

# Final Capstone Presentation

Rhea Nair, Samiksha Sarda, Sana Sawant, Vedika Khandelwal

MGT 5905 - Business Analytics Capstone

Spring 2025

# Business Understanding

Our goal is to build a comprehensive customer intelligence study that not only describes what is happening in the business, but diagnoses why it's happening, prescribes what should be done next, and predicts what lies ahead.

Key strategic questions include:

1. What products should be sold to which countries, and during which time periods, to maximize sales?
2. How do customer behaviors vary by region, season and product type?
3. Which products are often purchased together, how can this support effective bundling strategies?
4. How can we better understand and serve different customer segments based on their value contribution?

# Data Cleaning and Preprocessing

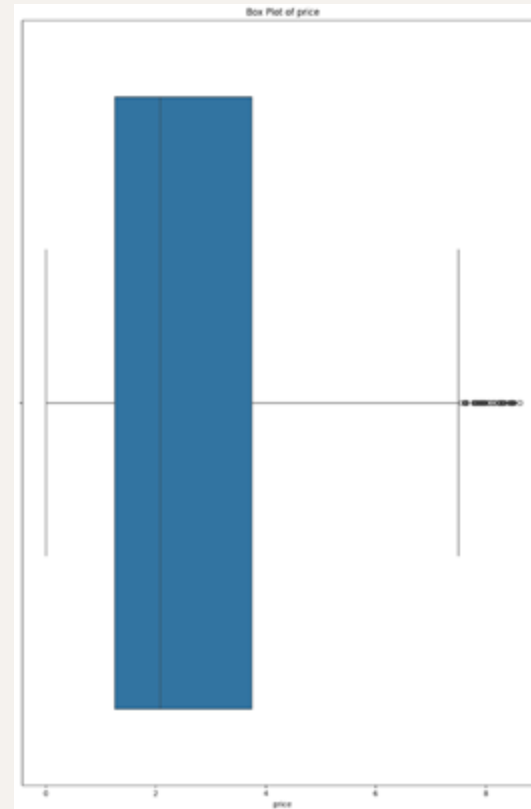
This Online Retail II data set contains comprises of 1,033,036 transactions occurring for a UK-based and registered, non-store online retail between Dec'09 and Dec'11. The company mainly sells unique all-occasion giftware to wholesalers.

Challenges of the dataset included missing values in Customer ID and product descriptions, illegible product descriptions ("missing"), non-Product ID codes, negative quantity and prices.

We removed rows with Illegible Product Descriptions for e.g., "????", "?damage??", "missing", non-product Stock Codes for eg,B (bad debt), C (canceled) and negative or zero Quantity/Price

We then created a new variable,  $Sales = Quantity \times UnitPrice$ , and filtered for positive-value transactions to focus on valid purchases

# Exploratory Data Analysis



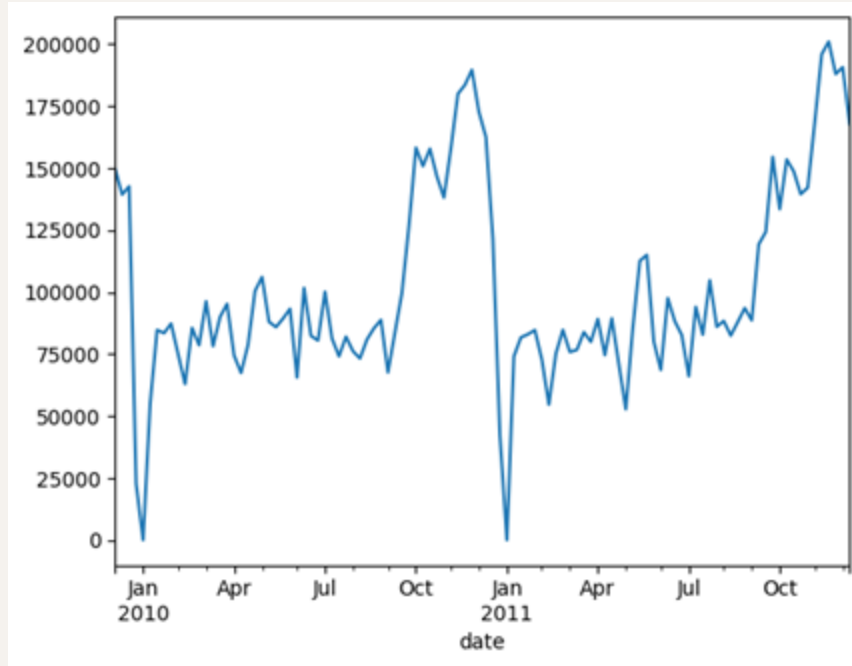
Outliers removed: 140,870

# Identifying High-Grossing Countries by Sales Volume

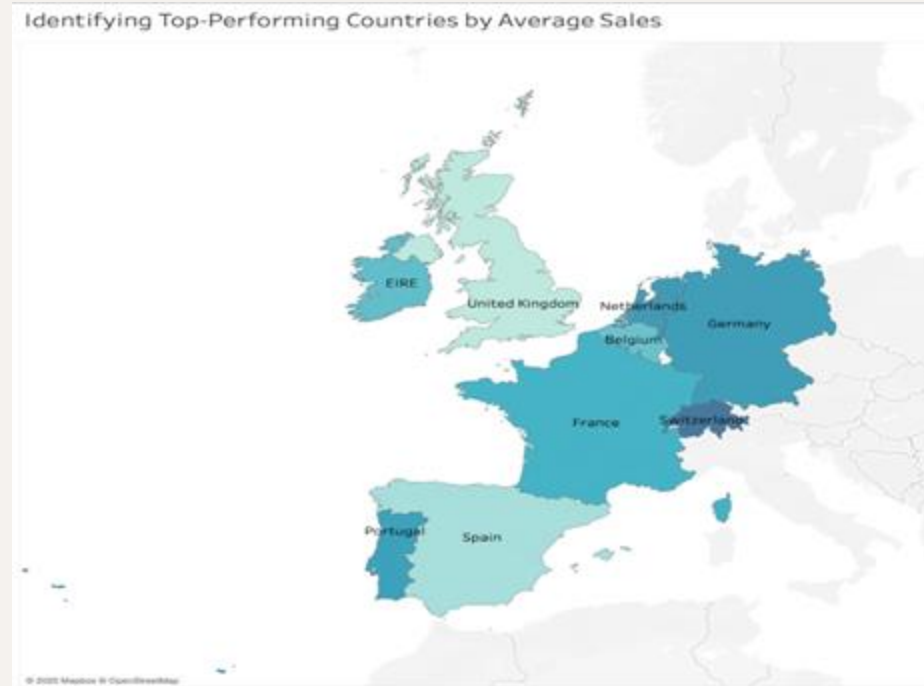


# Data Analysis & Strategy, Understanding the Why?

Avg Daily Sales by Month



Avg Sales by Country



## Popular Product in All At A Glance

Country

WOODLAND STICKERS  
RED/WHITE DOT MINI CASES  
RED RETROSPOT BOWL  
SET OF 3 HEART COOKIE CUTTERS  
DOORMAT RED SPOT CHARLOTTE BAG  
ASSORTED COLOUR BIRD ORNAMENT  
RED SPOTTY BISCUIT TIN  
NO SINGING METAL SIGN  
FELTCRAFT 6 FLOWER FRIENDS  
OFFICE MUG WARMER PINK  
SPACEBOY BIRTHDAY CARD  
JUMMO BAG PINK VINTAGE PAISLEY  
ASSORTED COLOUR MINI CASES  
RED SPOTTY CHILD'S UMBRELLA  
JUMMO BAG PINK POLKADOT  
RETROSPOT CHILDREN'S APRON  
SET/6 RED SPOTTY PAPER CUPS  
ROUND CONTAINER SET OF 5 RETROSPOT  
VINTAGE HEADS AND TAILS CARD GAME  
CHILDREN'S APRON SPACEBOY DESIGN  
PINK 3 PIECE POLKADOT CUTLERY SET  
LUNCH BOX WITH CUTLERY RETROSPOT  
STRAWBERRY CERAMIC TRINKET BOX  
CHILDREN'S CUTLERY POLKADOT PINK  
COOK WITH WINE METAL SIGN  
LUNCH BAG PINK RETROSPOT  
OFFICE MUG WARMER (CHOC+BLUE)  
LUNCH BAG PINK RETROSPOT  
OFFICE MUG WARMER POLKADOT  
ZINC METAL HEART DECORATION  
BATHROOM METAL SIGN  
PACK OF 6 SKULL PAPER PLATES  
LUNCH BAG SPACEBOY DESIGN  
DOORMAT UNION FLAG  
GREEN SPOTTY PLATE  
JUMMO BAG TOYS  
PARTY BUNTING  
WALL TIDY RETROSPOT  
RED SPOTTY CHILD'S APRON  
JUMMO STORAGE BAG WOODLANDS  
BLUE POLKADOT PLATE  
JUMMO BAG OWLS  
RED RETROSPOT PICNIC BAG  
GREEN SPOTTY BOWL  
BLUE SPOTTY PLATE  
PAPER BUNTING RETROSPOT  
PINK HAPPY BIRTHDAY BUNTING  
BLUE HAPPY BIRTHDAY BUNTING  
PLEASE ONE PERSON METAL SIGN  
PAINTED METAL PEARLS ASSORTED HEART  
IVORY TRELLIS LARGE  
CHARLOTTE BAG, SUKI DESIGN  
WASH BAG VINTAGE ROSE PAISLEY  
BAKING SET 9 PIECE RETROSPOT  
RETRO SPOT TEA SET CERAMIC 11 PC  
HOT BATHS METAL SIGN  
SINGLE HEART ZINC T-LIGHT HOLDER  
SWEETHEART CERAMIC TRINKET BOX  
POTTERING IN THE SHED METAL SIGN  
GREEN POLKADOT PLATE  
PINK CREAM FELT CRAFT TRINKET BOX  
PINK 3 PIECE MINI DOTS CUTLERY SET  
RINGO SET  
RED SPOTTY PLATE  
HANGING HEART ZINC T-LIGHT HOLDER  
HANGING HEART JAR T-LIGHT HOLDER  
BLACK/BLUE DOTS RUFFLED UMBRELLA  
STRAWBERRY CHARLOTTE BAG  
GIN + TONIC DIET METAL SIGN  
DOORMAT SPOTTY HOME SWEET HOME  
RETROSPOT TEA SET CERAMIC 11 PC  
LUNCH BAG SPACEBOY DESIGN  
JUMMO BAG RED WHITE SPOTTY  
ROUND SNACK BOXES SET OF 4 WOODLAND  
11 PC CERAMIC TEA SET POLKADOT  
BLACK/BLUE POLKADOT UMBRELLA  
STRAWBERRY LUNCHBOX WITH CUTLERY  
CREAM HANGING HEART T-LIGHT HOLDER  
JUMMO BAG STRAWBERRY  
JUMMO SHOPPER VINTAGE RED PAISLEY  
JUMMO BAG BAROQUE BLACK WHITE  
CHARLOTTE BAG SUKI DESIGN  
LOVEBIRD HANGING DECORATION WHITE  
SPOTTY HOME SWEET HOME DOORMAT  
RED HANGING HEART T-LIGHT HOLDER  
CHILDREN'S CUTLERY RETROSPOT RED  
SET OF 3 BUTTERFLY COOKIE CUTTERS  
SMALL GLASS HEART TRINKET POT  
DOOR MAT UNION FLAG  
STRAWBERRY CERAMIC TRINKET POT  
LUNCH BAG SUKI DESIGN  
JUMMO BAG SCANDINAVIAN PAISLEY  
RED 3 PIECE MINI DOTS CUTLERY SET  
SCANDINAVIAN PAISLEY PICNIC BAG  
CHARLOTTE BAG PINK POLKADOT  
LUNCH BAG PINK POLKADOTS  
LUNCH BAG SUKI DESIGN  
GIN AND TONIC DIET METAL SIGN  
PAPER CHAIN KIT 50'S CHRISTMAS  
PLEASE ONE PERSON METAL SIGN  
STRAWBERRY SHOPPER BAG  
BLUE POLKADOT BOWL  
JUMMO BAG SCANDINAVIAN PAISLEY  
RED 3 PIECE MINI DOTS CUTLERY SET  
SCANDINAVIAN PAISLEY PICNIC BAG  
CHARLOTTE BAG PINK POLKADOT  
LUNCH BAG SUKI DESIGN  
GIN AND TONIC DIET METAL SIGN  
PAPER CHAIN KIT 50'S CHRISTMAS  
PLEASE ONE PERSON METAL SIGN  
STRAWBERRY SHOPPER BAG  
BLUE POLKADOT BOWL  
JUMMO BAG TOYS  
6 RIBBONS RUSTIC CHARM  
LUNCH BAG SUKI DESIGN  
DOOR MAT RED SPOT  
VINTAGE SNAP CARDS  
JUMMO BAG SCANDINAVIAN PAISLEY  
RED 3 PIECE MINI DOTS CUTLERY SET  
SCANDINAVIAN PAISLEY PICNIC BAG  
CHARLOTTE BAG PINK POLKADOT  
LUNCH BAG SUKI DESIGN  
GIN AND TONIC DIET METAL SIGN  
PAPER CHAIN KIT 50'S CHRISTMAS  
PLEASE ONE PERSON METAL SIGN  
STRAWBERRY SHOPPER BAG  
BLUE POLKADOT BOWL  
JUMMO BAG TOYS  
6 RIBBONS RUSTIC CHARM  
LUNCH BAG SUKI DESIGN  
DOOR MAT RED SPOT  
VINTAGE SNAP CARDS

# Descriptive Statistics

Item-level analysis data

	invoiceId	stockCode	description	quantity	date	price	customerId	country	sales
0	489434	85048	15CM CHRISTMAS GLASS BALL 20 LIGHTS	12	2009-12-01 07:45:00	6.95	13085.0	United Kingdom	83.4
1	489434	79323P	PINK CHERRY LIGHTS	12	2009-12-01 07:45:00	6.75	13085.0	United Kingdom	81.0
2	489434	79323W	WHITE CHERRY LIGHTS	12	2009-12-01 07:45:00	6.75	13085.0	United Kingdom	81.0
4	489434	21232	STRAWBERRY CERAMIC TRINKET BOX	24	2009-12-01 07:45:00	1.25	13085.0	United Kingdom	30.0
5	489434	22064	PINK DOUGHNUT TRINKET POT	24	2009-12-01 07:45:00	1.65	13085.0	United Kingdom	39.6

Quartile for sales

	sales
0.00	0.001
0.25	3.400
0.50	8.300
0.75	15.900
1.00	237.160

Descriptive statistics for sales

	sales_mean	sales_median	quantity_mean	quantity_median	price_mean	price_median
stockCode						
10002	7.220564	6.800	8.141066	6.0	1.002978	0.85
10080	4.095185	4.680	9.888889	12.0	0.509259	0.39
10109	1.680000	1.680	4.000000	4.0	0.420000	0.42
10120	0.909000	0.735	4.066667	3.0	0.246667	0.21
10125	7.296196	3.320	8.791411	3.0	0.940859	0.85



# Customer Categorization- Recency Frequency and Monetary Analysis

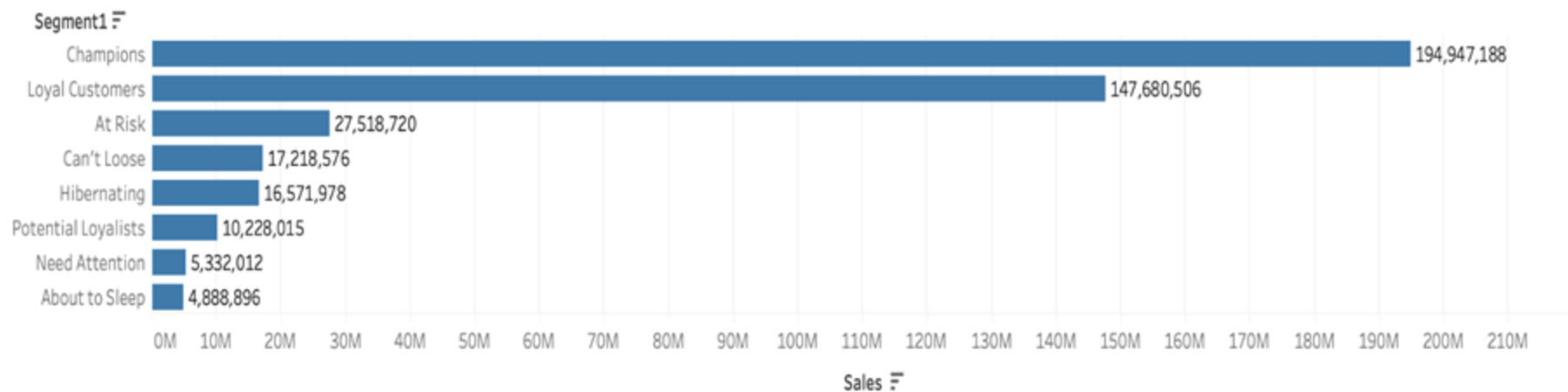
RFM Analysis is a customer value based segmentation technique used in marketing and analytics. It categorizes customers based on their purchasing behavior using three key factors- Recency, Frequency and Monetary.

*We chose RFM because it is well-suited for transactional retail data, requires minimal assumptions, and provides intuitive, actionable results for marketing teams. The diagnostic nature of RFM allowed us to pinpoint both strengths (loyal buyers) and opportunities (lapsed or newly acquired customers) at a very granular level, helping inform who to target.*

# Business Insights From RFM

- In the UK, we identified a high-engagement customer base characterized by frequent and recent purchases, but with a lower average order value. This suggests a loyal customer segment with potential to increase basket size through cross-sell or volume incentives.
- In Ireland and the Netherlands, purchasing behavior was consistent, but transaction value remained low. This indicates an opportunity to drive revenue growth through strategic bundling or targeted upselling campaigns.
- In Switzerland, we observed a wave of growing potential loyalists exhibiting recent but infrequent purchase activity. This signals an early-stage engagement window that can be leveraged through onboarding journeys and retention-focused offers.
- The RFM segmentation uncovered both high-performing customer cohorts and untapped growth segments, enabling a data-driven foundation for precision marketing, reactivation efforts, and segment-specific campaign planning.

## Champions Lead the Way: Segment-Wise Sales Performance



# Number of Customers By Segment



# Product Bundling- Apriori Market Basket

The Apriori algorithm is used for association rule mining that identifies frequent itemsets and association rules in a dataset, particularly useful for applications like market basket analysis.

It identifies which items are frequently bought together by analyzing transactions using three key concepts- Support (How often items appear together), Confidence (How often item B appears after item A) and Lift – How likely is the bundle an association and not a coincidence.

*We opted for Apriori due to its interpretability, wide adoption in retail analytics, and ability to surface meaningful product affinities across thousands of SKUs. It highlights the cultural depth of buyer preferences and the prescriptive recommendations are especially valuable when linked back to RFM segments, allowing us to suggest the right bundle for the right customer.*

# Business Insights

- In the UK, customers frequently purchased decorative and gift items such as *t-light holders*, *patterned lunch bags*, and *novelty storage bags*. However, these were often bought individually, not as sets. This reflects a high engagement, low basket value pattern. Given this behavior, there's a strong opportunity to introduce thematic bundles or multi-buy offers to increase average order value without disrupting existing buying habits.
- In France and Germany, purchasing patterns leaned toward practical home goods—including *night lights*, *snack boxes*, and *utility tins*—often bought in sets of 4 or 6. This indicates a value-conscious, utility-driven customer base, where promoting multi-unit packs or function-based kits could enhance perceived value and margin.
- By combining Apriori associations with RFM segmentation, we crafted bundling strategies that matched both product affinity and customer value tier. This allowed us to deploy geography-specific, segment-personalized product recommendations—increasing the likelihood of conversion, basket lift, and long-term retention.

## Top Product Combinations for Cross-Selling in United Kingdom

Country

United Kingdom

Stock Code	Description	Associated Product	Country X United Kingdom
85099B	JUMBO BAG RED RETROSPOT	JUMBO BAG RED RETROSPOT,CHARLOTTE BAG SUKI DESIGN	
		JUMBO BAG RED RETROSPOT,HEART OF WICKER SMALL	
		JUMBO BAG RED RETROSPOT,JUMBO BAG ALPHABET	
		JUMBO BAG RED RETROSPOT,JUMBO BAG APPLES	
		JUMBO BAG RED RETROSPOT,JUMBO BAG BAROQUE BLACK WHITE	
		JUMBO BAG RED RETROSPOT,JUMBO BAG OWLS	
		JUMBO BAG RED RETROSPOT,JUMBO BAG PINK POLKADOT	
		JUMBO BAG RED RETROSPOT,JUMBO BAG PINK VINTAGE PAISLEY	
		JUMBO BAG RED RETROSPOT,JUMBO BAG SPACEBOY DESIGN	
		JUMBO BAG RED RETROSPOT,JUMBO BAG STRAWBERRY	
		JUMBO BAG RED RETROSPOT,JUMBO BAG TOYS	
		JUMBO BAG RED RETROSPOT,JUMBO BAG VINTAGE LEAF	
		JUMBO BAG RED RETROSPOT,JUMBO BAG WOODLAND ANIMALS	
		JUMBO BAG RED RETROSPOT,JUMBO SHOPPER VINTAGE RED PAIS..	
		JUMBO BAG RED RETROSPOT,JUMBO STORAGE BAG SKULLS	
		JUMBO BAG RED RETROSPOT,JUMBO BAG BLACK SKULLS	

# Holt-Winters Analysis

Holt-Winters Analysis is a time series forecasting technique used in business and analytics. It predicts future values based on past sales trends, seasonal patterns, and overall growth or decline in the data. It uses three key components- Level, Trend, Seasonality.

Prior to modeling, we aggregated sales at a monthly level by product and country, filtered out series with fewer than 24 months of data, and resampled to ensure consistency.

*We selected Holt-Winters because it handles seasonal cycles effectively, is computationally lightweight for thousands of product-country combinations, and does not require external features, making it ideal for SKU-level demand forecasting in an e-retail setting.*



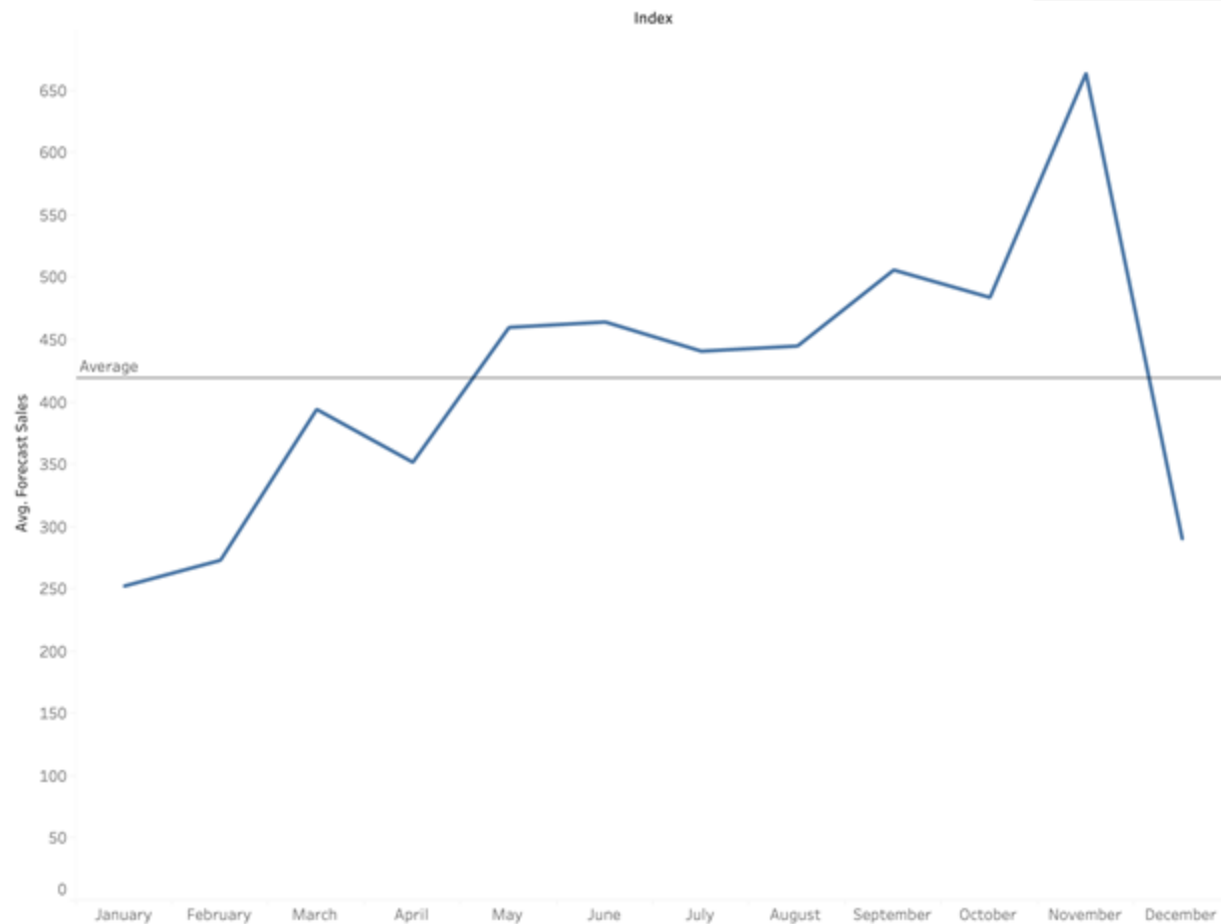
# Business Insights

- In the UK and Germany, demand consistently peaked in November, particularly for gifting and decorative products like festive t-light holders and holiday-themed bags. This underscores the importance of front-loading inventory and marketing spend in Q4 to capitalize on seasonal buying behavior.
- In Spain, a predictable summer dip in sales suggests a need to scale back promotions or reallocate marketing resources during that period to maximize ROI.
- Although smaller in volume, Netherlands and Australia demonstrated stable seasonal demand curves, highlighting them as candidates for localized campaigns and precision-driven inventory planning.
- Additionally, early-stage growth trajectories in emerging markets for select SKUs point to new product-market fit opportunities and strategic expansion potential.
- These insights allowed us to move from reactive sales planning to a proactive, demand-aligned strategy, ensuring that bundles, discounts, and advertising are deployed at the right time, in the right market, to the right segment.

## Anticipating the Market: Monthly Sales Forecast

Country

United Kingdom ▼



# Smart Product Bundling: Matching the Right Products to the Right Customers at the Right Time

- By combining RFM, Apriori, and Holt-Winters analyses, we turned raw data into a fully actionable customer strategy.
- We mapped Apriori-identified product bundles to specific RFM customer segments, and then used Holt-Winters forecasts to time those bundles for maximum seasonal impact.
- For example, a holiday-themed gift bundle was frequently purchased by high-frequency, moderate-value UK customers, with demand forecasted to peak in November making it ideal for a targeted Q4 campaign.
- In Germany, summer bundles aligned with formerly high-value but inactive customers, creating a timely opportunity for reactivation during seasonal dips.

# Targeting the Right Segment with the Right Bundle

				Country X	
				United Kingdom	
Stock Code	Description	Associated Product	Segment1		
85099B	JUMBO BAG RED RETROSPOT	JUMBO BAG RED RETROSPOT,JUMBO STORAGE BAG SUKI	Champions		
			Loyal Customers		
			At Risk		
			Can't Loose		
			Potential Loyalists		
			Hibernating		
			About to Sleep		
			Need Attention		

Country

United Kingdom

# Our Recommendations

1. Champions & Loyal Customers: These are your top-value customers engaged, consistent buyers who respond well to exclusivity and early access. The goal is to retain loyalty and increase basket size.

- Launch premium decorative bundles (e.g., T-light holders, patterned lunch bags) as exclusive early-access offers.
- Time campaigns to hit seasonal peaks like October in the UK or early summer in Germany to maximize ROI.
- Use targeted email campaigns and VIP-only landing pages to make them feel recognized and prioritized.

2. Potential Loyalists & Promising Customers: This group is on the verge of becoming loyal they need a small incentive or nudge. The goal is to convert them into repeat, high-value buyers.

- Promote mid-range practical bundles (e.g., kitchen sets, snack boxes) with offers like “buy more, save more.”
- Align campaigns with moderate-demand months like May–June or September to catch them before they churn.
- Use retargeting ads, loyalty perks, and personalized newsletter offers to drive conversions and increase engagement.

# Our Recommendations

3. At Risk & About to Sleep Customers: These are customers slipping away. They've shown strong behavior in the past, so the goal is to re-engage them with relevance and urgency.

- Offer bundles featuring their past favorite items, paired with a small free add-on to encourage re-trial.
- Launch these offers 2–3 weeks ahead of seasonal peaks (e.g., early November in the UK) to catch their attention when purchase intent may return.
- Use personalized, urgency-based messaging like “We miss you” via email, SMS, or account-based follow-up to prompt action.