Sweet Data: The Story Behind Chocolate Sales Trends

by Magali Mougey
April 2025

OUTLINE

- **Executive Summary**
- **Introduction**
- **Methodology**
- **Results**
- **Key Insights**
- **Discussion**
- **Conclusion**

EXECUTIVE SUMMARY

This project leverages a Kaggle dataset to uncover key insights into **global chocolate sales trends**, identifying consumer preferences and **market dynamics**. By analyzing sales performance across regions, products, and sales representatives, the findings provide **a strategic roadmap** for optimizing inventory, marketing, and pricing strategies.

INTRODUCTION

Chocolate is more than just a delightful indulgence—it represents a global industry driven by diverse consumer preferences and seasonal demand. This analysis leverages a Kaggle dataset to explore chocolate sales trends across various countries, identifying key market insights that shape business strategies.

The objective of this study is to:

- **Analyze global sales trends**
- **Identify regional preferences**
- **Evaluate salesperson performance**
- **Explore forecasting models**

This report is designed for **Retail & sales strategists** aiming to refine their market positioning. **Business analysts & data scientists** seeking insights into global chocolate consumption trends. **Product managers & marketers** looking to tailor offerings based on regional preferences. **Supply chain & logistics experts** optimizing inventory distribution strategies.

By utilizing data-driven storytelling and advanced visualizations, this analysis provides a comprehensive understanding of **chocolate sales dynamics**, helping businesses make informed decisions to enhance their market impact.

METHODOLOGY

This project analyzes chocolate sales trends using a **Kaggle dataset (.csv)** licensed under MIT. The data was processed and explored in a **Jupyter Notebook** (Lite version), with some limitations in Python library compatibility.

Data Processing & Verification
- The dataset was loaded into a Pandas DataFrame for transformation and validation.
- No major challenges were encountered in data cleaning and preparation, ensuring **a smooth preprocessing** phase.

Visualization & Geographic Analysis
- Matplotlib was used to generate insightful plots to visualize sales trends.
- The main challenge involved creating the choropleth map using **NaturalEarth** data, where the .shp file was converted to GeoJSON for compatibility with **GeoPandas**.

Machine Learning (ML) & Forecasting
- Sales forecasting models were explored using ML techniques to predict future trends.
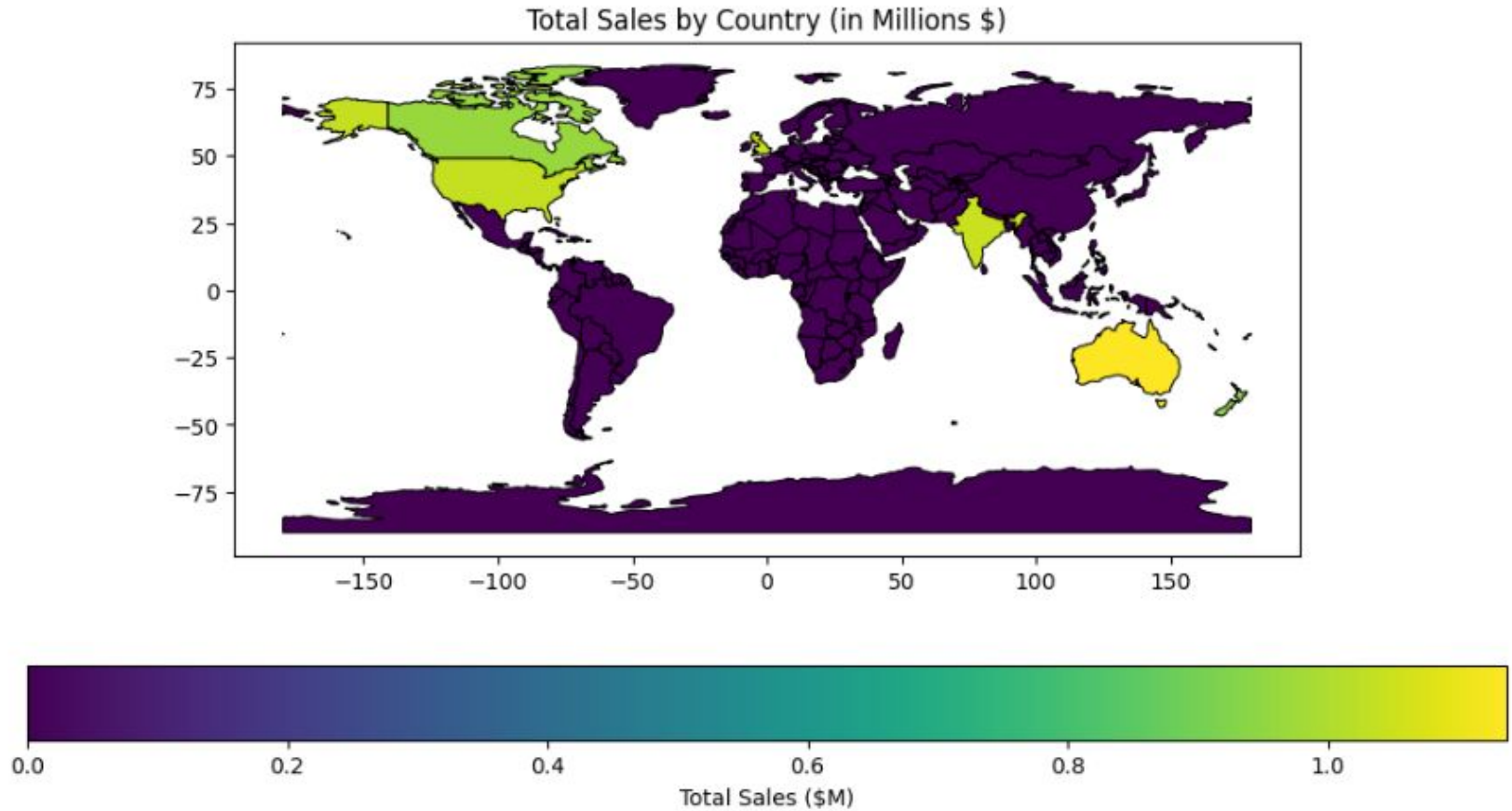
Dashboard & AI-Assisted Optimization
- An interactive dashboard was developed in **IBM Cognos**, enabling real-time sales tracking and business intelligence insights.
- **Generative AI** enhanced every step of the data analysis workflow, optimizing processes from data cleaning to visualization and forecasting.
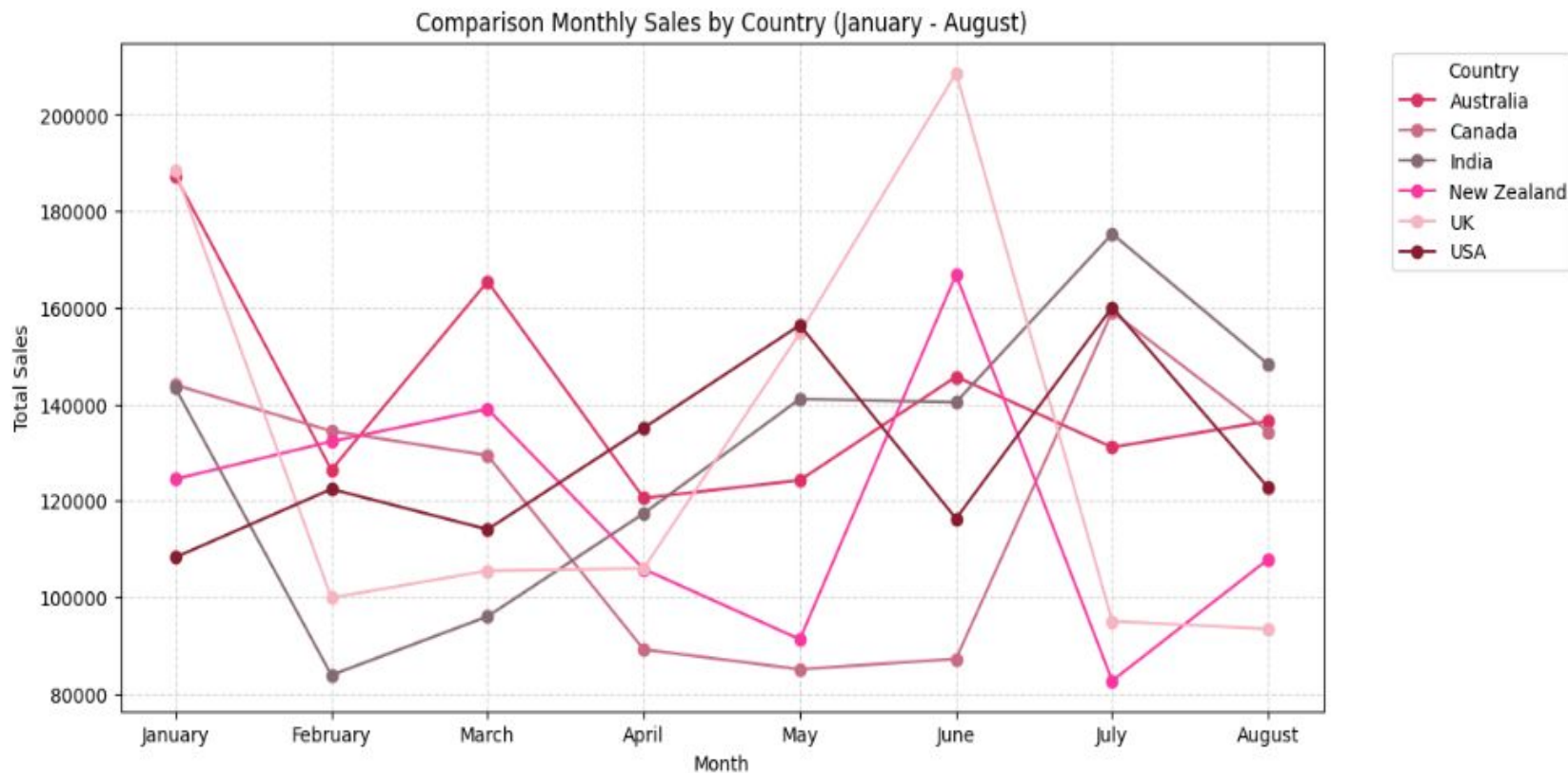
RESULTS & KEY INSIGHTS

Total Sales ($M) by Country

Total Sales by Country (in Millions $)

Monthly Sales by Country

Comparison Monthly Sales by Country (January - August)

Key Insights : Global Market & Monthly Sales Trends

Australia leads chocolate consumption with **$1.13M**, followed by the **UK** and **India** as strong markets.

Understanding why Australia tops the list can help refine regional marketing strategies.

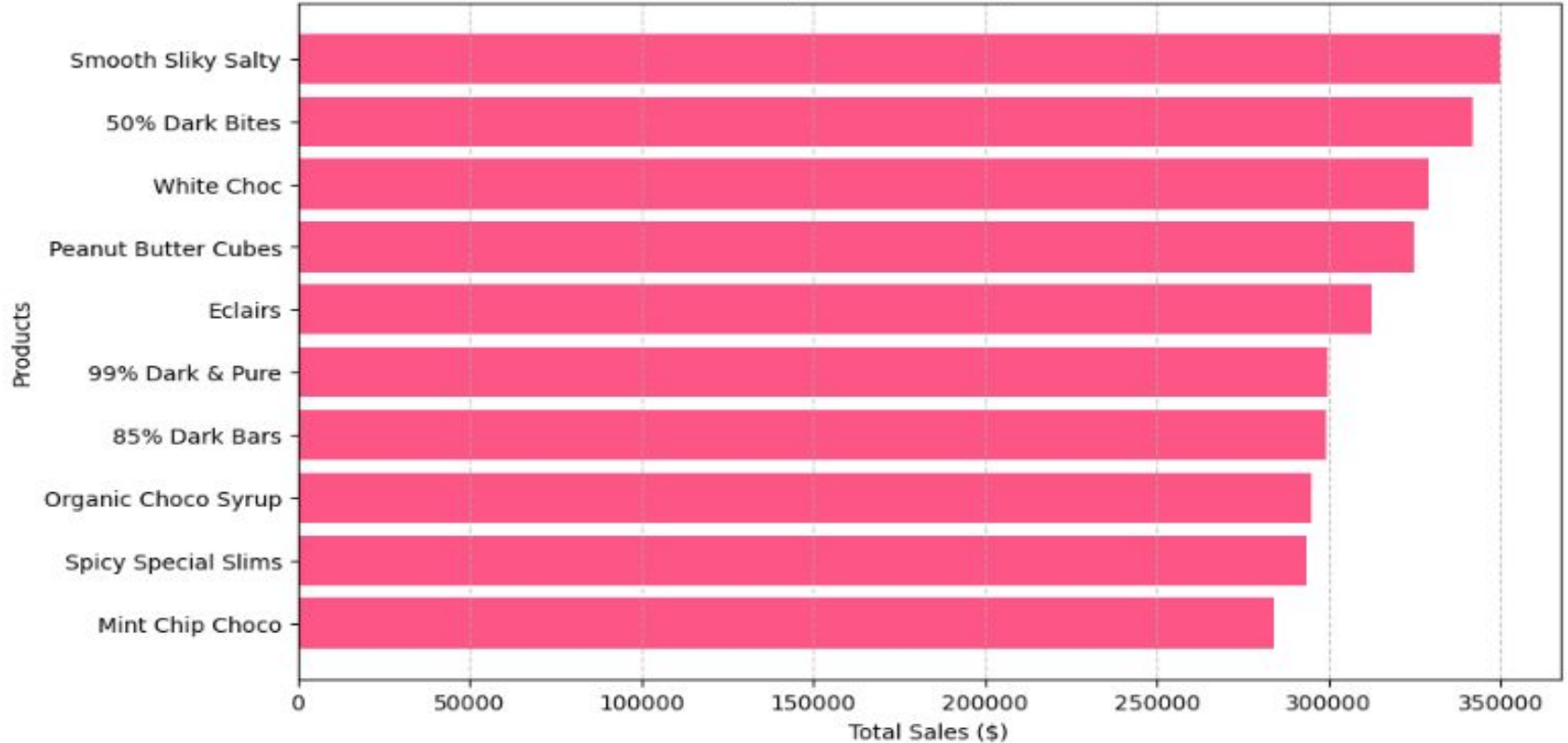Sales peak globally in **January** and **March**, likely due to **Valentine's Day and Easter preparations**.

UK & New Zealand show an unexpected **June spike**, while USA, India, and Canada peak in July.

Seasonality plays a significant role—strategic promotions in peak months can boost revenue.
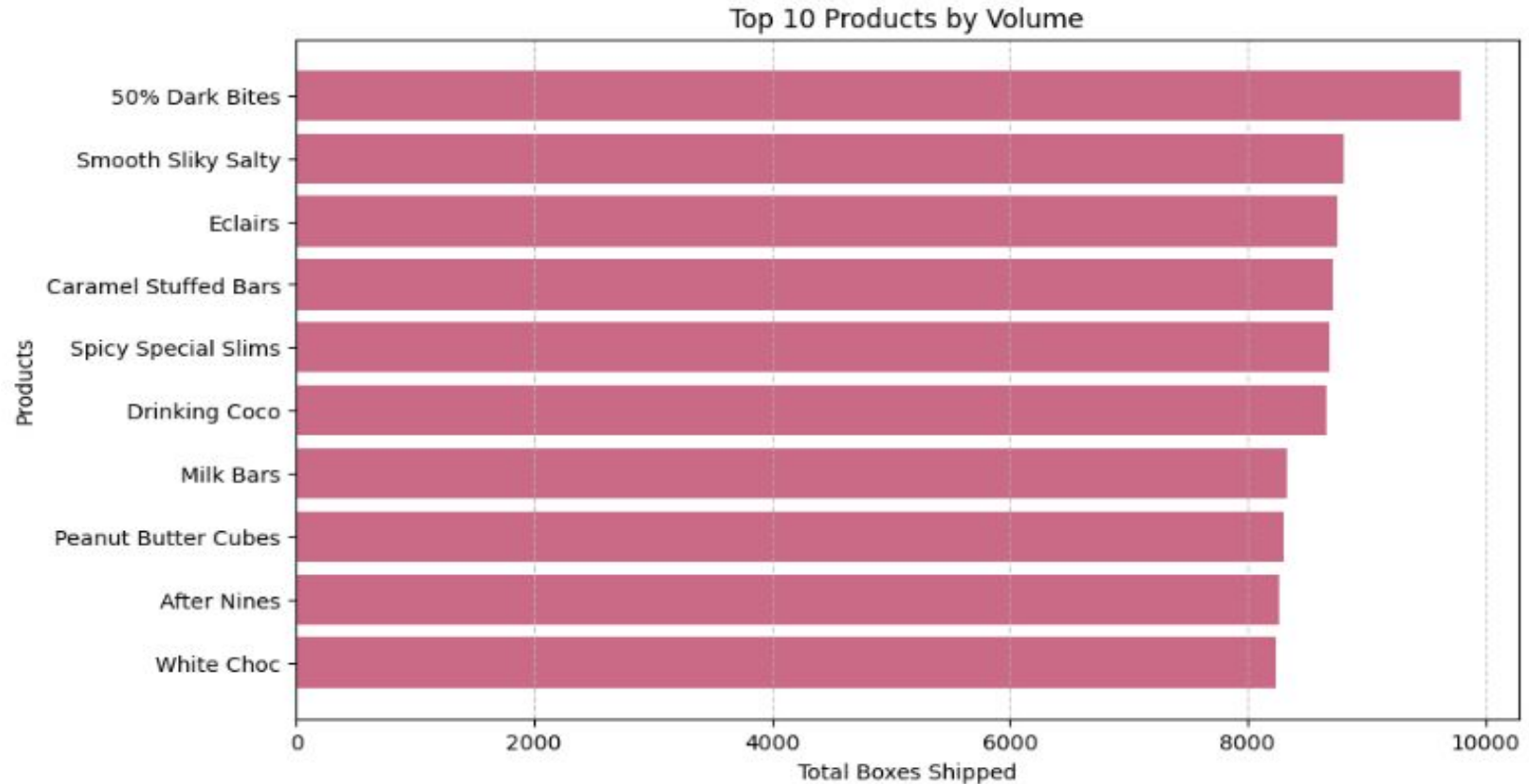
Top 10 Products by Sales ($)

Top 10 Products by Sales

Top 10 Products by Volume (Boxes Shipped)

Top 10 Products by Volume

Key Insights : Products Best Seller

The biggest revenue drivers:
Smooth Silky Salty, 50% Dark Bites, and White Chocolate dominate.
High sales suggest strong brand loyalty or successful product positioning.

In sheer popularity (volume), the leading products differ slightly:
50% Dark Bites, Smooth Silky Salty,and Eclairs are top-selling in quantity.
White Chocolate, despite high revenue, ranks 10th in volume, indicating **higher pricing** or **lower shipment numbers**.
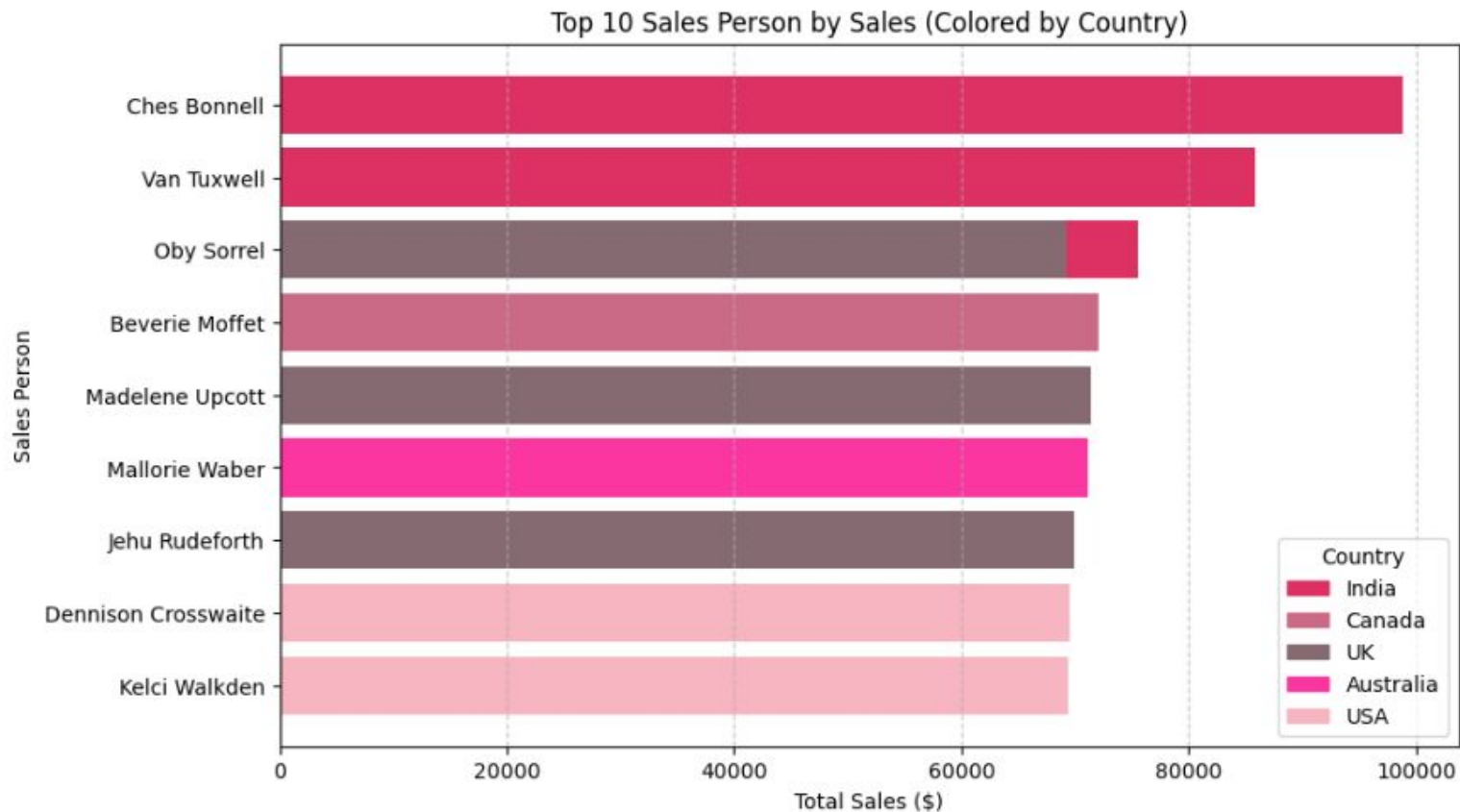
Performance Across All 25 Sellers

Sales Performance of All 25 Sales Persons (Grouped by Country)
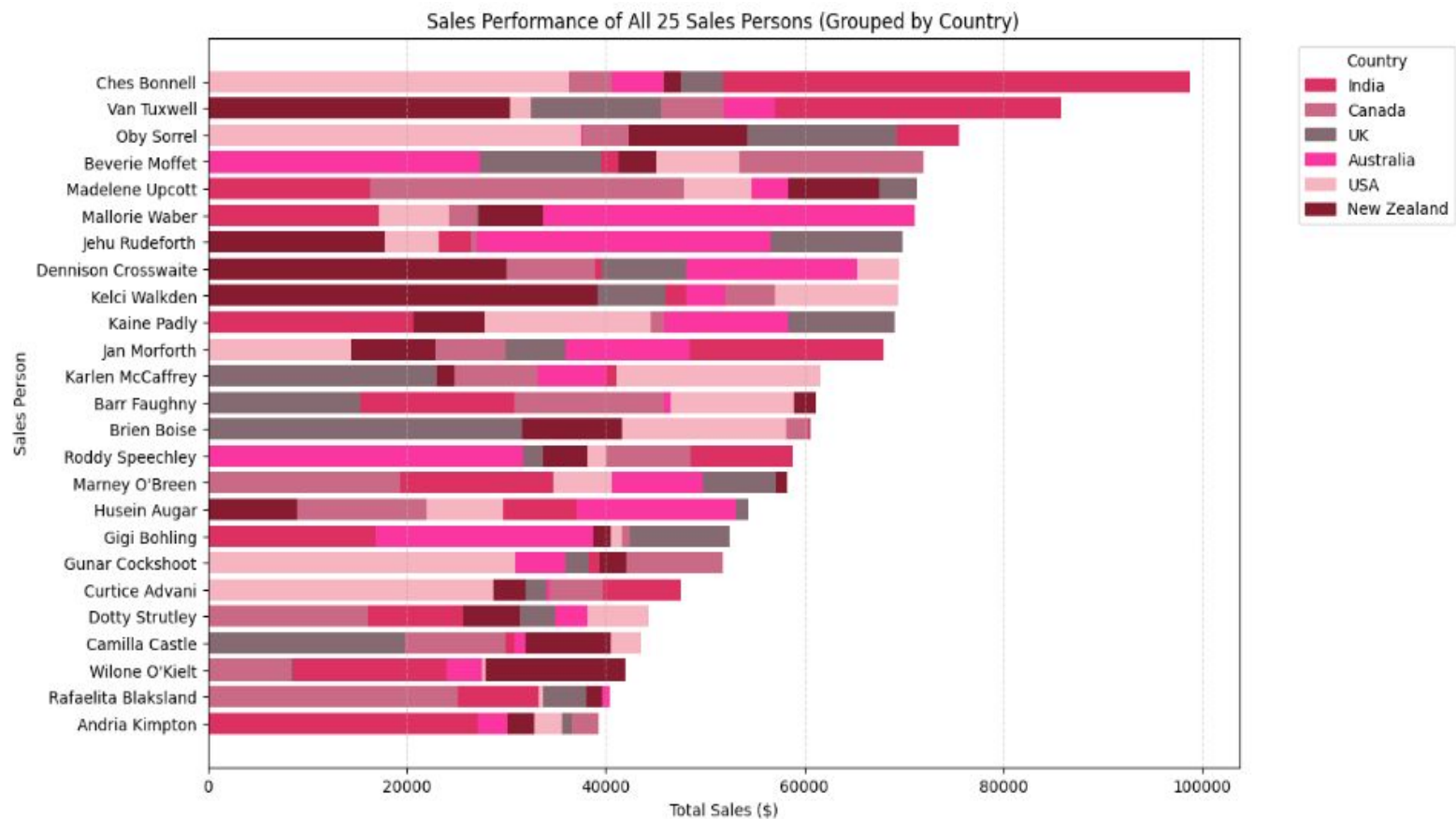
Top 10 Performers as Sellers

Top 10 Sales Person by Sales (Colored by Country)

Key Insights : Sales Performances

The Top 10 Bar Chart reveals **strong India-driven sales** performances at the top and **highlights USA's weak performance**, placing its salespersons at the lower end—a key strategic finding.

The Full 25 Salesperson Chart gives **a broader view of salesperson contribution**s, including middle-tier sellers who still play vital roles in multi-country sales efforts.

Top Performers & Country Distribution:
The #1 salesperson **drives nearly 50% of India's total sales**—a dominant player in the market.
The #2 salesperson contributes 1/3 to India's revenue, reinforcing India's strong market performance but with slightly broader diversification.

Regional Powerhouses:
Among the Top 9 sellers, New Zealand's market presence is significant, ranging from 1/3 to nearly half of total sales for these individuals.
This highlights **New Zealand as a profitable yet diverse market**, warranting further strategic investment.

**Middle & Lower-Tier Salespersons:**
Sales among the remaining sellers are more evenly distributed across different countries, meaning they serve multiple regions rather than a single dominant country.
This suggests **a diverse sales force**, where only a handful of salespersons specialize heavily in certain countries.

Sales Correlation Matrix: Best-Selling Chocolate by Country
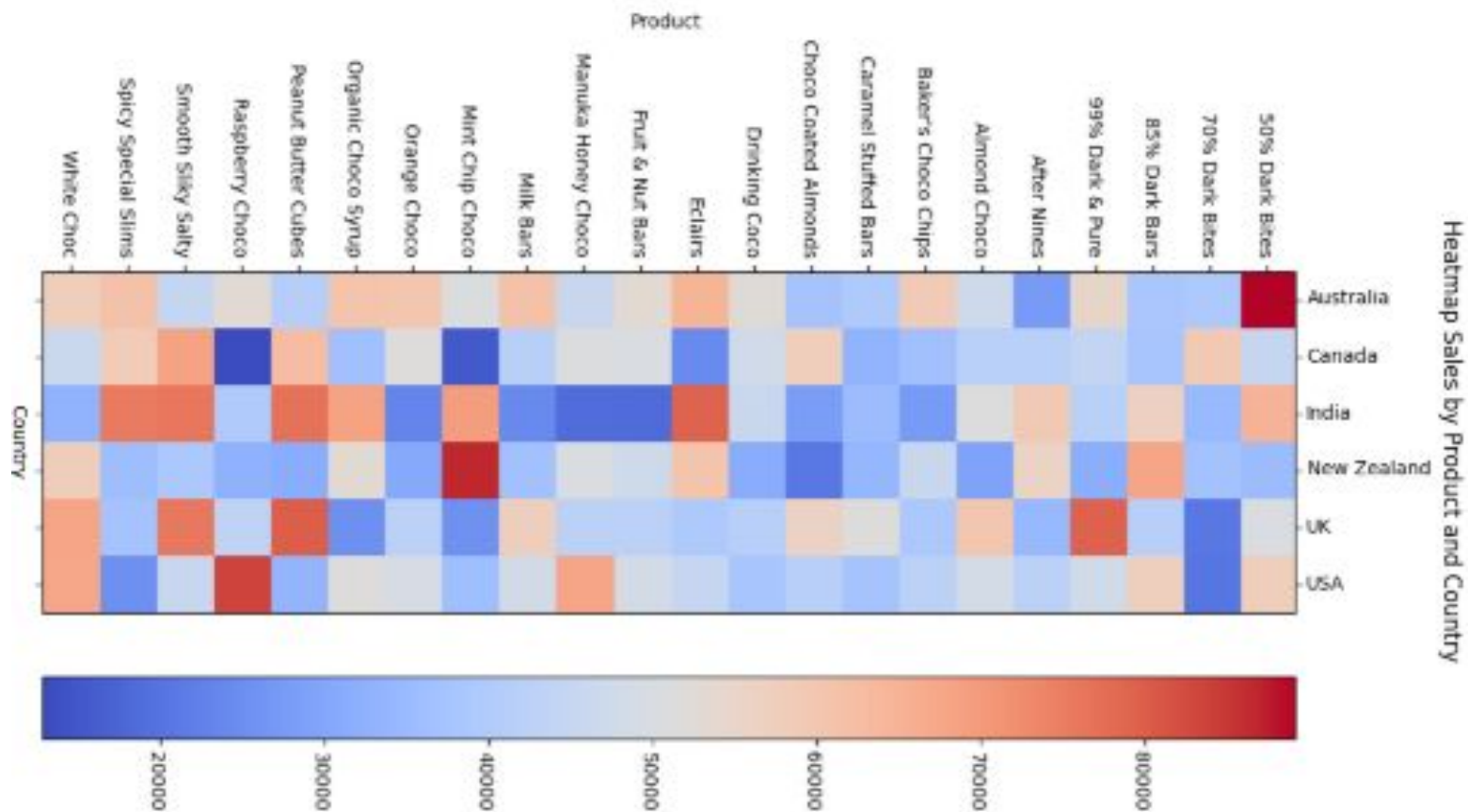
Top 5 Best-Seller Products by Country (Volume)
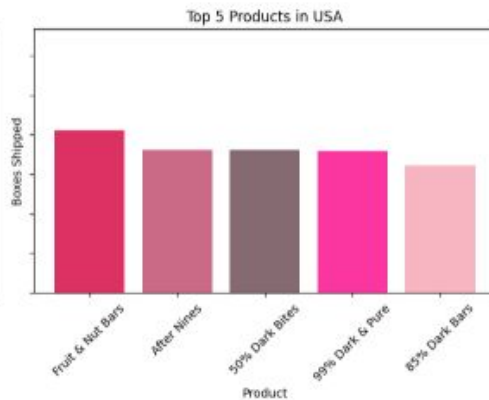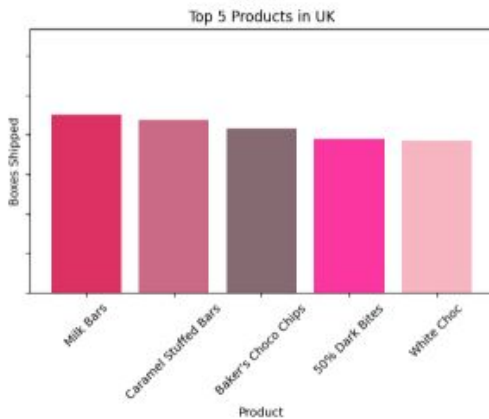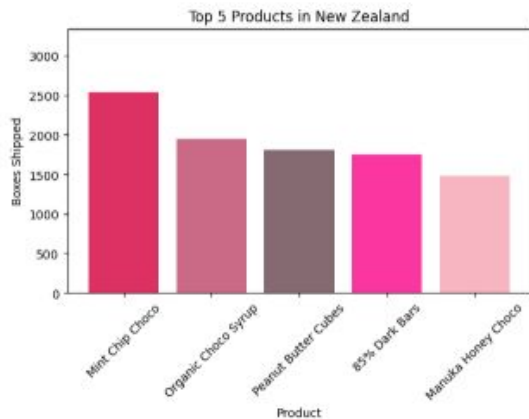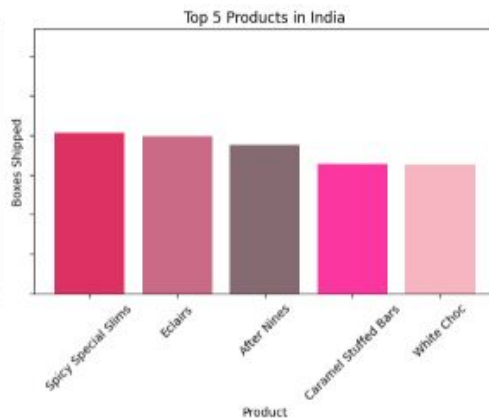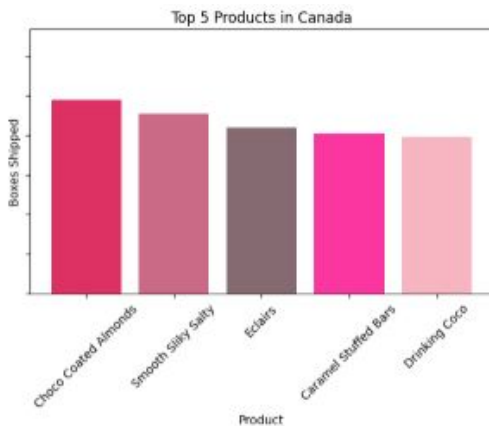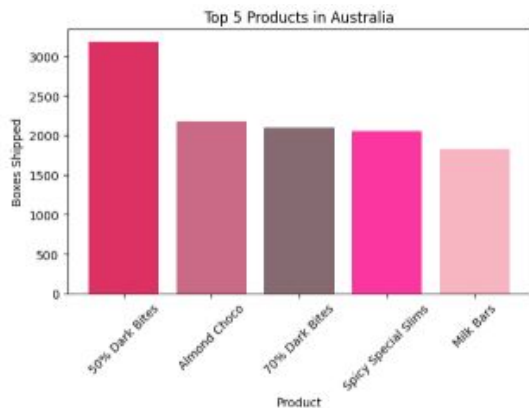
Key Insights : Best-selling Products by Country (Volume & Revenue)

Australia's top revenue-generating product is 50% Dark Bites, reflecting a strong preference for mid-range dark chocolate. In **New Zealand, Mint-flavored chocolate leads** both in sales and shipments, confirming a distinct consumer taste profile. India sees Eclairs as its highest revenue driver, while Canada favors Smooth Silky Salty as its best-selling product. The UK market is dominated by 99% Dark & Pure, indicating a strong demand for ultra-dark chocolate. In the USA, Raspberry chocolate generates the most revenue.

Dark chocolate remains a **key preference for US consumers**, with 3 out of the top 5 products containing 50%, 85%, or 99% cocoa. Australia follows a similar trend, with 50% dark chocolate leading shipment volume, while 70% dark chocolate ranks third. **New Zealand** stands out with **a unique taste profile**, favoring Mint, Peanut Butter, Manuka Honey, and Organic Chocolate Syrup, differentiating it from other markets.

These variations in product preference **highlight the influence of regional taste trends** on chocolate sales performance.

DISCUSSION

The analysis of chocolate sales across six countries from January to August revealed **clear regional preferences** and **significant disparities in both revenue and shipment volumes**. These variations suggest **potential marketing and pricing strategies** tailored to consumer demand. Additionally, some products generated revenue outliers, warranting further investigation into the reasons behind their exceptional sales performance. The insights presented here provide valuable guidance for decision-making in marketing, promotions, and sales planning.

However, **attempts to forecast future sales using machine learning techniques proved unsuccessful despite employing advanced tools.** The challenges encountered using different models suggest that **sales are influenced by external factors not captured in the dataset**, such as market dynamics, competitor strategies, and unpredictable shifts in consumer demand.

Ultimately, the failure of forecasting methods underscores the need for a more comprehensive approach that integrates additional variables, such as marketing efforts, competitor pricing, and seasonality trends, to enhance predictive accuracy and improve decision-making in future sales strategies. Further research and data collection are recommended to refine forecasting models and develop more reliable predictions for chocolate sales performance.

CONCLUSION

To optimize chocolate sales, the following strategies should be pursued:

Leverage seasonality: Adjust marketing efforts for **peak months—January, March, June, July.**

Tailor pricing strategies: Align premium pricing with high-revenue but lower-volume products.

Expand product offerings: Introduce regional flavors like Mint for New Zealand or Dark Chocolate for USA & Australia.

Strengthen weak sales regions: USA's sales performance is significantly lower—adjust incentive structures and reinforce strategic sales training. Optimize inventory distribution: Ensure dark chocolate-heavy markets (USA & Australia) receive targeted stock allocations for maximum profitability.

This analysis provides actionable insights to refine chocolate sales strategies, ensuring higher revenue, optimized regional distribution, and consumer-driven product positioning.