



ONLINE PRODUCT REVIEW ANALYSIS USING POWER BI

Prof. Namrata Jangam¹, Prerna Ahirrao²

¹Faculty, Computer Engineering

VPS College of Engineering and Technology, Lonavala

²Student, Computer Engineering

VPS College of Engineering and Technology, Lonavala

ABSTRACT

The rapid growth of e-commerce platforms has made online product reviews one of the most influential factors in customer purchase decisions. Consumers increasingly rely on product ratings, written reviews, and verified purchase indicators before selecting items from online marketplaces. However, the large volume of customer-generated feedback makes manual analysis inefficient and often impractical. This paper presents an internship-based analytical study titled “Online Product Review Analysis Using Power BI,” which focuses on transforming raw product and customer review data into actionable business insights through interactive dashboarding and review-oriented analytics. The project integrates structured product attributes such as product category, price, discount percentage, rating, review count, brand, city, age group, delivery days, order status, and verified review status. Power BI was used as the primary business intelligence tool for data import, transformation, modeling, DAX-based KPI generation, and dashboard creation. The analytical framework supports category-wise product analysis, rating distribution, customer segmentation, verified review evaluation, and delivery-performance impact assessment. The internship dataset indicates approximately 3892 total reviews, an average rating of 3.24, and nearly 85% verified reviews, suggesting a moderately positive and comparatively reliable review environment. Results show that customer satisfaction is influenced not only by product quality and ratings but also by pricing strategy, discounts, delivery timelines, and category-specific expectations. The study demonstrates that even without complex machine learning pipelines, a well-designed Power BI dashboard can provide practical sentiment-oriented insights and decision support for e-commerce businesses. This paper highlights the importance of combining descriptive analytics with review intelligence to improve customer understanding, product strategy, and operational performance in digital marketplaces.

Keywords

Online Product Review Analysis, Power BI, Business Intelligence, E-commerce Analytics, Sentiment Analysis, Customer Reviews, Data Visualization, Product Ratings.

1. INTRODUCTION

Online shopping has become an essential part of modern consumer behavior. Platforms such as Amazon, Flipkart, and other digital marketplaces host thousands of products and receive large volumes of customer feedback in the form of ratings and textual reviews. These reviews significantly influence purchasing decisions because they act as a public source of product experience and perceived quality.

As a result, review data has become a valuable resource for businesses seeking to understand customer satisfaction, improve products, and strengthen market positioning.

Despite its importance, online review data is difficult to interpret manually due to its scale and diversity. A single product may receive hundreds or thousands of reviews, and these reviews often vary in tone, length, usefulness, and reliability. In addition, product performance is shaped by multiple business factors such as price, discount percentage, brand reputation, category competition, and delivery efficiency. Therefore, product review analysis should not be limited to textual opinion extraction alone; it should also consider the broader business context in which customer feedback is generated.

This internship project addresses this challenge by developing an analytical dashboard using Power BI to evaluate product review behavior and associated business variables. The project transforms raw e-commerce product data into interactive visual insights, allowing users to understand review volume, average ratings, category-wise performance, verified review distribution, customer segmentation, and delivery impact. The study also incorporates a simplified sentiment-oriented classification strategy based on ratings and review behavior to group customer opinion into positive, neutral, and negative patterns.

2. PROBLEM STATEMENT

E-commerce systems generate large amounts of product and review data every day, but most of this information remains underutilized because it is stored in raw tabular form. Manual interpretation of thousands of reviews, ratings, discounts, and product records is time-consuming and unsuitable for timely decision-making. Businesses require a practical framework that can clean, structure, and visually analyse this information to identify customer satisfaction patterns, product trends, and operational issues. The problem addressed in this paper is how to design an effective Power BI-based analytical system for online product review analysis that can support faster and more reliable decision-making in an e-commerce environment.

3. OBJECTIVES

- To analyse online product and customer review data from an e-commerce dataset.
- To clean and transform raw product data using Power Query in Power BI.
- To evaluate product performance using ratings, review counts, discounts, and categories.
- To classify customer feedback into positive, neutral, and negative sentiment-oriented groups.
- To analyse customer behavior by city and age group.
- To examine the impact of delivery days and order status on customer ratings.
- To create an interactive dashboard for business-oriented review analysis.

4. LITERATURE REVIEW

Research in opinion mining and sentiment analysis has established that online reviews are a major source of consumer intelligence. Pang and Lee provided one of the most influential foundational surveys on opinion mining, showing how sentiment-bearing text can be computationally analyzed for polarity and subjectivity. Bing Liu further expanded the field by presenting sentiment analysis as a practical framework for extracting opinions, aspects, and sentiment from user-generated content. In product-review contexts, Hu and Liu introduced an early and important approach for mining and summarizing customer reviews by identifying product features and associated opinion expressions.

For applied analytics projects, full-scale natural language processing is not always necessary. Business intelligence tools such as Power BI are increasingly used to transform structured and semi-structured datasets into interactive dashboards that support descriptive analytics, KPI monitoring, and decision support. Microsoft's official Power BI guidance emphasizes data transformation, semantic modeling, DAX measures, and report optimization as essential practices for effective analytics development. This makes Power BI suitable for internship-level research where practical usability and clear visual insight are often more valuable than highly complex modeling pipelines.

5. DATASET DESCRIPTION

The internship dataset used in this project contains product and review-related fields such as product name, product category, product price, discount percentage, product rating, number of reviews, brand name, product availability, city, age group, delivery days, verified review percentage, and order status. Each row represents a product or review-related record, while each column represents a measurable attribute that contributes to product performance analysis. The dataset supports multi-dimensional analysis by combining customer feedback with business variables such as pricing and logistics.

The project report indicates that the dashboard includes approximately 3892 total reviews, an average rating of 3.24, and a verified review share of 85%, suggesting that a large portion of the available customer feedback comes from comparatively reliable review sources. These figures make the dataset suitable for review-focused business analysis and dashboard-driven insight generation.

6. METHODOLOGY

The methodology followed a structured analytics workflow. First, the dataset was imported into Power BI from Excel/CSV format. Next, Power Query Editor was used for preprocessing tasks such as removing duplicate records, handling missing values, correcting data types, renaming columns, and filtering irrelevant fields. This step ensured that the dataset was clean, consistent, and ready for analysis.

After cleaning, the data was transformed into a report-ready format. Derived fields such as sentiment labels, rating groups, and discount bands were created where required. Power BI data modeling was then used to structure relationships and build dynamic measures. DAX was used to create KPIs such as total reviews, average rating, average delivery days, average return days, and verified review percentage. These calculations allowed the dashboard to respond dynamically to filters and slicers.

Finally, the dashboard was designed using cards, bar charts, pie charts, line charts, scatter plots, and detailed tables. Separate views were created for overall review insights, product-wise performance, customer-wise analysis, and delivery impact analysis. The methodology prioritised practical interpretability and business relevance over algorithmic complexity, making the system suitable for academic internship research.

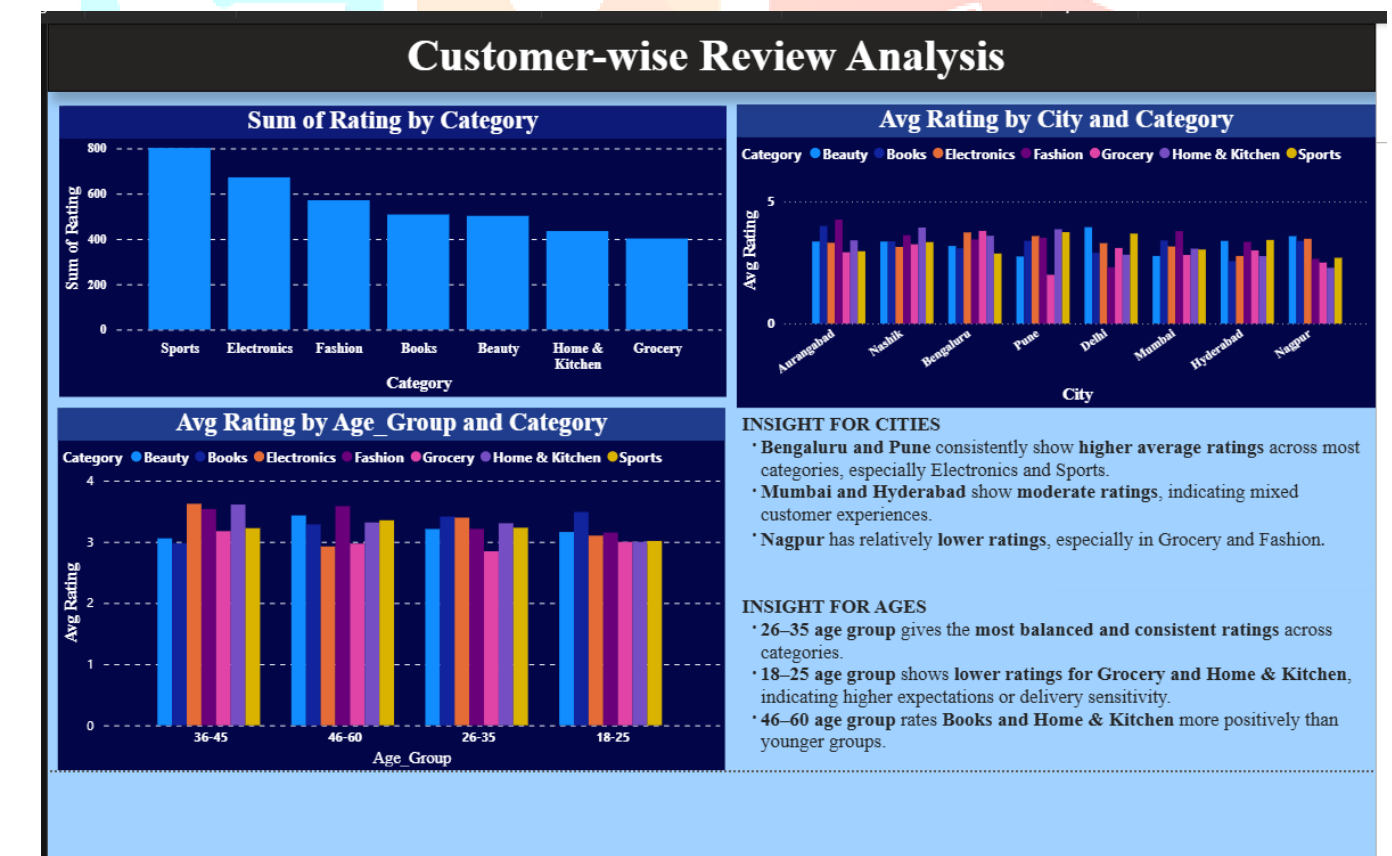
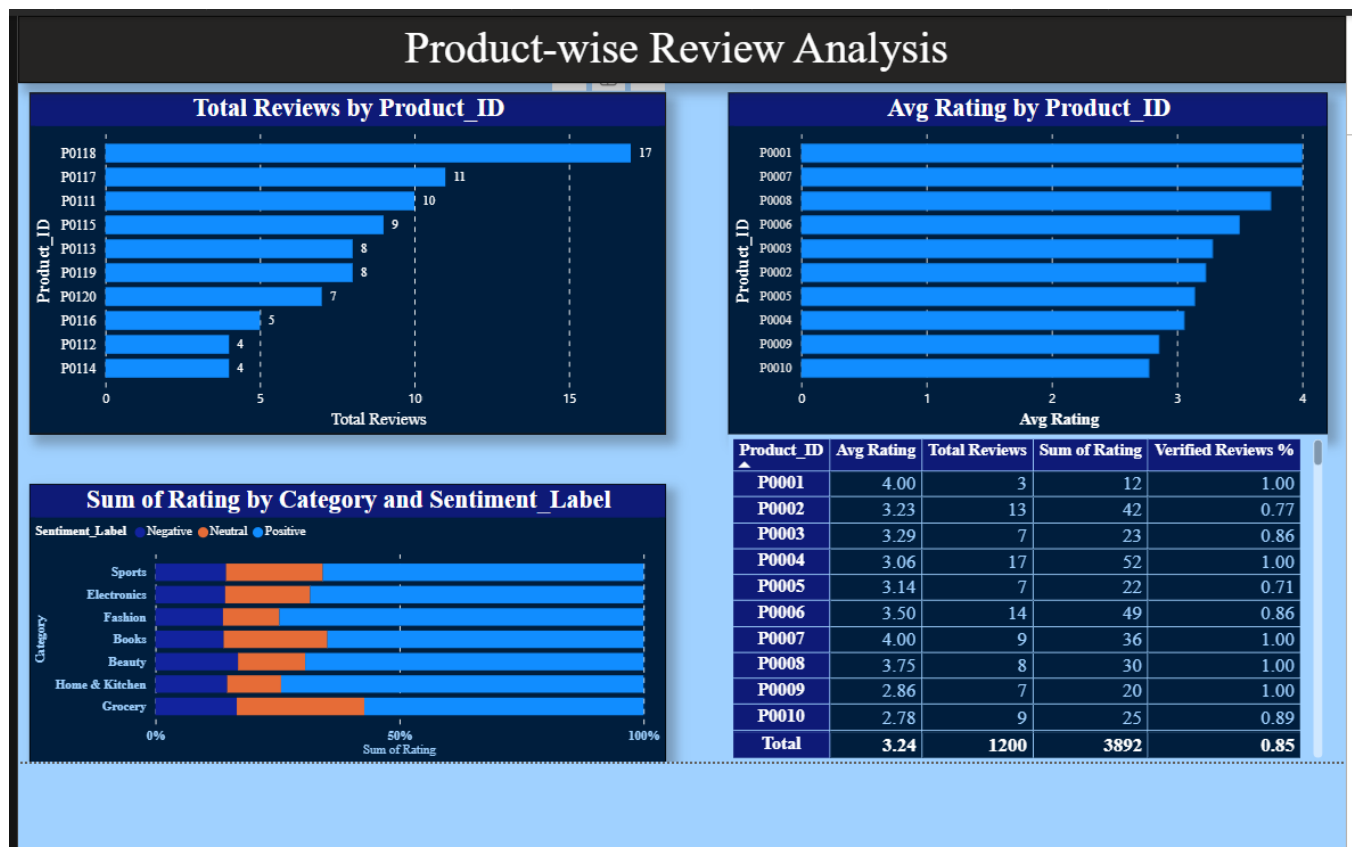
7. DASHBOARD MODULES

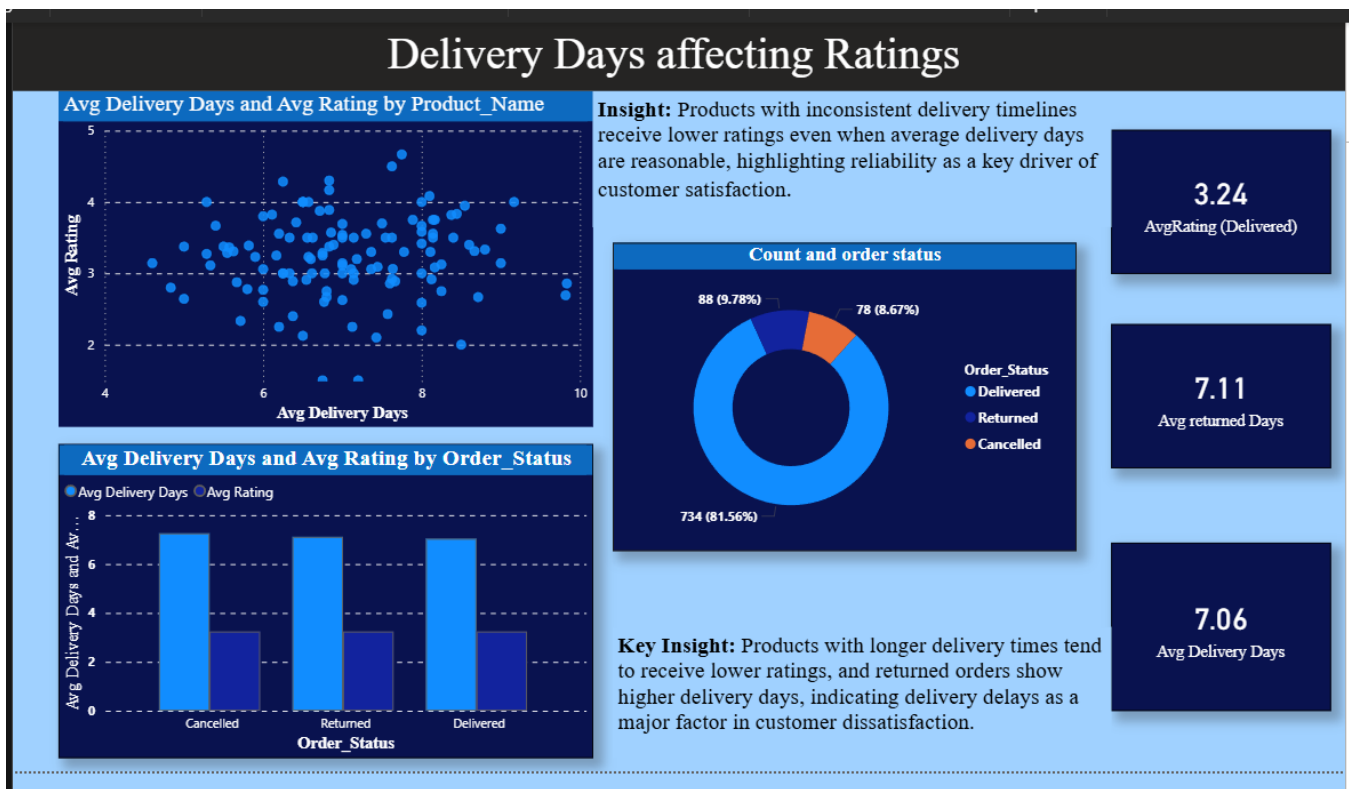
The dashboard developed in this project contains multiple analysis modules. The first module provides an overview of core KPIs such as total reviews, average rating, verified review percentage, and sentiment distribution. The second module focuses on product-wise analysis, highlighting products

with high review counts, high average ratings, and stronger category performance. The third module studies customer behavior across cities and age groups to identify demographic and regional differences in satisfaction patterns.

A dedicated delivery-performance module examines the relationship between delivery days and customer ratings. The internship report notes that average delivery days are around 7.06 and average return days are around 7.11, allowing the dashboard to explore whether longer delivery timelines correspond to lower ratings or higher return probabilities. This adds operational value beyond basic review analytics.

8. DASHBOARD SCREENSHOTS





9. RESULTS AND DISCUSSION

The project results show that the dataset contains approximately 3892 reviews and an average product rating of 3.24, which indicates a moderately positive customer environment. Verified reviews account for nearly 85% of the total review set, increasing the trustworthiness of the observed patterns.

Category-wise analysis shows that product segments such as sports and electronics demonstrate stronger engagement and higher rating concentration, while other categories show more mixed customer responses.

Customer-wise analysis suggests that city and age-group segmentation can reveal meaningful differences in product perception. The internship report highlights that cities such as Bengaluru and Pune tend to show relatively better average rating behavior in the dataset. Similarly, the 26–35 age group appears to provide balanced ratings across multiple categories, while younger groups may be more sensitive to service quality or delivery mismatches.

The delivery analysis is one of the most practical outcomes of the project. The dashboard indicates that longer or inconsistent delivery times can contribute to slightly lower ratings and a higher chance of returns or cancellations. This finding demonstrates that customer satisfaction in online shopping is influenced not only by product quality but also by fulfillment performance. Therefore, product review analysis should be interpreted as a combination of customer sentiment, product quality signals, and operational service quality.

10. ADVANTAGES AND LIMITATIONS

The proposed system offers several advantages: it reduces manual analysis effort, provides interactive and visually understandable insights, supports multi-dimensional filtering, and helps identify patterns in customer satisfaction, category performance, and delivery efficiency. Because the dashboard is built in Power BI, it is easy to extend and suitable for non-technical users such as managers, faculty reviewers, or business analysts.

However, the study also has limitations. The sentiment-oriented classification is simplified and primarily based on ratings and review behavior rather than advanced natural language processing. As a result, it cannot fully capture sarcasm, context, or aspect-level sentiment. In addition, the dashboard is based on historical dataset snapshots rather than live API integration, which limits real-time applicability. These limitations do not reduce the practical value of the project, but they indicate clear directions for future enhancement.

11. CONCLUSION

This paper presented an internship-based analytical study on Online Product Review Analysis Using Power BI. The project successfully transformed raw product and review data into a meaningful business intelligence dashboard that supports product evaluation, customer understanding, and operational analysis. Through data cleaning, transformation, modeling, DAX-based KPI generation, and dashboard development, the project demonstrated how Power BI can be used to analyse ratings, review counts, discounts, customer segments, and delivery impact in a structured and interactive manner.

The findings confirm that review analytics becomes more useful when combined with business variables such as category, pricing, and delivery performance. The developed system provides a practical, publishable example of how academic internship work can be converted into a relevant e-commerce analytics paper. It is especially suitable for educational and entry-level research environments where clarity, usability, and business relevance are essential.

12. FUTURE SCOPE

- Use advanced NLP models for true text-based sentiment classification.
- Implement aspect-based sentiment analysis for features such as quality, packaging, delivery, and durability.
- Integrate live data from e-commerce APIs or automated scraping pipelines.
- Add predictive analytics for product success, return risk, and rating trends.
- Build alert systems for sudden drops in customer satisfaction.

REFERENCES

1. Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135. DOI: 10.1561/1500000011
2. Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167. DOI: 10.2200/S00416ED1V01Y201204HLT016
3. Hu, M., & Liu, B. (2004). Mining and Summarizing Customer Reviews. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. DOI: 10.1145/1014052.1014073
4. Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. DOI: 10.1609/icwsm.v8i1.14550
5. Microsoft. (2025). Power BI guidance documentation. Microsoft Learn. <https://learn.microsoft.com/power-bi/guidance/overview>
6. Microsoft. (2025). Transform, shape, and model data in Power BI. Microsoft Learn. <https://learn.microsoft.com/power-bi/transform-model/>
7. Microsoft. (2024). DAX overview. Microsoft Learn. <https://learn.microsoft.com/dax/dax-overview>
8. Microsoft. (2025). Semantic models in the Power BI service. Microsoft Learn. <https://learn.microsoft.com/power-bi/connect-data/service-datasets-understand>