

Braintoy Individual report

By Neha Tararr



July 30, 2023

Briantoy

Table of Contents

[Introduction 2](#_Toc141638289)

[Data Analyzation 2](#_Toc141638290)

[Analyzing Dataset for Identifying Problems in the Dataset: 2](#_Toc141638291)

[Business Perspective 4](#_Toc141638292)

[Empowering Entrepreneurs and Investors with Data-Driven Insights: 4](#_Toc141638293)

[Python Coding 5](#_Toc141638294)

[Data Cleaning and Transformation Process: 5](#_Toc141638295)

[PowerBI 6](#_Toc141638296)

[Data Transformation with Power BI: 6](#_Toc141638297)

[Unveiling the Top-Selling Products for International Ventures 7](#_Toc141638298)

[Focused Analysis: Top-Selling Products in Canada 8](#_Toc141638299)

[Analyzing Top-Selling Countries: A Strategic Insight 11](#_Toc141638300)

[Conclusion 12](#_Toc141638301)

# Introduction

Welcome to the comprehensive individual report for Braintoy. In this report, we will embark on an exciting journey of data analysis and exploration, aiming to extract valuable insights from the dataset of an E-commerce site operated by a UK retailer. The dataset contains a wide array of products and orders, reflecting real-world challenges that an actual E-commerce site may encounter.

# Data Analyzation

The dataset contains orders from an E-commerce site belonging to a UK retailer. It encompasses a wide range of products and orders. To gain insights and uncover potential issues, we will begin with Analyzing Dataset through Python and PowerBI and apply various functions to analyze the data. The dataset might reflect real-world challenges that an actual E-commerce site could encounter.

## Analyzing Dataset for Identifying Problems in the Dataset:

* Duplicates Detection and Removal:

We will address the issue of duplicate records by using Python to identify and exclude them. The primary key in this case is the "Invoice number." Any duplicates under this key will be removed, ensuring the dataset remains clean and accurate.

* Handling Missing Values:

We will investigate the dataset for missing values. Based on the identified missing values, we will employ suitable techniques like averaging or using the median to fill in the gaps, ensuring a complete dataset for further analysis.

* Negative Unit Price Analysis:

The dataset includes instances of negative unit prices, which requires exploration. We will attempt to identify the reasons behind these negative values to better understand their impact on the business.

**Prescriptive Analysis Approach to Improve Sales:**

By applying a prescriptive analysis approach, we aim to uncover actionable insights that can help boost sales and improve business intelligence. Through this fascinating task, we can build valuable skills in data-driven decision-making to optimize the retailer's performance.

Exploring the E-commerce dataset from a UK retailer promises to be an engaging and insightful task. employing various data analysis techniques, we will gain valuable knowledge about the dataset and discover potential areas for improvement. Ultimately, the prescriptive analysis will empower us to suggest actionable strategies to enhance sales and enhance business intelligence for the retailer.

In 'Figure 1.1: Negative Quantities,' it is apparent that negative quantities are associated with printing smudges or items thrown away. We will explore these occurrences in-depth to understand their implications and devise strategies to mitigate such situations.

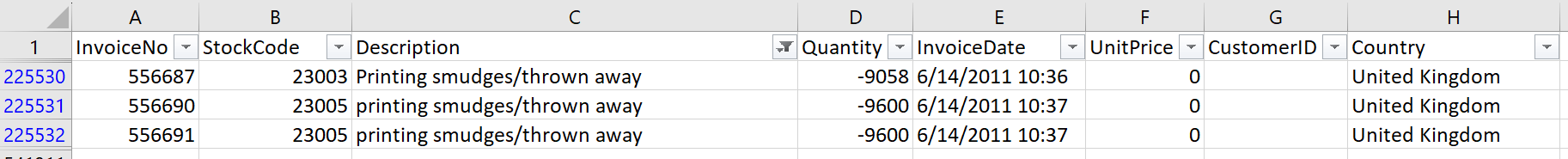


Figure 1.1: Negative Quantities

Items with the description ‘?’ mostly had negative quantities and little to no unit price, as shown in ‘Figure 1.2: Description Discrepancies.’ Further investigation is required to understand and resolve these discrepancies in the dataset.

A screenshot of a computer

Description automatically generated

Figure 1.2: Description Discrepancies

Consistently, all C InvoiceNo entries have negative quantities for the some certain items, as depicted in 'Figure 1.3: InvoiceNo Discrepancies.' Further investigation is needed to understand and address these discrepancies in the dataset.

A screenshot of a computer

Description automatically generated

Figure1.3: InvoiceNo Discrepancies

# Business Perspective

## Empowering Entrepreneurs and Investors with Data-Driven Insights:

Our mission is to empower entrepreneurs and investors with data-driven insights that will enable them to make well-informed decisions in the dynamic world of business. We have set our sights on three key areas to achieve this:

* Smart Product Selection:

Through meticulous data analysis, we will identify trends and patterns to guide entrepreneurs in selecting the most promising products. By considering factors such as purchased goods, customer preferences, and quantities of products, we aim to offer valuable guidance on which products hold the greatest potential for success.

* Uncovering Top-Performing Countries:

Understanding global sales dynamics is critical for strategic expansion. Our analysis will delve into sales data from different countries, revealing lucrative markets where products are thriving. This information will empower entrepreneurs and investors to focus their efforts on regions with the highest potential for growth and profitability.

* Canadian Market Perspective:

Canada, being a crucial local market, will receive special attention in our analysis. We will thoroughly examine the Canadian market landscape, evaluating consumer behavior, high selling products, and frequencies of highest selling products. This perspective will provide valuable context for those looking to establish or strengthen their presence in the Canadian market.

By leveraging comprehensive data analysis, our insights will be actionable and relevant, supporting entrepreneurs and investors in their pursuit of success. We aim to facilitate strategic decision-making, minimize risks, and maximize opportunities for growth and prosperity.

# Python Coding

## Data Cleaning and Transformation Process:

During the initial stages of our analysis, we encountered a sizable dataset of approximately 5 million rows, which proved too large to be accommodated within the storage capacity of my laptop. Consequently, to work with a manageable dataset, I had to filter out countries with excessive data, ultimately focusing on Australia, Canada, France, Hong Kong, United Arab Emirates, and the USA. This refined selection yielded a dataset of 10,615 rows.

To ensure data integrity, we took measures to eliminate any potential duplicates from the dataset using Python. The code snippet used for this process is shown in 'Figure 2.1: Python Code.'

Moreover, addressing null values in the dataset was a vital step in our cleaning process. Given the dataset's substantial size and the relatively low occurrence of null values, we opted to remove the rows containing these null values. The code used to accomplish this task can also be found in 'Figure 2.1: Python Code.'

After completing the data cleaning process, our dataset was reduced from 10,615 rows to 10,261 rows. This streamlined and cleaned dataset serves as an excellent foundation for our subsequent data transformations and analyses.

With the data now in a more structured and refined state, we are well-prepared to proceed with further data transformations and explore valuable insights to support our analysis and decision-making processes.

A screen shot of a computer program

Description automatically generated

Figure 2.1: Python Code

# PowerBI

### Data Transformation with Power BI:

To facilitate data transformation and gain deeper insights, I turned to the powerful capabilities of Power BI. During this process, I made some intriguing observations regarding the quantities of certain products in Australia, France, and the USA, as they were recorded in negative values.

Upon closer examination, I discovered a consistent pattern - all negative quantity records were associated with invoice numbers that began with the letter 'C'. This observation suggested that these negative quantities might be indicative of stock issues or product returns.

Given that our project's primary focus is on identifying the best-selling products, it was essential to address these negative quantities. Therefore, in Power BI's Data Transformation section, I took the necessary steps to exclude the negative quantity entries from the analysis. This action is clearly illustrated in 'Figure 2.2: Data Transforming.'

By removing the negative quantities from the dataset, we ensure that our analysis remains focused on the products that are truly driving sales and performance in the selected countries. This data transformation enhances the accuracy and reliability of our findings, enabling us to make more informed decisions and strategic recommendations.

With Power BI's robust capabilities, we have successfully transformed the data, enabling us to proceed with an in-depth exploration of the best-selling products and uncover further insights to support our project's objectives. The combination of advanced data transformation and powerful visualization tools positions us for a comprehensive and impactful analysis.

A screenshot of a computer

Description automatically generated

Figure 2.2: Data Transforming

## Unveiling the Top-Selling Products for International Ventures

With our dataset now successfully transformed and refined, we can delve into the exciting realm of visualizations to uncover valuable insights. Our focus lies in identifying the top 5 products across all selected countries for aspiring international sellers.

As illustrated in 'Figure 2.3: Most Sold Products,' the visualizations reveal the standout performers that have secured the top positions in Australia, Canada, France, United Arab Emirates (UAE), and the USA. These top-selling products possess significant potential for international success and profitability.

The top five most sold products across these countries are as follows:

1. **Rabbit Night Light**
2. **Mini Paint Set Vintage**
3. **Red Toadstool Led Night Light**
4. **Red Harmonica In Box**
5. **Spaceboy Lunch Box**

For entrepreneurs seeking to venture into international markets, this information is immensely valuable. Understanding which products consistently top the sales charts empowers sellers to make informed decisions about their international expansion strategies. By focusing on these high-performing products, sellers can capitalize on their popularity and ensure maximum profitability.

Furthermore, being aware of the products that sell the most allows sellers to plan for frequent restocking, ensuring a smooth and uninterrupted supply chain. This proactive approach ensures that customer demands are promptly met, leading to enhanced customer satisfaction and loyalty.

All in all, the visualization of the top 5 most sold products in Australia, Canada, France, UAE, and the USA provides indispensable guidance for international sellers. Armed with this knowledge, entrepreneurs can confidently venture into new markets, capitalize on high-demand products, and build a thriving global business.

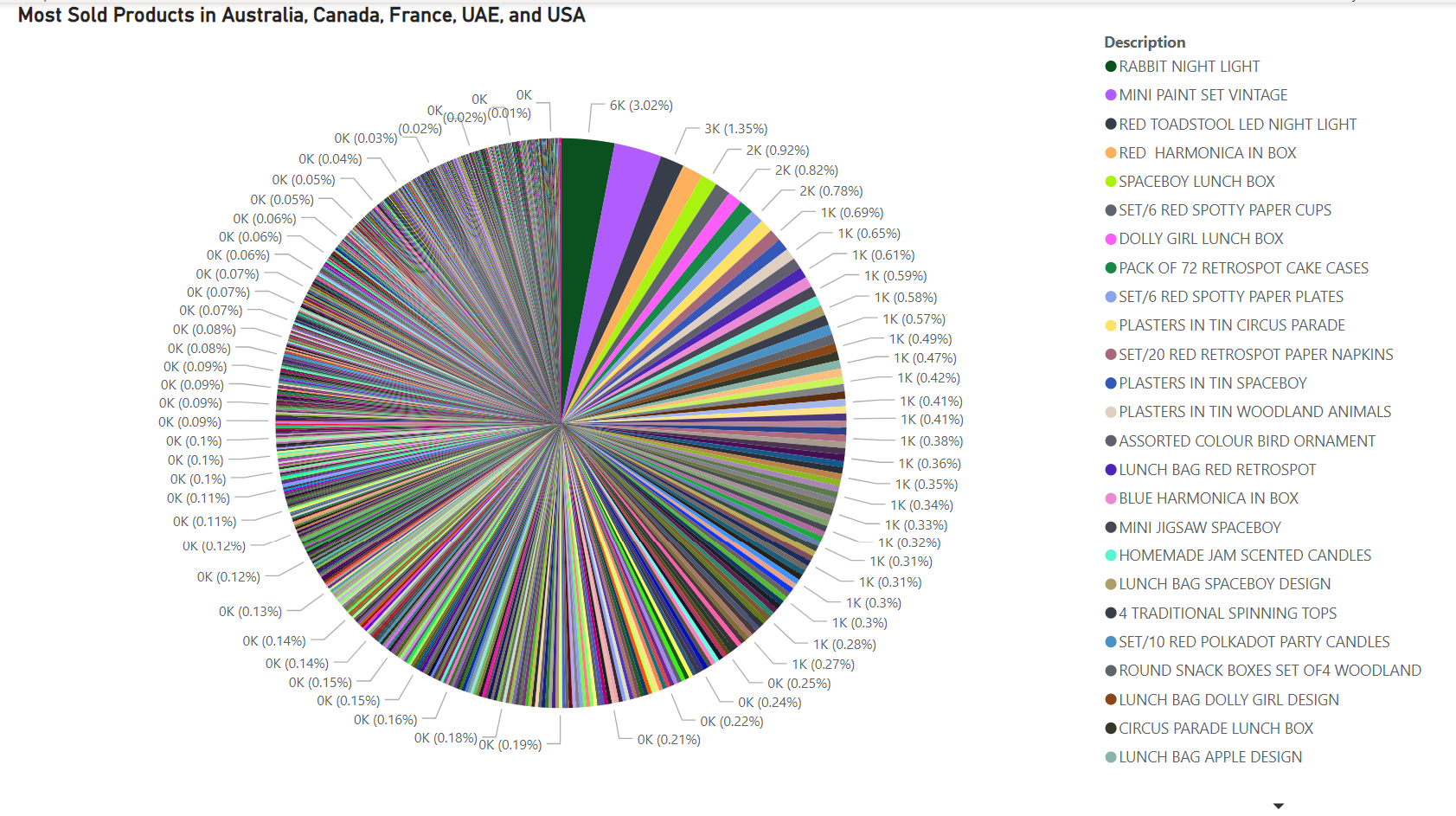


Figure 2.3: Most Sold Products

## Focused Analysis: Top-Selling Products in Canada

In our pursuit of comprehensive insights, we recognize the significance of catering to sellers with a specific interest in the Canadian market. To achieve this, we embarked on a separate comparison, honing our attention to the top 3 most sold products in Canada.

Given the abundance of products in the dataset, we sought to streamline our analysis by concentrating solely on the most sold items. To accomplish this, we utilized the filtering capabilities of our data tool, as demonstrated in 'Figure 2.4: Data Filtering.'

By employing this approach, we narrowed down our focus to the cream of the crop - the products that have consistently achieved exceptional sales figures within the Canadian market. This specialized analysis aims to assist entrepreneurs with a keen interest in establishing a successful presence in Canada.

The top 3 most sold products in Canada, as shown in ‘Figure 2.5: Top 3 Products Sold in Canada’, carefully extracted from the filtered data, hold immense potential for sellers seeking to thrive in this market:

1. **Colouring Pencils Brown Tube**
2. **Bundle of Three Alphabet Exercise Books**
3. **Bundle of Three School Exercise Books**

This targeted comparison empowers sellers to channel their efforts into promoting these high-performing products, optimizing sales strategies, and establishing a strong foothold in the Canadian market.

By providing a clear and concise overview of the most sought-after products in Canada, we aim to equip entrepreneurs with the knowledge and confidence needed to navigate this specific market successfully.

In conclusion, our tailored analysis of the top 3 most sold products in Canada offers invaluable guidance to sellers with a distinct focus on this vibrant and lucrative market. With this information in hand, sellers can take bold strides towards achieving their goals and attaining long-term success in the Canadian business landscape.

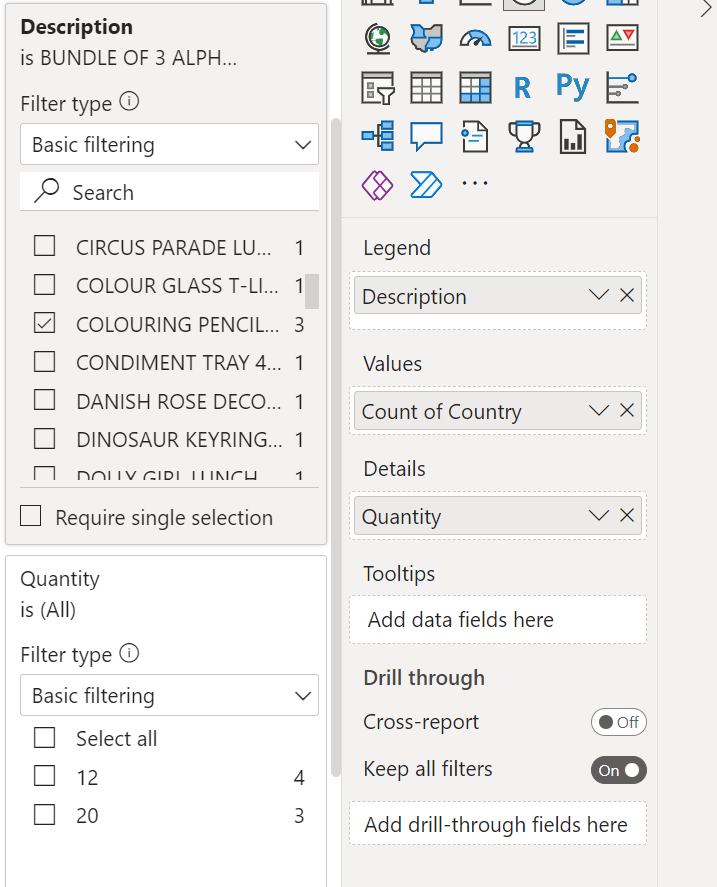


Figure 2.4: Data Filtering

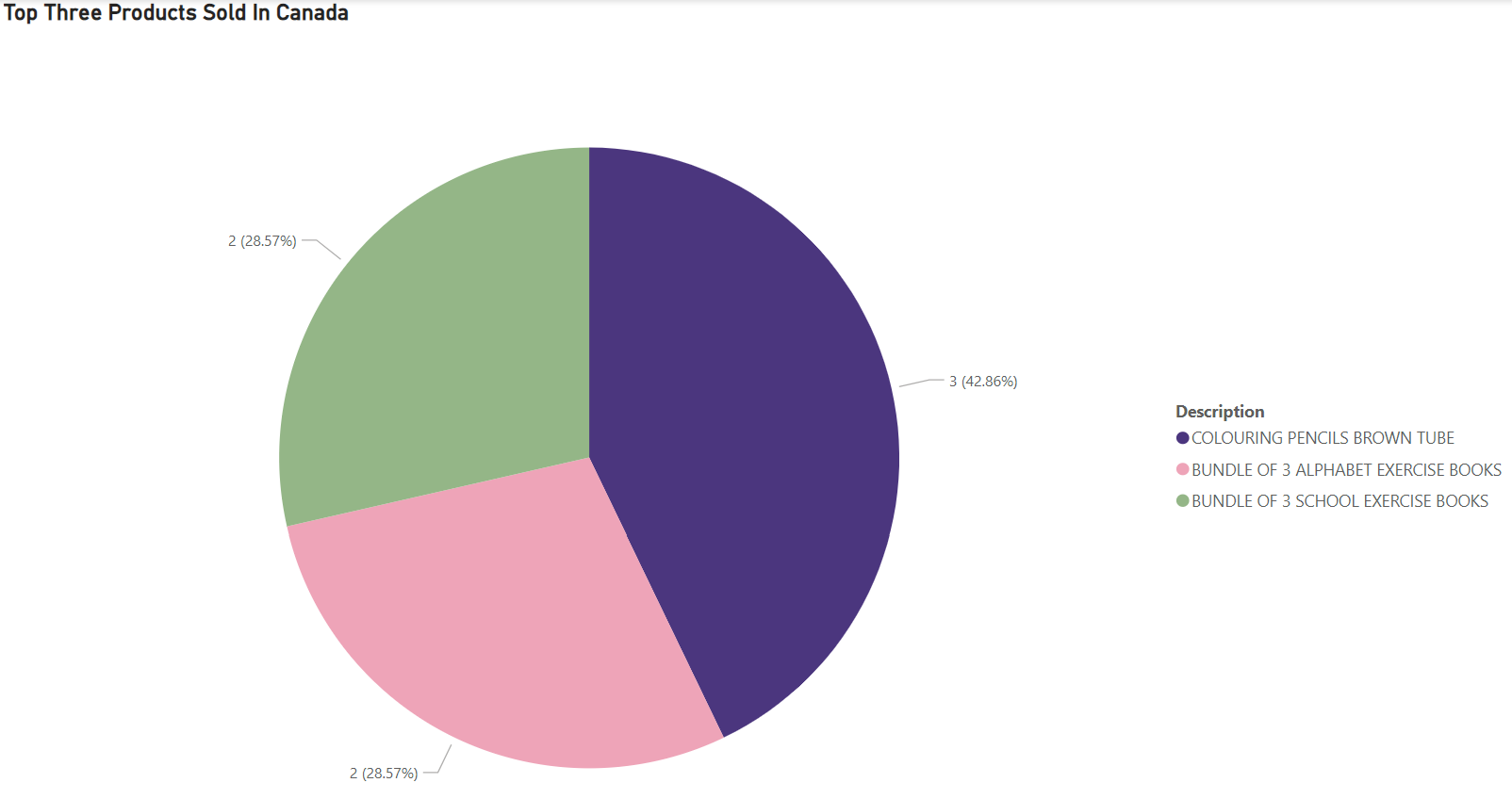


Figure 2.5: Top 3 Products Sold in Canada

## Analyzing Top-Selling Countries: A Strategic Insight

Expanding our analysis to include the selected countries, we sought to determine which ones have emerged as frontrunners in terms of product sales. The insightful visualization presented in 'Figure 2.6: Most Products Sold By Country' unveils the countries that have excelled in selling the products from our dataset.

The findings reveal that France, Australia, and Canada have taken the lead as the top three countries, demonstrating remarkable sales performance. For entrepreneurs and investors seeking targeted markets for their products, these insights are a treasure trove of valuable information.

Understanding which countries have sold the most products is an indispensable strategic asset. It enables sellers to focus their efforts on these high-performing markets, thereby maximizing the potential for profitability and business growth. By honing their attention on these top three countries, sellers can make informed decisions regarding their Supply Chain support requirements, such as product storage, logistics, and transportation logistics.

Moreover, identifying these leading markets empowers entrepreneurs and investors to allocate their marketing investments strategically. By directing resources towards countries with high product demand, they can effectively reach and engage their target audience, ultimately boosting brand awareness and sales.

In summary, the analysis of top-selling countries, with France, Australia, and Canada taking the lead, offers significant advantages to those seeking targeted markets for their products. With this data-driven knowledge, entrepreneurs and investors can chart a course for success, tailor their strategies to specific markets, and drive their business to new heights of prosperity.

A screenshot of a computer

Description automatically generated

Figure 2.6: Most Products Sold By Country

# Conclusion

The journey of data analysis and exploration has brought us valuable insights, equipping entrepreneurs, and investors with the knowledge to make informed decisions. By leveraging data-driven strategies, we are confident that respective audience will thrive in the competitive E-commerce landscape.