

Customer and Product Analytics: Analyzing Modecraft Retail Sales



Introduction

This project explores transactional data to uncover key business insights around customer behavior, product performance, and sales trends.

Using tools like Tableau and Python, we analyzed patterns across countries, product categories, and customer segments.

The goal was to identify growth opportunities, improve retention, and support data-driven decision-making for marketing and operations.

Dataset Used

This project analyzes anonymized e-commerce transaction data from Modecraft, provided for the Dubstech x University of Washington Datathon 2025. The dataset includes order-level details like invoice number, product code, description, quantity, price, date, customer ID, and country.

- Context: Provided by Dubstech x University of Washington Datathon 2025
- Fields: InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID, Country
- Raw Data (CSV): https://github.com/maicorebong/Datathon2025-DubsTechUW/blob/main/modecraft_ecommerce_compressed_data.gz

Data Pipeline Architecture

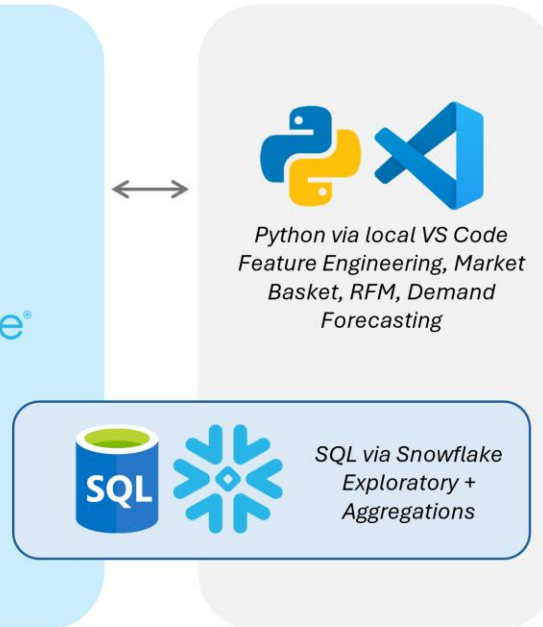
GENERATION + STORAGE



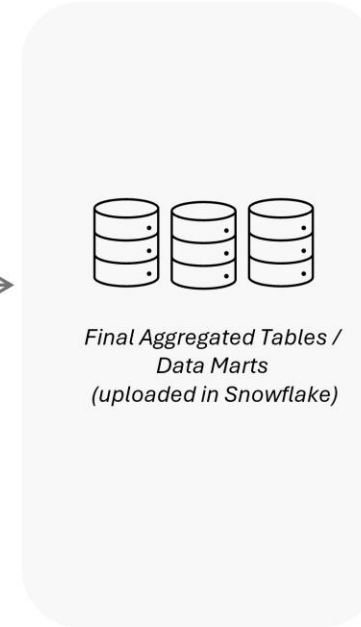
STAGING



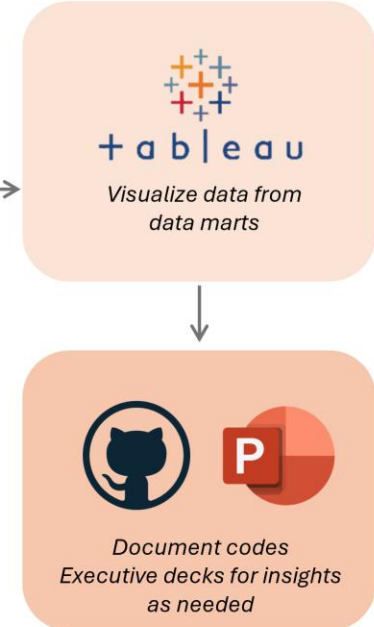
TRANSFORMATION



SERVING

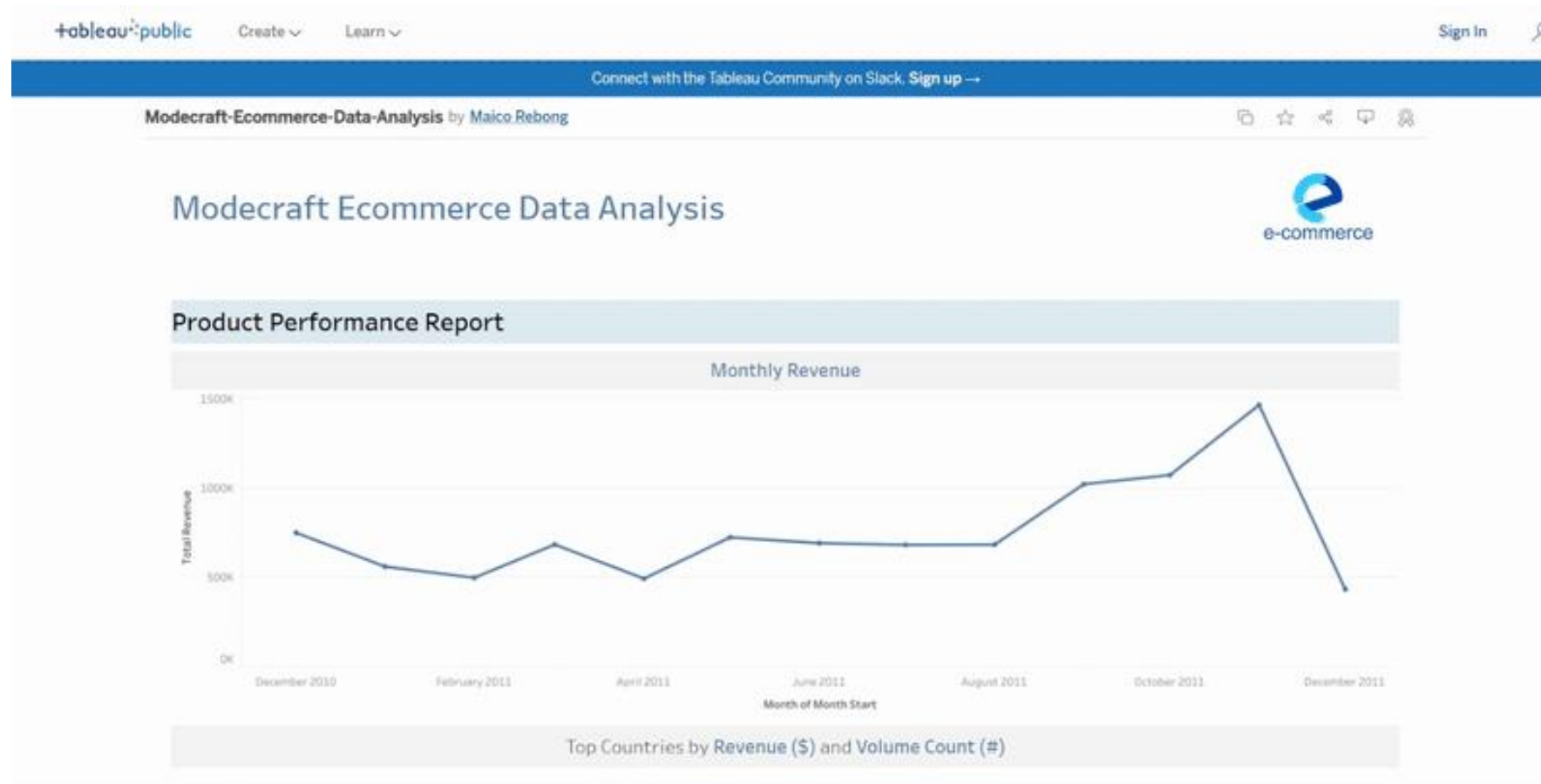


BI & INSIGHTS

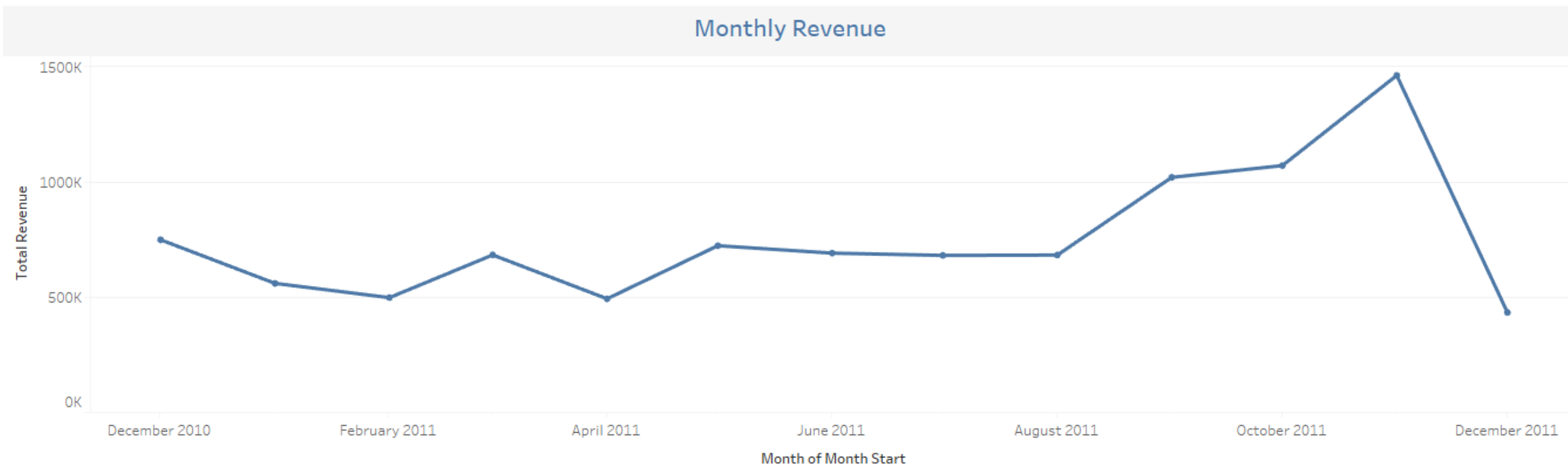


Dashboard

Dashboard Demo: [Tableau link](#)



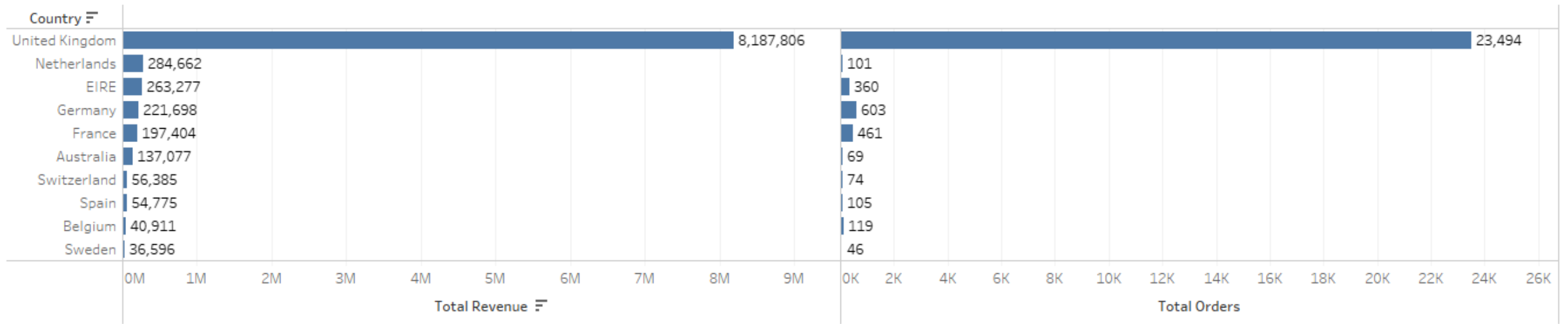
Seasonal Performance Patterns



- Revenue stayed relatively flat from January to August 2011, averaging around \$500K–\$600K
- Strong increase began in September with revenue peaking at nearly \$1.5M in November
- December saw a sharp decline, likely due to lower post-holiday demand or limited inventory

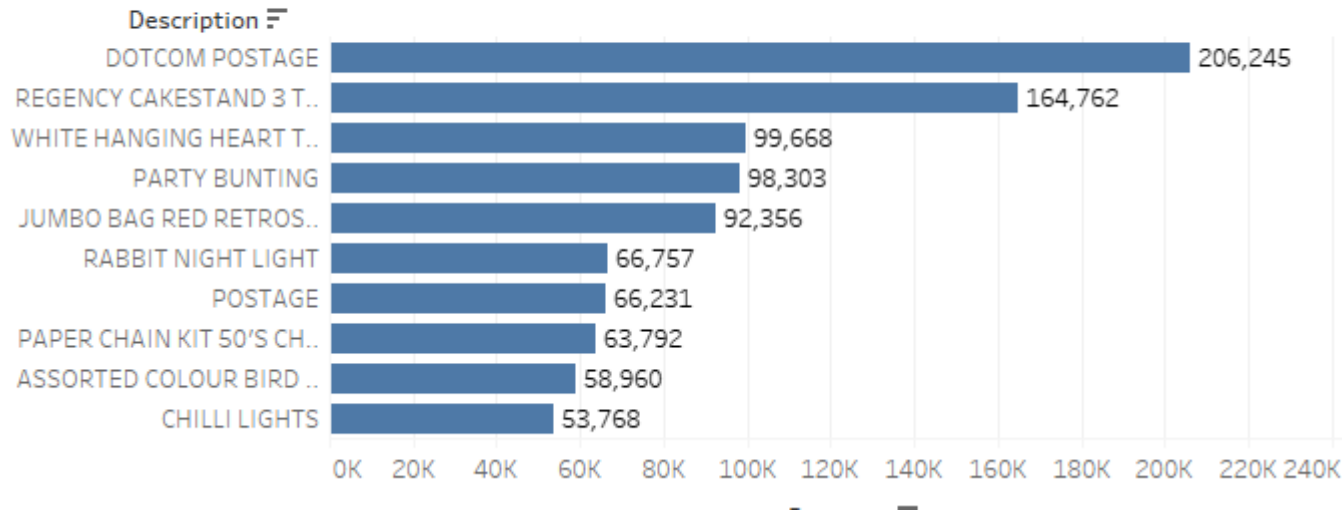
Focus Regions on for Sales and Marketing

Top Countries by Revenue (\$) and Volume Count (#)

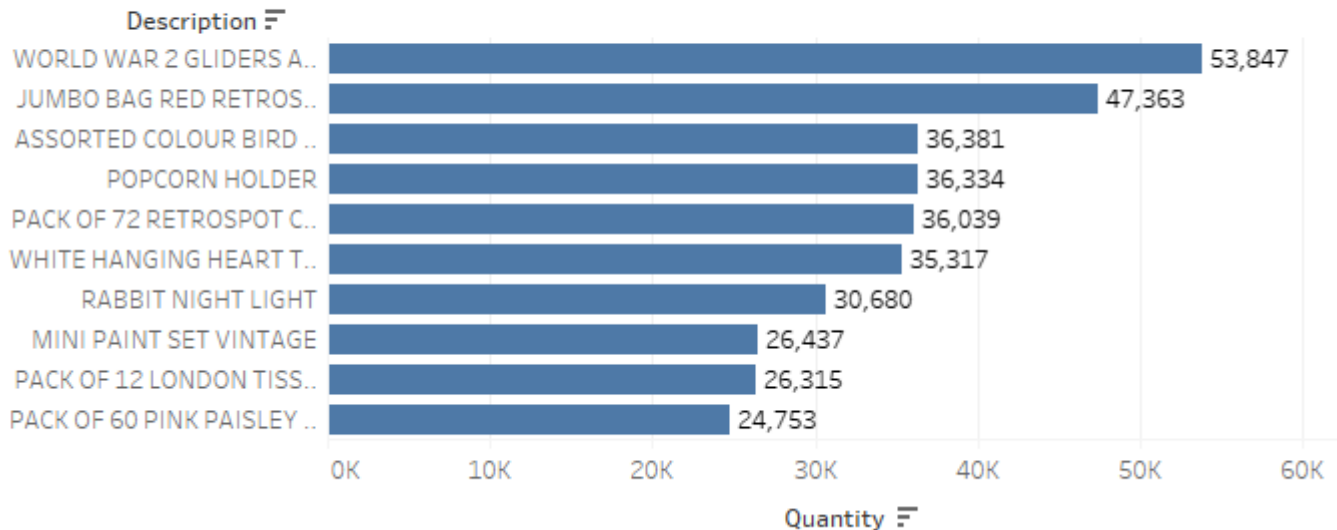


- United Kingdom generated the highest sales with over \$8M in revenue and 23K+ orders
- Germany and EIRE shows higher average sales order value, based on its strong revenue despite fewer orders. Possible growth market.

Top Products by Revenue



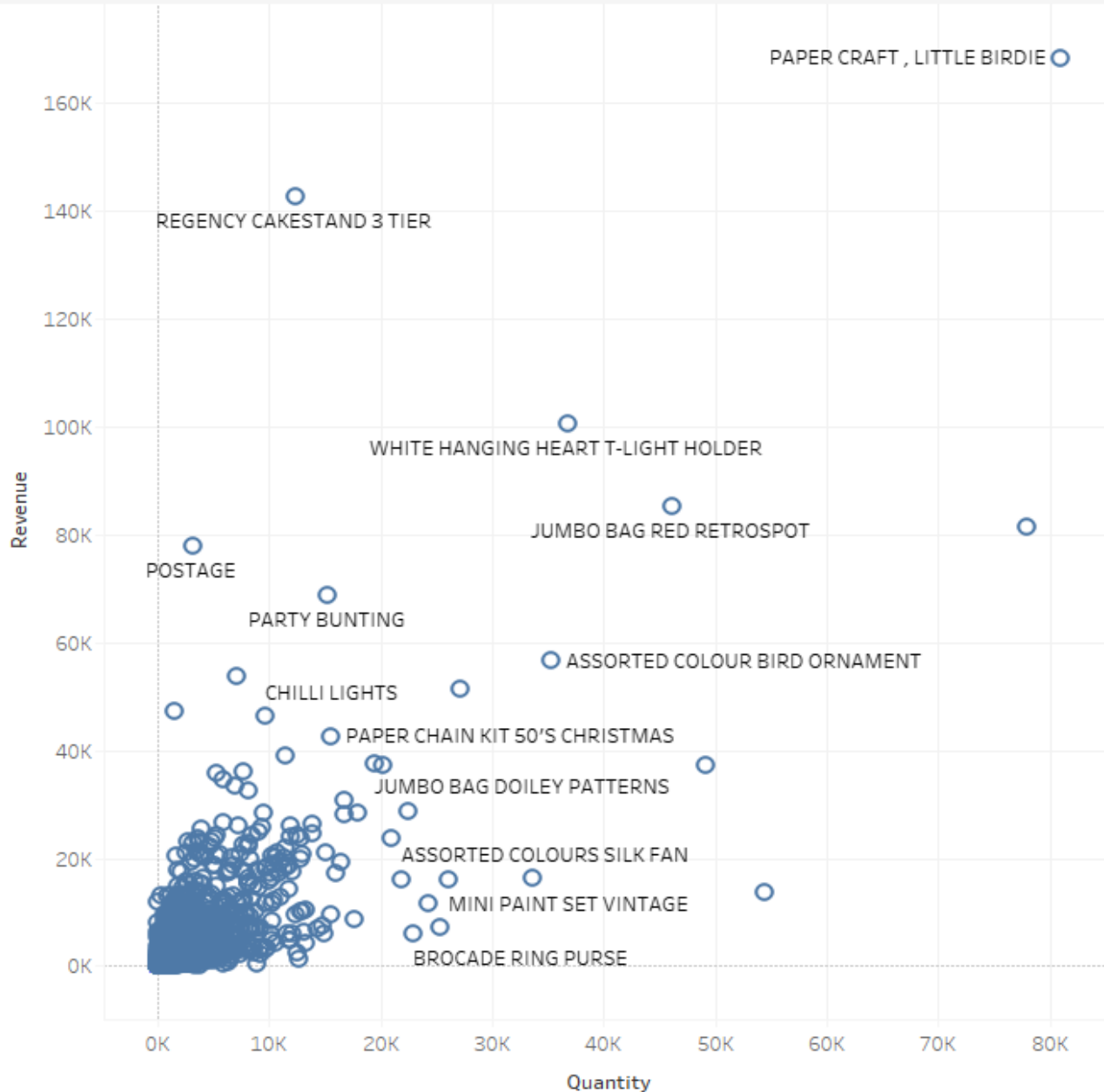
Top Products by Volume Count



Revenue & Volume Drivers

- Dotcom Postage delivered the highest revenue at over \$200K, showing strong earnings from logistics-related charges
- Regency Cake Stand and White Hanging Heart T-Light Holder also performed well in revenue even with lower sales volume
- Products like World War 2 Gliders and Jumbo Bag Red Retrospot had high sales volume but lower revenue which may suggest lower price points or margins

Product Quadrant Analysis



Revenue vs. Quantity Analysis

- 'Paper Craft, Little Birdie' is a clear top product which performed well in both sales volume and revenue
- Items like 'Regency Cake Stand' and 'T-Light Holder' had high revenue with fewer units, showing potential as premium products and may be considered for better promotions
- Products identified with low sales and revenue should be reviewed for possible overstock or phase out

Bundling and Cross-sell Recommendations

Frequent Pairs Analysis

546 JUMBO BAG PINK POLKADOT	523 LUNCH BAG PINK POLKADOT	468 JUMBO BAG DOILEY	465 LUNCH BAG BLACK SKULL. -	464 LUNCH BAG BLACK SKULL. -	463
541 GREEN REGENCY TEACUP AND	519 LUNCH BAG RED RETROSPOT -	468 WOODEN FRAME	460 GREEN REGENCY		457 RED
530 ALARM CLOCK BAKELIKE GREEN	517 LUNCH BAG BLACK SKULL. -	467 LUNCH BAG RED	459 LUNCH BAG CARS		

- Multiple "lunch bag" variants appear together in baskets which suggests style-based buying behavior
- Items like Jumbo Bag Pink Polkadot and Lunch Bag Pink Polkadot were repeatedly bought together indicating customers value matching sets
- Cross-pairings with items like wooden frames and teacups suggest customers are also combining decorative and practical items

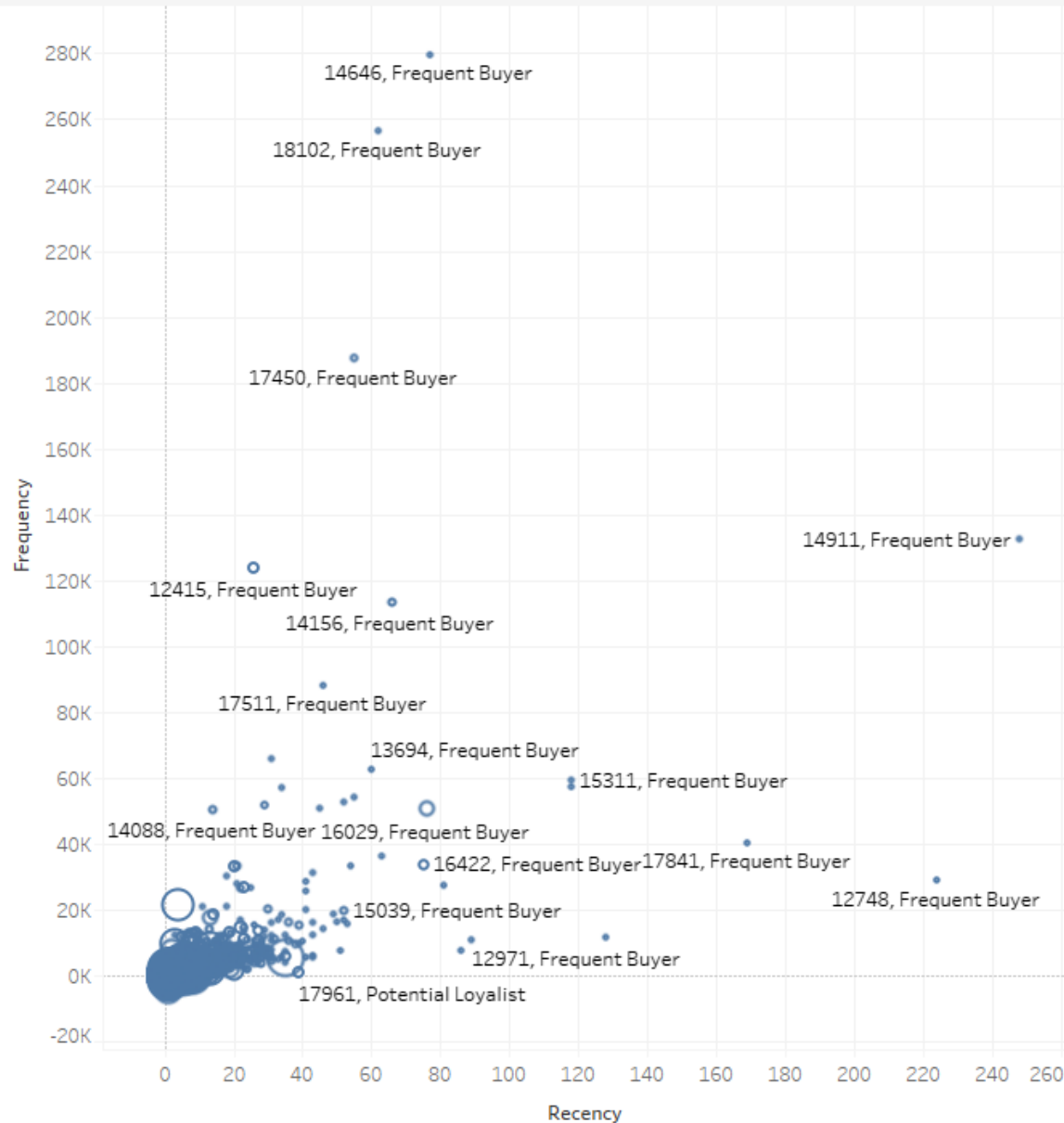
Product Ranking

Products by Health Score

Description	Rank	Healthscore	Revenue	Quantity
PAPER CRAFT , LITTLE BIRDIE	1	2	168,470	80,995
MEDIUM CERAMIC TOP STORAGE JAR	2	2	81,417	77,916
JUMBO BAG RED RETROSPOT	4	2	85,221	46,181
WHITE HANGING HEART T-LIGHT HOLDER	4	2	100,604	36,782
ASSORTED COLOUR BIRD ORNAMENT	5	2	56,580	35,362
RABBIT NIGHT LIGHT	6	2	51,346	27,202
SMALL POPCORN HOLDER	7	2	37,226	49,183
JUMBO BAG DOILEY PATTERNS	8	2	37,684	19,516
JUMBO BAG PINK POLKADOT	9	2	37,290	20,165
PARTY BUNTING	10	2	68,844	15,291
PAPER CHAIN KIT 50'S CHRISTMAS	12	2	42,661	15,617
VICTORIAN GLASS HANGING T-LIGHT	12	2	28,777	22,433
JUMBO BAG STRAWBERRY	13	2	30,644	16,807

- “Paper Craft, Little Birdie” consistently led in both revenue and quantity as a well-rounded performer
- Some products like “Small Popcorn Holder” and “Jumbo Bag Doiley Patterns” had good sales but may not rank high in profitability which suggests need for margin review
- Lower-ranked products still showed solid volume, indicating room for repositioning or strategic discounts

RFM Customer Scatterplot



RFM Analysis

- Most high-frequency buyers have low recency values which indicates they made purchases recently and consistently
- A few outliers with high frequency but high recency may be disengaging, despite previously being strong customers
- The cluster in the lower-left shows a large number of customers who bought recently but infrequently similar to one-time or trial buyers
- Customer ID 14911 stands out for high recency and frequency, signaling strong past loyalty but potential risk of churn

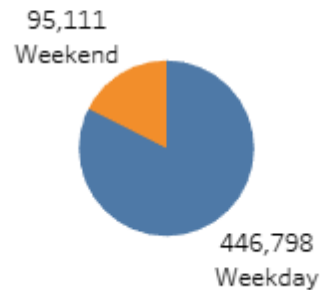
Customer Types and Order Timing Patterns

Customer Segmentation

Segment	
VIP	17
Frequent Buyer	636
Potential Loyalist	1,569
Lost Customer	2,150

- Lost Customers are the largest group, showing high churn
- Only 17 VIPs exist but likely generate strong value per order
- Potential Loyalists are promising and worth targeting for reactivation

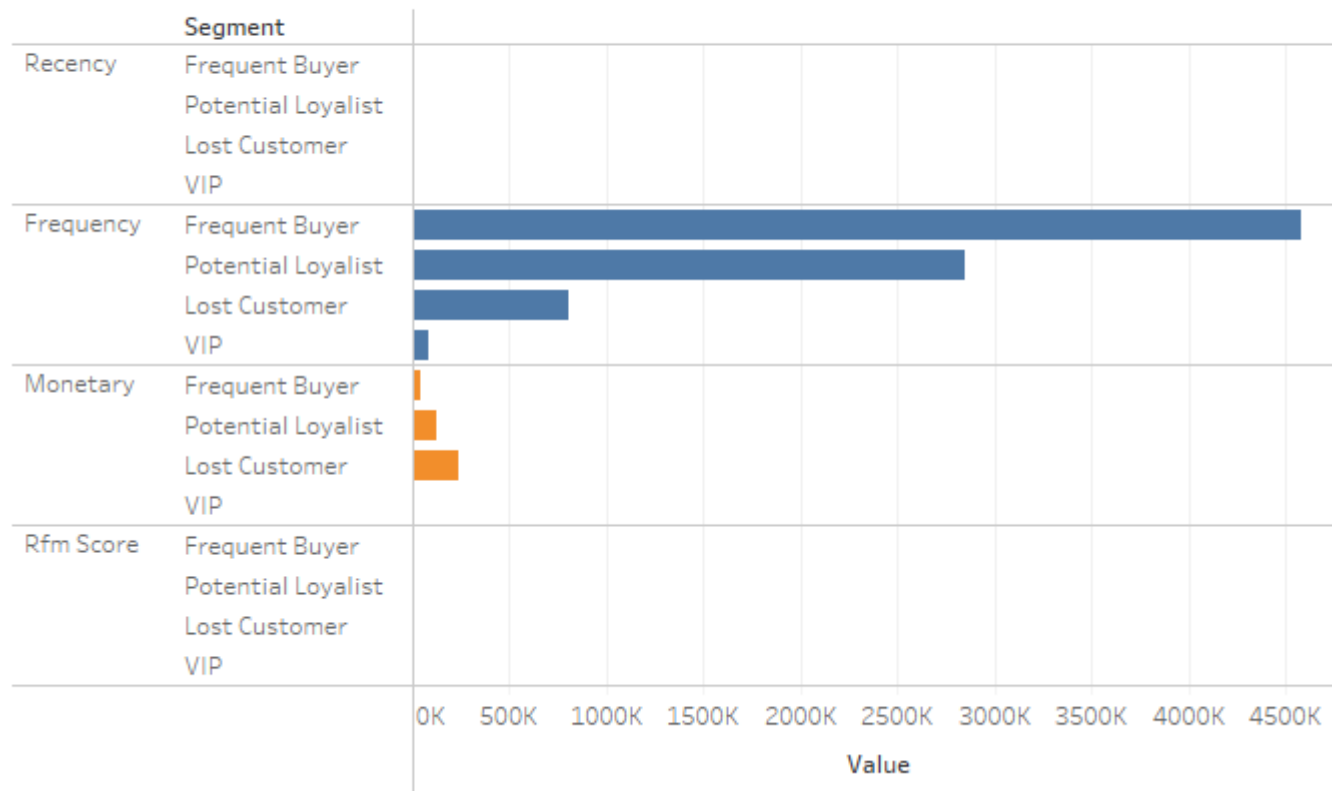
Orders by Weekday vs. Weekend



- Over 80% of orders happen on weekdays with strong traffic
- Weekend sales are lower with less business or browsing activity
- Weekday-focused campaigns could drive higher conversions

RFM Trends

Recency, Frequency, and Monetary Trends by Segment



- Frequent Buyers lead all segments which confirms repeat-purchase behavior
- Lost Customers surprisingly lead in monetary value indicating they once had high-value transactions
- VIP customers are few but show the highest average revenue per customer even with lower frequency
- Potential Loyalists offer the best opportunity for growth as they already show moderate frequency and value

End