

**Investigating the Potential Use of Data Analysis in E-commerce for Improved Decision-Making**

**Student Name:** Ha Phuc Hung

**Student ID:** @00740880

**Module:** Data Analysis for Business

**Module Leader:** Dave Lowe

**Date:** 15 April 2025

**Word Count:** 2050 words

# 

1. Abstract

This report examines how data analysis can support better decision-making within an organisation. It explores the differences between business intelligence and data analytics, discusses key analytical methods and highlights potential issues like data bias. By using an e-commerce retail store dataset, the report uncovers insights that can help drive business strategies. Finally, it offers recommendations on how the organisation can leverage data more effectively to improve their future outcomes.

Investigating the Potential Use of Data Analysis in E-commerce for Improved Decision-Making

Table of Contents

[1. Abstract 1](#_Toc198156705)

[2. Introduction 3](#_Toc198156706)

[2.1. Background 3](#_Toc198156707)

[2.2. The importance of data 3](#_Toc198156708)

[2.3. What is data Analysis? 3](#_Toc198156709)

[3. Data Analysis 4](#_Toc198156710)

[3.1 Business Intelligence and Data Analytics 4](#_Toc198156711)

[4. Findings 5](#_Toc198156712)

[4.1 Data Structuring and Preparation 5](#_Toc198156713)

[4.2 Sales Revenue Distribution and Category Insights 5](#_Toc198156714)

[4.2.1 Sales Performance by Time and Region 6](#_Toc198156715)

[4.2.2 Product Category Trends and Seasonal Impact 7](#_Toc198156716)

[4.3 Customer Demographics and Sales Analysis 8](#_Toc198156717)

[4.3.1. Sales by Gender and Age Bracket 8](#_Toc198156718)

[4.3.2 Target Customer 9](#_Toc198156719)

[4.4 Correlation analysis 10](#_Toc198156720)

[4.5 Analysis of Data Bias in Correlation 11](#_Toc198156721)

[5. Conclusion and Recommendation 12](#_Toc198156722)

[References 13](#_Toc198156723)

2. Introduction

2.1. Background

This report is based on a mid-sized e-commerce retail online store that operates in multiple regions by selling a wide range of products through online platforms. Comprehensively analysing this dataset can help improve business performance by enhancing customers insights, optimising pricing strategies, and increasing sales efficiency.

2.2. The importance of data

In today’s digital age, data has become a critical asset for businesses across all industries by driving informed decision-making and strategic planning. Effective data analysis allows organisations to identify new opportunities, optimise operations, and gain a competitive edge in their respective markets. According to Provost and Fawcett (2013), data science encompasses principles, processes, and techniques for understanding phenomena via data analysis, with the goal of improving decision-making.

2.3. What is data Analysis?

Data Analysis is the systematic process of collecting, cleaning, transforming, describing, modelling, and interpreting data to discover useful insights, draw conclusions, and support decision-making (Eldridge, 2025). This process is crucial in both scientific research and business, where the demand for data-driven decision-making has grown in recent years. It enables organisations to interpret trends, predict outcomes accurately, and make evidence-based decisions.

3. Data Analysis

3.1 Business Intelligence and Data Analytics

Business Intelligence (BI) and Data Analytics (DA) are both essential for data-driven decision-making but serve different purposes. According to Bajaj (2025), BI primarily focuses on descriptive analytics by using historical and current data to generate reports, dashboards, and visualisations that help organisations monitor performance and support strategic decision-making. It allows analysts to explore valuable insights from structured data, assisting businesses in enhancing overall business performance and long-term growth.

Data Analytics involves examining raw data to identify patterns, correlations and trends that drive informed decision-making. Unlike BI, which mainly relies on predefined queries and structured reporting for historical analysis, DA employs predictive and prescriptive techniques, including machine learning and statistical models, to forecast future trends and drive precise decision-making (Bajaj, 2025).

While Business Intelligence (BI) and Data Analytics (DA) serve different primary functions, they share several key similarities. Both BI and DA involve data analysis, utilise advanced statistical models, and require expertise in statistical analysis and data science to provide insights for decision-making (GeeksforGeeks, 2024).

4. Findings

4.1 Data Structuring and Preparation

Before conducting any analysis, it is essential to refine the dataset to ensure clarity and efficiency in extracting insights. To preserve data integrity, a copy of the raw dataset was created in a separate worksheet before modifications were applied. Several changes were made in the retail online shop dataset to improve data structure and analytical effectiveness.

One key addition involved introducing an age bracket classification to categorise users into distinct demographic groups. For instance, individuals aged 10 to 19 were classified as "Teen," those between 20 and 39 as "Adult," while the 40 to 59 and 60 plus groups were labeled "Middle Age" and "Old" respectively (Integris Health, 2015). This transformation allowed a more structured analysis of purchasing behaviours across demographics, facilitating trend identification.

Additionally, the dataset was converted into a table format to improve organisation, filtering, and visualisation. Structured formatting, including color-coded tables, enhanced readability and interpretation of key variables. These changes streamlined the data analysis process, reducing errors and improving efficiency when comparing different data segments.

Further refinements were made to improve data consistency and clarity. The TotalAmount column was formatted to display values in pound currency (£), ensuring financial data was more interpretable. User identifiers were also standardised by converting formats such as “User\_1” to “U\_1”, to enhance consistency and readability.

4.2 Sales Revenue Distribution and Category Insights

This section illustrates the total sales and revenue performance for the given period from January 2021 to January 2022, focusing on monthly performance, sales by country and product sales.

Overall, there was an increase in total sales, with January 2021 generating the highest revenue, while the following month witnessed a decline. In addition, sales distribution across categories remained balanced, with no category dominating the market.

4.2.1 Sales Performance by Time and Region

The total sales for the period are £126,085,235.97, with mild monthly fluctuations. After the first month of the period, the store reached its peak by generating £11,150,317.15, after which it saw a sharp decline to £9,586,040.96 in February. It then recovered to £10,797,449.91 in the following month, before fluctuating in the range of £10,246,559.39 and £10,770,383.65 toward the end of the period, with a final figure of £10,759,422.83 *(See figure 1.1)*. Additionally, sales were evenly distributed across countries, with leading contributors such as Canada (16.79%), the USA (16.78%), and the UK (16.76%). Germany recorded the lowest sales, generating £20,813,741 during the period (*See figure 1.2*).



*Figure 1.1. Total Sales by Month*

|  |  |  |
| --- | --- | --- |
| **Country** | **Total Sales** | **Total Sales in %** |
| Canada | £ 21,173,596 | 16.79% |
| USA | £ 21,159,329 | 16.78% |
| UK | £ 21,133,151 | 16.76% |
| Australia | £ 20,945,706 | 16.61% |
| India | £ 20,859,713 | 16.54% |
| Germany | £ 20,813,741 | 16.51% |

*Figure 1.2. Total Sales by Country*

4.2.2 Product Category Trends and Seasonal Impact

In terms of sales by category, the distribution was nearly equal across Accessories (25.03%), Apparel (24.99%), Books (25.01%), and Electronics (24.97%). This suggests no single category dominates the market, indicating a balanced product demand.



*Figure 1.3. Sales by Category (March 2021 to May 2021)*

However, Apparel sales increased to 25.42% during the Spring months (March 2021 to May 2021), which aligns with consumer buying lighter clothing as the temperature rises. On the other hand, Books saw a slight increase to 25.96% in Winter (November 2021 to January 2022), likely due to academic preparation for the new semester at schools. *(See figure 1.3 and 1.4)*



*Figure 1.4. Sales by Category (November 2021 to January 2022)*

4.3 Customer Demographics and Sales Analysis

This section provides an analysis of customer demographic, including gender, age brackets, and other key attributes. The goal is to explore patterns in the customer base, identify the target customer group, and understand how these factors influence sales performance.

4.3.1. Sales by Gender and Age Bracket

Sales by gender show a fairly balanced distribution, with Male and Female customers each contributing 33%, while Non-Binary customers account for 34%. In terms of sales by category, the distribution was nearly equal across all genders. However, the data reveals distinct preferences, Apparel had the highest percentage of sales among Men (8.42%), while Women were most inclined to purchase Accessories, contributing 8.44%. Non-Binary customers showed stronger preference for Books (8.39%) and Electronics (8.41%) *(See figure 2.1, where conditional formatting highlights high values in blue and lower values in red)*.

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Female** | **Male** | **Non-Binary** |
| Accessories | 8.44% | 8.28% | 8.31% |
| Apparel | 8.27% | 8.42% | 8.30% |
| Books | 8.33% | 8.29% | 8.39% |
| Electronics | 8.29% | 8.26% | 8.41% |

*Figure 2.1. Category by Gender*

Sales distribution across different age groups reveals a relatively balanced spread. The Adult group (ages 25-44) contributed the highest percentage at 38.66%, followed closely by the Middle Age group (ages 45-54) at 38.50%. The Old group (55+) represented 19.07% of total sales, while the Teen group (under 18) made up the smallest share at 3.77% *(See figure 2.2)*.



*Figure 2.2. Sales by Age Bracket*

4.3.2 Target Customer

Analysis of referral sources shows that the UK, Germany, the USA, and Canada saw the highest engagement through email campaigns. This suggests that customers from these countries are likely employed and check their emails regularly, which aligns with a more professional, working-age demographic. In contrast, social media had relatively lower engagement across the chart, especially in India and Canada, indicating the business is less effective at attracting younger, social-media-driven audiences like teenagers or young adults *(See Figure 2.3).*



*Figure 2.3. Referral Source by Country*

4.4 Correlation analysis

To explore customer behaviour, a correlation analysis was conducted between Discount Rate and Review Score, as many businesses aim to examine whether higher discount rates lead to higher review scores. The result showed a weak positive correlation, with the CORREL function in Microsoft Excel showing a value of 0.1529 and an R2 of 0.0234 *(See figure 3.1)*. This suggests a slight upward trend, indicating that customers may respond somewhat positively to discounts. However, the impact on review scores is minimal. The scatter chart below shows a dispersed set of points with a slight upward-sloping trendline, indicating that other factors may influence customer satisfaction more significantly than discount levels alone.



*Figure 3.1. Correlation between Discount Rate and Review Score*

4.5 Analysis of Data Bias in Correlation

The weak positive correlation between discount rate and review score (CORREL = 0.1529) may reflect confirmation bias. It was initially believed that higher discounts would lead to better reviews, but the data only revealed a slight trend. This suggests that expectations may have influenced the way analysts interpreted the outcome. Confirmation bias occurs when analysts unconsciously favour data that supports their beliefs while overlooking evidence to the contrary. This highlights the importance of maintaining objectivity and critically assessing data patterns rather than relying on assumptions, especially when the correlation is weak or statistically insignificant.

5. Conclusion and Recommendation

The analysis demonstrated steady sales performance throughout the year, with January 2021 generating the highest revenue. Product sales and category contributions were balanced, indicating a well-diversified portfolio. Sales were evenly distributed across countries, with Canada, the USA, and the UK slightly outperforming others. Demographic insights showed that the primary customer base consists of adult and middle-aged individuals, with a balanced gender distribution, supporting broad marketing strategies. Seasonal trends were noted, such as increased book sales in winter and apparel in spring, indicating shifts in consumer demand. A weak correlation between discount rates and review scores was also identified.

Based on these findings, the business should leverage seasonal trends and purchasing behaviour through timely promotional campaigns, maintain a balanced product range, and tailor marketing strategies to specific age groups. Furthermore, addressing underperforming markets like Germany and India by investing in targeted advertising and region-specific campaigns could enhance overall performance. Finally, incorporating qualitative data, such as customer feedback, production costs, and import expenses, would offer deeper insights into customer expectations and operational efficiency, revealing more informed, data-driven decisions.

The End

References

Fawcett, T., & Provost, F. (2019, July 25). *Data science and its relationship to big data and data-driven decision making*. Foster Provost  <https://fosterprovost.com/publication/data-science-and-its-relation-to-big-data-and-data-driven-decision-making/>

Eldridge, S. (2023, August 1). *Data analysis | Definition, research, & methodology*. Encyclopedia Britannica. <https://www.britannica.com/science/data-analysis>

Bajaj, S. (2025, March 13). *Business intelligence vs data analytics: Key differences.* SoftwareSuggest Blog. <https://www.softwaresuggest.com/blog/business-intelligence-vs-data-analytics/>

*Difference between business intelligence and data analytics*. (2024, August 13). GeeksforGeeks. <https://www.geeksforgeeks.org/difference-between-business-intelligence-and-data-analytics/>

*Stages of Life: Health for Every Age*. (2015, October 9). Integris Health. <https://integrishealth.org/resources/on-your-health/2015/october/stages-of-life-health-for-every-age>

McLeod, S. (2024, September 2). *Gibbs reflective cycle*. Simply Psychology. <https://www.simplypsychology.org/gibbs-reflective-cycle.html>

Cecilia, K. Y., & Katherine, K. W. (2021, February). *Reflection literacy: A multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review*. Science Direct. <https://www.sciencedirect.com/science/article/pii/S1747938X20308368>