



“Online Product Review Analysis Using Power BI Software”

¹Prof. Nagesh Patil , ² Diksha Satpute

¹ Faculty of computer engineering Vidya Prasarini Sabhas college of Engineering and technology Lonavala,

² Student Computer Engineering Vidya Prasarini Sabhas college of Engineering and Technology , Pune , India

Abstract : In today’s digital marketplace, people rarely buy anything without checking reviews first. But the problem is there are **too many reviews**, and not all of them are easy to interpret. This project focuses on analyzing online product reviews using data analytics techniques to extract meaningful insights such as customer sentiment, common issues, and overall product perception.

The goal of this study is to transform raw, unstructured review data into **clear, visual, and actionable information** using tools like Power BI and basic natural language processing techniques. This helps both customers make better decisions and businesses understand their products more effectively.

After data pre-processing, **Exploratory Data Analysis (EDA)** is performed to understand the structure, distribution, and behaviour of the data. EDA helps in identifying trends, relationships, outliers, and patterns within the dataset. Statistical methods such as mean, median, variance, and correlation are used along with graphical representations like bar charts, line graphs, histograms, scatter plots, and heatmaps. These visualizations make complex data easier to interpret and help in gaining meaningful insights at a glance. The project further demonstrates how Python can be effectively used for data analysis and visualization to support decision-making. By analysing the dataset, key factors influencing the data are identified, and important trends are highlighted. The insights obtained from the analysis can be used to make data-driven decisions, improve strategies, and predict future outcomes.

In conclusion, this project emphasizes the importance of data analytics in extracting useful information from large datasets.

Keywords :- Online Product Reviews, Sentiment Analysis, Data Analytics, NLP, Text Mining, Data Visualization, Power BI, Customer Feedback Analysis

Introduction : Online shopping platforms like Amazon and Flipkart generate massive amounts of customer feedback daily. While this feedback is valuable, manually reading thousands of reviews is impractical. This data includes information about products, customer preferences, ratings, pricing, sales trends, and customer reviews. Analysing this data helps companies understand consumer behaviour, improve their products, and make better business decisions. The objective of the **Online Product Analysis project** was to analyse product data collected from an online marketplace and extract useful insights from it. By analysing product performance, ratings, pricing patterns, and category-wise trends, businesses can identify which products perform well in the market and which ones need improvement. To perform this analysis effectively, we used **Power BI**, a powerful data visualization and business intelligence tool developed by **Microsoft**. Power BI allows users to transform raw data into interactive dashboards and visual reports that help in understanding complex data easily. During the internship, this project provided practical exposure to data analytics concepts such as **data cleaning, data transformation, data visualization, and business insight generation.**

“What if we could automatically understand what people are saying about a product?”

By analyzing reviews, we can:

- Identify whether customers feel **positive, negative, or neutral**
- Detect frequently mentioned issues (like battery, delivery, quality)
- Compare product performance based on user opinions This makes review analysis a powerful tool in **data analytics and business intelligence.**

Objectives One of the primary objectives of the project was to analyse online product sales data. Online marketplaces generate large volumes of data related to product listings, sales performance, customer interactions, and purchasing behaviour. By analysing this data, it becomes possible to identify patterns such as which products are selling more frequently, which categories are gaining popularity, and which products may require improvement or better marketing strategies. Understanding sales data helps businesses track product demand and make decisions related to inventory management and marketing.

- To collect and analyze product review data
- To classify reviews into **positive, negative, and neutral sentiments**
- To identify common keywords and trends in customer feedback
- To create an interactive dashboard for visualization
- To provide insights that help in decision-making

Methodology

The methodology of this study consists of four major stages: data collection, data preprocessing, sentiment analysis, and data visualization.

1. Data Collection

Product review data was collected from online sources such as e-commerce platforms or publicly available datasets (e.g., Kaggle). The collected dataset consists of textual customer reviews along with associated ratings, which serve as the primary input for analysis.

2. Data Cleaning and Preprocessing

The collected data was preprocessed to improve its quality and suitability for analysis. This step included removing duplicate entries, handling missing or null values, and cleaning the text data by eliminating unwanted symbols, punctuation, and irrelevant words. Basic natural language processing techniques such as tokenization and stopword removal were applied to prepare the text for sentiment analysis.

3. Sentiment Analysis

Sentiment analysis was performed to classify customer reviews into positive, negative, or neutral categories. This was achieved using a keyword-based approach and/or predefined sentiment analysis libraries. The analysis helped in identifying the overall opinion of customers and detecting patterns in user feedback.

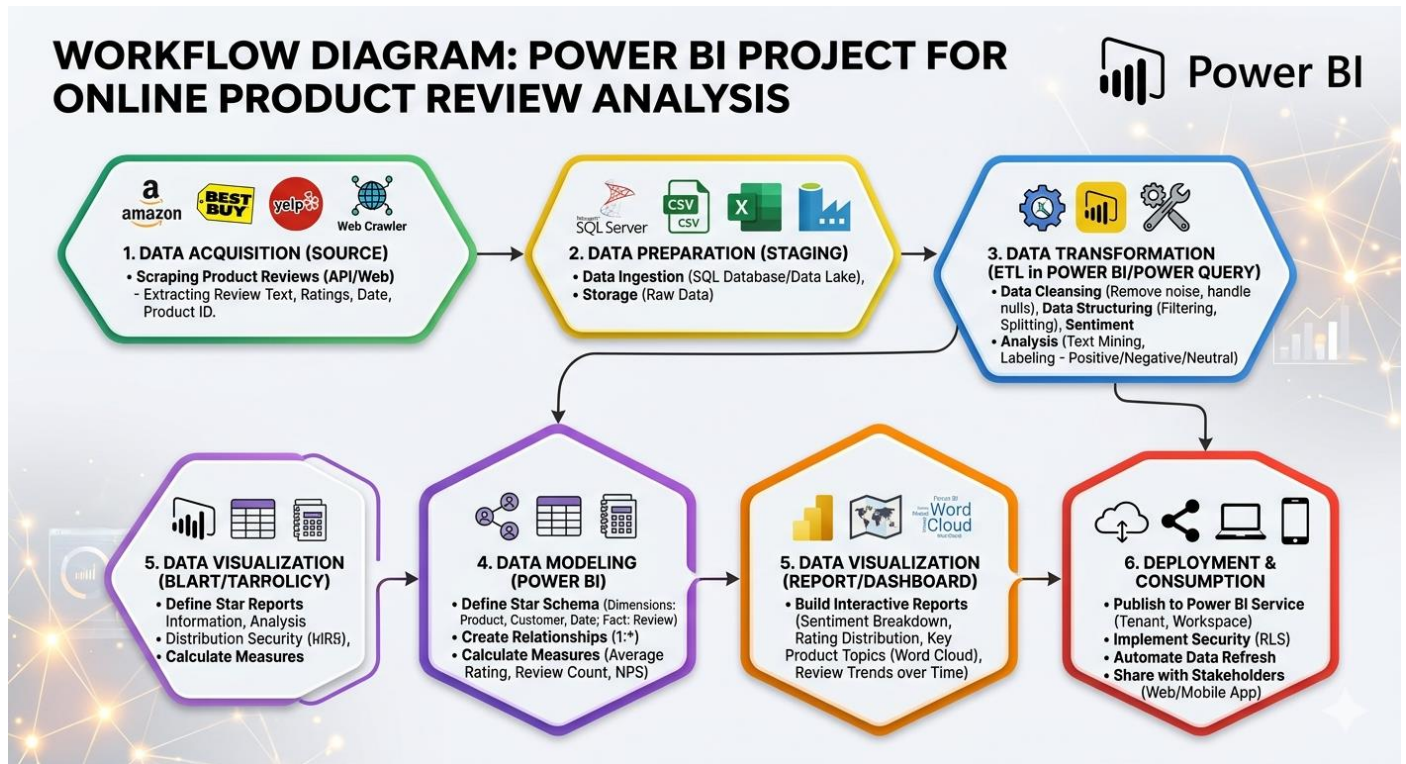
4. Data Visualization

The processed data was visualized using Power BI to generate interactive dashboards. Various visual elements such as bar charts, pie charts, and word frequency visuals were used to represent sentiment distribution, commonly used keywords, and rating trends. These visualizations enabled easier interpretation of the results and supported data-driven decision-making.

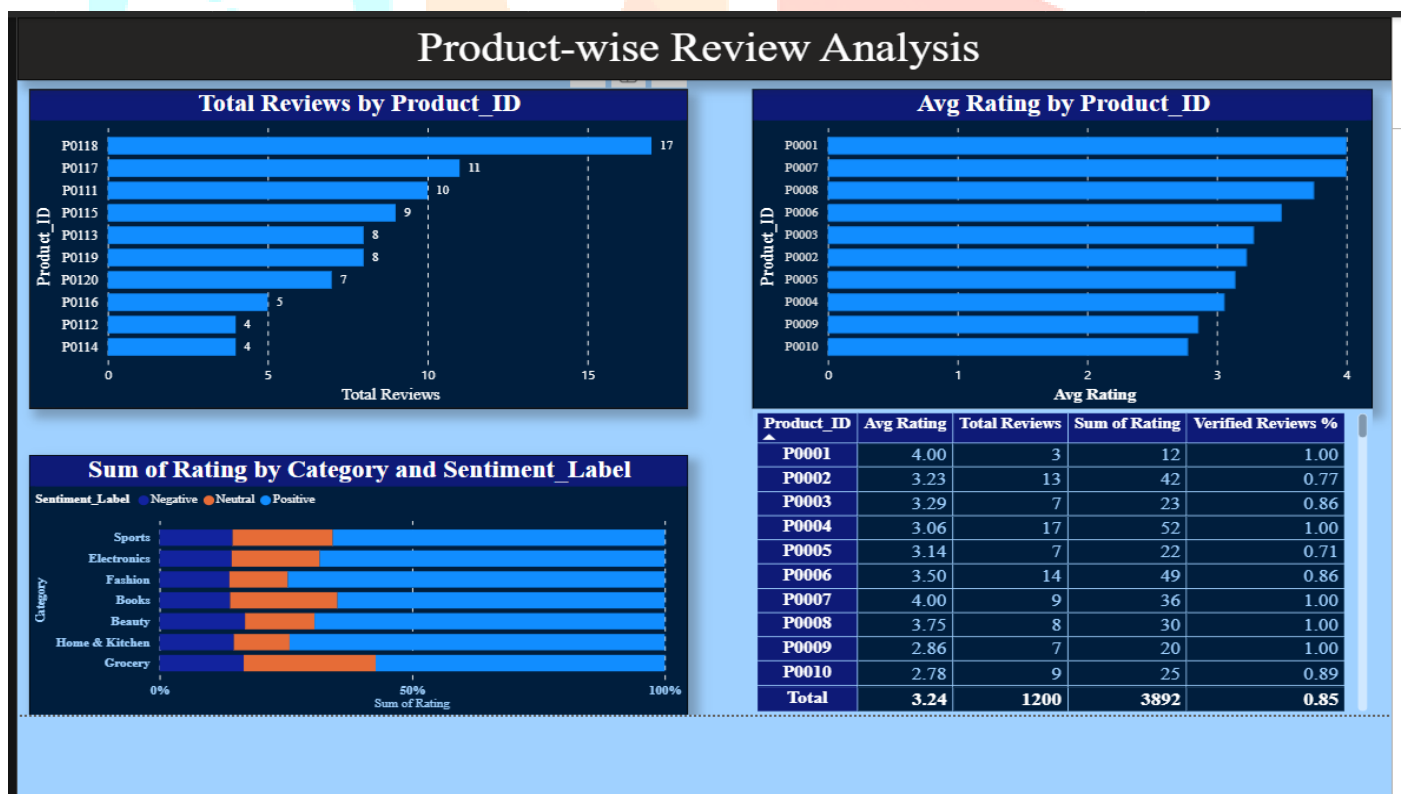
Tools and Technologies

The implementation of this project involved the use of multiple tools and technologies. Python was optionally used for data preprocessing and basic text analysis. Microsoft Power BI served as the primary tool for data visualization and dashboard creation. The dataset was managed in Excel or CSV format. Additionally, basic natural language processing (NLP) techniques were employed for text cleaning and sentiment classification.

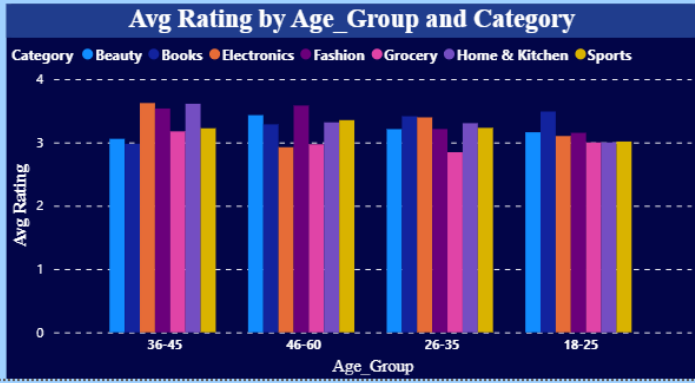
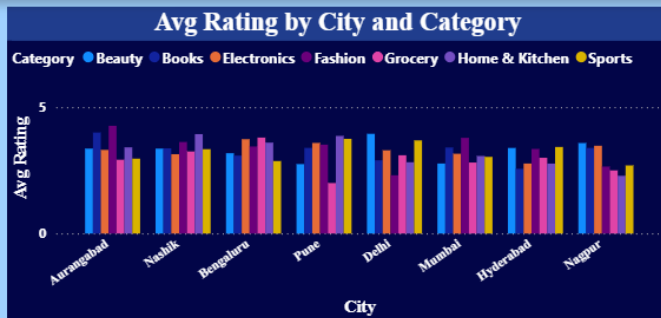
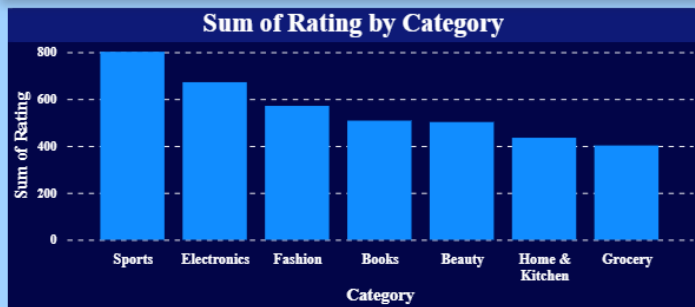
Work flow Diagram



Dashboard Screenshots



Customer-wise Review Analysis



INSIGHT FOR CITIES

- Bengaluru and Pune consistently show higher average ratings across most categories, especially Electronics and Sports.
- Mumbai and Hyderabad show moderate ratings, indicating mixed customer experiences.
- Nagpur has relatively lower ratings, especially in Grocery and Fashion.

INSIGHT FOR AGES

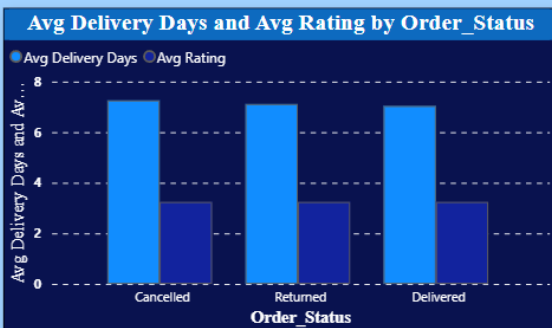
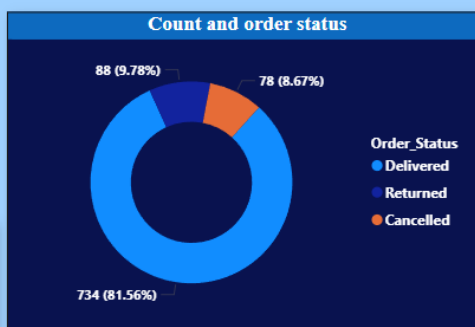
- 26-35 age group gives the most balanced and consistent ratings across categories.
- 18-25 age group shows lower ratings for Grocery and Home & Kitchen, indicating higher expectations or delivery sensitivity.
- 46-60 age group rates Books and Home & Kitchen more positively than younger groups.



Delivery Days affecting Ratings



Insight: Products with inconsistent delivery timelines receive lower ratings even when average delivery days are reasonable, highlighting reliability as a key driver of customer satisfaction.



Key Insight: Products with longer delivery times tend to receive lower ratings, and returned orders show higher delivery days, indicating delivery delays as a major factor in customer dissatisfaction.

3.24
AvgRating (Delivered)

7.11
Avg returned Days

7.06
Avg Delivery Days

Results & Findings

- Majority of reviews were found to be **positive**, indicating customer satisfaction
- Negative reviews mostly focused on specific issues (e.g., delivery delays, product durability)
- Word frequency analysis highlighted key features customers care about

The dashboard helped convert raw data into **easy-to-understand visuals**, making analysis faster and clearer.

Conclusion

This project demonstrates how data analytics can simplify decision-making by converting large volumes of text data into meaningful insights. Even with basic tools, it is possible to uncover patterns in customer behavior and improve product understanding.

This project on **Data Analytics Using Python** successfully demonstrates how raw data can be transformed into meaningful and actionable insights through systematic analysis. In the present data-driven world, organizations generate large volumes of data, but the true value of this data can only be realized when it is properly analyzed and interpreted. This project highlights the importance of data analytics in converting complex datasets into understandable information that supports effective decision-making.

Future improvements could include:

- Advanced machine learning models
- Real-time data integration
- More accurate sentiment detection

References

- [1] Kaggle Dataset – Amazon Product Reviews
- [2] Microsoft Power BI Documentation
- [3] Bing Liu, “Sentiment Analysis and Opinion Mining”
- [4] Research article on NLP techniques (Google Scholar)