Decoding Consumer Spending: Strategic Retail Insights That Drive Growth

by Olumide Balogun

Table of Contents

1. Executive Summary

2. Objectives of the Project

3. Business Questions

4. Tools Used throughout the Project

5. Methodology

6. Key Findings: Insights, EDA Visualizations, and   Model-Driven Interpretations

7. Strategic Recommendations

8. Expected Business Impact

9. Conclusion

Executive Summary

This end-to-end, Python-powered analytics pipeline was developed to extract, analyze, and transform U.S. retail sales data into strategic, real-time business intelligence. By decoding the complexities of customer behavior, product performance, and regional sales trends, the project empowers data-driven decision-making at scale.

✓ **Key methodologies included**: **Exploratory Data Analysis (EDA)** for trend discovery, **Prophet forecasting** to anticipate sales movements, **Market Basket Analysis** for cross-selling opportunities, and **KMeans clustering** to segment customers and uncover behavioral patterns

Rather than stopping at "what happened," the analysis dove into the why, and most importantly, the what's next.

✓ **Impact:** Streamlined marketing strategies tailored to customer segments, Data-informed inventory decisions to reduce overstock and stockouts, and Forecasts that prioritized profitability, not just revenue

As a result, *Chris & Kris* gained a competitive, predictive edge, fuelled by insight, not guesswork.

Key Insights

1. **Sales & Performance**

➤ Revenue peaked in **December ($6.41M)**; February saw a sharp **73.39% dip**, confirming strong seasonal swings. **November sales surged +226.47%** due to holiday campaigns, timing is everything.

Executive Summary (continued)

2. Product & Consumer Behavior

➤ **The Top sellers are AAA Batteries, Charging Cables, Galaxy Buds, and accessories are the unsung heroes. Tuesdays, especially 11 AM–2 PM and 6–9 PM, are the sweet spots for customer traffic.**

➤ **Bundle purchases (e.g., phone + cable) reveal solid cross-sell potential.**

3. Market Insights

➤ **San Francisco, LA, NYC lead in revenue; Portland (ME) and Austin show untapped potential.**

➤ **Clustering analysis uncovered 3 market tiers for tailored regional approaches.**

4. Forecasting & Strategy

➤ **Prophet model shows Q4 growth and early 2025 volatility; consistent >$1M monthly sales supports solid budgeting.**

➤ **Timing campaigns with peak shopping hours can boost ROI significantly.**

Strategic Recommendations

➤ **Maximize Q4**: Double down on marketing, inventory, and staffing.

➤ **Promote Bundles**: Elevate AOV with high-frequency product pairings.

➤ **Target Peak Hours**: Optimize promotions for Tuesdays and key shopping windows.

➤ **Invest in Top Cities**: Scale up efforts in SF, LA, NYC.

➤ **Revive Low-Performers**: Localized campaigns for Portland and Austin.

➤ **Forecast-Driven Planning**: Leverage Prophet to guide resource allocation.

Objectives of this Project

The primary objective of this project was to **transform raw transactional sales data into actionable business insights** that support **data-driven decision-making and strategic growth** for *Chris & Kris*, a U.S.-based retail company. Specifically, the project aimed to:

➢ **Uncover hidden sales patterns** and highlight untapped growth opportunities.

➢ **Identify top-performing products and cities** contributing the most revenue.

➢ **Understand customer purchasing behavior** to better time and target marketing efforts.

➢ **Forecast future sales trends** for smarter budgeting, inventory planning, and resource allocation.

➢ **Recommend tailored strategies** using advanced analytics like customer segmentation, market basket analysis, and sales forecasting.

Business Questions

1. Which month achieved the highest total sales, and what was the total sales figure recorded during that period?

2. What is the percentage growth in total monthly sales over time, and how does it vary across different months?

3. Which city achieved the highest total sales, and what was the total sales amount for that city?

4. Which product had the highest sales, and what factors do you believe contributed to its success?

5. How do product sales vary across different days of the week, and what insights can be drawn from the distribution of order volumes by product and day?

6. On which day of the week do customers make the most purchases?

7. What is the optimal timing for advertisements and promotions to maximize customer purchases, based on historical purchase behavior?

8. What will our monthly sales look like over the next two years based on historical purchasing patterns?

9. What products are frequently bought together?

10. How can we group cities based on sales performance to uncover patterns for better sales strategy?

Tools Used throughout the Project

1. **Python**: The foundation of my analysis, enabling efficient data manipulation and insight extraction. Key libraries included:

➤ Pandas for data cleaning and analysis

➤ Matplotlib for basic data visualizations

➤ Seaborn for enhanced and detailed plotting

2. **Machine Learning**: To uncover patterns and make predictive insights, I incorporated:

➤ scikit-learn for model development and evaluation

➤ XGBoost for high-performance gradient boosting

➤ Prophet for time series forecasting

3. **Jupyter Notebooks**: Provided an interactive coding environment, allowing for seamless integration of code, analysis, and documentation.

4. **Git & GitHub**: Used for version control, code management, and collaboration, ensuring transparency and reproducibility throughout the project.

Methodology

Phase I: EDA – Descriptive & Diagnostic Analysis

1. **Data Cleaning & Preprocessing**: Removing duplicates, handling missing values, creating more columns, and formatting variables

2. **Monthly Sales Analysis**: Identified the highest-grossing months and overall trends.

3. **City-Level Analysis**: Mapped sales performance geographically.

4. **Product & Weekly Trends**: Uncovered product demand patterns by weekday.

5. **Insights Visualization**: Created clear visuals (bar charts, heatmaps) for storytelling.

## Phase II: Machine Learning – Predictive & Prescriptive Models

1. **Time Series Analysis**: Determined optimal times for promotions and customer engagement.

2. **Sales Forecasting**: Used Prophet models to predict future sales.

3. **Market Basket Analysis**: Revealed frequent product pairings using Apriori algorithm or any other analysis.

4. **Clustering**: Applied KMeans to segment cities for targeted sales strategies.

Key Findings: Insights, EDA Visualizations, and Model-Driven Interpretations

Which month achieved the highest total sales, and what was the total sales figure recorded during that period?

Monthly Sales
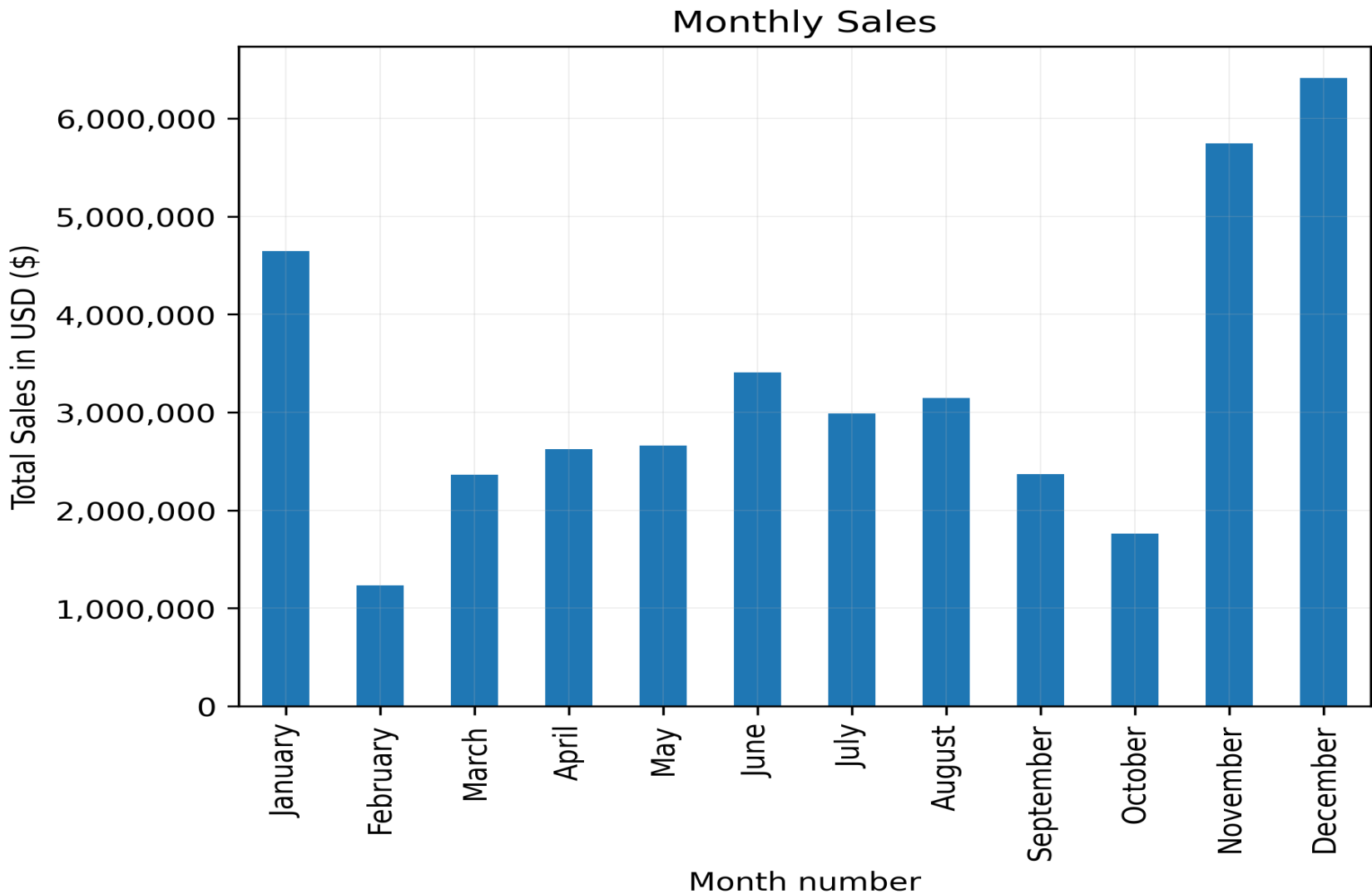
Key Insights

December reigns supreme with **$6.41M** in revenue, proving the unmatched power of **Q4 and holiday-driven consumer spending.**

➤ **February is a red flag**, recording the lowest sales at **$1.24M**, a signal to rethink early-year strategies.

➤ **Sales momentum builds from June onward**, showing a consistent lead-up to year-end peaks.

What is the percentage growth in total monthly sales over time, and how does it vary across different months?

## Monthly Total Sales and Growth Rate (%)
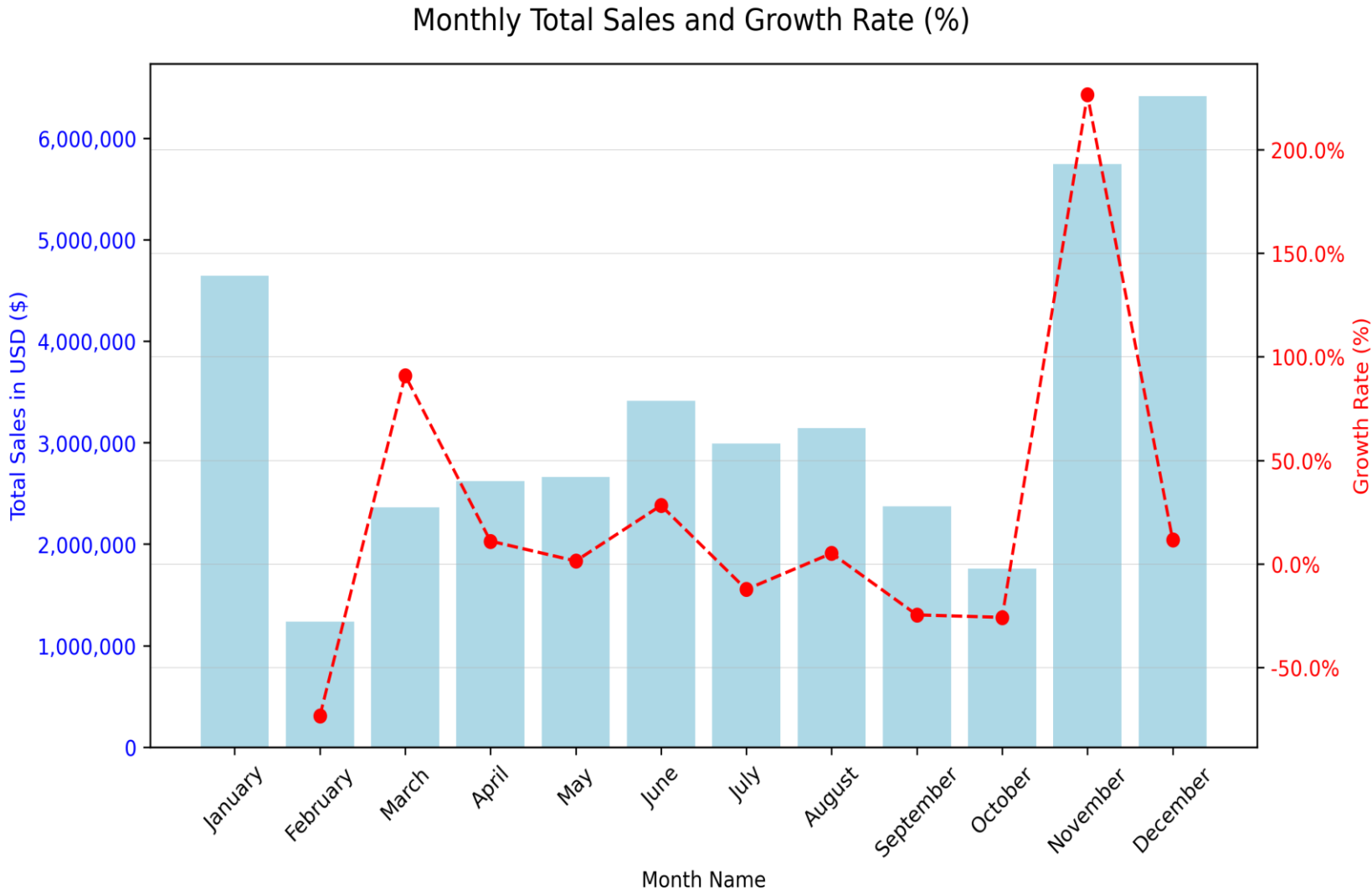
Key Insights

November shows explosive growth at **+226.47%,** driven by **pre-holiday promotions**, double down on this period.

➢ **February (-73.39%) and October (-25.77%) reveal volatility,** highlighting the need for buffer campaigns and agile planning.

➢ **Growth is uneven**, a reminder that demand planning must adapt to both calendar seasons and consumer mood shifts.

Which city achieved the highest total sales, and what was the total sales amount for that city?

City Sales
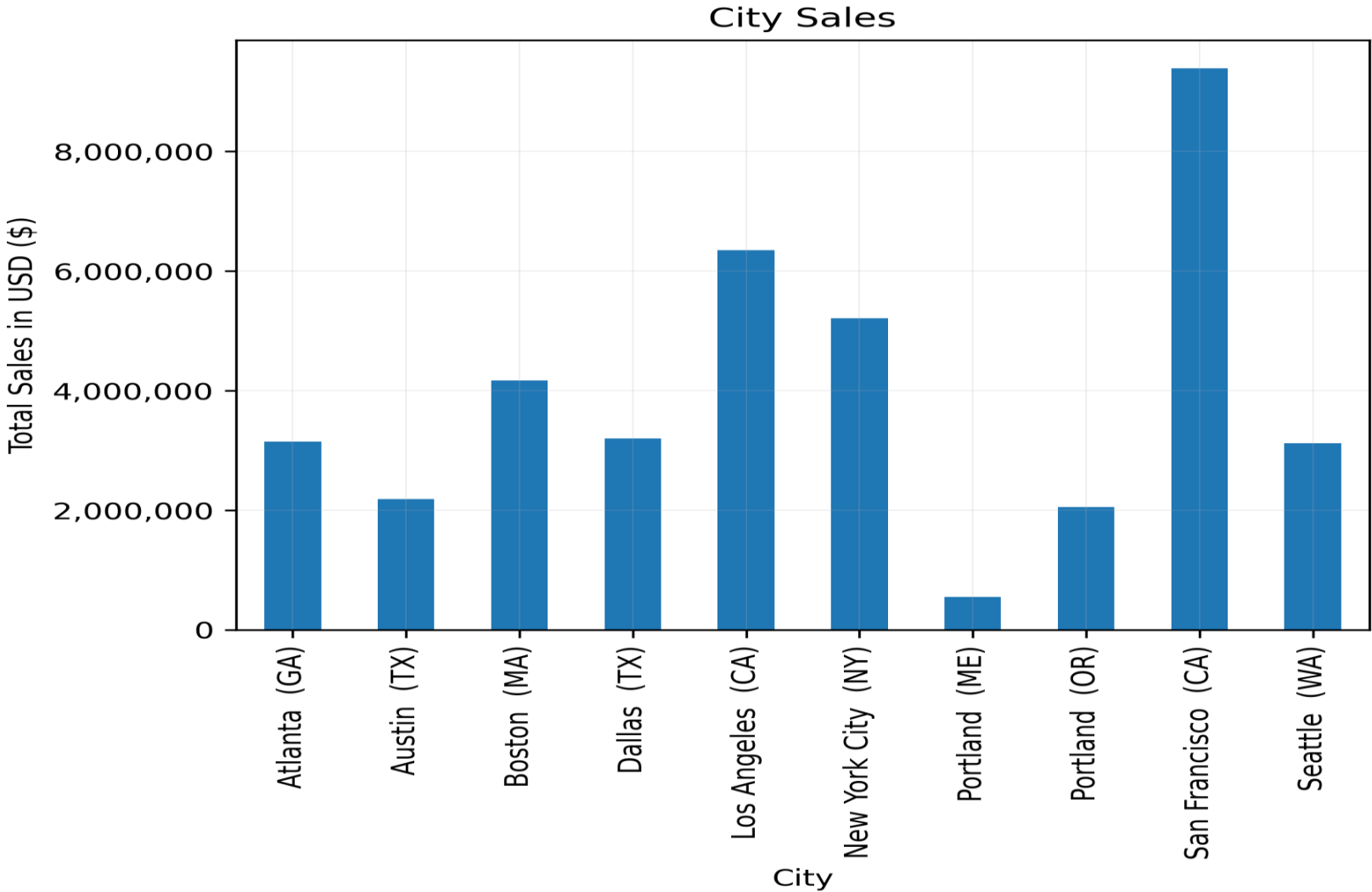
Key Insights

San Francisco dominates the leaderboard with **$9.4M** in total sales, **over 17x more** than Portland (ME).

➢ **LA and NYC are revenue magnets with $6.35M and $5.21M** respectively, solidifying the value of investing in metro hubs.

➢ **Mid-tier cities like Austin and Portland (OR)** represent untapped potential with the right local strategies.

Which product had the highest sales, and what factors do you believe contributed to its success?

Units Purchased and Unit Price by Product
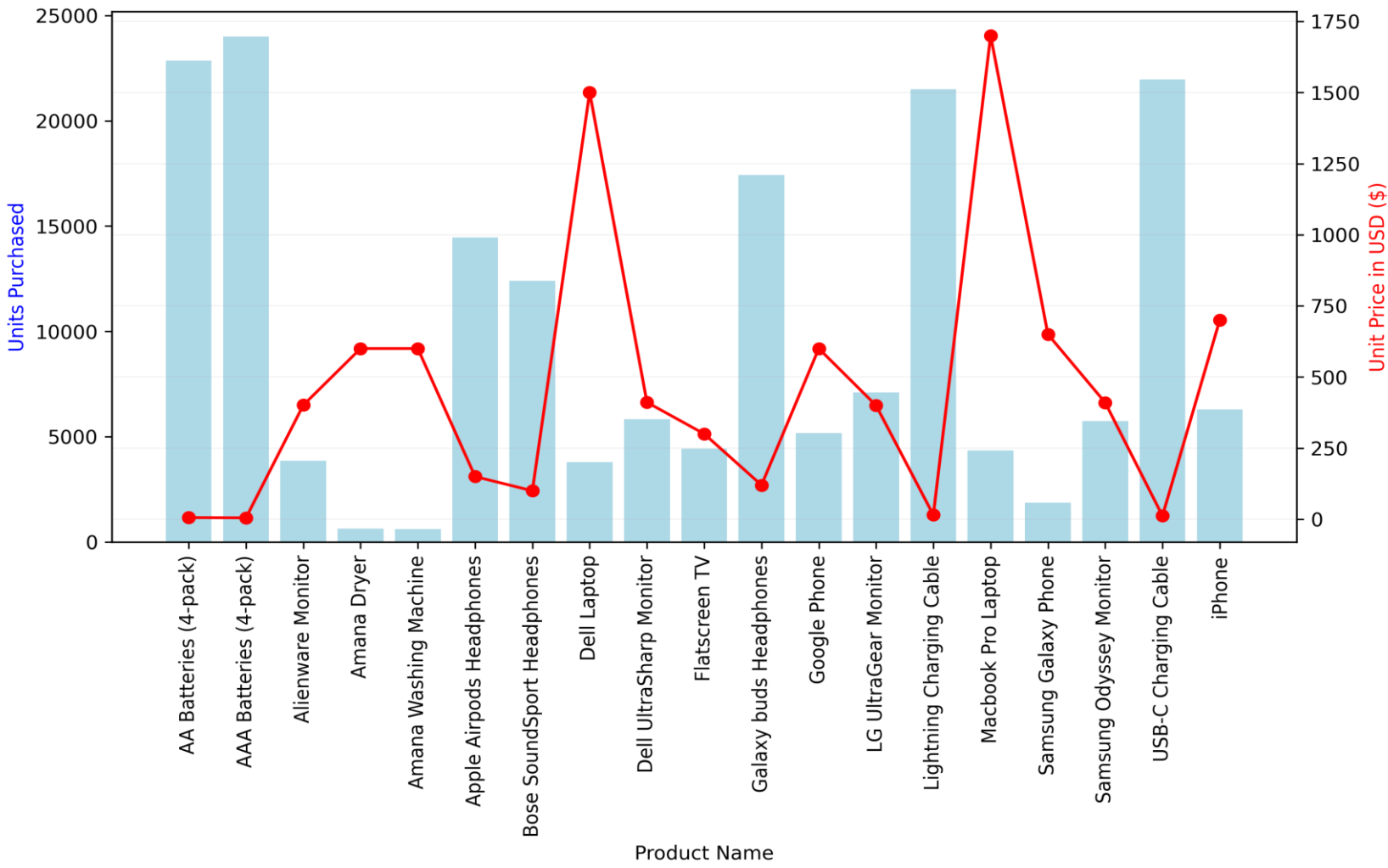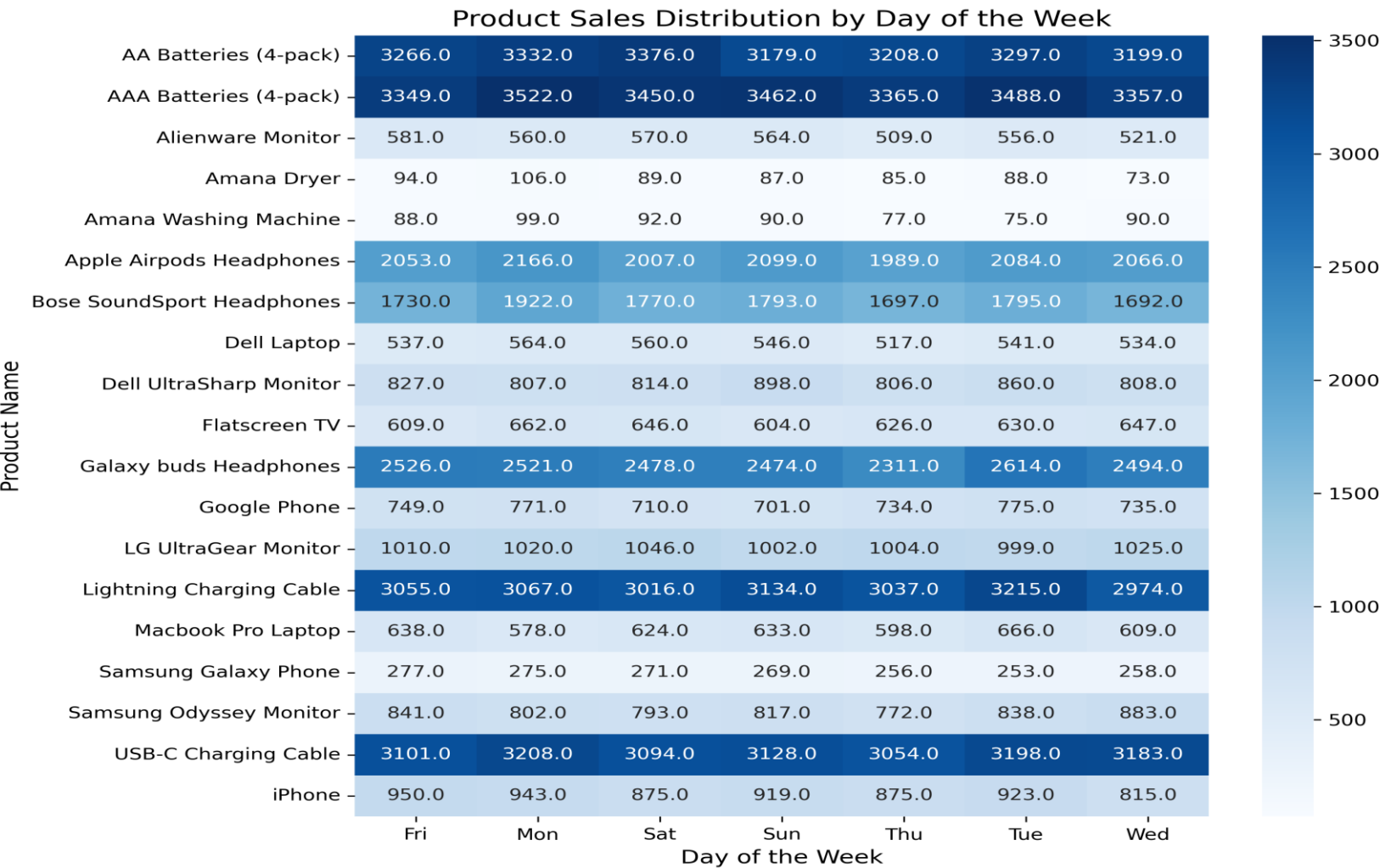
Key Insights

Accessories are the unsung heroes, AAA Batteries (**23,993 units**) outsell iPhones by **4x** and MacBooks by **5.5x.**

➢ **Galaxy Buds outshine AirPods and Bose**, suggesting pricing and value perception trump brand loyalty.

➢ **High-value electronics underperform**, likely due to price sensitivity and purchase frequency, needs repositioning or bundle support.

How do product sales vary across different days of the week, and what insights can be drawn from the distribution of order volumes by product and day?

Product Sales Distribution by Day of the Week

| Product Name	Fri	Mon	Sat	Sun	Thu	Tue	Wed
AA Batteries (4-pack)	3266.0	3332.0	3376.0	3179.0	3208.0	3297.0	3199.0
AAA Batteries (4-pack)	3349.0	3522.0	3450.0	3462.0	3365.0	3488.0	3357.0
Alienware Monitor	581.0	560.0	570.0	564.0	509.0	556.0	521.0
Amana Dryer	94.0	106.0	89.0	87.0	85.0	88.0	73.0
Amana Washing Machine	88.0	99.0	92.0	90.0	77.0	75.0	90.0
Apple Airpods Headphones	2053.0	2166.0	2007.0	2099.0	1989.0	2084.0	2066.0
Bose SoundSport Headphones	1730.0	1922.0	1770.0	1793.0	1697.0	1795.0	1692.0
Dell Laptop	537.0	564.0	560.0	546.0	517.0	541.0	534.0
Dell UltraSharp Monitor	827.0	807.0	814.0	898.0	806.0	860.0	808.0
Flatscreen TV	609.0	662.0	646.0	604.0	626.0	630.0	647.0
Galaxy buds Headphones	2526.0	2521.0	2478.0	2474.0	2311.0	2614.0	2494.0
Google Phone	749.0	771.0	710.0	701.0	734.0	775.0	735.0
LG UltraGear Monitor	1010.0	1020.0	1046.0	1002.0	1004.0	999.0	1025.0
Lightning Charging Cable	3055.0	3067.0	3016.0	3134.0	3037.0	3215.0	2974.0
Macbook Pro Laptop	638.0	578.0	624.0	633.0	598.0	666.0	609.0
Samsung Galaxy Phone	277.0	275.0	271.0	269.0	256.0	253.0	258.0
Samsung Odyssey Monitor	841.0	802.0	793.0	817.0	772.0	838.0	883.0
USB-C Charging Cable	3101.0	3208.0	3094.0	3128.0	3054.0	3198.0	3183.0
iPhone	950.0	943.0	875.0	919.0	875.0	923.0	815.0

Day of the Week

Key Insights

Top-selling items (batteries, cables, Galaxy Buds) maintain strong, consistent daily sales, ideal for automation and stock stability..

➢ **Entertainment products (TVs, headphones)** spike over weekends, perfect for weekend campaign pushes.

➢ **Laptops peak early-week**, especially Mondays, cater to "back-to-work" energy with productivity promos.

# On which day of the week do customers make the most purchases?

Daily Sales
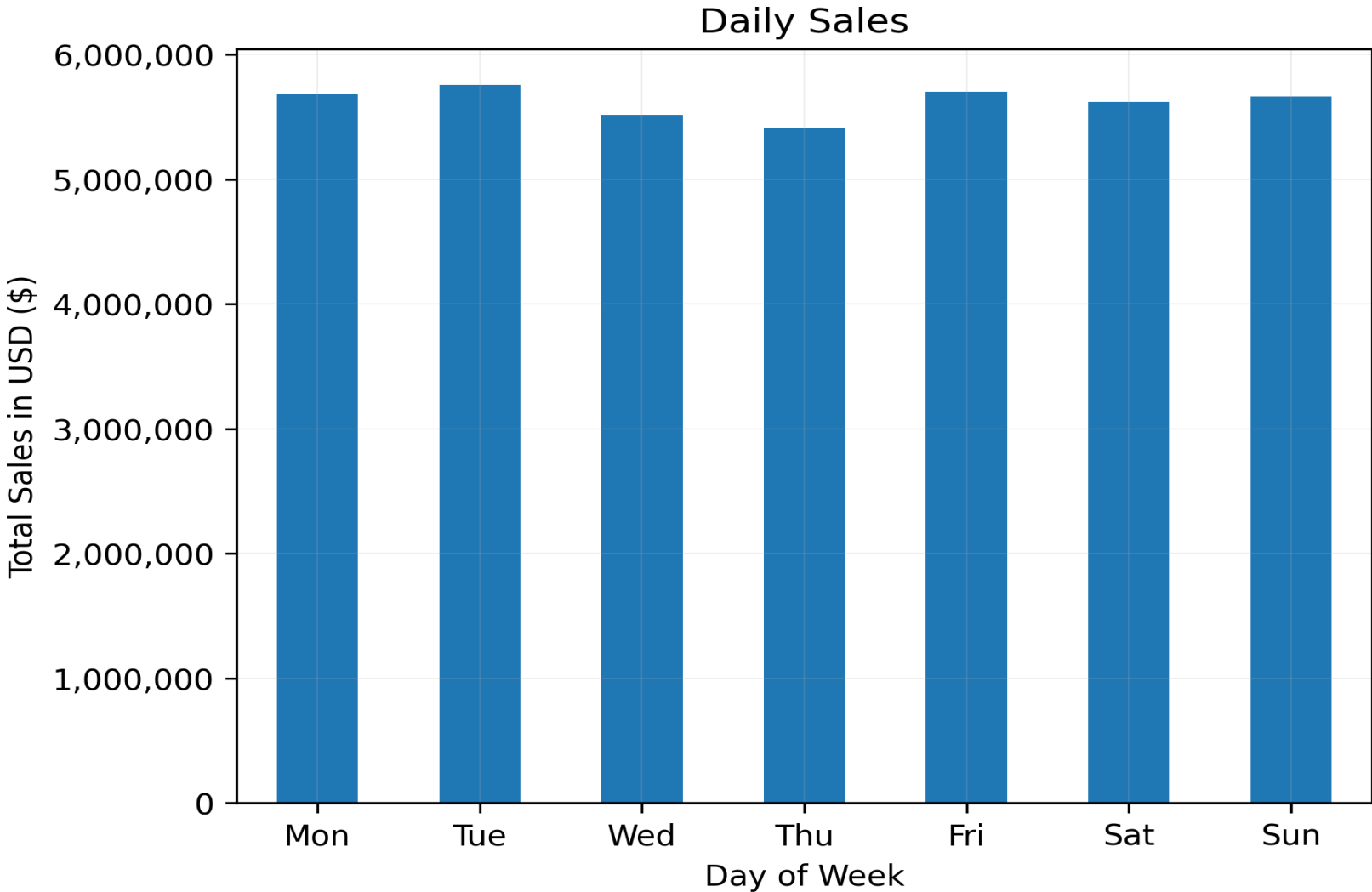
Key Insights

Tuesday is king, with the highest total sales (**$5.75M**), the best day for major campaign launches.

➢ **Thursday is the quietest** with total sales of **$5.41M**, revealing an opportunity for mid-week flash sales or loyalty nudges.

➢ **Sunday remains strong** with total sales of **$5.66M**, which may reflect increased leisure-time shopping behavior aligning with leisure-time browsing and impulse purchases.

What is the optimal timing for advertisements and promotions to maximize customer purchases, based on historical purchase behavior?

Hourly Purchase Trend
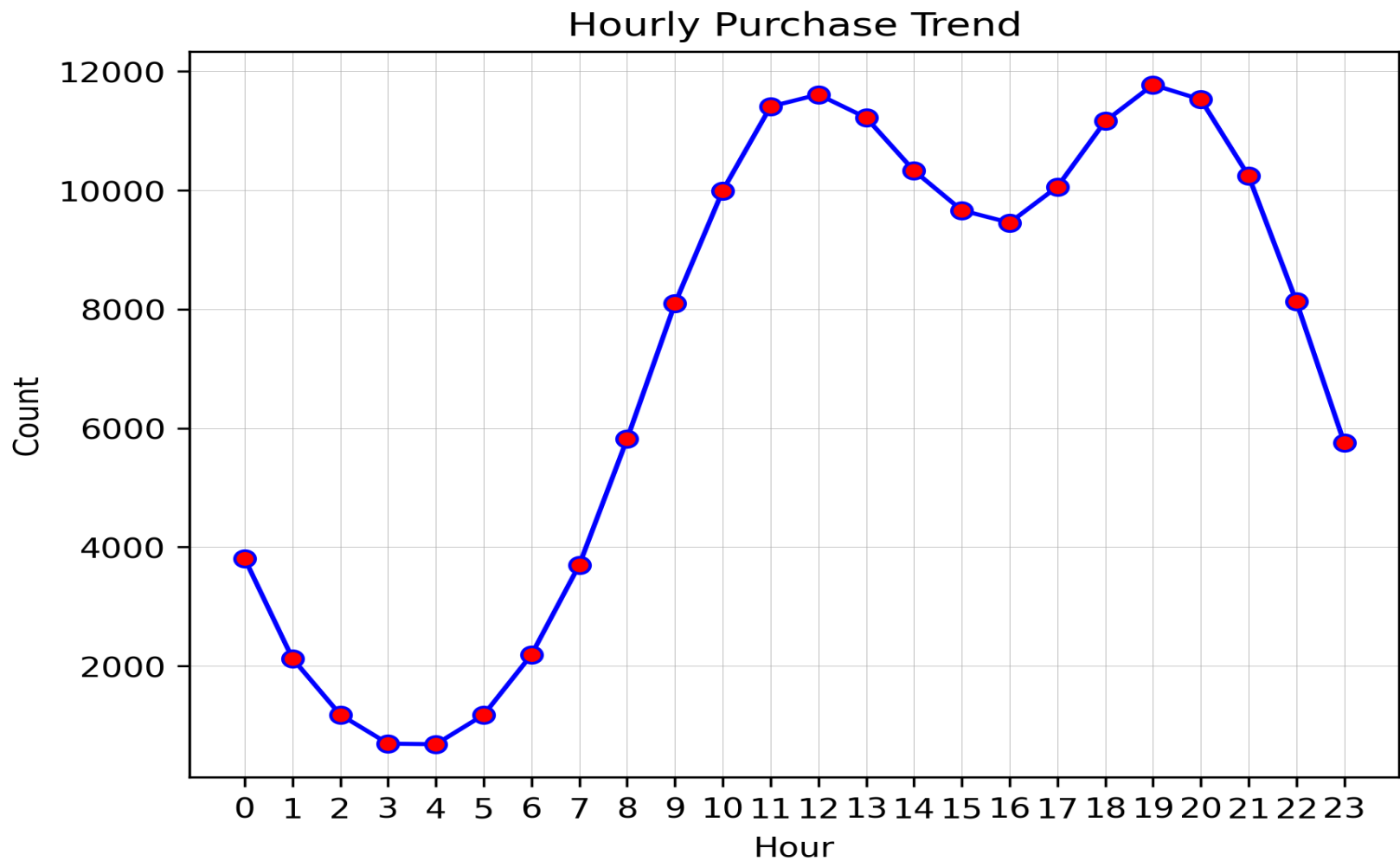
Key Insights

Peak activity falls between 11 AM–1 PM and 6–8 PM, making these golden hours for ad spend and campaign drops.

➢ **Lowest engagement happens between 2 AM–5 AM**, confirming minimal ROI for overnight marketing.

➢ **Purchasing behavior follows a daily rhythm**, ideal for scheduling messaging that mirrors buyer energy levels.

What will our monthly sales look like over the next two years based on historical purchasing patterns?

Prophet Forecast: Monthly Sales (Next 2 Years)
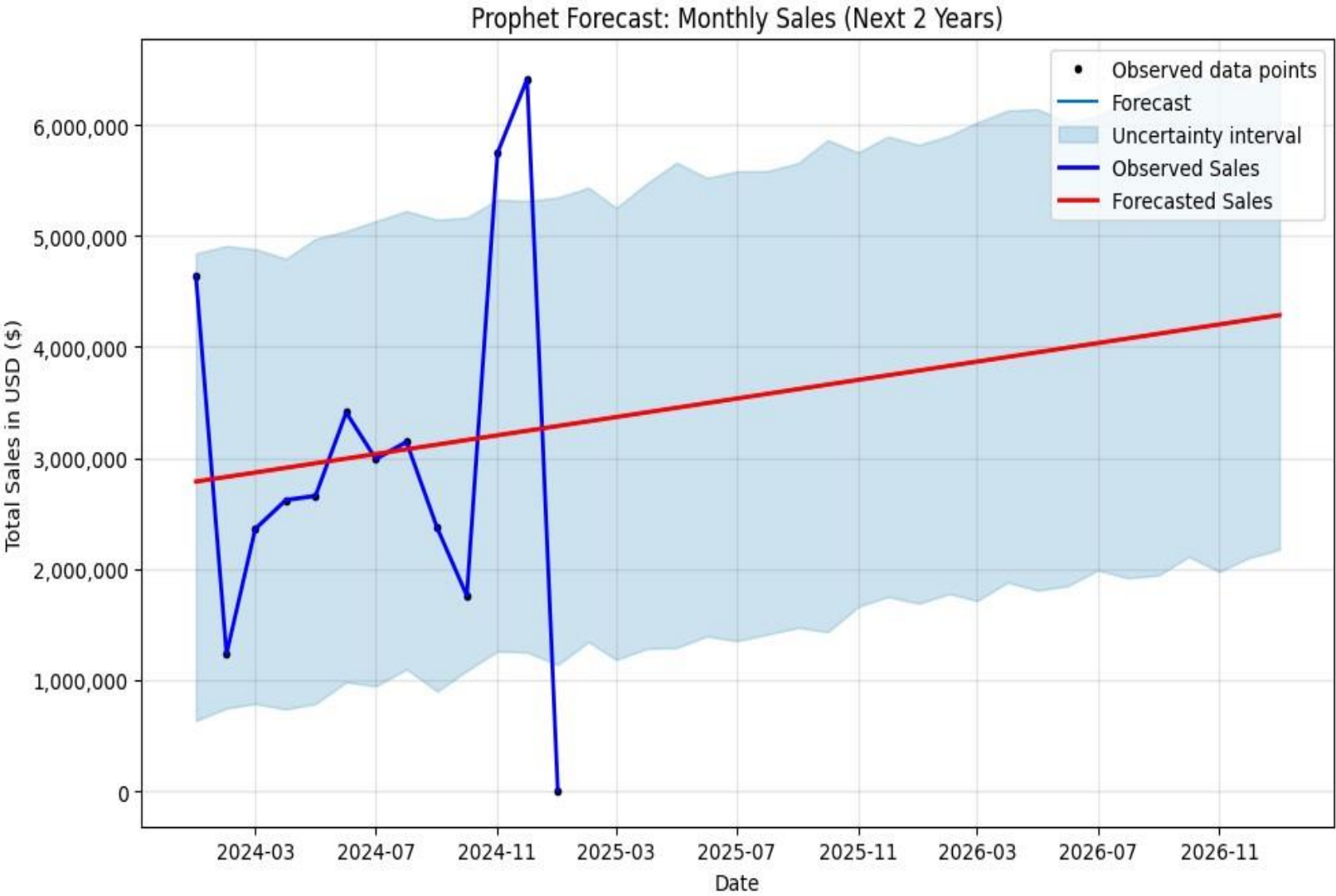
Key Insights

Monthly sales show a healthy upward trajectory, with approximately **$50K** average growth, momentum is strong.

➢ **Q4 spikes expected**, confirming need for early inventory and staff scaling.

➢ **Volatility projected in early 2025**, confidence intervals widen in **Feb–Apr**, suggesting agility will be critical.

What products are frequently bought together?

```
[(('Google Phone', 'USB-C Charging Cable'), 473),
 (('iPhone', 'Lightning Charging Cable'), 465),
 (('USB-C Charging Cable', 'Google Phone'), 464),
 (('Lightning Charging Cable', 'iPhone'), 453),
 (('Galaxy buds Headphones', 'iPhone'), 235),
 (('iPhone', 'Galaxy buds Headphones'), 221),
 (('Galaxy buds Headphones', 'Google Phone'), 197),
 (('Google Phone', 'Galaxy buds Headphones'), 176),
 (('USB-C Charging Cable', 'Samsung Galaxy Phone'), 175),
 (('Samsung Galaxy Phone', 'USB-C Charging Cable'), 172),
 (('Apple Airpods Headphones', 'iPhone'), 164),
 (('iPhone', 'Apple Airpods Headphones'), 157),
 (('Bose SoundSport Headphones', 'Google Phone'), 107),
 (('Google Phone', 'Bose SoundSport Headphones'), 98),
 (('Galaxy buds Headphones', 'USB-C Charging Cable'), 89),
 (('USB-C Charging Cable', 'Galaxy buds Headphones'), 79),
 (('Galaxy buds Headphones', 'Samsung Galaxy Phone'), 72),
 (('Galaxy buds Headphones', 'Lightning Charging Cable'), 70),
 (('Samsung Galaxy Phone', 'Galaxy buds Headphones'), 63),
 (('Lightning Charging Cable', 'Galaxy buds Headphones'), 62)]
```
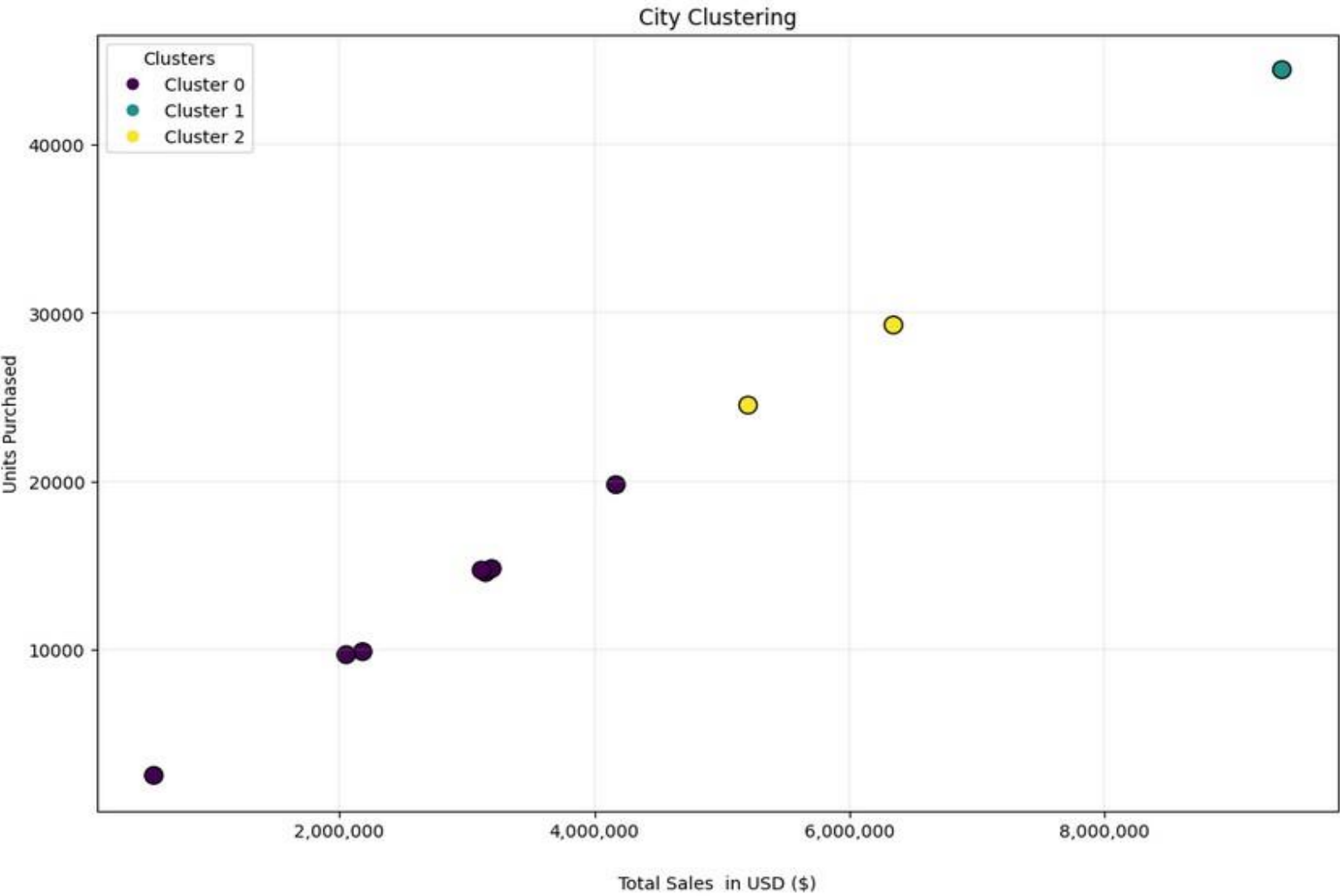
Key Insights

Accessory-first purchasing behavior (cables, chargers, earbuds) creates a **goldmine** for cross-sell strategies.

Strong Accessory Dependency:

➢ **Google Phone** is frequently bought with the **USB-C Charging Cable** (**473 times**), and **Samsung Galaxy Phone** with the same cable (**172 times**), confirming a strong dependency on accessories.

➢ A similar pattern was observed with the iPhone, frequently purchased alongside the Lightning Charging Cable (465 times) and Apple AirPods (157 times).

How can we group cities based on sales performance to uncover patterns for better sales strategy?

City Clustering

Key Insights

San Francisco (CA), the only city in **cluster 1**, demonstrates strong performance in both revenue (**$9.4M**) and **44.42K** unit sales.

➢ **Cluster 2 (Mid-tier)**: Stable, with room to grow, apply localized promotions and performance-based incentives.

➢ **Cluster 0 (Low-tier)**: Low sales, low volume, either rethink strategy or cut losses to improve ROI.

Strategic Recommendations

1. **Double down on Q4 strategies**: Scale marketing, inventory, and staffing to maximize November–December peak seasons.

2. **Create targeted product bundles**: Leverage high-frequency pairings (e.g., iPhone + Cable + AirPods) to boost Average Order Value (AOV).

3. **Optimize campaign timing**: Run high-impact promotions during peak activity windows (Tuesdays, 11 AM–2 PM, and 6–9 PM).

4. **Expand in top-tier cities**: Increase investment in high-performing cities like San Francisco, LA, and NYC.

5. **Revive low-performing regions**: Roll out localized campaigns in underperforming cities, and reassess long-term viability for continued investment.

6. **Forecast-driven planning**: Use Prophet-based sales forecasting to inform budgeting, procurement, and staffing plans.

7. **Leverage clustering insights**: Customize strategy by cluster to improve performance and resource allocation.

Expected Business Impact

1. Growth & Profitability

➢ Achieve **15–20% revenue uplift** through strategic scaling during Q4 and bundling strategies.

➢ Improve profit margins by **8–12%** via product mix optimization and smarter inventory alignment.

2. Operational Efficiency

➢ Boost forecasting accuracy by **25–30%** using Prophet modeling, enhancing demand planning and budgeting.

➢ Reduce profit volatility by **10–15%** through better timing of promotions and cost control.

3. Market Expansion & Sales Optimization

➢ Increase regional market penetration by **10–14%** through localized promotions in mid- and low-tier clusters.

➢ Raise cross-/upsell revenue by **12–15%** using market basket insights to personalize product recommendations and bundles.

Conclusion

This project demonstrates how the fusion of **Exploratory Data Analysis** and **Machine Learning** can transform raw sales data into **clear, actionable strategies** that fuel business growth.

By uncovering patterns across **time, geography, products, and customer behavior**, the project effectively answers key business questions, identifying **top-performing months, cities, and products**, while spotlighting hidden opportunities in **product bundling**, **promotion timing**, and **customer segmentation**.

But this project goes beyond just visualizing what happened, it **strategically connects insights to execution**, using **Python-powered analytics, predictive modeling, and real business acumen**. It equips stakeholders with **intelligent, scalable recommendations** that optimize marketing, streamline inventory, and future-proof decision-making.

BALOGUN OLUMIDE CHRIS CONTACTS

Tel: +234(0)8065060691

Email: krisbalo11@gmail.com

Github - https://github.com/olumidebalogun1

LinkedIn -https://www.linkedin.com/in/olumide-balogun1/

Medium - https://medium.com/@Olumide-Balogun

X - https://x.com/IAmOluBalogun

Olumide Balogun   +234-8065060691