International Review of Management and Marketing

ISSN: 2146-4405

available at http: www.econjournals.com

International Review of Management and Marketing, 2026, 16(1), 267-274.

What Drives Satisfaction in Fresh E-Commerce? Evidence from Review-Based Topic and Sentiment Analysis

Xiaomei Han[1,2]*, Hamrila Abdul Latif[1], Chin-Hong Puah[1]

[1]Anhui Finance and Trade Vocational College, Hefei, Anhui, China, [2]Faculty of Economics and Business, Universiti Malaysia Sarawak (UNIMAS), Kota Samarahan, Sarawak, Malaysia. *Email: 21010239@siswa.unimas.my

Received: 17 June 2025 **Accepted:** 04 October 2025 **DOI:** https://doi.org/10.32479/irmm.21172

ABSTRACT

This study explores key determinants of customer satisfaction in China's fresh e-commerce sector by analyzing large-scale user-generated reviews from JD Fresh. Unlike prior research relying on surveys, this study integrates semantic segmentation, high-frequency textual data, and multivariate modeling to offer a data-driven and fine-grained understanding of customer experience. Nine service dimensions were identified: Freshness, Affordability, Size, Taste, Ice Pack, Delivery, Quality, Storage, and Platform Functionality. Sentiment scores were calculated for each dimension and used in a multiple linear regression model to assess their impact on overall satisfaction. Results indicate that all nine factors significantly affect satisfaction, with Taste, Quality, and Storage being the most influential. These findings demonstrate the effectiveness of combining text mining and sentiment analysis in identifying service quality dimensions and predicting satisfaction. The study also highlights the value of online reviews as a scalable source for service evaluation and provides actionable insights for improving user experience and repurchase intention in the highly competitive fresh e-commerce industry.

Keywords: Fresh E-Commerce, Customer Satisfaction, Online Reviews, Sentiment Analysis, Topic Modeling, Text Mining
JEL Classification: M31, L81, C55, D12

1. INTRODUCTION

Fresh e-commerce refers to online retail and delivery of perishable food products such as fruits, vegetables, meat, seafood, and dairy through digital platforms. In China, the rapid advancement of internet technologies and the ongoing digital transformation of consumer shopping habits have driven significant growth in this sector over the past decade. The COVID-19 pandemic further accelerated this trend, as consumers increasingly turned to online grocery platforms for convenience and safety. By 2024, the total transaction volume of China's fresh e-commerce industry reached approximately RMB 736.8 billion, representing a 14.7 percent year-over-year increase. This data, reported by the E-commerce Research Center of 100EC.CN (NetEconomy, 2024), highlights the sector's expanding role in the modern retail landscape.

Despite its rapid growth, the fresh e-commerce industry in China continues to grapple with significant structural challenges that threaten its long-term sustainability. The perishable nature of fresh products, along with their lack of standardization and high time sensitivity, contributes to elevated logistics costs and ongoing difficulties in maintaining consistent quality. These issues directly affect customer satisfaction and loyalty. For example, although JD Fresh has established a strong market presence through its self-operated logistics network, customer feedback frequently points to problems such as delivery delays and a lack of product freshness (Chen et al., 2023).

Service-related challenges have contributed to a broader growth bottleneck in China's fresh e-commerce industry. Recent studies show that customer loyalty remains low, with repurchase rates significantly below those of general e-commerce platforms (Pu et al., 2025). Improving customer satisfaction has thus become a key strategy for overcoming stagnation. However, two major shortcomings persist in current management practices: an overreliance on manual review of customer feedback, which is inefficient and limited, and a tendency to implement fragmented improvements without a system-wide approach. This study uses large-scale review data to identify core customer pain points and guide more effective resource allocation.

2. LITERATURE REVIEW

Research on customer satisfaction began with Oliver's (1980) expectation-disconfirmation model, which explains that satisfaction depends on the gap between what customers expect and what they experience. Later, Parasuraman et al. (1985) developed the SERVQUAL model, identifying five key service quality dimensions: reliability, responsiveness, assurance, empathy, and tangibles. This model has become a key reference in e-commerce studies. In fresh e-commerce, unique factors like delivery speed and the visual appeal of perishable goods also play an important role (Xu and Chen, 2020).

To better understand customer experiences, researchers have increasingly used online reviews. Earlier methods, such as word matching and frequency counts, were easy to use but often missed context and subtle meanings like sarcasm or emotion (Cambria et al., 2013). With the rise of deep learning, models like BERT (Devlin et al., 2019) have improved how text is analyzed.

For topic modeling, LDA was widely used but had limits, such as requiring the number of topics to be set in advance and ignoring word context. Newer methods like BERTopic (Grootendorst, 2022), which use BERT and clustering, offer better topic detection and more meaningful groupings. While effective in areas like social media, BERTopic has rarely been applied to fresh e-commerce.

Despite growing interest, existing literature presents three key limitations: (1) studies on fresh e-commerce remain heavily dependent on survey data, lacking in-depth analysis of authentic consumer language; (2) most attempts to combine topic modeling with sentiment analysis are limited to superficial associations, without constructing robust quantitative frameworks; and (3) there is a notable lack of case studies focusing on major platforms such as JD Fresh, raising concerns about the generalizability of existing findings. This study aims to systematically analyze customer review data from the JD Fresh platform by integrating topic modeling with fine-grained sentiment analysis. The objectives are as follows:

RO1: Identify key experience-related themes reflected in customer reviews of fresh e-commerce

RO2: Evaluate consumers' emotional tendencies (positive/negative) across different themes

RO3: Identify which themes have a significant impact on overall customer satisfaction.

3. METHODOLOGY

This study focuses on JD Fresh, one of China's leading fresh e-commerce platforms, which holds a 25.8% market share. Its supply chain spans over 300 cities, with a cold chain logistics coverage rate of 95%, reflecting strong operational capabilities. JD Fresh also offers the most complete product range in the industry, with 16 main categories and 218 subcategories, providing a rich foundation for analyzing consumer behavior. The platform receives over 50,000 user reviews daily, making it ideal for large-scale text analysis. As a representative case, JD Fresh offers insights that can inform broader strategies in the fresh e-commerce sector.

3.1. Data Collection

This study uses customer reviews from best-selling products on the JD Fresh platform. Data were collected using the Octoparse web scraping tool (version 9.1) from January 2024 to January 2025, covering all four seasons.

A total of 22,563 raw reviews were gathered, including information such as review ID, product category, star rating, content, posting date, and province-level location. The analysis focused on four key product categories: fruits, vegetables, meat and eggs, and seafood—core areas of fresh food consumption. To ensure sample quality, only the top 200 best-selling products were included, excluding items with monthly sales below 1,000 units. Table 1 shows sample entries from the dataset.

The raw data underwent several cleaning steps. First, duplicate reviews were removed using the Levenshtein distance algorithm with a similarity threshold above 90% to reduce redundancy. Second, short reviews with fewer than 10 characters (e.g., "Good," "666") were filtered out for being semantically uninformative. Third, a custom noise dictionary of 327 regular expressions was applied to remove system messages, promotional content, special characters, and meaningless symbols. Finally, non-Chinese reviews were filtered using the langdetect Python library, which identifies Chinese text with 98.7% accuracy. After these steps, 19,147 high-quality reviews remained. Table 2 summarizes the dataset's statistical characteristics.

3.2. Topic Modeling

This study uses a hierarchical semantic segmentation method to perform topic modeling on customer reviews with the BERTopic framework (Figure 1). BERTopic has clear advantages over traditional methods like LDA and NMF (Blei and Lafferty, 2006), especially for short and fragmented texts. Studies show it produces more coherent and meaningful topics by leveraging Transformer-based embeddings and cTFIDF keyword extraction (de Groot et al., 2022; Liu, 2024).

Before analysis, each review was split into smaller semantic clauses using Chinese punctuation rules. This approach reduces topic overlap and improves the detail of the analysis. From 19,147 reviews, a total of 92,115 clauses were extracted, averaging 12.6 characters and 4.81 clauses per review. For example, the review "The apples are sweet, large in size, but the packaging was damaged, and customer service quickly resolved the issue" is divided into four parts: ["The apples are sweet," "Large in size," "But the packaging was damaged," "Customer service quickly resolved the issue"].

For semantic embedding, the study uses the "sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2" model

Figure 1: Topic modeling workflow

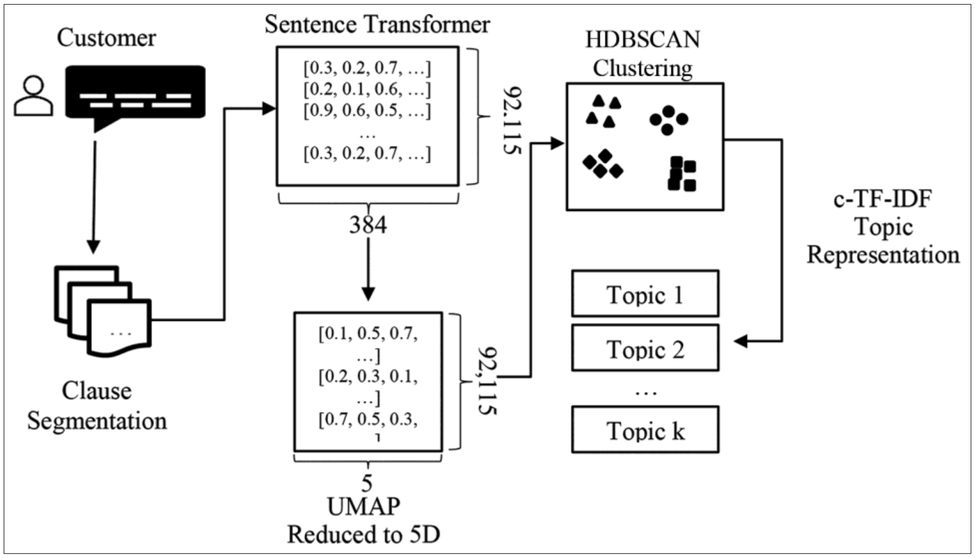


Table 1: Sample customer review data from JD fresh

Review ID	Product category	Star rating	Review text	Date posted	Location
REV_202401_ 001	Fruits	5	Strawberries are fresh and large. Cold-chain packaging is professional. Will repurchase!	2024/1/12	Beijing
REV_202403_ 045	Vegetables	3	Spinach had yellow leaves, and delivery was delayed by a day.	2024/3/5	Jiangsu
REV_202405_ 128	Meat and Eggs	4	Steak was of moderate thickness, but had some blood after thawing.	2024/5/18	Guangdong
REV_202408_ 302	Seafood	5	The live shrimp were still moving on arrival. Ice packs were sufficient.	2024/8/22	Zhejiang
REV_202410_ 517	Fruits	2	Some cherries were rotten. Customer service was polite but compensation was limited.	2024/10/11	Shanghai
REV_202412_ 689	Seafood	4	Salmon slices were neatly cut, but weight was 10% less than described.	2024/12/30	Sichuan

Table 2: Multidimensional feature distribution of JD fresh customer reviews

Feature dimension	Category	Count	Percentage
Time distribution	Q1 (January-March)	4,614	24.1
	Q2 (April-June)	3,791	19.8
	Q3 (July-September)	4,902	25.6
	Q4 (October-December)	5,840	30.5
Product category	Fruits	7,123	37.2
	Seafood	5,457	28.5
	Meat and eggs	4,231	22.1
	Vegetables	2,336	12.2
Rating distribution	5 Stars	8,099	42.3
	4 Stars	6,070	31.7
	3 Stars and below	4,978	26.0
Geographic location	Beijing	3,025	15.8
	Guangdong	2,776	14.5
	Jiangsu	2,144	11.2
	Zhejiang	1,838	9.6
	Shanghai	1,532	8.0
	Sichuan	1,340	7.0
	Shandong	1,168	6.1
	Other provinces	5,324	27.8
Total		19,147	100

(Reimers and Gurevych, 2019). This lightweight, efficient multilingual model supports over 50 languages, including Chinese and English, encoding each clause into a 384-dimensional vector. Compared to traditional BERT models, it offers faster computation while maintaining strong semantic accuracy. Its compact size and strong cross-lingual ability make it well suited for tasks like semantic similarity, short-text matching, clustering, and information retrieval across multiple languages.

Typical clauses like “professional cold chain packaging” and “some strawberries were rotten” are converted into high-dimensional vectors using the sentence embedding model. These form a 92,115 × 384 semantic matrix, which serves as input for clustering. To improve efficiency while preserving meaning, the vectors are first reduced to five dimensions using Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018).

Next, the reduced vectors are clustered using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm (Campello et al., 2013). A minimum cluster size of 15 is set, meaning each topic must contain at least 15 clauses. This helps remove noise and ensures more coherent topics. The total number of topics is determined dynamically based on data density and clustering settings.

Using this automated process, initial topic clusters such as “cold chain logistics” and “product freshness” were identified from the semantic embeddings and clustering results. Representative terms for each topic were extracted using the class-based TF-IDF (c-TF-IDF) method, which treats all documents in a topic as one class and highlights terms frequent within that topic but rare elsewhere

Table 3: Hierarchical topic schema based on BERTopic

Topic ID	Sample size	Topic label	Representative terms
0	2385	0_Freshness_Clean_Fruit_Tasty	[“fresh,” “clean,” “fruit,” “tasty,” “healthy,” “pomegranate,” “vibrant,” “purple skin,” “reddish,” “green”]
1	7136	1_Affordability_Cost- effectiveness_ Shopping_Superm arket	[“affordable,” “cost-effective,” “shopping,” “supermarket,” “value for money,” “recommend,” “in- store,” “physical store,” “worth it,” “store”]
2	7086	2_Size_Fullness_Surprise_Porti on	[“size,” “full,” “surprise,” “portion,” “description,” “tender’, “review’, “quite big,” “pomegranate,” “peeled”]
3	9933	3_Taste_Texture_Delicious_Fla vorful	[“taste,” “texture,” “delicious,” “flavorful,” “flavor,” “home,” “very sweet,” “cooking,” “breakfast,” “tasty’]
4	1591	4_Ice Pack_Frozen_Thawing_Cold Chain	[“ice pack,” “ice,” “thawing,” “cold chain,” “frozen,” “dry ice,” “frozen solid,” “insulated box,” “winter,” “pre-frozen”]
5	4152	5_Delivery_Logistics_Service Attitude_Packaging	[“delivery,” “logistics,” “service attitude,” “foam,” “doorstep delivery,” “quite fast,” “SF Express,” “box,” “outer packaging,” “pandemic”]
6	5137	6_Quality_Positive Review_ Satisfaction	[“quality,” “positive review,” “good quality,” “satisfaction,” “recommend,” “quite good,” “worth it,” “delicate,” “aesthetic,” “nice looking”]
7	7282	7_Refrigerator_Disappointment _Days_Weight	[“refrigerator,” “disappointed,” “days,” “weight,” “garbage,” “worried,” “too small,” “worst,” “five jin,” “twice”]
8	1683	8_JD_Mall_Online Shopping_ Logistics	[“JD,” “mall,” “shopping,” “logistics,” “trustworthy,” “subsidy,” “billion discount,” “online,” “brand,” “convenient”]

(Grootendorst, 2022). Key expressions like “ice packs,” “delivery timeliness,” and “defrosted blood water” clearly capture major user concerns about delivery in fresh e-commerce.

After topic modeling, each semantic clause was assigned to its most relevant topic using a probabilistic approach based on HDBSCAN’s soft clustering. This resulted in a dataset linking clause IDs, topic IDs, and assignment probabilities—for example, the clause “professional cold chain packaging” was assigned to “logistics service quality” with a 0.92 probability. The final output included 9 primary topics, detailed in Table 3, each described by representative keywords reflecting core user concerns. For instance, Topic 0 focuses on “freshness” with terms like “fresh” and “clean,” Topic 3 on “taste and flavor” with words such as “tasty” and “delicious,” and Topic 4 on “cold chain assurance” with terms like “ice pack” and “cold chain.” The number of clauses per topic varied widely, from 1,591 to 9,933, showing different levels of user attention.

To validate topic assignments, a stratified random sample of 200 clauses was drawn from the full 92,115, with weighting by topic size and assignment probability (30% above 0.8, 30% below 0.5). Three domain experts independently rated each clause’s topic relevance on a three-level scale: fully accurate, partially accurate, or inaccurate.

Accuracy was calculated using a weighted formula: (fully accurate + 0.5 × partially accurate)/total samples × 100%. Results showed that 162 clauses (81.0%) were fully accurate, 34 (17.0%) were partially accurate, and 4 (2.0%) were inaccurate, yielding a weighted accuracy of 89.3%.

3.3. Sentiment Analysis

Sentiment analysis is a key task in natural language processing (NLP) that detects emotions, sentiment polarity, and attitudes in text using computational methods. This study uses Baidu Intelligent Cloud’s Sentiment Classification API, which analyzes Chinese short texts and labels them as positive, neutral, or negative, along with confidence scores.

Tang et al. (2020) compared seven popular Chinese sentiment tools and found Baidu’s API to be the most accurate, especially for short, user-generated content like e-commerce reviews. It handles strong emotions and colloquial language better than others, making it ideal for commercial sentiment analysis.

Building on the clause-level segmentation done during topic modeling, each clause is fed into Baidu’s API (https://ai.baidu.com/tech/nlp_apply/sentiment_classify), which assigns a sentiment score from −1 (negative) to +1 (positive), with 0 as neutral. These scores are then combined at the review level to create structured sentiment indicators, shown in Figure 2.

To ensure accuracy, the study used two validation methods. First, 500 randomly selected clauses were manually labeled, confirming the API’s accuracy at 89.7% (Cohen’s Kappa = 0.81). Second, a time-series validation showed the API had a much lower error rate (3.2%) than traditional lexicon-based methods (12.1%) when detecting fresh food-specific terms like “defrosted” or “yellowing leafy greens.”

3.4. Topic–sentiment Mapping

In the topic–sentiment mapping phase, each semantic clause is assigned a topic ID based on earlier topic modeling (e.g., Topic 0 for “product freshness,” Topic 5 for “logistics timeliness”). Since a clause can relate to multiple topics, a semantic overlap method using Jaccard similarity (Nwattanakul et al., 2013) compares the clause’s keywords with each topic’s keywords.

The topic with the highest similarity is chosen as the final assignment. This creates a set of topic–sentiment score pairs for each clause. Aggregating these scores gives an overall sentiment for each topic—for example, Topic 0 (“product freshness”) has a mean sentiment score of +0.62, showing general satisfaction. If a review doesn’t mention a topic, it is assigned a neutral score of 0, following common practices in multi-label aspect-based sentiment analysis (Ma et al., 2018).

To validate this mapping, the review's star rating (1-5), which reflects overall satisfaction, is used as a reference. Star ratings are collected alongside reviews and help check the consistency between topic-specific sentiments and overall user ratings.

Table 4 shows a sample sentiment matrix where each row is a review, each column is a topic's sentiment score, and the last column is the star rating. Most reviews only express sentiment for some topics, while unmentioned topics are marked with 0, indicating neutrality.

The boxplot in Figure 3 shows the distribution of sentiment scores across nine main topics (T0-T8) for JD Fresh reviews.

Most topics have positive sentiment, with median scores above 0.5—especially Freshness (T0, median ≈ 0.84), Affordability (T1, ≈ 0.87), Taste (T3, ≈ 0.91), and Delivery (T5, ≈ 0.86)—indicating strong customer satisfaction. In contrast, Ice Pack (T4) and Storage (T7) display wider sentiment ranges (from about –1.0 to 1.0) and lower medians (T4 ≈ 0.06, T7 ≈ –0.16), reflecting mixed to negative feelings. Storage (T7) shows particularly dispersed sentiment, suggesting uneven service quality. Outliers across nearly all topics reveal some users had extreme positive or negative experiences. Overall, the sentiment patterns highlight generally positive perceptions but expose key weaknesses in cooling and storage logistics that need managerial focus.

3.5. Regression Modeling

To quantitatively measure how topic-specific sentiments influence overall customer ratings, this study uses multiple regression. Sentiment scores for each topic—derived from clause-level aspect-based sentiment analysis—serve as independent variables, while the overall star rating (1-5) is the dependent variable. This approach identifies which aspects, such as "product freshness," "logistics timeliness," or "pricing," most affect customer satisfaction.

Table 4: Example of topic–sentiment score matrix with star ratings

Review ID	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Star rating
123	0	0	0.909	0.990	0	0	0	0	0.998	5
785	0.995	0.989	0	0	0	0	0	0	0	5
841	0	0	0	−0.969	0	0	0	0.764	0	1
916	0	0.992	0	0	0	0	0	0	0	5
1345	0	0.997	0	0	0	0	0	0	0	5
2592	0.999	0.986	0	0.980	0	0	0	0	0	5
3451	0	0.997	0	0	0	1	0	0	0	5
4125	0	0	0	0	0	0	0.966	0.590	0	4
8916	0	0	0	0.959	0	0	0.767	0.410	0	5

Figure 2: Sentiment score extraction using baidu sentiment classification API

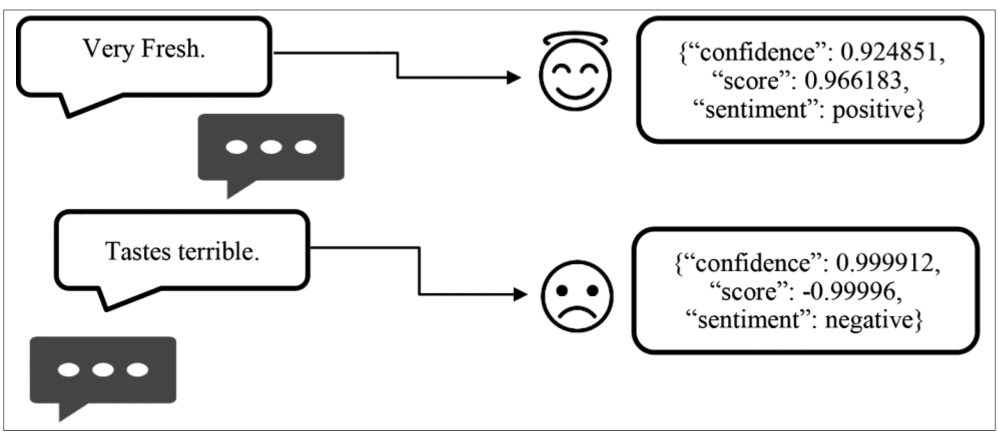


Figure 3: Sentiment score distribution across topics

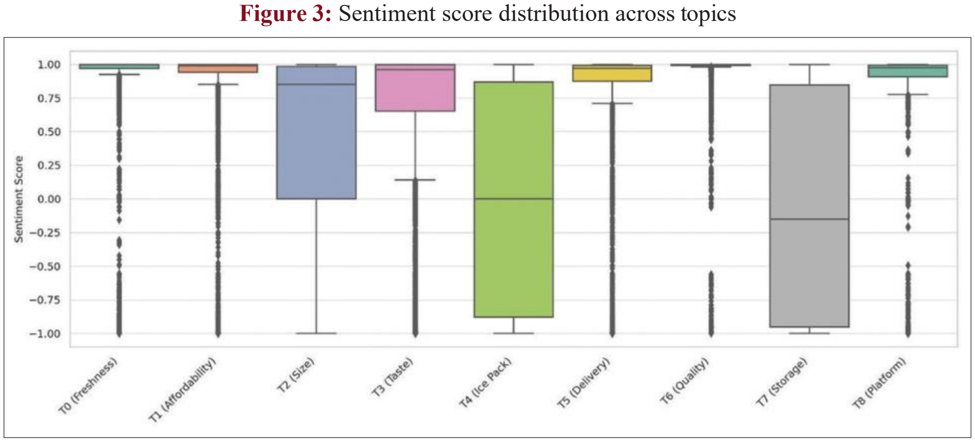


Table 5: Multiple linear regression coefficients for overall satisfaction rating

Variable	B	Std. Error	β	t	P	Tolerance	VIF
(Constant)	4.200	0.010		418.642	0.000	-	-
T0 (Freshness)	0.183	0.014	0.080	12.890	<0.001	0.965	1.037
T1 (Affordability)	0.233	0.012	0.124	19.493	<0.001	0.914	1.095
T2 (Size)	0.229	0.012	0.124	19.540	<0.001	0.912	1.096
T3 (Taste)	0.393	0.011	0.223	35.370	<0.001	0.927	1.078
T4 (Ice Pack)	0.096	0.026	0.023	3.710	<0.001	0.996	1.004
T5 (Delivery)	0.23	0.013	0.114	17.966	<0.001	0.915	1.093
T6 (Quality)	0.348	0.012	0.182	28.828	<0.001	0.922	1.084
T7 (Storage)	0.388	0.013	0.198	30.868	<0.001	0.892	1.121
T8 (Platform)	0.079	0.020	0.024	3.949	<0.001	0.976	1.024

***P<0.001, Dependent Variable: Overall Satisfaction Rating (1-5 scale), Sample Size: N=19,147

Table 6: ANOVA results for overall regression model

Source	Sum of squares (SS)	df	Mean square (MS)	F-value/ P-value
Regression	5,186.90	9	576.322	F=893.35, P<0.001
Residual	12,345.74	19,137	0.645	-
Total	17,532.63	19,146	-	-

This method builds on established ABSA research, which shows that sentiment toward different aspects can predict overall attitudes (Pontiki et al., 2016; Zhang et al., 2021). Using structured topic-aligned sentiment scores captures detailed user feedback, and treating star ratings as continuous variables aligns with prior work (Mu et al., 2021).

Multiple linear regression is chosen for its simplicity and proven effectiveness in estimating how each aspect sentiment impacts overall ratings (Zhang et al., 2023; Regitz et al., 2024). The model's coefficients reveal which service dimensions most strongly influence perceived customer value.

4. DISCUSSION

Using 19,147 customer reviews from JD Fresh, this study applied a multiple linear regression model to assess how nine thematic dimensions affect overall satisfaction ratings. Table 5 presents the regression coefficients, standard errors, t-values, and significance levels. All predictors are statistically significant (P < 0.001), demonstrating their individual impact on overall ratings. Low VIF values confirm there is no multicollinearity among the variables.

Table 6 presents the ANOVA results, showing the regression model is statistically significant (F = 893.35, P < 0.001). The model explains a large portion of the variance in customer satisfaction, with the mean square of the model (MS = 576.322) greatly exceeding that of the residuals (MS = 0.645). Degrees of freedom are 9 for the regression and 19,137 for the residuals, reflecting the number of predictors and total sample size (N = 19,147). These findings confirm that the model effectively predicts customer satisfaction, supporting further analysis of individual coefficients (Field, 2024).

Table 7 shows the collinearity diagnostics, confirming no harmful multicollinearity in the regression model. The maximum condition

Figure 4: Standardized regression residuals histogram

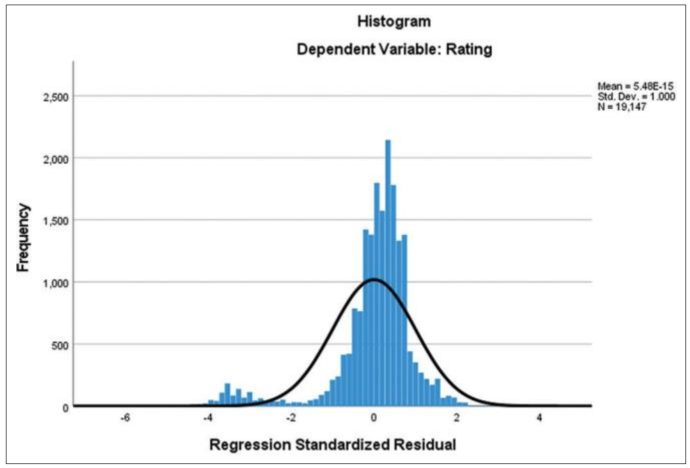
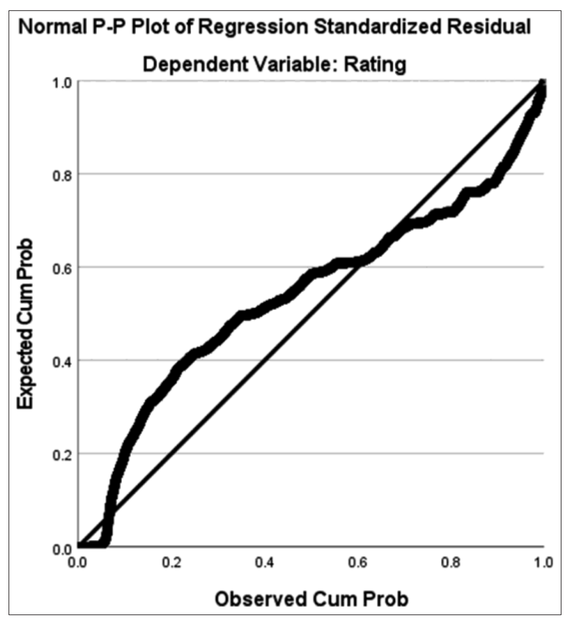


Figure 5: Normal probability-probability (P-P) plot of standardized regression residuals



index of 3.831 (Dimension 10) is well below the threshold of 30, indicating no significant collinearity issues. All eigenvalues exceed 0.238, ruling out perfect linear dependence among predictors. Variance proportion analysis reveals no dimension with multiple predictors having high variance proportions (>0.5), confirming variable independence (O'Brien, 2007).

Table 7: Collinearity diagnosis results

Dim	Eigenvalue	Cond. index	Const.	T0	T1	T2	T3	T4	T5	T6	T7	T8
1	3.487	1.0	0.02	0.02	0.03	0.02	0.02	0.0	0.03	0.02	0.0	0.01
2	1.076	1.801	0.02	0.0	0.0	0.05	0.01	0.14	0.0	0.01	0.55	0.0
3	0.986	1.88	0.0	0.0	0.0	0.01	0.0	0.84	0.0	0.0	0.13	0.0
4	0.899	1.969	0.0	0.12	0.0	0.03	0.05	0.0	0.01	0.0	0.0	0.74
5	0.794	2.096	0.0	0.18	0.05	0.08	0.13	0.01	0.14	0.07	0.05	0.23
6	0.758	2.145	0.0	0.5	0.01	0.39	0.04	0.0	0.05	0.0	0.04	0.01
7	0.691	2.246	0.01	0.06	0.02	0.42	0.39	0.0	0.13	0.0	0.03	0.0
8	0.584	2.444	0.0	0.06	0.14	0.01	0.25	0.0	0.61	0.06	0.03	0.0
9	0.487	2.676	0.02	0.0	0.66	0.0	0.0	0.0	0.03	0.38	0.02	0.0
10	0.238	3.831	0.92	0.05	0.08	0.0	0.11	0.0	0.01	0.45	0.14	0.0

Residual analysis indicates the model meets key linear regression assumptions. Figure 4 shows the histogram of standardized residuals approximates normality (skewness = –0.13, kurtosis = 0.98). Figure 5’s P-P plot aligns closely with the diagonal, and the Kolmogorov–Smirnov test is non-significant ($P = 0.082$), supporting normality. Additionally, 95.2% of standardized residuals fall within ±1.96, confirming homoscedasticity.

The regression analysis shows that all nine sentiment topics (T0-T8) significantly affect overall customer satisfaction ($P < 0.001$). These topics—including product taste, quality, size, freshness, affordability, delivery, cold-chain packaging, storage, and platform perception—each play an important role in shaping customer evaluations. The results highlight that satisfaction in fresh e-commerce depends on multiple aspects of the consumer experience. While the impact of each topic varies, none are redundant, emphasizing the comprehensive nature of service quality in this sector.

These findings support previous research highlighting the role of multidimensional feedback in understanding consumer sentiment. For example, Zhang et al. (2021) showed that perceived value in service, pricing, and quality strongly influences satisfaction and repeat purchases. Chen et al. (2019) found that structured review data can effectively predict customer satisfaction. Additionally, Lee and Kim (2018) emphasized that both platform features and customer service qualities, like convenience and responsiveness, play key roles in shaping satisfaction in digital marketplaces.

Importantly, these findings align well with classic service quality theory. For instance, Taste and Freshness relate to the Tangibles dimension in the SERVQUAL model; Delivery and Ice Pack correspond to Reliability and Responsiveness; Customer Service reflects Empathy; and Platform Functionality aligns with Assurance (Parasuraman et al., 1985). This alignment supports the validity of the topic-based results and connects data-driven insights with established service frameworks.

Based on these findings, this study suggests several practical recommendations for fresh e-commerce platforms. First, prioritize improvements in Taste, Product Quality, and Storage Conditions, as these have the strongest influence on customer satisfaction. Enhancing cold-chain logistics—such as ensuring adequate ice packs and controlling warehouse freshness—is especially important. Second, consistently maintain high standards in Freshness, Pricing, and Delivery Efficiency to boost positive customer perceptions and reduce complaints. Third, optimize Platform Usability by improving the user interface and streamlining transaction processes to encourage repeat purchases. Finally, implement real-time sentiment monitoring of customer reviews to quickly identify issues and refine services. Together, these strategies can help fresh e-commerce platforms deliver a more consistent, satisfying, and loyalty-driven shopping experience in a competitive market.

5. CONCLUSION

This study aimed to identify and measure the key factors influencing customer satisfaction in China’s fresh e-commerce market by analyzing large-scale customer reviews from JD Fresh. For RO1, topic modeling revealed nine major service dimensions: Freshness, Affordability, Size, Taste, Ice Pack, Delivery, Quality, Storage, and Platform Functionality. These themes represent the core elements of consumers’ service experiences. For RO2, sentiment analysis showed that most themes—especially Freshness, Affordability, Taste, Delivery, and Quality—received predominantly positive feedback. In contrast, Ice Pack and Storage had more negative and varied sentiments, highlighting important areas of dissatisfaction. For RO3, multiple linear regression demonstrated that all nine sentiment dimensions significantly predicted overall customer satisfaction. Taste, Quality, and Storage had the strongest impact on satisfaction ratings.

Using JD Fresh as a case study, this research offers actionable insights to optimize its “211 Time-Limited Delivery” service and enhance its product traceability system. Beyond JD Fresh, the advanced text mining and sentiment analysis framework can be applied to other leading fresh e-commerce platforms like Miss Fresh and Hema Fresh, providing a systematic method to improve service quality industry-wide. As the fresh e-commerce sector shifts from rapid expansion to quality-driven competition, this study addresses a critical need by developing a dynamic, data-driven system for evaluating customer satisfaction. It not only expands the use of NLP in management science but also tackles the industry’s key challenge: balancing fast growth with improving customer loyalty.

6. FUNDING

This research was supported by the 2023 Natural Science Research Project of Universities in Anhui Province, China (Key Project, Grant No. 2023AH052071), entitled “*Aspect-Level Fine-Grained*

Sentiment Analysis of Product Reviews on E-commerce Platforms," the 2025 Training Program for Young and Middle-aged University Teachers in Anhui Province (Grant No. YQYB2025206), and the 2022 Annual Education and Teaching Research Planning Project of Anhui Vocational and Adult Education Society (Grant No. Azcj2022002).

REFERENCES

Blei, D.M., Lafferty, J.D. (2006), Dynamic Topic Models. In: Proceedings of the 23rd International Conference on Machine Learning.

Cambria, E., Schuller, B., Xia, Y., Havasi, C. (2013), New avenues in opinion mining and sentiment analysis. IEEE Intelligent Systems, 28(2), 15-21.

Campello, R.J., Moulavi, D., Sander, J. (2013), Density-based clustering based on hierarchical density estimates. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Berlin, Heidelberg: Springer.

Chen, S., Wang, Y., Han, S., Lim, M.K. (2023), Evaluation of fresh food logistics service quality using online customer reviews. International Journal of Logistics Research and Applications, 26(8), 917-933.

Chen, W., Gu, B., Ye, Q., Zhu, K.X. (2019), Measuring and managing the externality of managerial responses to online customer reviews. Information Systems Research, 30(1), 81-96.

De Groot, M., Aliannejadi, M., Haas, M.R. (2022), Experiments on generalizability of BERTopic on multi-domain short text. arXiv preprint arXiv:2212.08459.

Devlin, J., Chang, M.W., Lee, K., Toutanova, K. (2019), BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Vol. 1. p4171-4186.

Field, A. (2024), Discovering Statistics Using IBM SPSS Statistics. California: Sage Publications Limited.

Grootendorst, M. (2022), BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.

Lee, S., Kim, D.Y. (2018), The effect of hedonic and utilitarian values on satisfaction and loyalty of Airbnb users. International Journal of Contemporary Hospitality Management, 30(3), 1332-1351.

Liu, Y. (2024), Comparison of LDA and BERTopic in news topic modeling: A case study of The New York Times' reports on China. Pacific International Journal, 7(3), 47-51.

Ma, Y., Peng, H., Cambria, E. (2018), Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1), 5876-5883.

McInnes, L., Healy, J., Melville, J. (2018), UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

Mu, R., Zheng, Y., Zhang, K., Zhang, Y. (2021), Research on customer satisfaction based on multidimensional analysis. International Journal of Computational Intelligence Systems, 14(1), 605-616.

NetEconomy. (2024), 2024 China fresh e-commerce development report. Available from: https://www.100ec.cn/detail--6648767.html

Niwattanakul, S., Singthongchai, J., Naenudorn, E., Wanapu, S. (2013), Using of Jaccard Coefficient for Keywords Similarity. In: Proceedings of the International Multi Conference of Engineers and Computer Scientists.

O'Brien, R.M. (2007), A caution regarding rules of thumb for variance inflation factors. Quality and Quantity, 41(5), 673-690.

Oliver, R.L. (1980), A cognitive model of the antecedents and consequences of satisfaction decisions. Journal of Marketing Research, 17(4), 460-469.

Parasuraman, A., Zeithaml, V.A., Berry, L.L. (1985), A conceptual model of service quality and its implications for future research. Journal of Marketing, 49(4), 41-50.

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S.M., Eryiğit, G. (2016), SemEval-2016 Task 5: ASPECT based Sentiment Analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation. p19-30.

Pu, M., Cui, L., He, S. (2025), The factors influencing customer loyalty in fresh food e-commerce based on fuzzy sets of qualitative comparative analysis. Journal of Industrial and Management Optimization, 21(3), 1673-1685.

Regitz, D., Höpken, W., Fuchs, M. (2024), UGC-Based Factors Influencing Customer Satisfaction Pre and Post COVID-19: The Case of Lake Constance. In: Paper Presented at the ENTER e-Tourism Conference.

Reimers, N., Gurevych, I. (2019), Sentence-BERT: Sentence embeddings using siamese BERT-networks. arXiv preprint arXiv:1908.10084.

Tang, T., Huang, L., Chen, Y. (2020), Evaluation of Chinese Sentiment Analysis APIs Based on Online Reviews. In: 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). p530-534.

Xu, X., Chen, J. (2020), Research on the influence of e-commerce service quality of fresh agricultural products on customer satisfaction. E3S Web of Conferences, 217, 03008.

Zhang, N., Liu, R., Zhang, X.Y., Pang, Z.L. (2021), The impact of consumer perceived value on repeat purchase intention based on online reviews: By the method of text mining. Data Science and Management, 3, 22-32.