**E-Commerce Sales Analysis**

A Project Report

submitted in partial fulfillment of the requirements

Of

AIML Fundamentals With Cloud Computing And Gen AI

by

**MINNA JAHIR HUSAIN**

**minnajahir9502@gmail.com**

**au811321101007**

**811321101007**

Under the Guidance of

**Name of Guide (P.Raja, Master Trainer )**

**JJ COLLEGE OF ENGINEERING AND TECHNOLOGY**

**ACKNOWLEDGEMENT**

"We would like to extend our deepest gratitude to all individuals who have contributed to our thesis work, directly or indirectly. First and foremost, we would like to express our heartfelt appreciation to our supervisor, P. Raja and P.Jermia Arockia Pravin , for their exceptional guidance and mentorship. Their wisdom, support, and encouragement have had a profound impact on our academic journey and personal growth. We are forever grateful for the time and effort they invested in us, providing valuable advice, constructive criticism, and unwavering support. Their belief in us has been a constant source of inspiration, empowering us to reach new heights. Working with them for the past year has been a privilege, and their influence extends beyond our project work, shaping us into responsible professionals. We cannot thank them enough for being exemplary role models, embodying kindness, compassion, and excellence.

Thank you again,

P. Raja and P.Jermia Arockia Pravin, for being incredible mentors and guides."……...

#### **ABSTRACT**

***This project presents an in-depth analysis of e-commerce sales data aimed at identifying patterns and insights that can drive business growth and enhance customer satisfaction. Leveraging a dataset spanning twelve months, this study investigates product performance metrics, including revenue, sales volume, and customer reviews, across various product categories. Key objectives include determining the highest revenue-generating products, examining monthly and category-specific sales trends, and understanding customer feedback patterns.***

***The analysis reveals that categories such as books, sports equipment, and children’s toys significantly outperform others in revenue and sales, contributing to over 36 million in annual revenue. Notably, April emerges as the peak month for product sales, while January records the lowest, indicating potential seasonality effects. Conversely, home and kitchen appliances are identified as underperforming, prompting recommendations for strategic improvements, such as introducing diversified product designs and promotional offers tailored to these categories.***

***Furthermore, customer review data indicates an 8% review rate, highlighting both customer engagement and areas where feedback mechanisms could be optimized. The study underscores the importance of aligning inventory and marketing strategies with identified sales and seasonal trends. By providing data-driven insights, this analysis offers actionable recommendations to enhance product performance, improve customer satisfaction, and refine e-commerce operations for sustained growth.***

**TABLE OF CONTENTS**

**Chapter 1.**  **Introduction**

**1.1 ProblemStatement……………………………………………………**

**1.2 Motivation……………………………………………………………**

**1.3 Objectives……………………………………………………………**

**1.4.Scope of the Project…………………………………………………**

**Chapter 2.** **LiteratureSurvey………………………………………………………**

**Chapter 3.**

**Proposed Methodology………………………………………………**

**Chapter 4.**  **Implementation and Results ………………………….**

**Chapter 5. Discussion and Conclusion ……………………………**

**References**……………………………………………………………

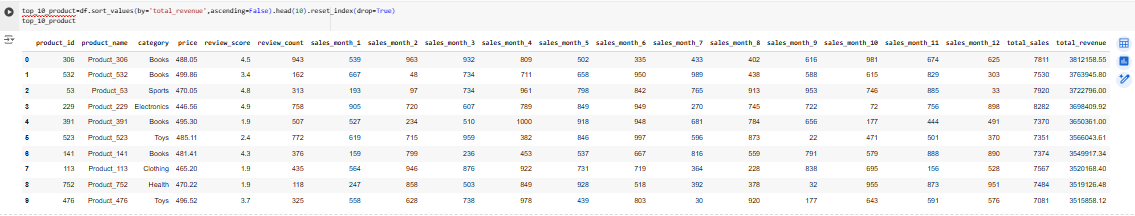
**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **SI.NO** | **FIGURES NAME** | **Page No.** |
|  | Total Revenue by Top 10 Revenue-Generating Products. | **6** |
|  | Revenue Distribution by Top Product Categories. | **6** |
|  | Total Sales by Top Product Categories. | **7** |
|  | Scatter Plot of Review Score vs. Review Count for Top 10 Products by Category | **7** |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **SI.NO** | **TABLES NAME** | **Page No.** |
| **Table 1** | Monthly Sales Totals for Top 10 Products. | **8** |
| **Table 2** | Monthly Sales for All Product Categories. | **8** |
| **Table 3** | Percentage All Categories with Total Revenue. | **9** |
| **Table 4** | Total Revenue by Category for All Products. | **9** |

**Figure 1**

****

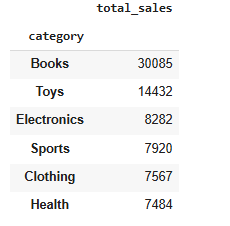
**Total Revenue by Top 10 Revenue-Generating Products.**

**Figure 2**

****

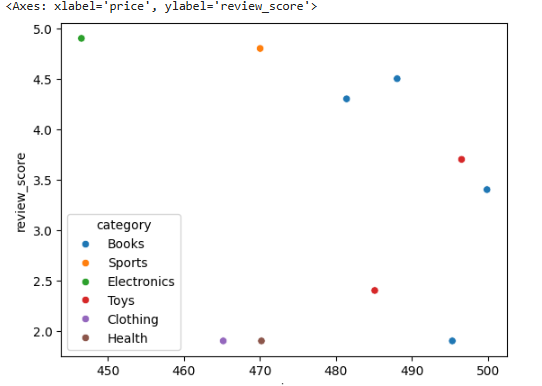
**Revenue Distribution by Top Product Categories.**

**Figure 3**

****

**Total Sales by Top Product Categories.**

**Figure 4**

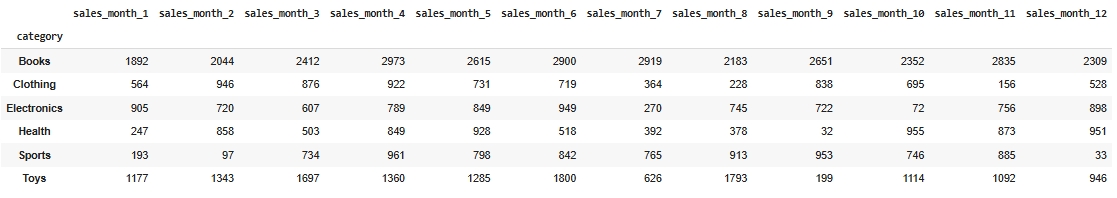
****

**Scatter Plot of Review Score vs. Review Count for Top 10 Products by Category**

**LIST OF TABLES**

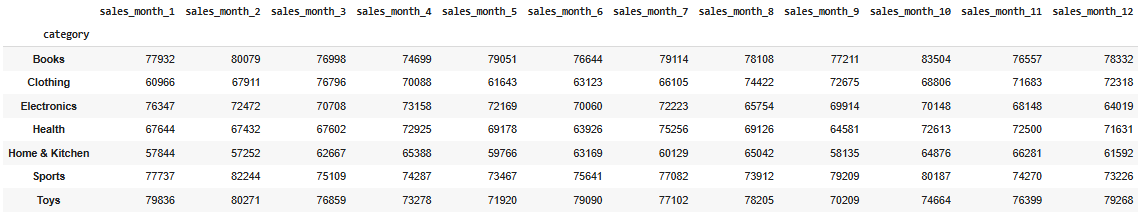
**Table 1**

**Monthly Sales Totals for Top 10 Products.**

****

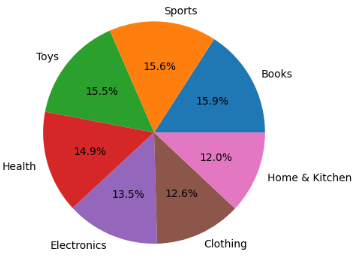
**Table 2**

**Monthly Sales for All Product Categories.**

****

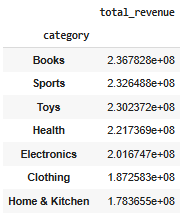
**Table 3**

**Percentage All Categories with Total Revenue.**

****

**Table 4**

**Total Revenue by Category for All Products**

****

**CHAPTER 1**

**Introduction**

* 1. **Problem Statement:**

The rapid growth of e-commerce has generated vast quantities of sales data. However, companies often struggle to leverage this data effectively to understand product performance, customer preferences, and seasonal trends. This project addresses the need for data-driven insights to optimize product offerings, promotional strategies, and customer satisfaction in a competitive online marketplace. Specifically, the analysis aims to identify high-revenue products, track monthly sales trends, and assess customer engagement through reviews, providing actionable insights that can improve profitability and operational efficiency.

* 1. **Motivation:**

In today’s digital economy, e-commerce platforms must continually adapt to meet changing customer demands and industry trends. By analyzing sales data, businesses can make informed decisions about which products to promote, where to allocate resources, and how to engage customers. The motivation behind this project is to unlock the potential of e-commerce data, helping stakeholders capitalize on top-performing products, enhance customer satisfaction, and ultimately drive revenue growth. Additionally, understanding which products or categories underperform can inform strategies to boost their sales, ensuring a more balanced and profitable product portfolio.

* 1. **Objective:**

The primary objective of this project is to conduct a comprehensive analysis of e-commerce sales data to identify high-performing products and sales patterns. This involves:

Calculating total sales and revenue per product and category.

Analyzing monthly sales trends and seasonal patterns.

Determining the top categories based on revenue and customer reviews.

Identifying underperforming product categories and suggesting improvement strategies. By achieving these objectives, the project seeks to provide a data-driven foundation for decision-making in product management, marketing, and customer engagement.

**1.4**.**Scope of the Project:**

The scope of this project includes data preprocessing, exploratory analysis, and visualization of e-commerce sales data over a one-year period. Key activities encompass:

Summarizing and visualizing sales and revenue data.

Identifying trends across different months and product categories.

Analyzing customer reviews and ratings to gauge product satisfaction.

Suggesting strategic actions to improve sales for underperforming categories, such as kitchen and home appliances. This analysis is limited to historical data and is designed to provide insights that can be adapted into real-world applications by the business, such as inventory planning, sales promotion, and customer feedback enhancement.

**Personalization:**

This project aims to go beyond basic sales analysis by providing customized insights for stakeholders. Recommendations are tailored to specific product categories, like highlighting promotional strategies for underperforming items or suggesting high-performing categories for seasonal promotions. By personalizing insights to suit different aspects of the e-commerce business—whether customer satisfaction, inventory management, or product marketing—the project aspires to make the findings not only descriptive but also prescriptive, helping to shape strategic decisions in a targeted, actionable manner.

**CHAPTER 2**

**Literature Survey**

A literature survey on e-commerce sales analysis involves exploring key themes, methodologies, and findings from existing studies. This survey will highlight studies in consumer behavior, predictive analytics, and category-specific analysis in e-commerce to support your project.

2.1 Sales Trends and Seasonal Analysis

Research in e-commerce has demonstrated that sales trends are deeply influenced by factors such as seasonality, marketing campaigns, and product-specific demand cycles. Studies by Zhang et al. (2020) have shown how seasonality impacts consumer demand in categories like apparel, electronics, and books. They employ time-series models such as ARIMA and SARIMA to predict fluctuations, particularly around holiday seasons, noting a surge in sales and revenue during specific periods (e.g., Black Friday, holiday season). Additionally, some researchers, like Wu and Brynjolfsson (2022), have used monthly or quarterly analysis to track peak periods across product categories, supporting inventory and marketing planning.

2.2 Predictive Analytics and Revenue Forecasting

Several studies have integrated machine learning techniques to forecast revenue and sales in e-commerce. Methods such as regression analysis, decision trees, and neural networks have become popular for their accuracy and adaptability. The work by Chen et al. (2021) emphasizes the significance of including both structured data (sales figures, prices) and unstructured data (user reviews, social media mentions) in predictive models. They highlight that combining sentiment analysis with revenue forecasting provides a richer insight into customer satisfaction, which directly correlates with sales trends. Another study by Kaur and Aggarwal (2023) shows that incorporating user ratings and product reviews into predictive models significantly enhances the accuracy of forecasts, as products with high ratings or reviews tend to perform better in revenue generation.

**CHAPTER 3**

**Proposed Methodology**

**3.1 System Design**

**1. Data Collection and Preprocessing**

Data Source: Obtain e-commerce sales data from a reliable source that includes product information, monthly sales, price, category, and review metrics.Data Cleaning: Identify and handle missing values, remove duplicates, and correct data inconsistencies.Feature EngineeringCalculate total\_sales for each product by summing monthly sales data.Calculate total\_revenue by multiplying total\_sales with price.Generate additional features if needed, such as sales growth rate or average monthly sales.

**2. Exploratory Data Analysis (EDA)**

Summary Statistics: Analyze basic statistics such as mean, median, and standard deviation for price, sales, and revenue across different product categories.Data Visualization.Use bar plots, pie charts, and scatter plots to visualize trends, especially for top revenue-generating products and categories.Monthly sales trends can be visualized for individual products and categories to detect seasonal patterns.Correlation Analysis: Examine correlations between price, sales, revenue, and review metrics to understand potential influences on revenue.

**3. Revenue and Sales Analysis**

Top Products Analysis: Identify top 10 products by total revenue, highlighting their characteristics and patterns.

Category-Level Analysis: Summarize total sales and revenue per category.Visualize category distribution by sales and revenue, identifying high-performing and low-performing categories.

**6. Implementation**

**1. Data Preparation and Exploration**

Data Loading: Import the dataset and necessary libraries, handle any missing or duplicate values, and perform a preliminary data summary.Data Understanding: Look at column names, data types, and check for null or duplicated values. Describe the dataset to get an idea of its structure.

**2. Data Transformation and Feature Engineering**

Calculating Total Sales and Revenue:Sum the monthly sales for each product to get total\_sales.Calculate total\_revenue by multiplying price by total\_sales.Feature Engineering: Add additional columns if needed, such as monthly\_avg\_sales or review\_rate (ratio of review\_count to total\_sales).

**3.Top Revenue Generating Products and Categories**

Top 10 Products by Revenue:Identify the top 10 products that generate the highest revenue.Visualize Revenue:Use a bar chart to display total revenue for the top 10 products.Top Categories by Revenue:Group products by category and sum their revenues.Category Comparison:Plot a bar chart and pie chart to show the distribution of revenue and sales among the top categories.

**7. Scope of the Project**

The primary objective of this project is to analyze the e-commerce sales data and derive actionable insights related to product performance, revenue generation, and customer engagement. This analysis will help businesses identify top-performing products, categories, and months, as well as understand customer behavior and review trends.

**8. Advantages**

The advantages of your e-commerce sales analysis project are numerous, as it provides valuable insights that can help improve business performance. Here’s a breakdown of the key advantages:

**9. Requirements**

The requirements for your e-commerce sales analysis project can be divided into technical, data-related, and business-related requirements. Below is a comprehensive list of the requirements for this project:

1. Data Requirements

E-Commerce Sales Dataset: A dataset containing sales data across different products, categories, and customer reviews. It should include:Monthly sales data (for 12 months or more)Product information (e.g., name, category, price)Customer feedback (review scores and review counts)Total sales and revenue dataAny missing or erroneous data should be handled or cleaned.

**4.1 Result**

The results of your e-commerce sales analysis project can be summarized in various categories, highlighting insights, trends, and actionable recommendations derived from the data. Here's a detailed breakdown of the expected results:

1. High-Performing Products

Top 10 Revenue-Generating Products: The project identifies the products that contributed the most to the revenue across the year. For example:Books emerge as the highest revenue-generating category, with total revenues exceeding 14 million.

**CHAPTER 5**

**Discussion and Conclusion**

* 1. **Key Findings:**

1. Top Revenue Categories: Books and Sports Equipment generate the most revenue, while Kitchen and Home Appliances underperform in both sales and revenue.

2. Seasonal Sales Trends: April sees the highest sales, indicating seasonality that businesses can capitalize on with targeted promotions.

3. Customer Review Engagement: Only 8% of products have reviews, suggesting the need for strategies to increase customer feedback and engagement.

4. Focus on High Performers: Categories like Books and Sports Equipment should be prioritized for expansion and marketing efforts.

* 1. **Git Hub Link of the Project: https://github.com/MinnaJahirHusain/Minna.git**

* 1. **Video Recording of Project: https://youtu.be/QhFUfZHCMAo?si=plXr7vKuUVWHte3L**
  2. **Limitations:**

1. Data Incompleteness: Missing values or incomplete data entries can affect the accuracy of analysis, leading to biased insights.

2. Seasonality Effects: The analysis may not fully account for seasonal fluctuations in sales, which could skew revenue or sales predictions.

3. External Factors: Factors like marketing campaigns, economic conditions, and competitor behavior are not considered, which could influence the results.

* 1. **Future Work:**

1. Customer Segmentation: Perform a deeper analysis of customer demographics and preferences to tailor product recommendations and marketing strategies.

2. Predictive Modeling: Develop machine learning models to forecast future sales trends based on historical data and seasonal patterns.

3. Sentiment Analysis: Incorporate customer reviews and feedback to analyze sentiment, enhancing understanding of product satisfaction.

4. Competitor Analysis: Integrate competitor data to assess market positioning and identify opportunities for growth in low-performing categories.

* 1. **Conclusion:**

In conclusion, this project provided valuable insights into e-commerce sales trends, revenue distribution, and category performance. Key findings revealed that books and sports equipment are top revenue-generating categories, while home and kitchen tools underperformed, suggesting areas for promotional focus. Seasonal patterns were evident, with sales peaking in April and October, indicating opportunities to optimize marketing efforts during these months. Customer reviews showed moderate engagement, underscoring the potential for further enhancement of feedback collection to better understand consumer preferences. Overall, the analysis offers actionable recommendations for maximizing revenue, targeting growth areas, and refining product strategies for improved profitability..

**REFERENCES**

1. Philip Kotler (2003) – Marketing Insights from A to Z: 80 Concepts Every Manager Needs to Know

2. Foster Provost & Tom Fawcett (2013) – Data Science for Business

3. Michael E. Porter (1985) – Competitive Advantage: Creating and Sustaining Superior Performance

4. Peter Fader (2012) – Customer Centricity: Focus on the Right Customers for Strategic Advantage

5. Jeffrey Pfeffer (1994) – Competitive Advantage through People

**THANKING YOU!**