by a high degree of concentration, with Flipkart and Amazon together accounting for a substantial majority of total e-commerce GMV. Flipkart leads the market, while Amazon maintains a strong second position with around one-third of the market, highlighting the dominance of these two platforms in shaping seller participation, assortment breadth, and consumer reach.

A similar concentration pattern is observed in the Q-commerce segment, where leadership is shared between Blinkit and

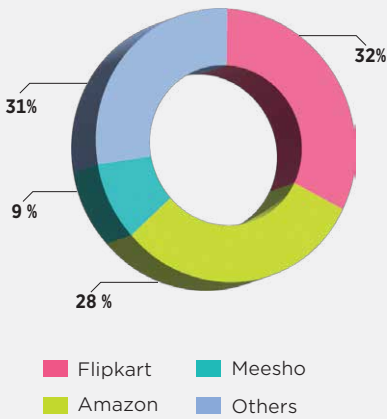
Zepto, with Swiggy Instamart emerging as strong third player⁸. Together, these platforms represent the majority of activity within this fast-growing operating model, underscoring the rapid consolidation of ultra-fast delivery into a small number of scaled operators.

Taken together, the data indicates that while India’s digital commerce market is expanding rapidly, market leadership is consolidating within each operating model, with a limited number of platforms exerting outsized influence across categories, sellers, and consumer experiences.

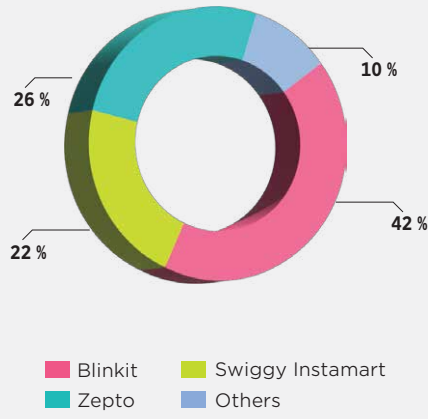
Exhibit 3

Flipkart & Blinkit hold a dominant share in Indian E-Commerce and Q-commerce sector respectively

E-Commerce players share \$77.1 Bn in 2024



Q-Commerce players share \$3.6 Bn in 2024



⁸ Unlocking India's E-commerce Boom, Merchantspring

2.5 Category dynamics in Indian E-commerce are closely tied to platform specialisation, with different marketplaces leading by category

India's e-commerce landscape shows strong category-level concentration, where platform leadership varies significantly depending on the product category. In fashion, the market is highly consolidated around Myntra and Ajio, which together account for around 80% of online fashion GMV⁹. This reflects the importance of category-specific capabilities such as curated assortments, brand partnerships, and content-rich product presentation in driving scale within fashion.

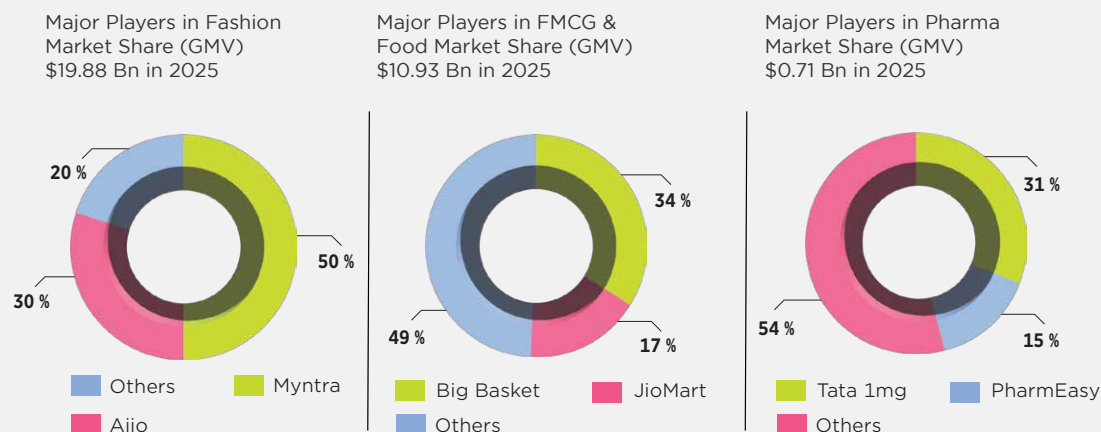
In contrast, leadership in FMCG and Food is centered on platforms built around grocery and essentials. BigBasket¹⁰ and JioMart¹¹ together capture approximately 51% of FMCG and Food GMV, highlighting the role of platforms optimized for high frequency purchases, broad SKU availability, and rapid fulfilment in these categories.

A similar pattern is visible in Pharma and Healthcare, where Tata 1mg and PharmEasy together account for around 46% of category GMV¹². Their leadership reflects the importance of category-specific requirements such as regulatory compliance, product information accuracy, and trust in healthcare-related purchases. Beyond category leaders, India's platform landscape also includes hybrid and value-focused players such as Tata Cliq and Meesho, which operate at the intersection of mass-market reach and category breadth.

Taken together, these patterns indicate that category growth in Indian e-commerce is not platform-agnostic, but instead closely aligned with marketplaces that have developed operating models and capabilities tailored to the specific needs of each category.

Exhibit 4

Category-level leaders have emerged across Fashion, FMCG, Food and Pharma



⁹ Flipkart dominates with 48% market share, mint, ¹⁰ Despite losing market share, why BigBasket is still a good bet, ET

¹¹ E-commerce statistics for India 2025, Bloomagency, ¹² Tata 1mg overtakes Pharmeasy as leaders in India's e-pharmacy market, BS

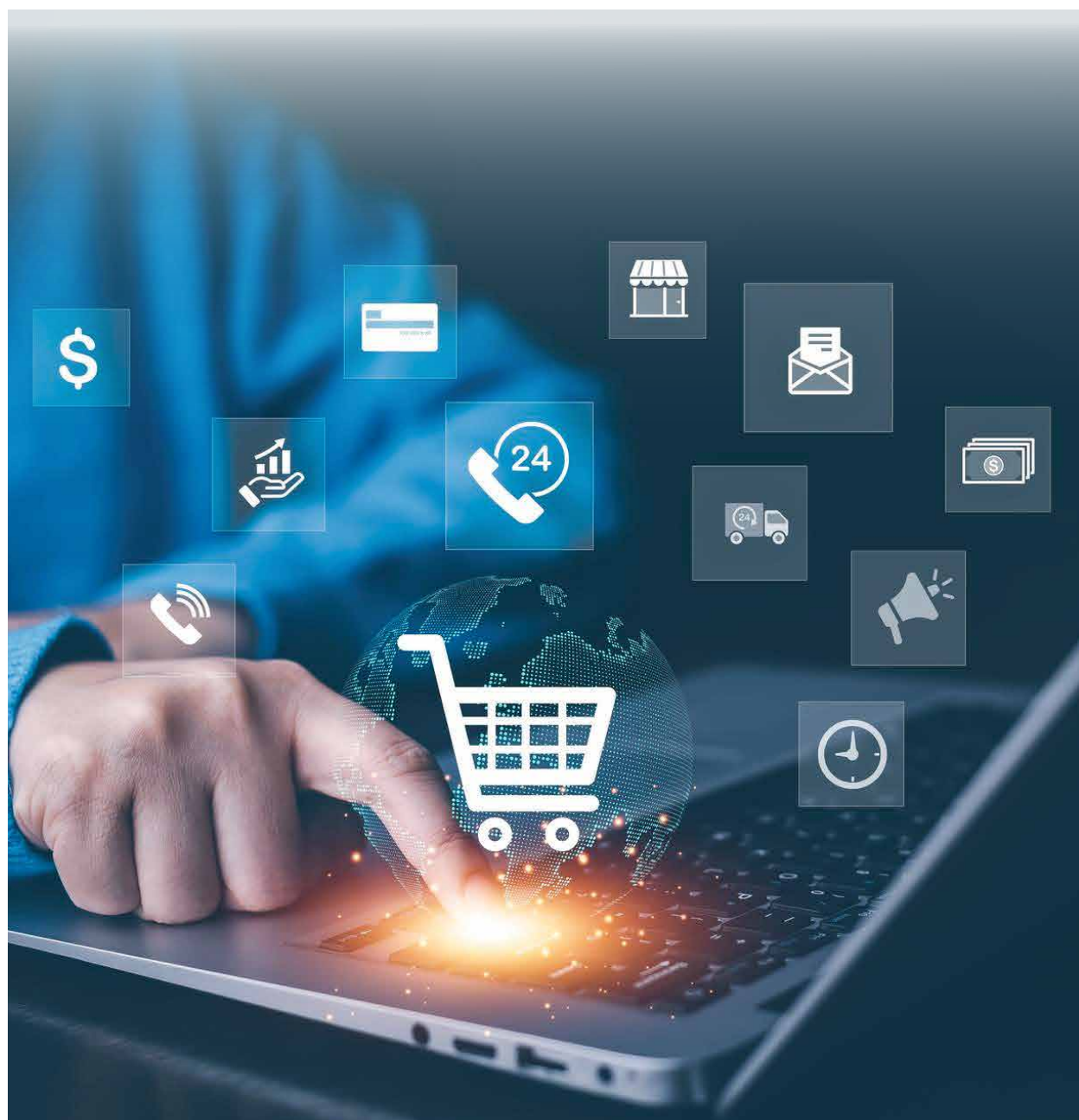
2.6 As E-commerce scales across categories and operating models, product information becomes a critical enabler

The diversity of categories, operating models, and platform specializations highlighted implies increasing complexity in how products are listed, discovered, fulfilled, and returned across India's e-commerce ecosystem.

As the market expands into high frequency categories and faster delivery models, the completeness, accuracy, and conformity of product information play an increasingly

central role in supporting effective marketplace operations and reliable consumer experiences.

In this context, an objective assessment of the current state of product data quality across platforms and categories provides an essential foundation for understanding how well the ecosystem is positioned to scale sustainably.





PRODUCT DATA QUALITY

3.1 Product data quality across Indian marketplaces

As Indian e-commerce scales rapidly across categories and operating models, product data quality has not kept pace with market growth. An assessment of SKUs across leading marketplaces and categories shows that most listings do not yet meet a minimum acceptable data quality standard.

On average, overall data quality score cluster around 90%, with approximately 14% of SKUs scoring below the 80% benchmark. This indicates that while products are broadly listed and visible, their underlying data is often incomplete, inaccurate, or inconsistently structured – creating friction across the consumer journey and downstream operations. These issues are not confined to the point of listing, but are introduced and compounded across the end-to-end product lifecycle, from SKU creation and marketplace syndication through fulfilment, returns, and post-purchase interactions.

The gap is not uniform. Performance varies significantly by marketplace, category, company size, and attribute type, suggesting that data quality challenges are structural rather than isolated. These weaknesses directly affect how products are discovered, understood, fulfilled, and governed – setting the foundation for measurable revenue loss, operational inefficiencies, and compliance exposure explored in the following sections.



3.2 E-commerce data processes are structurally fragmented, embedding data quality failures at every stage

Poor product data quality in Indian e-commerce is not an isolated execution issue – it is structurally embedded in how product information is created, distributed, and operationalized across the ecosystem. The end-to-end product listing process spans multiple disconnected stages, and weaknesses introduced early systematically propagate downstream.

The process starts with SKU creation within brand PIM/MDM systems, where core product identity is defined – product name, brand, category, GTIN, pack size, and regulatory fields. In practice, this data is often created across fragmented internal systems and geographies, with inconsistent standards and limited validation. As a result, SKUs frequently enter the ecosystem with incomplete, inaccurate, or non-standardized reference data, establishing a weak foundation for all subsequent steps.

Data Quality Challenges

Listing & syndication to marketplaces

- Different mandatory attribute requirements across marketplaces
- Data gaps or format issues during syndication
- Inadequate mapping between internal attributes and marketplace-specific schemas

SKU creation

- Manual data entry leads to inconsistencies and errors
- Limited validation at SKU creation stage
- Inconsistent definition of mandatory and optional attributes

Order, returns & post-purchase

- Higher return rates driven by gaps in images, descriptions and specifications
- Elevated customer service workload to resolve data-related complaints
- Downstream financial impact through returns and write-offs

Inventory, packaging & fulfilment

- Mismatch between declared and actual weight/volume leading to courier overcharges
- Poor alignment between product data and warehouse handling requirements



These records are then pushed through listing and syndication to marketplaces, where sellers must map their internal data to multiple marketplace-specific templates, taxonomies, and compliance rules. Because each platform requires different mandatory attributes and formats, the same SKU is repeatedly reworked across channels. Manual uploads, template mismatches, and inconsistent attribute definitions lead to partial listings, missing content, and conflicting versions of the same product across marketplaces.

This fragmentation is further amplified by inconsistent definitions of mandatory and optional attributes across marketplaces. Core consumer-facing fields such as product title, brand, images, size, colour, and material, along with operational attributes including weight, dimensions, and expiry date, and regulatory fields such as FSSAI licence number, net content, food type, and country of origin, are treated differently across platforms. Attributes that are compulsory on one marketplace may be optional or absent on another, forcing sellers to maintain multiple versions of product data for the same SKU. In the absence of a shared attribute standard, sellers often prioritize minimum platform

compliance rather than end-to-end data completeness, structurally embedding gaps in product content. Over time, this produces uneven and inconsistent product detail pages across marketplaces – not because of isolated errors, but because the ecosystem lacks a common language for product attributes.

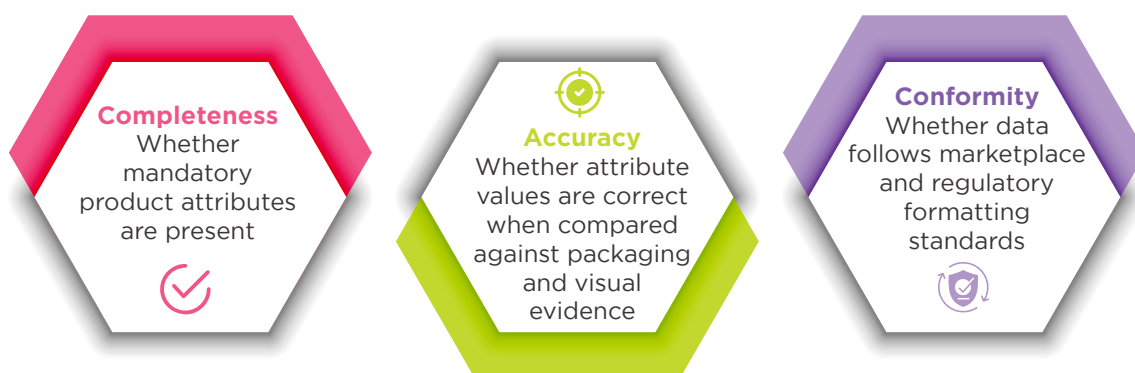
Once listed, product data is operationalized in inventory, packaging, and fulfilment processes. Errors in dimensions, weight, pack configuration, or shelf-life directly affect warehouse slotting, pick-pack accuracy, and courier billing. At this stage, data quality failures move beyond digital visibility and begin driving physical inefficiencies, excess handling costs, and logistics leakage.

Finally, inaccuracies surface most visibly in order execution, returns, and post-purchase processes. Misleading descriptions, incorrect specifications, or missing information increase customer dissatisfaction and return rates, while poor compliance data exposes sellers and platforms to regulatory risk. Because customer service, reverse logistics, and dispute resolution systems depend on listing data, upstream errors translate directly into recurring operational costs.

3.3 Measuring the product data quality

Product data quality was assessed through a structured, SKU-level evaluation designed to reflect how information quality affects real marketplace outcomes. The analysis covers a representative sample of SKUs across major Indian marketplaces and D2C online stores, spanning FMCG, Food, Pharma and Fashion, and including both large brands and SMEs¹³.

Each SKU was evaluated across three core data quality dimensions:



¹³ SMEs are small and medium sized companies/brands, selected on the basis of expert judgment.

To ensure the assessment reflects how data quality impacts real E-commerce operations, product attributes were grouped into three functional attribute types, based on their role in the customer journey and backend processes:



Consumer attributes, which influence product discovery and purchase decisions, including:

Brand name, product title, packaging images, size, colour, material, net content, food type, storage condition



Compliance attributes, which determine regulatory adherence and listing eligibility, including:

Country of origin, FSSAI information, net content, food type



Shipping attributes, which affect fulfilment accuracy, storage, and shipping costs, including:

Expiry date, storage condition

Each attribute was assessed at the individual field level, scored across the three dimensions, and then aggregated to produce dimension-level scores and an overall data quality score per SKU. An 80% threshold was used as a practical benchmark to distinguish acceptable from high risk product data, reflecting common marketplace enforcement practices and operational tolerance levels.

DQ Scoring Methodology

Table 2

DQ Dimension	Definition	Formula	Weight (%)
Completeness	Mandatory product attributes are present	$\frac{\# \text{ complete attributes }}{\# \text{ mandatory attributes }}$	50%
Accuracy	Attribute values are correct when compared against packaging and visual evidence	$\frac{\# \text{ accurate attributes }}{\# \text{ complete attributes }}$	40%
Conformity	Adherence to allowed formats & values	$\frac{\# \text{ conformed attributes }}{\# \text{ complete attributes }}$	10%

90 - 100

Excellent – data is market place ready

80 - 89

Good – small issues exist

70 - 79

Moderate – will cause listing/return issues

< 70

Poor – high commercial risk

This methodology allows product data quality to be evaluated not merely as a technical exercise, but as a business-relevant indicator directly connected to discoverability, conversion performance, returns, fulfilment efficiency, and regulatory risk across both marketplace and D2C channels.

3.4 Significant number of SKUs fail DQT

The assessment reveals that product data quality across Indian marketplaces is materially poor, with a significant number of SKUs failing to meet commonly accepted quality thresholds. When benchmarked against an 80% data quality score, 14% of the analysed SKUs fall below this level, indicating that poor quality product data is not an edge case but a systemic issue across platforms and categories.

At an aggregate level, the average overall data quality score across all marketplaces and categories stands at approximately 90%, masking significant underperformance beneath the surface. While some marketplaces and categories achieve high scores in isolated cases, these are outweighed by widespread gaps – particularly in Food where scores frequently remain in the low 80%.

Breaking the results down by brand size shows that this challenge is not limited to smaller sellers. Large brands achieve a

higher average data quality score (around 89%) compared to SMEs (approximately 87%), yet 11% and 16% of SKUs for Large brand and SMEs respectively fail the 80% (DQT). This suggests that scale and brand maturity alone do not guarantee robust product data governance.

Across marketplaces, performance varies, but no platform consistently delivers high-quality product data across all categories. Fashion, FMCG, and Healthcare platforms tend to score higher overall, while marketplaces focused on Food exhibit structural weaknesses, reflecting the higher complexity of the category.

Taken together, these findings highlight that sub-threshold data quality is the norm rather than the exception. As product assortments and categories such as groceries expand, the persistence of poor data quality creates a growing risk – undermining discoverability, operational efficiency, and consumer trust at scale.

3.5 Completeness and accuracy drive the majority of data quality risk

Table 3

Description	Completeness	Accuracy	Conformity
Average Data Quality Score	89%	87%	99%

The dimension-level results show that completeness and accuracy are the dominant contributors to data quality risk, while conformity performs consistently better across marketplaces and categories.

As shown in the analysis, completeness scores are the lowest-performing dimension (27% of SKUs fail DQT – table 4), materially below both completeness and conformity.

Accuracy scores are moderately higher, generally clustering in the high 80% range (table 3), but still falling short of the desired benchmark for 23% of SKUs (table 4).

This indicates that while many product listings contain a substantial portion of required fields, critical attributes are still missing with sufficient frequency to depress overall data quality scores.

Table 4

Description	Completeness	Accuracy	Conformity
Percentage SKUs scoring less than 80%	27%	23%	1%

In contrast, conformity scores are consistently the strongest dimension, with only 1% of SKUs failing DQT across market places and categories. This suggests that once data is present, it is more likely to follow required formats and structural standards, and that the primary challenge lies earlier in the data lifecycle – ensuring that correct and complete information is captured in the first place.

Overall, the results indicate that sub threshold data quality is driven less by formatting issues and more by missing or incorrect attribute values. Improvements in completeness and accuracy would therefore deliver the greatest uplift in overall data quality and reduce downstream risks linked to product discoverability, returns, fulfilment costs, and compliance exposure.

3.6 Consumer attributes drive the majority of product data quality failures

Consumer-facing attributes account for the largest share of data quality failures (40% of SKUs fail DQ test) across marketplaces and categories, materially underperforming shipping and compliance attributes.

From a completeness perspective, 39% of SKUs score below the 80% threshold on consumer attributes, indicating widespread missing or insufficient information in shopper-critical fields. The most common issues include incomplete image sets, partially populated titles, and missing size or material details.

These gaps directly reduce listing visibility and comparability, particularly in categories with high assortment depth and fast SKU turnover such as FMCG, Fashion, and Grocery-led quick commerce.

Accuracy issues further exacerbate this challenge. 31% of SKUs fail DQ test on consumer attribute accuracy, reflecting discrepancies between listed information and the actual product delivered. Inaccurate images, incorrect size or variant mapping, and misleading product descriptions contribute disproportionately to customer dissatisfaction and returns.

While compliance attributes tend to perform strongly due to regulatory enforcement, and logistics attributes are more operationally standardized, persistent weaknesses in consumer attributes represent the most significant and scalable risk to both growth and profitability across Indian e-commerce platforms.

3.7 Beyond a platform, data quality issues multiply across platforms

Product data problems rarely stay contained to a single channel – once a SKU is distributed across marketplaces, Q-commerce apps, and D2C sites, the issues multiply because each platform enforces different mandatory attributes, taxonomies, formats, and compliance fields, and the same product often has complex variants (packs, sizes, ingredients,

materials, weights, dimensions, regulated disclosures) that must be re-mapped repeatedly. That rework creates fragmentation and inconsistency resulting into duplicate effort, creating missing fields or wrong values that propagate into multiple listings and become harder to govern or correct at scale – especially as assortments and transaction volumes grow.

Product data quality varies across E-commerce and Q-commerce platforms

Comparison of Rice SKU across different Platforms

Table 5

ATTRIBUTES	Marketplace A	Marketplace B
Product Title	Sugarfit Sugar Control Rice, Low GI Telangana Sona Masoori Rice for Diabetic Care	Sugarfit Sugar Control Rice – Sona Masoori, lab-tested low glycemic, whole grain 100% natural, pesticide-free, diabetes-friendly, supports healthy weight & blood sugar – 1 Kg (Pack of 1)
Brand	Sugarfit	Sugarfit
Weight	1 kg	1000 Grams
Net Quantity	1 pack (1 kg)	1000.0 Grams
Diet Type	Veg	Vegetarian
Use by	NA	12 Jun 2026
Country of Origin	India	India
Storage Instruction	Store in a cool, hygienic, and dry place. Avoid exposure to direct sunlight	Keep in dry area, Tightly close container after each use

A comparison of SKUs across platforms presents a clear picture on how one SKU is presented with materially different levels of consumer-relevant information across platforms. For a rice SKU the listing is compact and optimized for speed, with a short product title and limited disclosure across attributes such as use-by date on Marketplace A. In contrast, Marketplace B provides a longer, more descriptive title and more complete nutritional and usage information, improving transparency but increasing content complexity. This highlights how Q-commerce prioritizes brevity and rapid decision-making, while traditional marketplaces emphasize completeness – creating inconsistency in consumer understanding and compliance coverage for identical products.



Product data quality varies across E-commerce and Q-commerce platforms

Comparison of Sweatshirt SKU across different Platforms

Table 6

ATTRIBUTES	Marketplace C	Marketplace D
Product Details	Men Typography Printed Cotton Sweatshirt	Monte Carlo Beige Regular Printed Sweatshirt
Product Details	Beige printed sweatshirt has a round neck, na pockets, long sleeves, closure, straight hem	Look your best by wearing this beige regular sweatshirt from Monte Carlo. Tailored to sartorial perfection from premium quality
Size & Fit	The model (height 6') is wearing a size M	Regular; Model Fit: Model is 6'0"/185 cms and is wearing size M
Pattern	NA	Printed
Color	NA	Beige
Material & Care	Cotton, Machine Wash	Cotton Blend, Machine wash
Sleeve Length	Long Sleeves	Long
Neck	Round Neck	Round Neck
Type	Pullover	NA
Number of Pocket	NA	NA
Hemline	Straight	NA



The sweatshirt example shows significant divergence in how fashion attributes are structured and populated across platforms. Marketplace C emphasizes rich, granular consumer attributes – fit, sleeve length, hemline, and styling details – supporting discovery and comparison, but with several fields left unstructured or marked as “NA”. Marketplace D, by contrast, provides a more narrative-driven description with fewer standardized attribute fields populated. These differences illustrate how inconsistent attribute definitions and completion standards across fashion marketplaces lead to uneven data quality, directly affecting comparability, search relevance, and downstream analytics for the same SKU.



ONLINE SHOPPING

BUSINESS AND ECONOMIC IMPACT

4.1 Poor product data quality translates directly into material financial losses across the E-commerce value chain

Despite the continued expansion of India's E-commerce and Q-commerce markets, poor product data quality continues to generate financial losses. Based on population-level analysis across marketplaces, D2C channels, and Q-commerce segments, poor product data quality is estimated to drive total revenue losses of ₹5,000 crore annually. These losses originate from gaps in how products are represented, discovered, and transacted – beginning at the point of product discovery and extending through fulfilment and post-purchase stages.

The financial impact is most acutely felt at the profitability layer. Of the total revenue loss, ₹2,000 crore translates directly into gross margin erosion, driven by lower conversion efficiency, suppressed listings, and slower sell-through resulting from incomplete or inaccurate product information. In parallel, product returns attributable to poor data quality generate an additional ₹1,900 crore in direct return costs, reflecting higher reverse logistics, handling, and processing expenses. Together, gross margin losses and return costs account for the majority of the quantified financial impact associated with poor product data quality.

Importantly, these losses are not isolated or episodic. They are embedded within everyday operating metrics – conversion rates, return ratios, inventory turnover, and fulfilment efficiency – making them difficult to isolate at the SKU or transaction level. As assortment depth and transaction volumes continue to scale across categories and channels, the economic consequences of poor product data quality scale proportionally, constraining both marketplace efficiency and seller profitability.

4.2 Financial impact was estimated by linking data quality failures to business-critical operational triggers

The financial impact of poor product data quality was estimated using a trigger-based methodology that links measurable data quality failures to specific operational and financial consequences across the e-commerce value chain. Rather than treating data quality as a standalone score, the approach identifies where and how deficiencies in product data translate into lost revenue, margin erosion, and incremental costs.

Attribute-level triggers were applied based on the type of data affected – consumer, shipping, or compliance attributes – recognising that different data failures lead to different business outcomes. For example, gaps in consumer-facing attributes primarily affect conversion and returns, shipping-related inaccuracies drive fulfilment inefficiencies and courier overcharges, and compliance gaps expose sellers to regulatory and write-off risk. Triggers were applied only to the portion of GMV demonstrably attributable to product data quality, ensuring impacts were not overstated.

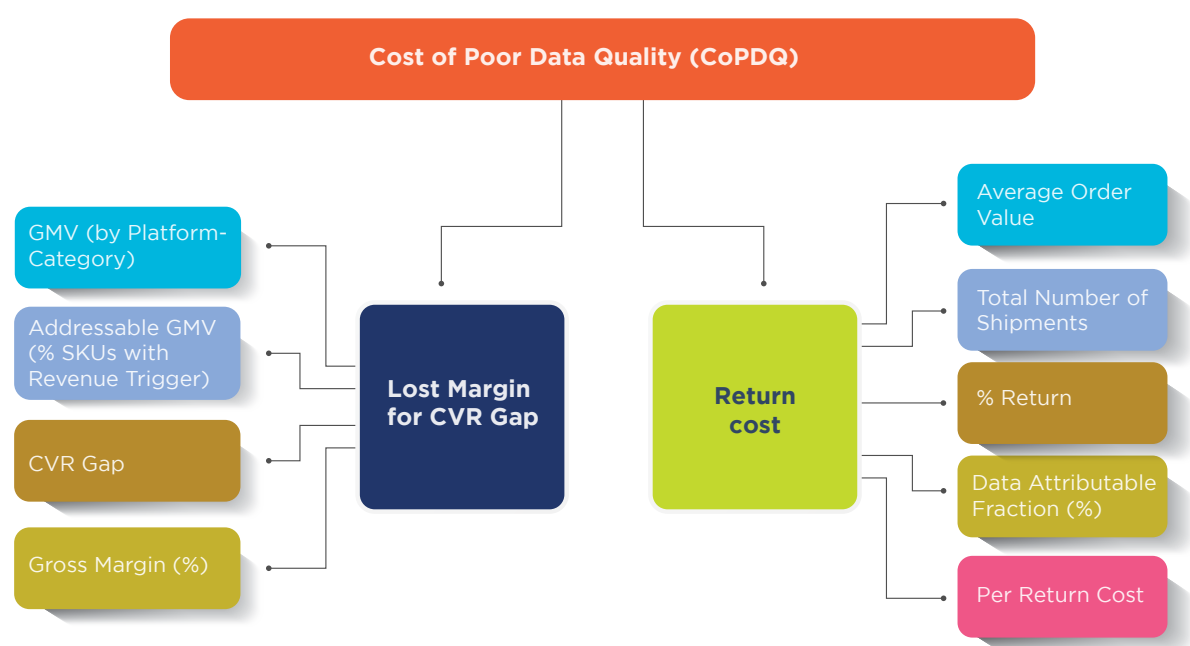
Once triggered, data quality failures were mapped to a set of business impact pathways covering the full product lifecycle. These include lost margin from conversion rate gaps and suppressed listings, higher return costs driven by misleading or inaccurate product information, excess inventory holding costs from slower sell-through, courier overcharges caused by incorrect weight or dimension data, write-offs resulting from compliance failures, and incremental customer service and manual rework costs. While all impact pathways were assessed, the analysis aggregates results into three primary financial dimensions to reflect how losses are realised at scale: total revenue loss, gross margin impact, and return costs.

At a high level, total revenue loss was estimated by linking conversion uplift potential to the share of GMV affected by data quality failures, gross margin impact was derived by applying category-specific margin structures to the lost revenue, and return costs were calculated based on shipment volumes, return rates, and per-return cost assumptions, adjusted for the portion attributable to product data. This structured approach ensures that financial impacts are grounded in observable operational behaviour while remaining comparable across channels and categories.

Exhibit 5

Revenue loss from CVR gap and returns are the major drivers of cost of poor data quality.

Methodology to estimate cost of poor data quality



Gross margin loss and returns represent the largest share of financial impact



₹5,000Cr

Total Revenue Loss

Annual value of unrealised sales caused by poor product data quality



₹2,000Cr

Gross Margin Erosion



₹1,900Cr

Direct Return Cost

Across Indian E-commerce and Q-commerce, the financial consequences of poor product data quality are concentrated in two dominant impact pathways: gross margin erosion and return-related costs. Together, these account for the majority of the total value leakage observed across the ecosystem.

Poor product data quality results in an estimated INR 5,000 crore in annual revenue loss, driven primarily by lower conversion and reduced effective demand across affected SKUs. When category-specific margin structures are applied, this translates into a gross margin impact of INR 2,000 crore. This margin loss reflects sales that could have been captured under normal operating conditions but are forfeited due to incomplete, inaccurate, or non-conformant product information at the point of purchase.

Returns also represent a substantial financial burden. The analysis estimates INR 1,900 crore in annual return-related costs attributable to poor product data quality, including reverse logistics, handling, and processing expenses. These costs are triggered when consumers receive products that do not match expectations set by the product detail page most commonly due to inaccurate descriptions, missing specifications, misleading images, or incorrect pack and size information. Importantly, return costs scale directly with shipment volumes, making them particularly material in high-frequency categories such as FMCG, food, and fashion.

Taken together, gross margin loss and return costs linked to poor data quality in Indian e-commerce value chain account for approximately 1.1% of the total E-commerce GMV. Other cost components such as excess inventory holding, courier over charges, write-offs, and customer service overhead remain important but secondary in magnitude. This concentration of impact highlights that data quality failures do not merely create operational friction; they directly undermine the two core economic levers of digital commerce: profitable demand capture and post-purchase efficiency.

4.3 Financial exposure varies significantly by marketplace, category, and brand size

The financial impact of poor product data quality is not evenly distributed across the Indian e-commerce ecosystem. Instead, exposure varies materially depending on channel dynamics Marketplace & D2C versus Q-Commerce and on category specific operating models, such as purchase frequency, return propensity, margin structure, and regulatory intensity. As a result, the same data quality failure can translate into very different financial outcomes depending on where it occurs in the value chain.

4.4 Marketplaces & D2C: High margin erosion driven by suppressed conversion and returns

Marketplace and D2C channels together account for the majority of the total financial impact observed. From a total estimated INR 5,000 crore in revenue loss, Marketplace & D2C contribute INR 4,200 crore, or roughly 84% of total lost revenue. When margin structures are applied, this results in a gross margin impact of INR 1,700 crore.

Returns further amplify this exposure. Marketplace & D2C account for INR 1,800 crore in return-related costs. This concentration reflects both higher shipment volumes and greater sensitivity to PDP inaccuracies in categories where consumers rely heavily on detailed product information prior to purchase.

4.5 FMCG & Food: Largest absolute financial exposure driven by scale

Within Marketplace & D2C, FMCG & Food represents the largest source of financial exposure from poor product data quality. The category accounts for ₹3,000 crore in revenue loss, translating into a gross margin impact of ₹1,000 crore. While return costs remain relatively limited at ₹49 crore, the sheer scale of transactions and SKU volumes amplifies the financial impact of even small data quality gaps. In FMCG & Food, issues related to missing product information primarily affect conversion and listing performance rather than returns, making revenue leakage and margin erosion the dominant loss vectors.

4.6 Fashion Apparel & Accessories: Returns-driven financial impact

Fashion Apparel & Accessories exhibits a distinct loss profile, where returns dominate the financial impact of poor data quality. The category records ₹800 crore in revenue loss, resulting in a gross margin impact of ₹300 crore. However, return-related costs reach ₹1,700 crore, making Fashion the largest contributor to return costs across all categories. This reflects high return propensity driven by gaps in consumer-facing attributes such as size, fit, material, colour, and images. In Fashion, data quality failures disproportionately materialise post-purchase, significantly increasing operational costs even when revenue loss is comparatively lower than in FMCG.

4.7 Beauty & Personal Care: Margin-sensitive exposure with moderate scale

Beauty & Personal Care shows high margin sensitivity to data quality failures despite a smaller overall scale than FMCG. The category experiences ₹470 crore in revenue loss, which converts into a gross margin impact of ₹260 crore. Return costs remain modest at ₹76 crore, indicating lower return intensity compared to Fashion. As a result, financial exposure in Beauty & Personal Care is driven primarily by lost conversion rather than post-purchase inefficiencies, making margin leakage the critical concern.

4.8 E-pharma: Lower absolute impact, high regulatory sensitivity

E-pharma contributes the smallest absolute financial loss among Marketplace & D2C categories, with ₹30 crore in revenue loss and a gross margin impact of ₹19 crore. Return costs stand at ₹12 crore. While the financial magnitude is limited, the category remains structurally sensitive to data quality failures due to strict compliance and regulatory requirements. Even isolated inaccuracies in mandatory attributes can lead to regulatory action, creating high-risk exposure disproportionate to the absolute financial size.



GLOBAL STANDARDS

As India's E-commerce ecosystem continues to expand across marketplaces, direct-to-consumer channels, and Q-commerce, the ability to sustain growth increasingly depends on the quality and reliability of product data. The analysis in the previous sections shows that deficiencies in product data quality are not isolated operational issues; they are systemic, recurring, and financially material. Addressing these challenges requires more than incremental fixes or platform specific controls. It requires global standards for how products are identified, described, and governed across the ecosystem.

Product data quality failures observed across channels and categories stem from fragmented data practices, inconsistent attribute definitions, and limited governance over how product information is created, maintained, and exchanged. Without common standards, sellers duplicate effort, marketplaces absorb operational inefficiencies, and regulators face challenges in enforcing compliance at scale. As transaction volumes and assortment complexity increases, these weaknesses compound, amplifying financial losses and undermining consumer trust.

Standardised product identification and data governance provide a structural solution to these challenges. By establishing a common language for product identity and attributes, standards enable consistency, interoperability, and accountability across the value chain. In doing so, they directly address the root causes of poor data quality highlighted in this report creating the conditions for improved discoverability, more accurate product representation, lower returns, and more efficient fulfilment. In a rapidly scaling digital commerce environment, such standards are not optional enablers but foundational infrastructure for sustainable growth.



5.1 GS1 standards directly address the root causes of data quality failures

The data quality failures identified across Indian e-commerce platforms are not random or isolated; they originate from structural gaps in how products are identified, described, and governed across the ecosystem. Fragmented identification systems, inconsistent attribute definitions, and limited controls over data creation and maintenance lead to the incompleteness, inaccuracies, and non-conformity observed across categories and channels.

GS1 standards are specifically designed to address these root causes at their source. At the foundation, unique and persistent product identification ensures that each product variant is referenced consistently across marketplaces, sellers, logistics providers, and regulatory systems. By eliminating ambiguity at the SKU level, standardised identifiers reduce duplication, mismatches, and errors that commonly arise when products are listed using non-standard or platform-specific identifiers.

This directly improves data accuracy and enables reliable linking of product information across systems. Beyond identification, GS1 provides a structured framework for product attributes, defining what information should be captured, how it should be formatted, and how it should be interpreted across trading partners. This addresses one of the most common drivers

of poor data quality highlighted in the analysis: inconsistent or incomplete attribute coverage. Standardised attribute definitions improve completeness by design and reduce discrepancies in how the same product information is represented across channels, supporting more consistent product detail pages (PDPs) and more reliable downstream operations.

Equally important, GS1 standards embed data governance principles that extend beyond initial product listing. By establishing clear ownership, validation rules, and lifecycle management for product data, standards help prevent degradation over time such as outdated dimensions, incorrect pack information, or missing compliance details. This governance layer is critical in mitigating the operational triggers identified earlier, including suppressed listings, elevated returns, fulfilment inefficiencies, and compliance exposure.

In combination, standardised identification, harmonised attribute models, and shared governance mechanisms address the underlying causes of data quality failures rather than their symptoms. This positions GS1 standards as a foundational enabler for improving product data quality at scale, directly aligning with the areas of financial leakage and operational risk identified across the e-commerce value chain.

5.2 GS1 enables better outcomes for marketplaces, sellers, and regulators

Improving product data quality at scale delivers tangible benefits across the e-commerce ecosystem, but the nature of those benefits differs by stakeholder. By providing a shared foundation for product identification, attributes, and data governance, GS1 standards enable structurally better outcomes for marketplaces, sellers, and regulators alike aligned with the financial and operational challenges identified in this analysis.

For marketplaces, standardised product data supports stronger commercial performance and operational efficiency.

Consistent and complete product information improves discoverability and conversion outcomes, directly addressing the margin erosion observed from suppressed or underperforming listings. At the same time, improved data accuracy reduces post-purchase friction, helping lower return volumes and associated handling and customer service costs.

As marketplaces scale assortment breadth and transaction volumes, standardisation reduces the operational burden of onboarding, validation, and ongoing data

correction supporting growth without a proportional increase in complexity or cost. For sellers, GS1 standards reduce the friction and inefficiency associated with managing product data across multiple channels.

A single, standardised product identity and attribute set enables faster and more accurate product listings, fewer manual interventions, and lower risk of inconsistencies that can trigger returns, relisting, or write-offs. By improving the quality and reliability of product information at source, sellers can protect margins, reduce avoidable costs, and improve conversion performance without relying solely on promotional or marketing levers.

For regulators, standardised product data enhances transparency, traceability, and

compliance oversight, particularly in regulated categories such as food, healthcare, and personal care. Clear and consistent product identifiers, combined with harmonised compliance attributes, make it easier to monitor adherence to regulatory requirements and reduce information asymmetry across the market. This not only supports enforcement but also contributes to safer, more trustworthy digital commerce environments for consumers.

Across all stakeholders, the common thread is that GS1 standards transform product data from a fragmented operational input into shared infrastructure. By aligning incentives across marketplaces, sellers, and regulators, standardisation helps address the systemic data quality failures identified in this report unlocking both commercial value and ecosystem-wide efficiency.

5.3 GS1 provides a scalable foundation for next-generation E-commerce

As Indian E-commerce evolves toward higher transaction density, faster fulfilment cycles, and deeper integration across online and offline channels, the demands placed on product data will continue to intensify. Emerging operating models such as Q-commerce, omnichannel retail, and cross-border digital trade – require product information that is not only accurate and complete, but also interoperable, machine-readable, and consistently governed across systems. In this context, fragmented or platform specific data practices are unlikely to scale sustainably.

By establishing globally aligned standards for product identification and data exchange, GS1 provides a foundation that is inherently suited to next-generation commerce. Standardised product identifiers enable seamless data sharing across marketplaces, D2C platforms, logistics providers, and regulatory systems, while structured attribute frameworks support automation, analytics, and AI-driven decision-making. This becomes increasingly critical as commerce shifts toward real-time inventory visibility, personalised product discovery, and hyper-local fulfilment models.

Looking ahead, GS1 standards also support the transition to next-generation data carriers, including 2D barcodes and digital product passports, which can unlock richer, more dynamic product information at the point of scan. These capabilities are particularly relevant in regulated and high velocity categories, where traceability, transparency, and accuracy are essential. By aligning today's product data practices with globally recognised standards, the ecosystem is better positioned to absorb future innovation without reintroducing fragmentation or data quality risk.

In an environment where growth is increasingly constrained by operational efficiency, consumer trust, and regulatory scrutiny, standardisation is not a constraint on innovation but an enabler of it. GS1's role as neutral, cross-industry infrastructure positions it as a critical enabler of scalable, resilient, and trustworthy e-commerce supporting sustainable growth as the sector enters its next phase.



WAY FORWARD

Strengthening Product Data Quality Across the E-commerce Ecosystem

Improving product data quality at scale requires coordinated action across brands, marketplaces, and regulators. The findings of this report show that data quality failures are structural, recurring, and economically material yet largely preventable. Addressing them demands not only better execution at individual touchpoints, but also stronger standards, clearer accountability, and shared infrastructure across the ecosystem.





6.1 For Brands: Treat product data as a strategic asset, not a listing task

- **Establish a single source of truth for product data** across consumer, shipping, and compliance attributes (e.g., PIM/MDM), before syndication to marketplaces, Q-commerce, and D2C channels.
- **Embed data quality governance upstream** by defining ownership, validation rules, and approval workflows for critical attributes (titles, images, net quantity, dimensions, regulatory disclosures).
- **Adopt standardized product identification and attribute structures** to reduce rework, inconsistencies, and platform-specific reformatting.
- **Automate data quality checks** (completeness, accuracy, format compliance) prior to listing and during assortment updates.
- **Link data quality to commercial outcomes** by tracking its impact on conversion, returns, margin leakage, and brand trust.
- **Integrate product data quality into go-to-market and brand protection processes**, rather than treating it as an operational afterthought.



6.2 For Marketplaces: Shift from reactive enforcement to proactive enablement

- **Move upstream in the seller journey** by embedding data quality requirements at onboarding, not only at listing or post-launch stages.
- **Standardize and clearly define mandatory vs. optional attributes** by category, reducing ambiguity and inconsistent seller interpretation.
- **Provide structured data templates and guided listing flows** that help sellers populate attributes correctly the first time.
- **Introduce early stage validation and real-time feedback** to flag completeness, accuracy, and compliance issues before listings go live.
- **Use data quality scores as a performance lever**, linking them to listing visibility, discoverability, and conversion outcomes.
- **Create closed feedback loops** by showing sellers how data quality improvements translate into higher conversion and fewer returns.
- **Align data quality standards across marketplace and Q-commerce formats** to reduce fragmentation for multi-channel sellers.



6.3 For Regulators: Strengthen enforcement through standardization and interoperability

- **Align regulatory disclosure requirements with standardized product data structures** to reduce ambiguity and inconsistent interpretation across platforms.
- **Mandate machine-readable formats for key compliance attributes** (e.g., FSSAI, net quantity, expiry, country of origin) to enable automated validation and monitoring.
- **Encourage adoption of common product identifiers and data standards** to improve traceability, transparency, and cross-platform consistency.
- **Shift enforcement from post-facto penalties to preventive compliance**, enabling earlier detection of data gaps before products go live.
- **Promote interoperability between regulators, marketplaces, and brands**, allowing compliance checks to scale with transaction volumes.
- **Prioritize high-frequency and high-risk categories** (FMCG, Healthcare) where even isolated data errors carry disproportionate consumer and financial risk.

CONCLUSION

India's e-commerce market is entering a phase of rapid scale, deeper category penetration, and increasing operational complexity. As this report demonstrates, however, product data quality has not evolved at the same pace. Gaps in completeness, accuracy, and conformity persist across platforms, categories, and operating models, embedding friction into the consumer journey and creating material inefficiencies across fulfilment, returns, and compliance processes.

By assessing product data quality across leading marketplaces, Q-commerce platforms, and D2C channels, and linking these gaps to measurable financial outcomes, this study quantifies the economic cost of poor product data quality across the e-commerce value chain. The analysis shows that data quality failures translate directly into lost revenue opportunities, gross margin erosion, and elevated return and operational costs effects that vary significantly by channel and category, but are consistently material at scale.

Improving product data quality is therefore not a cosmetic or compliance-only exercise. It is a foundational enabler of profitable growth, operational efficiency, and consumer trust in India's digital commerce ecosystem. As e-commerce continues to expand into high-frequency, regulated, and fulfilment-intensive categories, coordinated action across brands, marketplaces, and regulators supported by common standards and shared infrastructure will be essential to unlock the next phase of sustainable growth.

METHODOLOGY



Sampling Methodology

This study is based on a structured sample of 510 SKUs, designed to reflect the diversity and operating realities of India's E-commerce and Q-commerce ecosystem. The sample spans eight major platforms across e-commerce, Q-commerce, and e-pharmacy, including large horizontal marketplaces, category specialists, and rapid delivery platforms, as well as a selection of online D2C brands to capture seller-owned operating models. This approach ensures coverage across both marketplace-led and brand-led digital commerce environments.

SKUs were selected across five high-impact categories – Beauty & Personal Care, FMCG & Food, Healthcare, Fashion Apparel & Accessories, and E-pharma representing a significant share of online retail volume and operational complexity. The sampling strategy was intentionally balanced to capture differences in category requirements, fulfilment models, and regulatory exposure, allowing findings to be generalized at the platform, category, and channel level rather than reflecting isolated seller or SKU-specific issues.



Data Quality Score

Product data quality was assessed using a multi-dimensional scoring framework designed to reflect how data is evaluated and consumed across Indian e-commerce platforms. Attributes were selected based on two principles:

- (1) Mandatory and commonly required attributes across major marketplaces, and
- (2) Compliance critical attributes mandated by applicable regulations, particularly in food, and pharmaceutical categories.

Each SKU was evaluated across three core data quality parameters – completeness, accuracy, and conformity to capture not only whether information was present, but whether it was correct and aligned with platform and regulatory expectations. The resulting data quality scores were calculated at the SKU level and then aggregated across the 510 SKU sample to generate platform, category, and channel-level insights. This structure allows the analysis to isolate where data quality breaks down most materially across the e-commerce value chain.



Business and Economic Impact

The financial impact of poor product data quality was estimated using a top-down analytical framework, anchored in total market activity and systematically linked to SKU-level data quality outcomes. Starting from addressable GMV by channel and category, the analysis applies data quality triggers to identify the portion of economic activity exposed to data-related failures. These triggers reflect both global thresholds (e.g., overall data quality score below acceptable levels) and attribute-level thresholds, aligned with how platforms experience conversion and return degradation, suppress listings, or penalize non-compliance.

To translate data quality failures into business impact, the framework estimates losses across multiple dimensions, including lost gross margin from conversion gaps, returns driven by inaccurate or misleading product information, and suppressed listings¹⁴. The model is informed by a combination of secondary research and primary insights from more than 20 interviews with marketplaces, brands, logistics providers, and domain experts. Together, these inputs ensure that the resulting cost estimates reflect both observed data patterns and real-world operating behaviours across India's digital commerce ecosystem.

¹⁴ DQ model, Kanvic analysis and research, A+ Content, Amazon, Reverse Logistics, Horizon Industrial Parks, Order Value, Shiprocket Ecommerce Returns, Shopify

DISCLAIMER

This report has been prepared jointly by GS1 India and Kanvic Consulting for informational and analytical purposes only. The findings, insights, and estimates presented are based on a combination of primary research, secondary sources, expert interviews, and analytical modelling conducted as part of this study.

All financial impact figures are indicative estimates derived using defined assumptions and methodologies and are intended to illustrate the potential economic implications of product data quality issues at an aggregate level. Actual outcomes may vary materially depending on category, platform, operating model, and implementation context. This report does not constitute legal, regulatory, financial, or commercial advice, nor should it be relied upon as such.

The analysis reflects conditions and data available at the time of research and may not capture subsequent market developments or platform specific changes. References to marketplaces, brands, or platforms are for illustrative purposes only and do not imply endorsement, evaluation, or ranking.

While reasonable care has been taken to ensure accuracy, GS1 India and Kanvic Consulting make no representations or warranties, express or implied, regarding the completeness or accuracy of the information contained herein and accept no liability for decisions taken based on this report.

ABBREVIATIONS

AI	Artificial Intelligence
BNPL	Buy Now, Pay Later
CAGR	Compound Annual Growth Rate
CVR	Conversion Rate
DQ	Data Quality
DQT	Data Quality Test
D2C	Direct-to-Consumer
E-Commerce	Electronic Commerce
E-Pharma	Electronic Pharmacy
FMCG	Fast-Moving Consumer Goods
FSSAI	Food Safety and Standards Authority of India
FY	Financial Year
GMV	Gross Merchandise Value
GTIN	Global Trade Item Number
INR	Indian Rupee
MDM	Master Data Management
PDP	Product Detail Page
PIM	Product Information Management
PIN	Postal Index Number
SKU	Stock Keeping Unit
SME	Small and Medium Enterprises
UPI	Unified Payments Interface
USD	United States Dollar
2D	Two-Dimensional



BG00, Ground Floor, Tower B, World Trade Center,
Nauroji Nagar, New Delhi 110029

T: +91 (11) 42 890 890

E: info@gs1india.org

W: www.gs1india.org