Empowering Artisan Visibility:

Analyzing NOVICA's Amazon Reviews to Enhance Visibility & Growth

By: Rishab Harish Kumar

Introduction

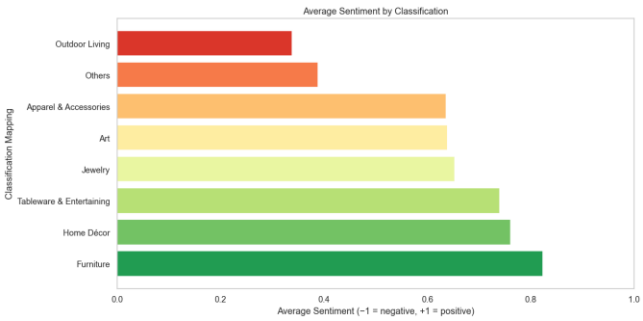
Novica is an e-commerce company that has supported thousands of artisans from developing countries across the world since 1999. They sell handcrafted products across the three main categories, “Jewelry”, “Home Decor”, and “Clothing".

The approach of working with this dataset paved the way for creating the three main objectives of this project: analyzing 50,000+ reviews to uncover patterns in customer satisfaction and dissatisfaction, detecting underperforming SKUs by uncovering negative sentiment, and translating findings into actionable business recommendations to improvise customer experience

This project is crafted by a three-step process. Firstly, scrap the data from Amazon reviews, product detail pages, and SKU-level sales metrics and compiled the data into a unified dataset. Secondly, in the cleaning process the duplicates and non-English reviews were removed. Finally, the modelling process is done by mapping the consumers' perception of the SKU performance. The objective of this project is to understand what consumer pain points are by analyzing consumer sentiment in Amazon reviews.

Results

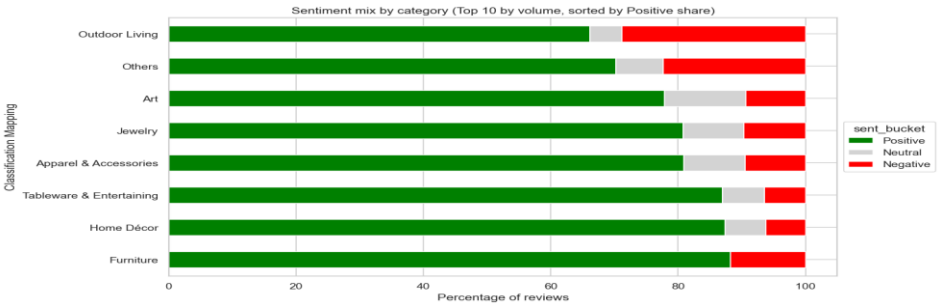
Analysis 1: Category Performance Overview



Analysis of the 17,265 customer reviews revealed significant heterogeneity in satisfaction across product categories. Sentiment scores ranged from 0.45 (“Outdoor Living”) to 0.87 (“Furniture”), demonstrating a performance gap of 0.42 points. This variation suggests that customer satisfaction of NOVICA is category-dependent rather than uniformly distributed. This is consistent with expectancy-disconfirmation theory wherein different product types elicit varying baseline expectations.

Methodology: Aspect-Based Sentiment Analysis (ABSA): 17,000 customer reviews were collected from NOVICA’s products on Amazon. The scrape data was cleaned text that preserved contextual punctuation. Using BERTopic with the all-MiniLM-L6-v2 transformer, latent topics were extracted and automatically labeled by combining product categories with top keywords. The sentiment was assessed nlptown's BERT model fine-tuned product reviews, which assigned each review a continuous score from -1 (negative) to +1 (positive).

Analysis 2: Sentiment Distribution Analysis



Sentiment composition analysis distinguished between consistent and polarized categories. High-performing categories—" Furniture" and "Home Decor"—exhibited 85-88% positive sentiment with minimal negative feedback, while lower-performing categories such as "Outdoor Living" showed bimodal distributions with 37% negative reviews.

Methodology: Deviation-Based Performance Measurement: Traditional absolute sentiment comparisons fail to account for category-specific expectations. A baseline adjustment metric was employed where each aspect's performance is evaluated relative to its category average, following comparative judgment theory. This approach controls varying expectation levels across product types. This analysis gives the entire picture of the categories that it may seem. For example, "Outdoor Living" is doing fine at 0.4 sentiment, but it has more than 25% negative reviews which signals that 1 in 4 customers are not happy with that category.

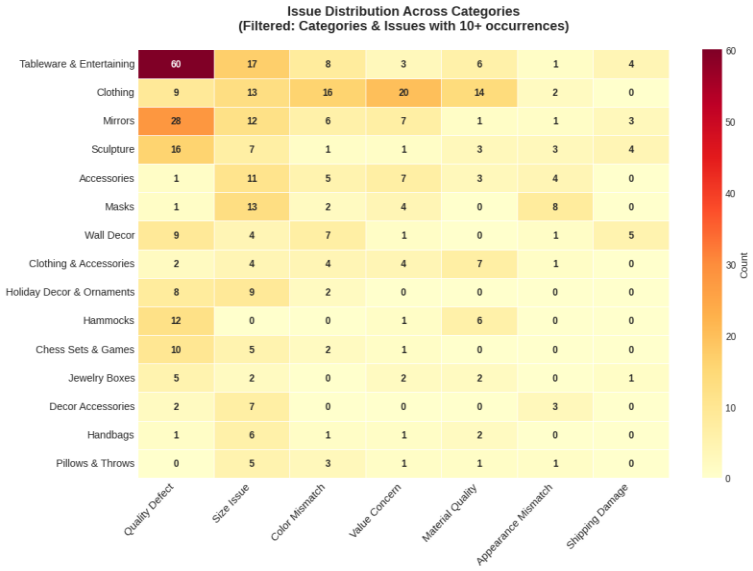
Analysis 3: Aspect-Level Performance Drivers

Systematic trends were found across categories using deviation-based aspect analysis. Shipping & Delivery consistently underperformed (-0.04 to -0.13), indicating operational bottlenecks regardless of product type, while Product Experience consistently outperformed category baselines (+0.07 to +0.14). Size & Dimensions revealed category-specific problems for categories such as "Apparel" (-0.04) and "Jewelry" (-0.05), indicating potential for focused intervention in line with attribute-level satisfaction models.

Methodology: Hybrid Topic Modeling integrates supervised business logic for organizational alignment with unsupervised machine learning for pattern recognition. While rule-based consolidation guarantees business

interpretability and transformer-based clustering finds semantic patterns, BERTopic strikes a balance between algorithmic impartiality and usefulness.

Analysis 4: Cross Category Issue Distribution



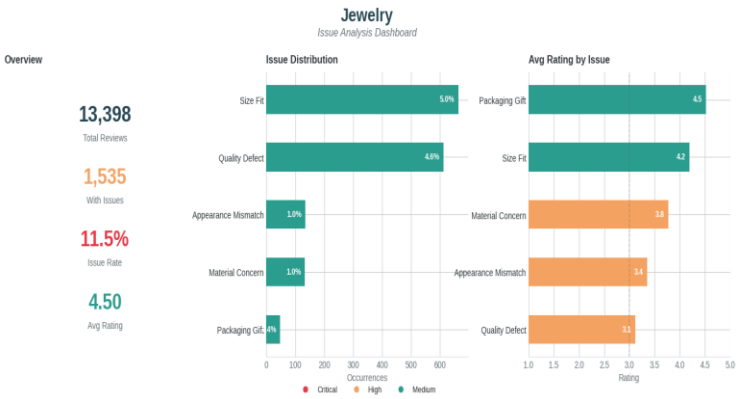
The heatmap visualization of issue distribution across product categories is filtered to display only categories and issue types with 10 or more occurrences to ensure statistical relevance. The color intensity scale ranges from light yellow—low frequency—to dark red—high frequency—with values ranging from 0 to 60 occurrences.

Methodology: The heatmap was generated by aggregating BERT-classified issue mentions across all reviews, then cross-tabulating with the product category. Seven distinct issue types were identified through the classification model: Quality Defect, Size Issue, Color Mismatch, Value Concern, Material Quality, Appearance Mismatch, and Shipping Damage. Categories with less than 10 total issue occurrences were excluded to minimize noise from low-sample categories.

Notably, “Jewelry” was deliberately excluded from this heatmap despite being NOVICA's largest selling category. “Jewelry” accounts for approximately 80% of all reviews in the dataset, and with that volume, it naturally accumulates far more negative remarks in absolute terms. Including “Jewelry” would have skewed the visualization with its values dominating the color scale and drowning out meaningful patterns in other categories. This

methodological decision enables clearer identification of issue concentrations across the remaining product categories.

Analysis 5: In-depth Category Analysis



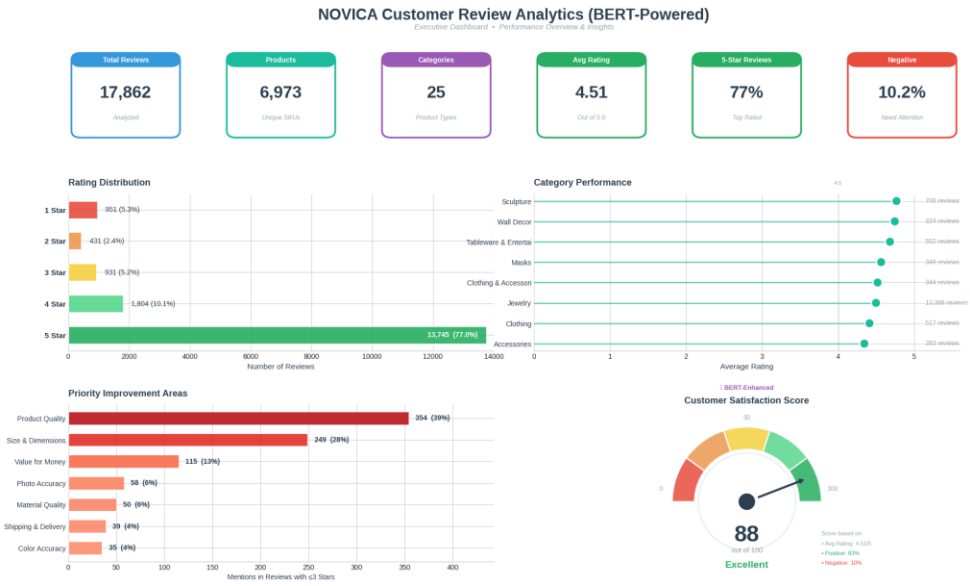
This visualization presents an issue analysis dashboard dedicated to “Jewelry” and is structured in three components. The left panel displays an *Overview* with aggregate metrics including total reviews, issue count, issue rate percentage, and average rating. The center panel shows *Issue Distribution* as a horizontal bar chart depicting the frequency of each issue type as a percentage of total reviews, with occurrence counts ranging from 0 to over 600. The right panel presents *Average Rating by Issue,* illustrating the mean customer rating associated with each issue type on a 1.0 to 5.0 scale. Issues across both bar charts are color coded by severity classification: Critical (red), High (orange), and Medium (teal), enabling immediate visual identification of problem areas that most significantly impact customer satisfaction.

Methodology: “Jewelry” warranted isolated analysis due to its saturation in the dataset. With 13,398 reviews, it constitutes approximately 75% of the total review corpus and represents NOVICA's core business segment.

Issue classification was performed using the fine-tuned BERT model, which identified five primary issue categories within “Jewelry” reviews: Size Fit, Quality Defect, Appearance Mismatch, Material Concern, and Packaging/Gift. For each issue type, two metrics were computed: Occurrence Rate and Average Rating. The occurrence rate was calculated as the percentage of total “Jewelry” reviews mentioning that specific issue. The

average rating was computed as the mean star rating across all reviews containing that issue to mention. This dual metric approach enables distinction between issues that occur frequently but cause minimal dissatisfaction versus issues that occur rarely but severely damage customer satisfaction, providing critical guidance for resource allocation decisions. This dashboard can be recreated for each NOVICA category, providing a more comprehensive overview of each.

Analysis 6: Executive Performance Dashboard



This visualization presents a comprehensive executive dashboard summarizing overall platform performance across all 25 product categories. The visualization incorporates six key performance indicator cards, rating distribution analysis, category performance comparison, priority improvement areas, and a Customer Satisfaction Score gauge.

Methodology: The executive dashboard aggregates data from the complete review corpus using BERT-powered sentiment analysis. The Customer Satisfaction Score (0-100 scale) was computed using a weighted formula incorporating average rating (normalized to percentage), proportion of positive reviews (77%), and inverse proportion of negative reviews (10%). Category performance metrics represent mean ratings weighted by review volume. Priority improvement areas were derived from frequency analysis of issue mentions within reviews rated 3 stars or below.

Interpretations & Business Implications

The analysis of customer sentiment and issue analytics divulges a dual narrative about the brand perception and product performance of NOVICA. On one hand, customers express high satisfaction with craftmanship, product experience and artisan quality, which directly meets with NOVICA's core goal of promoting and excelling in handmade, culturally rich goods which consequently drives strong brand affinity. This can be strongly seen in categories such as "Furniture" and "Home Decor" outperforming sentiment metrics and contribute positively to willingness to pay and conversion confidence.

Conversely, the analysis also undercovers the underlying pain points of customers that threaten reputation and operational efficiency. The two most enduring causes of disappointment were where customer expectations do not match, and product misinformation led to expensive returns and negative reviews. This dissatisfaction was prevalent across all categories, especially within "Jewelry" and "Apparel". A key insight that can be suggested from the analysis is that a communication gap exists rather than a flaw within their mission. This indicates that at the point of purchase NOVICA's value proposition may not be always fully understood.

The sentiment analysis and distribution indicate category specific risk zones. The analysis highlights that "Outdoor Living" and "Others" categories show higher than average negative sentiment. These negative sentiments are a consequence of delivery and durability concerns. When dealing with larger, more fragile items, return costs are higher and consumer confidence is hit harder, resulting in direct harm to NOVICA's margins.

A further concentration of product issues was highlighted in *Analysis 4 (Cross Category Issue Distribution)*, along with *Analysis 3 (Aspect-Level Performance Drivers):*

1) While NOVICA's highest-volume and fastest-selling categories, "Jewelry" accounts for over one-third of all complaints, driven largely by size and quality-related defects.

2) The packaging issues with "tableware" create an operational risk that might worsen as demand increases.

These patterns give insights for NOVICA to begin early risk detection and take charge in process improvements. From an operational perspective, these findings provide a roadmap for resource prioritization, directing quality assurance, packaging upgrades, and PDP enhancement efforts toward categories with the highest business upside.

Three crucial strategic functions are being conveyed through these insights:

1) Customer Experience Optimization: Addressing quality, fit, expectation-setting, and efficiency of delivery lowers dissatisfaction at the source.

2) Financial Efficiency & Margin Protection: Mitigating return-driven losses improves margins and overall performance, particularly in high-volume or fragile categories.

3) Brand Equity Enhancement: A handmade marketplace model relies on long-term retention and referrals, which are sustained by building trust in the purchasing experience.

Ultimately, this analysis helps connect directly to the decisions that NOVICA can make to strengthen their brand, operations, and growth.

Recommendations

During this semester-long analysis of Amazon reviews, the key insights led us to curate four actionable insights to help improve NOVICA's sales on Amazon using sentiment and topic analysis. By implementing these recommendations, NOVICA can increase their overall sales to help show the value of these artisans' pieces.

Building Trust Through Authentic Storytelling

The premium pricing should be explained to the customers by delivering the authenticity of the brand in a convincing story. By highlighting these points such as the labor, the high-quality materials used, and the cultural heritage the product carries. This helps to show *why* these products are priced the way they are.

Eliminating Size-Related Purchase Barriers

One of the biggest barriers to e-commerce shopping is the fear of sizing; this can be caused by unclear size guides. NOVICA can add a detailed size guide and images of models of various sizes wearing the products. This will help consumers gauge a better understanding of what would work best for them, which would also limit the number of returns.

Category-Specific Excellence Through Targeted Improvements

Each category has different ways of handling shipping, packaging, and product description to reduce returns. The analysis conducted shows that the products under "Jewelry" can have durability ratings for their metals and stones added to the product description. The "clothing" category can have more sample pictures of models of different sizes to check whether it suits different body types. For "home decor" pre-shipping inspections and post-purchase tips to take care of the product manual should be monitored and provided respectively, so that there is a reduction in the complaints about shipping and return rate.

Conclusion & Reflection

On Amazon, there is one critical aspect that helps products become visible to consumers: Customer Reviews. These Amazon reviews help buyers gain trust, become more visible on the Amazon digital shelf, and overall increase sales. Throughout this course, the technical, analytical, and the overall lessons learned helped shape the analysis and key insights. In the technical analysis, when exploring the data, there were sparse reviews amongst the SKUS, and most of the reviews had mostly positive sentiment. This was not a barrier but more of a reflection of how real-world e-

commerce data work. The analytical reflection observed is that few SKUs had the number of reviews directly proportional to the sales concentration; this led to skewed insights due to the imbalance in the sales of each category. It was also found that the keyword interpretation differs with respect to the context the review is written. One key point that truly impacted the analysis was a small but mighty statement: the quality of the reviews has more importance than the quantity of reviews. By fully being able to understand *why* a consumer enjoys or is dissatisfied with a product, though, the reviews can help NOVICA gain a holistic view of how their products are doing on Amazon.