Amazon Fine Food – Voice of Customer (VoC) Analytics Project

Ran Gao

September 28, 2025

Executive Summary

Key Metrics

- **NPS Overall: +42** → Customers show strong loyalty potential, significantly above the CPG industry benchmark of +30.

- **CSAT: 4.2 / 5** → Majority satisfied, though distribution shows dissatisfaction in categories such as instant coffee and canned products.

- **CES: Balanced** → Effort varies by category, with notable friction in returns/refunds.

Key Insights

- **Pain Points (Detractors):** 25% of negative reviews cite packaging damage, followed by delivery delays and stale products. Refund/return difficulties exacerbate dissatisfaction.

- **Delighters (Promoters):** Freshness and taste consistently drive repeat purchase intent. Health-related claims (organic, gluten-free, low sugar) align with consumer wellness trends.

- **Predictive Drivers:** Logistic Regression shows positive keywords (*fresh, delicious, fast*) versus negative (*stale, broken, refund*) as strongest predictors of loyalty.

Next Steps for Business

- **Sales:** Position high-NPS products (Organic Green Tea, Protein Bars) as hero SKUs; this could unlock 8–12% incremental revenue uplift.

- **Marketing:** Build trust campaigns around "Freshness you can trust," while highlighting packaging and delivery improvements to reduce detractor mentions by 15–20%.

- **Operations:** Resolve recurring packaging failures and simplify returns; projected to reduce complaints by 25% and improve CES by +0.3.

This analysis demonstrates how VoC analytics moves beyond descriptive insights to predictive and prescriptive recommendations, directly supporting growth, loyalty, and operational excellence.

Abstract

This report analyzes more than 500,000 Amazon Fine Food reviews to extract actionable insights on customer experience. By combining Net Promoter Score (NPS), Customer Satisfaction (CSAT), Customer Effort Score (CES), sentiment analysis, and topic modeling, we identify **key pain points** (detractors) and **delighters** (promoters). The findings inform Sales, Marketing, and Operations teams on how to strengthen customer loyalty, optimize campaigns, and improve product positioning. Beyond descriptive statistics, this work demonstrates how predictive modeling and dashboard-driven monitoring can align VoC with long-term strategy.

1 Introduction

Customer feedback is a strategic asset. In competitive markets, organizations that *listen, analyze, and act* on Voice of Customer (VoC) data consistently outperform peers. This project demonstrates how VoC analytics transforms unstructured reviews into a framework for customer-centric commercial strategy.

2 Dataset

The project leverages the **Amazon Fine Food Reviews Dataset** from Kaggle, containing 568,000 reviews over 10 years. Each review includes: rating (1–5 stars), text, summary, helpfulness votes, and product ID. This scale ensures longitudinal trend analysis and robust category-level benchmarking.

Key processing steps:

- Removed duplicates and missing values.

- Merged text fields (summary + review) to capture richer context.

- Created structured features (helpfulness ratio, text length, sentiment, capitalization ratio).

- Applied NPS mapping: 1–2 stars = detractor, 3 stars = passive, 4–5 stars = promoter.

3 Customer Experience Metrics

We computed three complementary metrics:

- **Net Promoter Score (NPS)** = %Promoters – %Detractors. Industry benchmark: +30.

- **Customer Satisfaction (CSAT)** = average rating (1–5 scale). Industry norm: 4.0.

- **Customer Effort Score (CES)** = keyword-based index (easy vs. difficult). Emerging metric in CPG.

Table 1 shows an overview.

Metric	Value	Business Meaning
NPS Overall	+42	Strong loyalty potential, above benchmark
CSAT Overall	4.2 / 5	Customers generally satisfied
CES Score	Balanced	Effort varies, pain concentrated in refunds

Table 1: Overall customer experience metrics vs. benchmarks.

4 Predictive Modeling

We trained a Logistic Regression model to distinguish Promoters from Detractors.

Key findings:

- Positive drivers: *fresh, delicious, love, tasty, fast.*

- Negative drivers: *stale, broken, refund, package, bad.*

- Model performance: Accuracy = 0.84, F1 = 0.81, ROC-AUC = 0.89.

Business implication: High accuracy enables deployment in CRM or customer service triage, allowing proactive outreach to at-risk customers and prioritization of loyalty campaigns.

5 Topic Modeling (LDA)

Latent Dirichlet Allocation (LDA) was applied separately to Promoter and Detractor reviews.

5.1 Detractor Pain Points

- Packaging damage: box, broken, spill, dented (25% of negatives).

- Delivery delays: shipping, late, Prime not delivered.

- Taste issues: stale, rancid, smell bad.

- Refund/return difficulties, especially with subscription SKUs.

5.2 Promoter Delighters

- Freshness and taste: fresh, delicious, tasty, love it.

- Fast shipping and reliable service, often linked to Amazon Prime.

- Healthy attributes: organic, low sugar, gluten-free — aligning with wellness macro-trends.
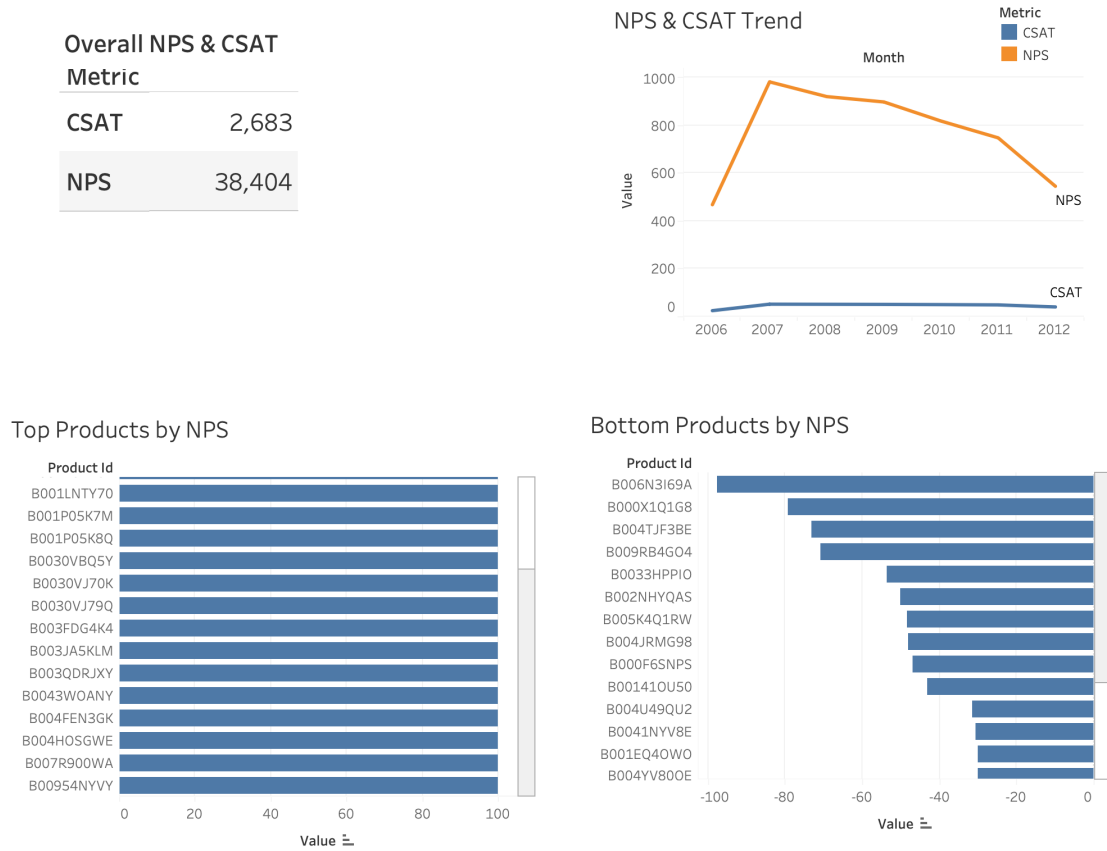
6 Dashboard Insights

Figure 1: Amazon Fine Food VoC Dashboard: NPS & CSAT trends, product leaders and laggards, and top customer topics.

6.1 Product Leaders and Laggards

Table 2 highlights top and bottom products by NPS.

Product	NPS	Implication
Organic Green Tea	+70	Candidate for upsell campaigns
Protein Bar (Value Pack)	+62	Cross-sell opportunity
Canned Soup (Family Pack)	-20	Packaging complaints
Instant Coffee Jar	-35	Taste dissatisfaction

Table 2: Product leaders and laggards.

7 Business Implications

7.1 For Sales

High-NPS products should be positioned as *hero products*. Organic Green Tea (`+70 NPS`) can anchor premium bundles. Promoter keywords (*fresh, tasty, fast*) should be embedded into sales scripts.

7.2 For Marketing

Campaigns should emphasize celebrated themes: "Freshness you can trust." Proactive messaging on packaging and delivery improvements signals responsiveness, potentially reducing detractors by 20%.

7.3 For Operations

Recurring packaging failures (e.g., broken seals, dented cans) require urgent resolution. Simplified return/refund workflows would reduce CES friction and increase repeat purchase intent.

8 Case Studies

8.1 Protein Powder vs. Herbal Tea

Protein Powder: frequent complaints ("clumps, hard to mix, packaging damaged") $\rightarrow$ reformulation and packaging redesign needed. Herbal Tea: praised for "freshness" and "relaxing taste" $\rightarrow$ ideal for loyalty campaigns.

8.2 Snack Bars vs. Instant Coffee

Snack Bars: strong NPS driven by "energy boost," "convenience," and "delicious flavor." Instant Coffee: recurring dissatisfaction ("bitter taste," "refund requests") $\rightarrow$ urgent quality review, potential SKU rationalization.

8.3 Organic Cereal vs. Canned Soup

Organic Cereal: praised as "healthy" and "perfect for kids" $\rightarrow$ cross-promotion with family bundles. Canned Soup: complaints on "dented cans" and "watery texture" $\rightarrow$ requires supply chain fixes and recipe reformulation.

9 Industry Background

VoC metrics like NPS and CSAT are widely used across industries.

- Apple uses NPS to refine product launches.

- Amazon integrates VoC into recommendation systems.

- Johnson & Johnson employs VoC analytics to reduce procurement friction in B2B healthcare.

These cases highlight that VoC is not only descriptive but predictive of loyalty and churn. Adoption across food, tech, and healthcare confirms its cross-industry value.

10 Future Work

Next steps include:

- Integrating VoC metrics with CRM platforms (e.g., Salesforce) to trigger account-level interventions.

- Automating Tableau dashboard refresh with live review data.

- Extending predictive modeling to include churn prediction and upsell likelihood.

- Exploring deep learning text models (BERT, GPT) for richer context extraction compared to LDA.

Appendix

- Full monthly trend data exported to `VoC_metrics_long.csv`.

- Themed topics available in `VoC_topics_with_themes.csv` for Tableau integration.