retail_business_analysis

August 21, 2025

0.1 # Retail Business Analysis

This analysis provides actionable insights to improve sales performance based on comprehensive retail transaction data analysis.

0.2 Executive Summary

Revenue has stalled: Sales plateaued in 2011 (–1.2% YoY). The issue is not order value, but stagnant customer growth and weakening repeat engagement.

Customer base is leaky: Acquisition was strong (37% new in 2011) but offset by churn (37% lost). Loyal customers (44% of base) drive ~88% of revenue, creating dependency on a small group.

Home Decor is the backbone of the product: Contributes ~40% of revenue and is the main entry funnel, but underconverts (only ~45% loyalty). Kitchen/Food and Beauty are stronger repeat drivers, while seasonal categories (Christmas, Toys) add short-term spikes but poor retention.

Sales are highly seasonal: Revenue peaks in November, driven by Home Decor and Christmas. Stable categories like Kitchen, Garden, and Stationery sustain baseline demand, but seasonal surges fail to convert into long-term buyers.

0.2.1 Recommendations

1. **Protect and grow Home Decor** as the flagship, while cross-selling into stickier categories (Kitchen/Food, Beauty) to improve repeat conversion.
2. **Close the retention gap** with campaigns that move one-time buyers into repeat customers, supported by a unified **Customer 360 view**.
3. **Balance the product mix** by revitalizing core everyday categories and using seasonal products mainly as acquisition funnels.
4. **Plan for seasonality** with a simple **demand forecasting framework** to align inventory and marketing around peak and off-peak months.

1 Objectives & Research Framework

1.1 Problem Statement

The company requires insights to improve sales performance. To do that, understanding product characteristic, customer behavior patterns, and market trend is key to identify growth opportunities in their business operation.

1.2 Business Questions

1. **KPI & Performance Overview**: How has overall revenue, customer base, and order behavior evolved over time?
2. **Customer Growth & Engagement**: What drives customer acquisition, retention, and churn? How engaged are different customer segments?
3. **Product Strategy & Assortment**: Which products or categories drive revenue, repeat conversion, and loyalty? How concentrated is sales contribution across SKUs?
4. **Seasonality & Demand Patterns**: How do sales vary by month/season, and which categories are most affected by seasonality?
5. **Geographic Performance**: Which regions/markets contribute most to sales, and where are the growth opportunities?

1.3 Approach

- **Data Quality Check:** Data loading and preprocessing to avoid noise and sanitize analysis
- **Data Analysis:** Using the defined business question, identify key metric and analyze results
- **Insight Visualization:** Visualize the findings into executive digestable format
- **Insights Summary:** Summarize the opportunity as recommendation for the stakeholders

2 Data Quality Check

Environment Setup Complete

```
DATASET OVERVIEW
--------------------------------------------------
Dataset Shape: 1,067,371 transactions with 8 features
Data Period: 2009-12-01 07:45:00 to 2011-12-09 12:50:00


Sample Data:

   order_id product_id                   product_description  quantity  \
0    489434      85048  15CM CHRISTMAS GLASS BALL 20 LIGHTS         12
1    489434     79323P                    PINK CHERRY LIGHTS         12
2    489434     79323W                   WHITE CHERRY LIGHTS         12
3    489434      22041            RECORD FRAME 7" SINGLE SIZE         48
4    489434      21232         STRAWBERRY CERAMIC TRINKET BOX         24


            order_date  unit_price  customer_id         country
0  2009-12-01 07:45:00        6.95      13085.0  United Kingdom
1  2009-12-01 07:45:00        6.75      13085.0  United Kingdom
2  2009-12-01 07:45:00        6.75      13085.0  United Kingdom
3  2009-12-01 07:45:00        2.10      13085.0  United Kingdom
4  2009-12-01 07:45:00        1.25      13085.0  United Kingdom

DATA QUALITY CHECK
--------------------------------------------------

          quantity  unit_price  customer_id
count  1067371.00  1067371.00     824364.00
```

```
mean            9.94           4.65       15324.64
std           172.71         123.55        1697.46
min        -80995.00      -53594.36       12346.00
25%             1.00           1.25       13975.00
50%             3.00           2.10       15255.00
75%            10.00           4.15       16797.00
max         80995.00       38970.00       18287.00
```

	Data_Type	Missing_Values	Missing_Percentage	\
order_id	object	0	0.00	
product_id	object	0	0.00	
product_description	object	4382	0.41	
quantity	int64	0	0.00	
order_date	object	0	0.00	
unit_price	float64	0	0.00	
customer_id	float64	243007	22.77	
country	object	0	0.00	

	Unique_Values
order_id	53628
product_id	5305
product_description	5698
quantity	1057
order_date	47635
unit_price	2807
customer_id	5942
country	43

```
Negative quantities: 22,950
Zero/negative prices: 6,207
Missing customer IDs: 243,007
```
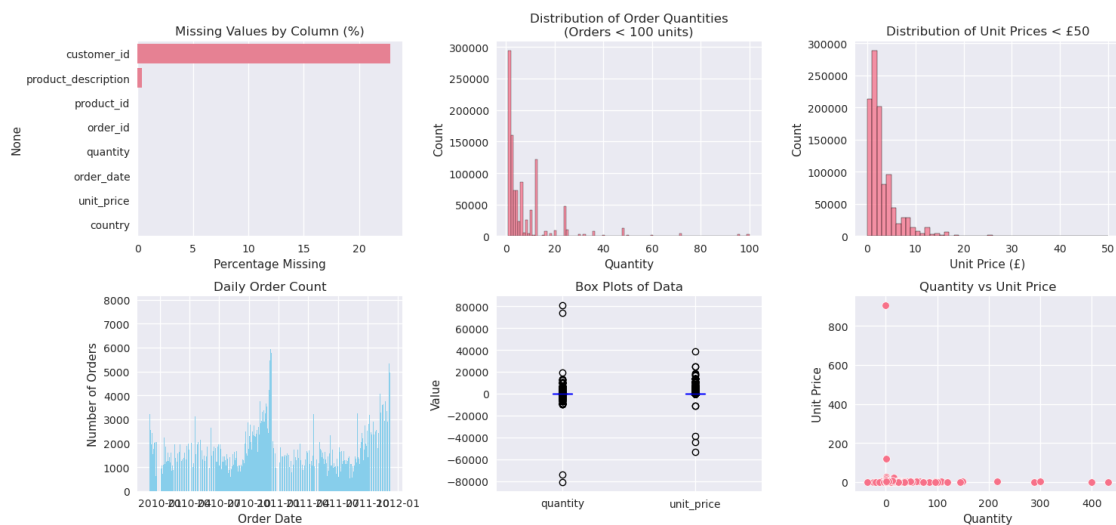
ADDITIONAL QUALITY METRICS

Price Range: £-53594.36 to £38970.00
Quantity Range: -80995 to 80995 units
Number of orders with quantity > 100: 10,561
Number of products priced > £50: 2,557

Based on the descriptive statistics and initial exploratory of the dataset, we found that:

1. Invalid data logic like negative quantities and zero prices is present in the dataset.
2. There are missing values in several columns, including customer_id and product_id.
3. The distribution of order quantities is highly skewed, with a small number of orders containing large quantities.
4. Most products have a unit price below £50, but a few outliers are priced significantly higher.
5. Invalid product IDs are found, which do not match the expected format.
6. order_ids starting with "C" is a refund transaction so we need to take out both.
7. Some product description shows non standardize value with questionable data quality, need to take out

So the data cleaning and preprocessing steps that we have to do:

1. Remove invalid transactions with negative quantities, zero prices, or missing customer IDs.
2. Remove missing values in critical columns like customer_id and product_id.
3. Remove product IDs that do not match the expected format.
4. Separate outliers transaction into different table, for different analysis. (anomaly exploration)

DATA PREPROCESSING & FEATURE ENGINEERING

Added features: 6 ['year', 'day_of_week', 'total_amount', 'month', 'month_year', 'quarter']
Original dataset: 1,067,371 transactions
Number of transactions with invalid product description: 82
Number of transactions with invalid data logic: 261,822
Number of transactions with invalid product_id: 6,093
Number of refund transactions: 19,494
Cleaned dataset: 757,349 transactions
Data removed: 310,022 transactions (29.0%)
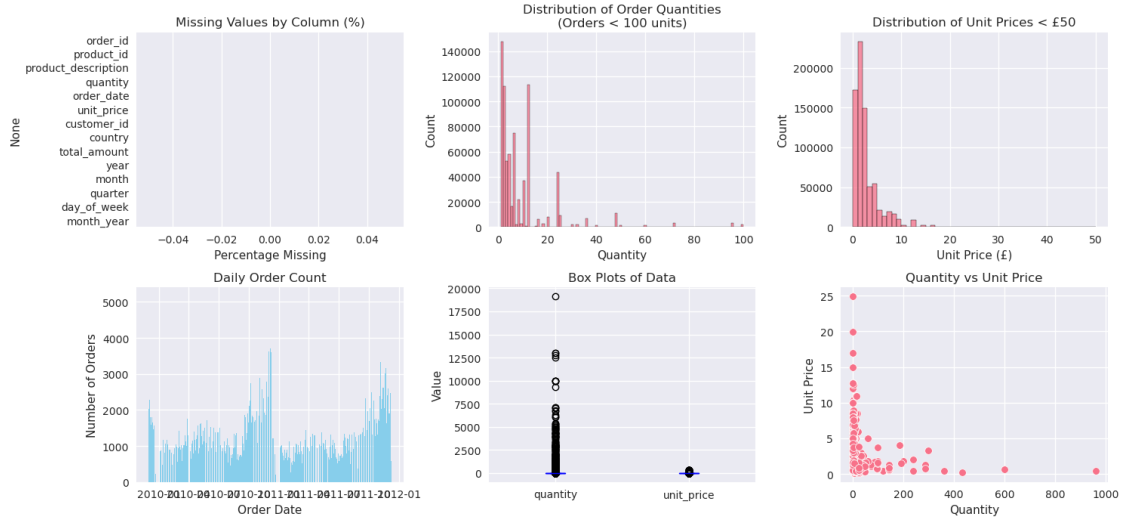
```
ADDITIONAL QUALITY METRICS
--------------------------------------------------
Price Range: £0.03 to £295.00
Quantity Range: 1 to 19152 units
Number of orders with quantity > 100: 8,325
Number of products priced > £50: 259
```
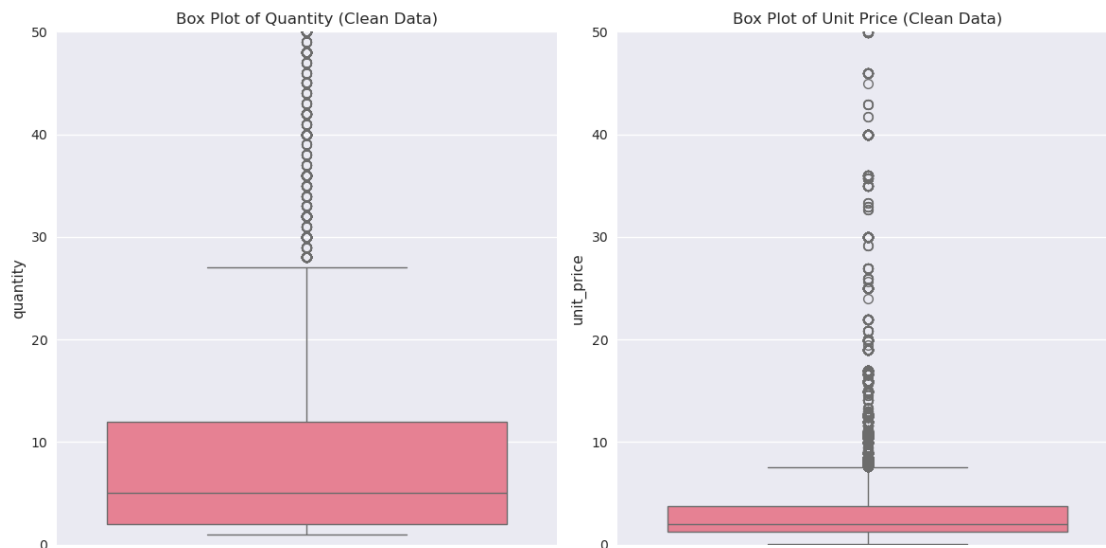


```
Final clean dataset shape: 757,349 transactions with 14 features
```

The cleaned data makes more reasonable, the quantity and unit price distributions are now more representative of typical pattern (lower unit price correlates with higher quantity sold). The box

plot reveal invalid data has been cleaned while we keep retaining reasonable outliers for further analysis.

We need to also create one table for outlier data for further outlier or anomaly analysis. (We might see hidden opportunities behind this data)

```
Created outlier only datasets

Outlier Analysis (df_outliers)
----------------------------------------
Total outlier transactions: 32,088 (3.01%)
Unit price range: £0.03 - £32.69
Quantity range: 43 - 19152
Total amount range: £2.08 - £8925.00

Percentage Unique products in outliers: 53.67%
Percentage Unique customers in outliers: 48.79%


Top 5 outlier transactions by value:
        order_id product_id                   product_description  quantity  \
228042    511465     15044A                    PINK PAPER PARASOL      3500
873786    567423      23243   SET OF TEA COFFEE SUGAR TINS PANTRY      1412
947062    573003      23084                    RABBIT NIGHT LIGHT      2400
358821    524181      21622      VINTAGE UNION JACK CUSHION COVER       648
769880    558526      23173                   REGENCY TEAPOT ROSES       540

                  order_date  unit_price  customer_id          country  \
228042   2010-06-08 12:59:00        2.55      18008.0   United Kingdom
873786   2011-09-20 11:05:00        5.06      17450.0   United Kingdom
947062   2011-10-27 12:11:00        2.08      14646.0      Netherlands
358821   2010-09-27 16:59:00        6.89      17450.0   United Kingdom
769880   2011-06-30 11:01:00        8.15      17949.0   United Kingdom

        total_amount  year  month  quarter day_of_week month_year
228042       8925.00  2010      6        2     Tuesday    2010-06
873786       7144.72  2011      9        3     Tuesday    2011-09
947062       4992.00  2011     10        4    Thursday    2011-10
358821       4464.72  2010      9        3      Monday    2010-09
769880       4401.00  2011      6        2    Thursday    2011-06
```

- The percentage of unique products and customers in the outlier group is significant.
- Outliers make up only about 3% of all transactions, but they involve a wide range of products and customers.
- This suggests that outlier transactions may reveal unique patterns or behaviors not seen in the main dataset.

3 Business Performance Analysis

3.1 Revenue Performance Overview

In this step, we will analyze the KPI of the business over the selected period. This will involve examining key metrics such as total revenue, transaction volume, and customer engagement levels to identify trends and areas for improvement.

```
KEY PERFORMANCE INDICATORS (KPIs)
---------------------------------------------------
Date Range: 2009-12-01 to 2011-12-09

                                 Metric          Value  \
0                         Total Revenue  $14,891,236.64
1                    Total Transactions         757,349
2                       Unique Customers           5,819
3                        Unique Products           4,604
4                    Average Order Value         $419.47
5    Average Unique Items per Transaction          20.35
6           Average Revenue per Customer        $2559.07
7  Average Number of Orders per Customer            6.10


                 Business_Impact
0        Primary revenue driver
1              Volume indicator
2                  Market reach
3             Product diversity
4  Customer spending behavior
5     Cross Sell Effectiveness
6           Value per Customer
7             Order Frequency
```

```
KEY PERFORMANCE INDICATORS (KPIs) - YoY Performance (November 2011)
-----------------------------------------------------------------------

   Year  Total Revenue Total Transactions Unique Customers Unique Products  \
0  2010  $6,656,431.27            340,173            4,098           3,864
1  2011  $6,576,989.56            333,923            4,139           3,560

  Average Order Value Average Unique Items per Transaction  \
0             $411.91                                20.38
1             $416.77                                20.64

  Average Revenue per Customer Average Number of Orders per Customer
0                     $1624.31                                  3.94
1                     $1589.03                                  3.81
```

Based on the KPI performance between 2010 and 2011, we found that:

1. **Revenue performance is stagnant.** Total revenue slipped slightly from **$6.65M to $6.57M ($-1.2\%$)**, with transactions also declining by **1.8%**. This indicates that topline

growth has stalled, with fewer orders placed overall.

2. **Customer base is not expanding.** Unique customers only increased marginally (**4,098 → 4,139**), showing that the company is essentially selling to the same pool of buyers. This lack of new acquisition limits growth potential.

3. **Product variety is shrinking.** Unique products sold declined from **3,864 to 3,560** (**− 7.9%**), reducing assortment breadth. This contraction directly impacts cross-sell opportunities and gives customers fewer reasons to engage.

4. **Customer value is weakening.** Average revenue per customer dropped from **$1,624 to $1,589**, driven by a fall in order frequency (**3.94 → 3.81 orders per customer**). Although average order value ticked up slightly (**+$5**), it was not enough to offset the decline in repeat purchases.

The data shows that growth is not constrained by order size , in fact, customers are spending slightly more per transaction. Instead, revenue is stagnant because **we are not acquiring new customers, our product range has contracted, and existing customers are buying less frequently.**

Having identified the areas of concern in revenue performance, the next step is to translate these findings into actionable directions for the business. Our focus will shift toward three strategic pillars:

1. Customer Growth & Engagement: The flat revenue in 2011 reflects stalled customer acquisition and fewer repeat purchases. We will examine where growth opportunities exist across segments and how to re-engage existing buyers to lift transaction frequency and recover the drop in average revenue per customer.

2. Product Analysis: The 8% decline in unique products sold reduced opportunities for cross-sell and weakened customer retention. Our analysis will focus on identifying which products lost traction and how refreshing the assortment—while reinforcing top performers—can restore momentum.

3. Seasonality & Demand Alignment: With stagnant number of orders, we need to better understand the underlying demand drivers and seasonal trends. This will involve analyzing historical sales data to identify patterns and adjust our inventory and marketing strategies accordingly.

4 Customer Growth & Engagement Analysis

Understanding the drivers behind customer acquisition and repeat purchase behavior is critical to reigniting growth. In this section, we segment customers, analyze new vs. repeat revenue, and assess order frequency and revenue concentration to identify actionable opportunities.

Customer Base Changes 2010-2011
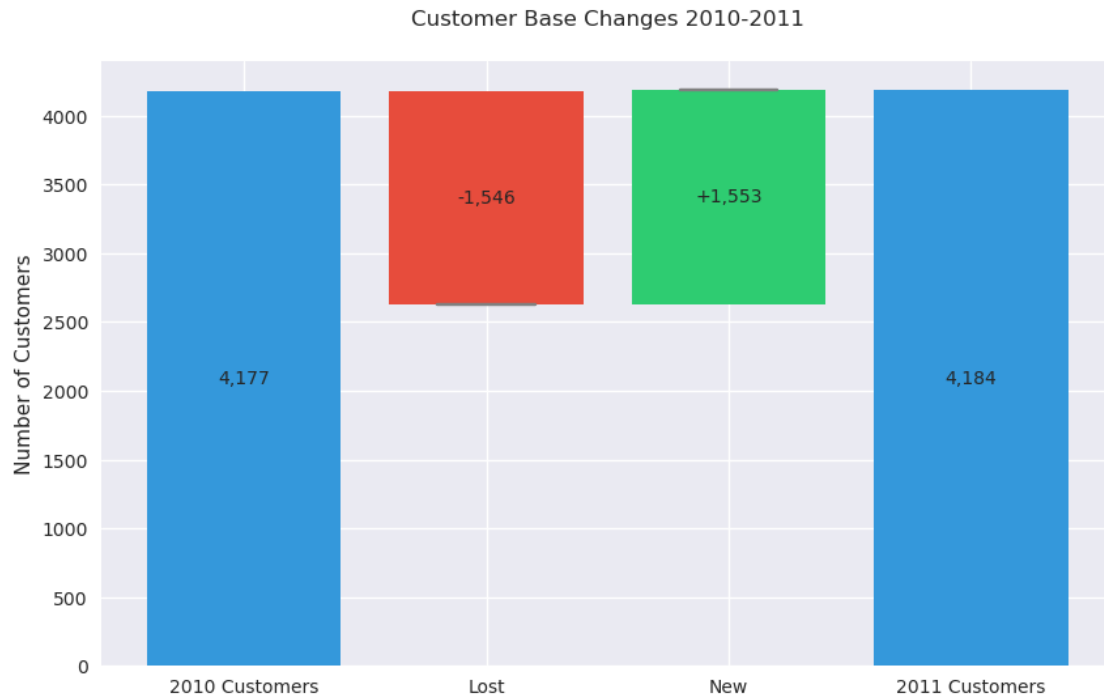

```
Number of retained customers: 2,631 (62.99%)
Number of lost customers: 1,546 (37.01%)
Number of new customers: 1,553 (37.12%)
Number of 2011 customers: 4,184
```

1. **High Acquisition, But Offset by Losses**

- The company added 1,553 new customers (37% of the 2011 base), indicating strong acquisition capability. However, this was almost completely offset by 1,546 customers lost from the 2010 base (37%), leaving net growth flat.

2. **Large Churn Rate**

- Only 63% of 2010 customers were retained into 2011. With nearly 4 in 10 customers churning each year, acquisition alone cannot sustain growth without stronger retention.

Our stagnant customer base was not due to weak acquisition (new customer inflow was strong) but because losses were high. Growth stalled because the business is running in place. The key opportunity lies in strengthening retention strategies to protect acquired customers

```
<Figure size 1200x600 with 0 Axes>
```

Customer Segments by Order Frequency
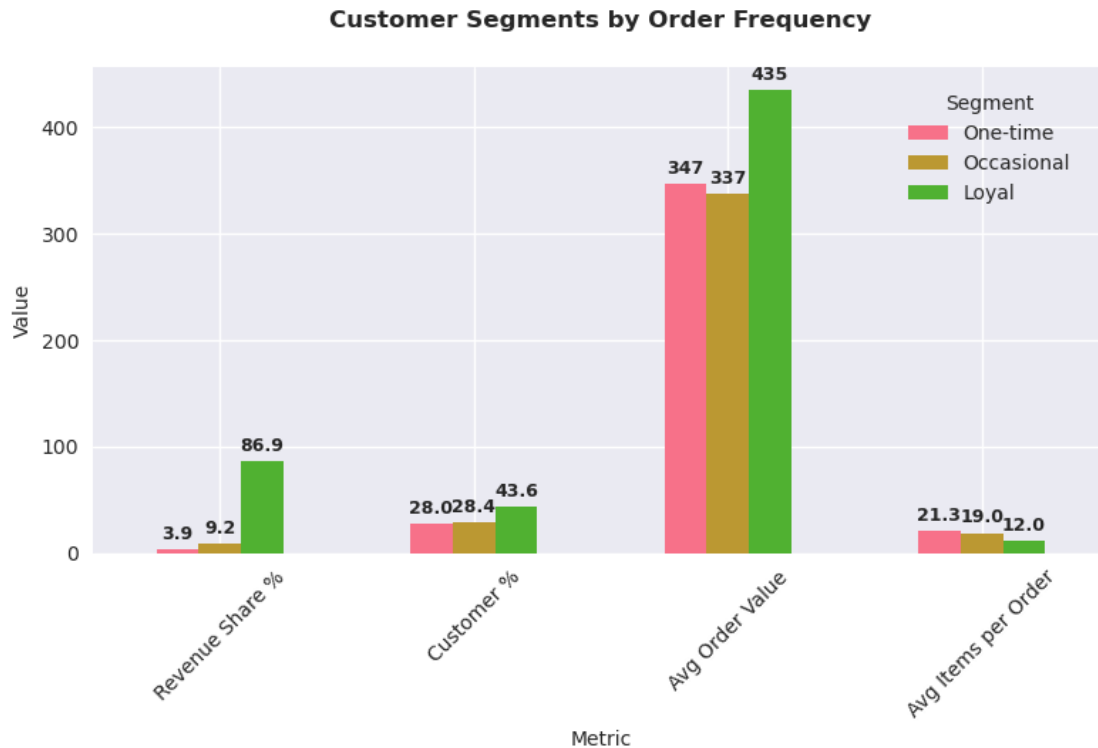

```
Order Frequency Segment Metrics:
--------------------------------------------------

                   Revenue Share %  Customer %  Avg Order Value  \
frequency_segment
One-time                       3.9        28.0           347.41
Occasional                     9.2        28.4           337.24
Loyal                         86.9        43.6           435.33

                   Avg Items per Order
frequency_segment
One-time                         21.29
Occasional                       18.98
Loyal                            12.02
```

1. **Revenue is Highly Skewed Toward Loyal Customers**

- Loyal buyers (44% of base) generate 87.7% of revenue, while one-time and occasional customers (56% of base) together contribute only 12.3%. This skew shows that growth depends on converting more of the base into loyal repeat customers, as acquisition without retention adds little revenue impact.
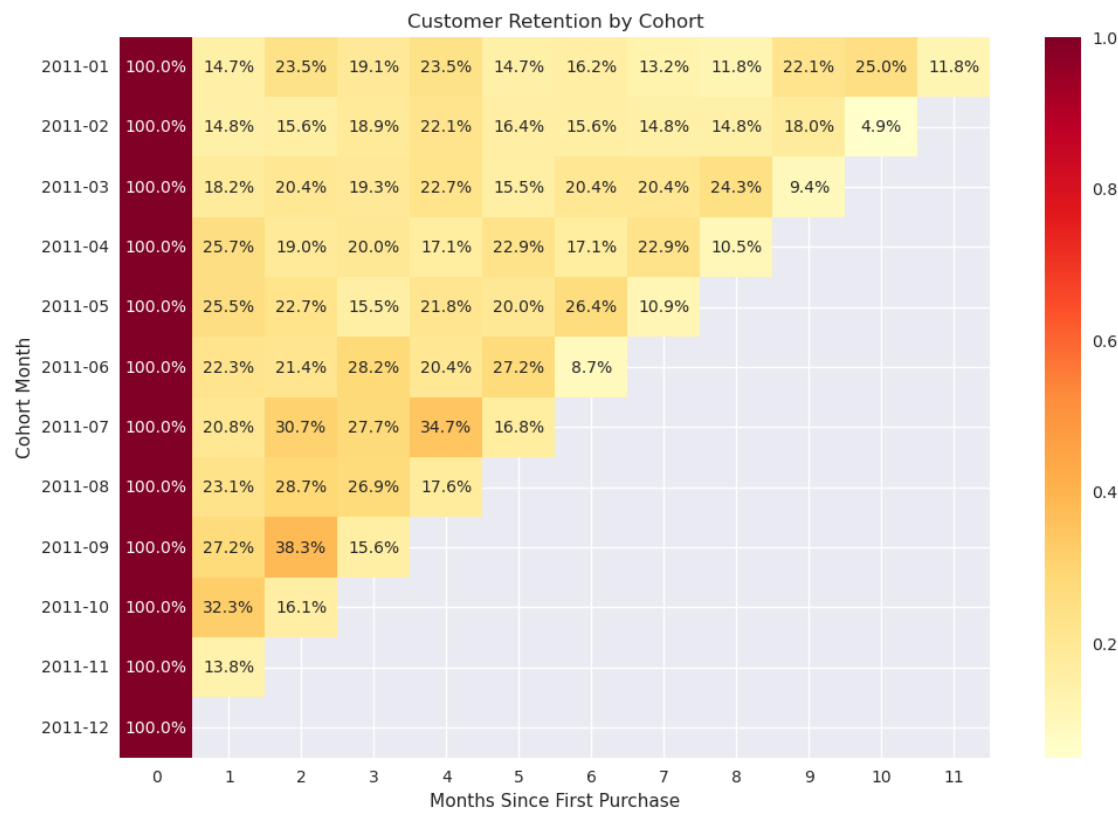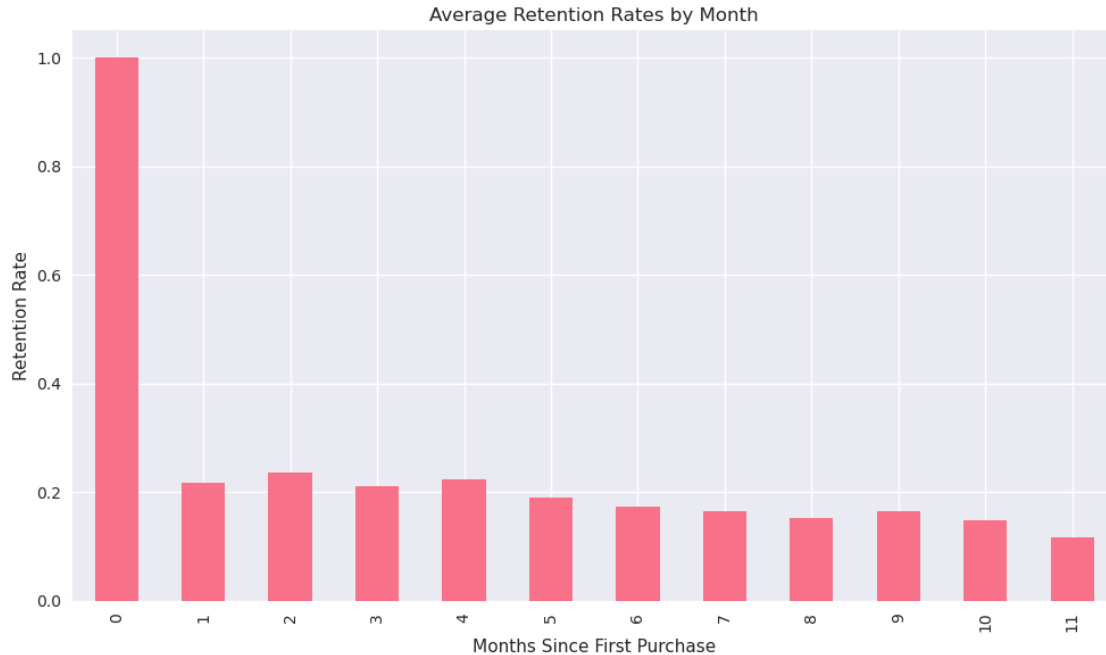
2. **Different Purchase Patterns Between Segments**

- Loyal customers spend more per order (~$435) but buy fewer unique items per transaction

(11.9 vs ~20 for one-time/occasional). This suggests loyal buyers focus on repeat, possibly bulk purchases of fewer SKUs, while first-time/occasional buyers experiment with more varied baskets.

The opportunity lies in reducing churn among new customers who already show strong purchase intent (showed by buying more varied product) but lack retention.

CUSTOMER COHORT ANALYSIS

Customer Retention by Cohort

Cohort Month	0	1	2	3	4	5	6	7	8	9	10	11
2011-01	100.0%	14.7%	23.5%	19.1%	23.5%	14.7%	16.2%	13.2%	11.8%	22.1%	25.0%	11.8%
2011-02	100.0%	14.8%	15.6%	18.9%	22.1%	16.4%	15.6%	14.8%	14.8%	18.0%	4.9%	
2011-03	100.0%	18.2%	20.4%	19.3%	22.7%	15.5%	20.4%	20.4%	24.3%	9.4%		
2011-04	100.0%	25.7%	19.0%	20.0%	17.1%	22.9%	17.1%	22.9%	10.5%			
2011-05	100.0%	25.5%	22.7%	15.5%	21.8%	20.0%	26.4%	10.9%				
2011-06	100.0%	22.3%	21.4%	28.2%	20.4%	27.2%	8.7%					
2011-07	100.0%	20.8%	30.7%	27.7%	34.7%	16.8%						
2011-08	100.0%	23.1%	28.7%	26.9%	17.6%							
2011-09	100.0%	27.2%	38.3%	15.6%								
2011-10	100.0%	32.3%	16.1%									
2011-11	100.0%	13.8%										
2011-12	100.0%											

Months Since First Purchase
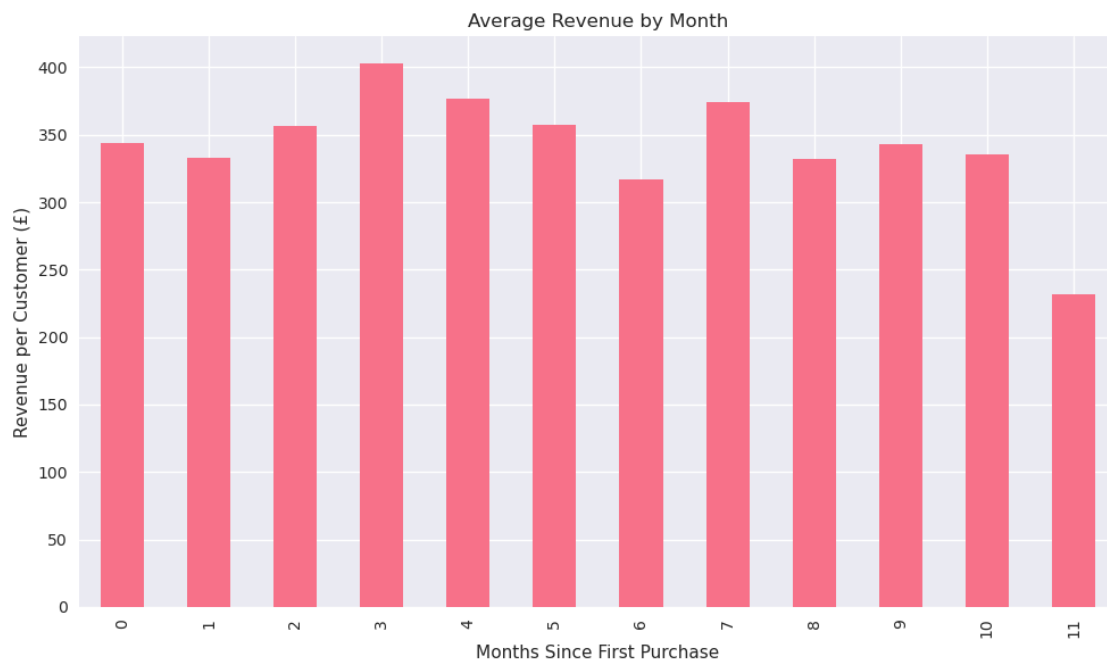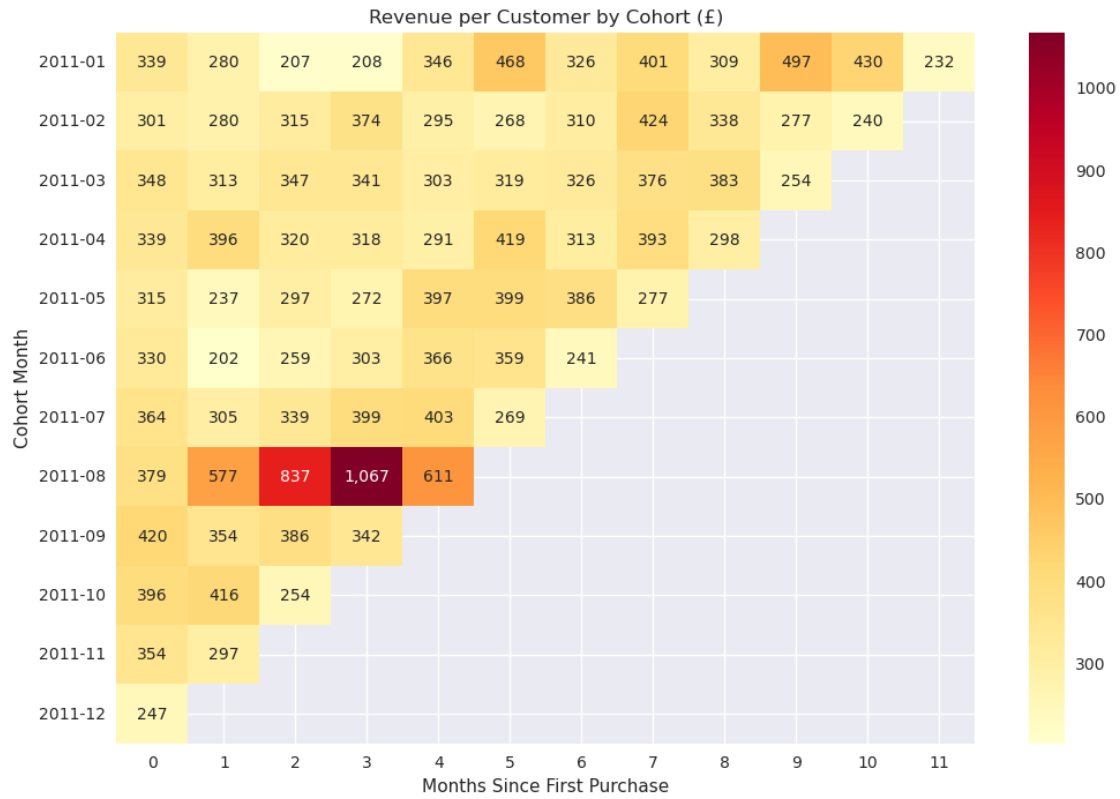
Average Retention Rates by Month

The business is highly leaky: new customer acquisition is strong, but most buyers churn immediately after the first purchase. Retention falls from 100% (Month 0) to just ~22% in Month 1. By Month 6, only ~17% remain active, and by Month 11, just ~12% of the original cohort is still buying.

But we can see interesting pattern where these customer usually come back in October - November

Revenue is therefore driven by a relatively small pool of repeat customers, though there is improvement of retention in the last 4 months (August 2011 - October 2011), The decay pattern is consistent.

Potential strategy: - **Onboarding & First-Repeat Push**: Focus efforts on converting customers from first to second order, since Month 1 retention is the biggest cliff. Promotions, reminders, or product bundling right after the first purchase could lift conversion.

- **Identify Repeat-Conversion Drivers**: Analyze what loyal customers bought on their first purchase vs. churners. If loyalists buy staples/bulk while one-timers buy novelty items, campaigns can nudge new buyers toward "sticky" product categories.

- **Protect the Core**: With only ~20% of customers sticking around, this group is disproportionately valuable and shown as the main driver of the revenue. Loyalty programs or tailored offers should focus here to extend lifetime value.

CUSTOMER COHORT ANALYSIS - REVENUE PER CUSTOMER

Revenue per Customer by Cohort (£)

Cohort Month	0	1	2	3	4	5	6	7	8	9	10	11
2011-01	339	280	207	208	346	468	326	401	309	497	430	232
2011-02	301	280	315	374	295	268	310	424	338	277	240	
2011-03	348	313	347	341	303	319	326	376	383	254		
2011-04	339	396	320	318	291	419	313	393	298			
2011-05	315	237	297	272	397	399	386	277				
2011-06	330	202	259	303	366	359	241					
2011-07	364	305	339	399	403	269						
2011-08	379	577	837	1,067	611							
2011-09	420	354	386	342								
2011-10	396	416	254									
2011-11	354	297										
2011-12	247											

Months Since First Purchase

Average Revenue by Month

Months Since First Purchase

Cohort revenue patterns show no meaningful uplift in spending. Customers spend roughly the same amount per order over time, suggesting basket size is stable. To note for the product analysis: growth therefore won't come from pushing higher order values, but focus on identify product to acquire more customers and improving retention to make them stickier.

5 Product Analysis

his section examines product category, top product contributions, long-tail performance, and identifies key decliners and gainers to inform assortment strategy.

```
Using device: GPU (CUDA)
Loading Hugging Face zero-shot classification model…

Device set to use cuda:0

Hugging Face model loaded successfully!

Defined 9 product categories:
    1. HOME_DECOR
    2. KITCHEN_FOOD_UTENSIL
    3. FURNITURE_STORAGE
    4. CHRISTMAS_HOLIDAY
    5. TOYS_GAMES
    6. TEXTILES_CLOTHING
    7. GARDEN_OUTDOOR
    8. STATIONERY_OFFICE
    9. BEAUTY_PERSONAL
Categorization functions defined

CATEGORIZING ALL PRODUCTS
--------------------------------------------------
Total unique products to categorize: 5,246
Using rule-based categorization

You seem to be using the pipelines sequentially on GPU. In order to maximize
efficiency please use a dataset

    Processed 500/5,246 products…
    Processed 1,000/5,246 products…
    Processed 1,500/5,246 products…
    Processed 2,000/5,246 products…
    Processed 2,500/5,246 products…
    Processed 3,000/5,246 products…
    Processed 3,500/5,246 products…
    Processed 4,000/5,246 products…
    Processed 4,500/5,246 products…
    Processed 5,000/5,246 products…
Categorization complete, 5,246 products categorized


CATEGORY DISTRIBUTION:
----------------------------------------
```

```
HOME_DECOR           : 261,591 transactions ( 34.5%) | 1,608 products
KITCHEN_FOOD_UTENSIL: 91,191 transactions ( 12.0%) | 624 products
GARDEN_OUTDOOR       : 82,077 transactions ( 10.8%) | 471 products
TEXTILES_CLOTHING    : 66,148 transactions (  8.7%) | 314 products
BEAUTY_PERSONAL      : 61,272 transactions (  8.1%) | 705 products
STATIONERY_OFFICE    : 50,274 transactions (  6.6%) | 318 products
CHRISTMAS_HOLIDAY    : 39,767 transactions (  5.3%) | 224 products
TOYS_GAMES           : 36,401 transactions (  4.8%) | 165 products
FURNITURE_STORAGE    : 35,969 transactions (  4.7%) | 186 products
FOOD_BEVERAGES       : 32,659 transactions (  4.3%) | 136 products

 REVENUE BY CATEGORY:
------------------------------
HOME_DECOR           : $ 5,387,527 ( 36.2%)
KITCHEN_FOOD_UTENSIL: $ 1,742,052 ( 11.7%)
GARDEN_OUTDOOR       : $ 1,704,608 ( 11.4%)
TEXTILES_CLOTHING    : $ 1,388,478 (  9.3%)
BEAUTY_PERSONAL      : $ 1,109,842 (  7.5%)
FURNITURE_STORAGE    : $   778,392 (  5.2%)
STATIONERY_OFFICE    : $   738,189 (  5.0%)
CHRISTMAS_HOLIDAY    : $   734,031 (  4.9%)
FOOD_BEVERAGES       : $   700,802 (  4.7%)
TOYS_GAMES           : $   607,315 (  4.1%)
```

CATEGORY PERFORMANCE ANALYSIS
----

product_category	total_revenue	avg_transaction_value	transaction_count
HOME_DECOR	5387527.23	20.60	261591
KITCHEN_FOOD_UTENSIL	1742052.21	19.10	91191
GARDEN_OUTDOOR	1704607.83	20.77	82077
TEXTILES_CLOTHING	1388477.71	20.99	66148
BEAUTY_PERSONAL	1109842.43	18.11	61272
FURNITURE_STORAGE	778392.30	21.64	35969
STATIONERY_OFFICE	738189.41	14.68	50274
CHRISTMAS_HOLIDAY	734031.18	18.46	39767
FOOD_BEVERAGES	700801.51	21.46	32659
TOYS_GAMES	607314.83	16.68	36401

product_category	avg_quantity	avg_unit_price	unique_customers
HOME_DECOR	12.86	2.78	5565
KITCHEN_FOOD_UTENSIL	11.98	2.94	4941
GARDEN_OUTDOOR	11.33	3.45	4952
TEXTILES_CLOTHING	11.74	2.42	4256
BEAUTY_PERSONAL	11.63	3.28	4819
FURNITURE_STORAGE	9.55	3.90	4159
STATIONERY_OFFICE	15.50	1.73	4379

```
CHRISTMAS_HOLIDAY          13.42            2.21                 3556
FOOD_BEVERAGES             13.60            3.62                 3872
TOYS_GAMES                 10.10            2.60                 3853


                      unique_products  revenue_share  \
product_category
HOME_DECOR                       1608           36.2
KITCHEN_FOOD_UTENSIL              624           11.7
GARDEN_OUTDOOR                    471           11.4
TEXTILES_CLOTHING                314            9.3
BEAUTY_PERSONAL                  705            7.5
FURNITURE_STORAGE                186            5.2
STATIONERY_OFFICE                318            5.0
CHRISTMAS_HOLIDAY                224            4.9
FOOD_BEVERAGES                   136            4.7
TOYS_GAMES                       165            4.1


                      avg_revenue_per_customer  transactions_per_product
product_category
HOME_DECOR                              968.11                     162.7
KITCHEN_FOOD_UTENSIL                    352.57                     146.1
GARDEN_OUTDOOR                          344.23                     174.3
TEXTILES_CLOTHING                       326.24                     210.7
BEAUTY_PERSONAL                         230.31                      86.9
FURNITURE_STORAGE                       187.16                     193.4
STATIONERY_OFFICE                       168.57                     158.1
CHRISTMAS_HOLIDAY                       206.42                     177.5
FOOD_BEVERAGES                          180.99                     240.1
TOYS_GAMES                              157.62                     220.6
```

Product Category Performance Dashboard
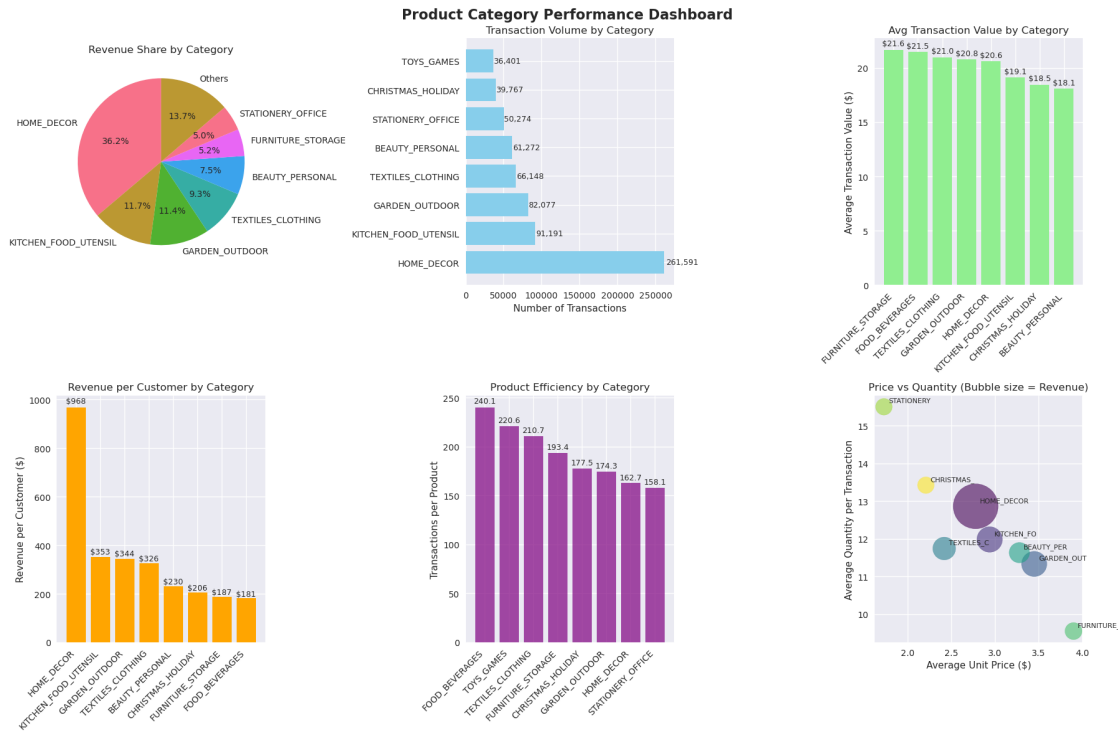
```
Revenue Leader: HOME_DECOR ($5,387,527, 36.2%)
Volume Leader: HOME_DECOR (261,591 transactions)
Highest AOV: FURNITURE_STORAGE ($21.64 per transaction)
Most Efficient: FOOD_BEVERAGES (240.1 transactions per product)
```

1. **Revenue Concentration**

- Home Decor (39.6%) and Kitchen/Food Utensils (20.6%) account for ~60% of total revenue
- Growth heavily dependent on these two key categories

2. **Customer Reach vs Monetization**

- Stationery/Office and Garden/Outdoor attract 4,000+ customers each but generate low revenue per customer (~$200)
- Home Decor shows stronger monetization with ~$1,051 revenue per customer vs Kitchen/Food at $578
- Some categories drive customer acquisition while others build deeper customer value

3. **Product Efficiency**

- Toys/Games and Textiles/Clothing lead in transactions per product (214-219)
- Home Decor has large assortment (1,700+ SKUs) but lower efficiency (167 transactions/product)
- Opportunity to optimize long-tail product assortment

Revenue is concentrated in Home Decor and Kitchen Food Utensil, but these categories have higher number of unique product (long tail). There is a good opportunity in optimizing the long-tail within

dominant categories (Need to do demand forecasting).

```
YEAR-OVER-YEAR CATEGORY PERFORMANCE
----------------------------------------------------
Year-over-Year Category Performance (Jan-Nov comparison):
============================================================================

                      2010_Revenue   2011_Revenue   Revenue_Growth_%  \
product_category
HOME_DECOR               2285927.9      2459789.9                7.6
KITCHEN_FOOD_UTENSIL      752702.0       751413.6               -0.2
GARDEN_OUTDOOR            843505.1       697191.0              -17.3
TEXTILES_CLOTHING         633712.4       636517.7                0.4
BEAUTY_PERSONAL           533722.2       428260.3              -19.8
CHRISTMAS_HOLIDAY         296045.0       349609.8               18.1
FURNITURE_STORAGE         355996.0       337361.5               -5.2
STATIONERY_OFFICE         355294.4       325500.3               -8.4
TOYS_GAMES                254578.4       303534.2               19.2
FOOD_BEVERAGES            344947.9       287811.4              -16.6


                      2010_Orders   2011_Orders   Order_Growth_%  \
product_category
HOME_DECOR                  14218         14190             -0.2
KITCHEN_FOOD_UTENSIL        10434          9987             -4.3
GARDEN_OUTDOOR              10990         10210             -7.1
TEXTILES_CLOTHING            8602          7914             -8.0
BEAUTY_PERSONAL              9743          9131             -6.3
CHRISTMAS_HOLIDAY            4148          4324              4.2
FURNITURE_STORAGE            7719          6837            -11.4
STATIONERY_OFFICE            7960          7637             -4.1
TOYS_GAMES                   6509          6334             -2.7
FOOD_BEVERAGES               6638          6045             -8.9


                      Customer_Growth_%
product_category
HOME_DECOR                          1.1
KITCHEN_FOOD_UTENSIL                1.1
GARDEN_OUTDOOR                     -2.1
TEXTILES_CLOTHING                  -6.7
BEAUTY_PERSONAL                    -0.4
CHRISTMAS_HOLIDAY                   8.2
FURNITURE_STORAGE                  -4.4
STATIONERY_OFFICE                  -1.6
TOYS_GAMES                          1.9
FOOD_BEVERAGES                     -2.5

CATEGORY WINNERS (>5% growth):
----------------------------------------
TOYS_GAMES: +19.2% revenue growth ($303,534)
```

```
CHRISTMAS_HOLIDAY: +18.1% revenue growth ($349,610)
HOME_DECOR: +7.6% revenue growth ($2,459,790)
CATEGORY DECLINERS (<-5% decline):
----------------------------------------
BEAUTY_PERSONAL: -19.8% revenue decline ($428,260)
GARDEN_OUTDOOR: -17.3% revenue decline ($697,191)
FOOD_BEVERAGES: -16.6% revenue decline ($287,811)
STATIONERY_OFFICE: -8.4% revenue decline ($325,500)
FURNITURE_STORAGE: -5.2% revenue decline ($337,362)
```


Year-over-Year Revenue Growth by Product Category (2010 vs 2011)
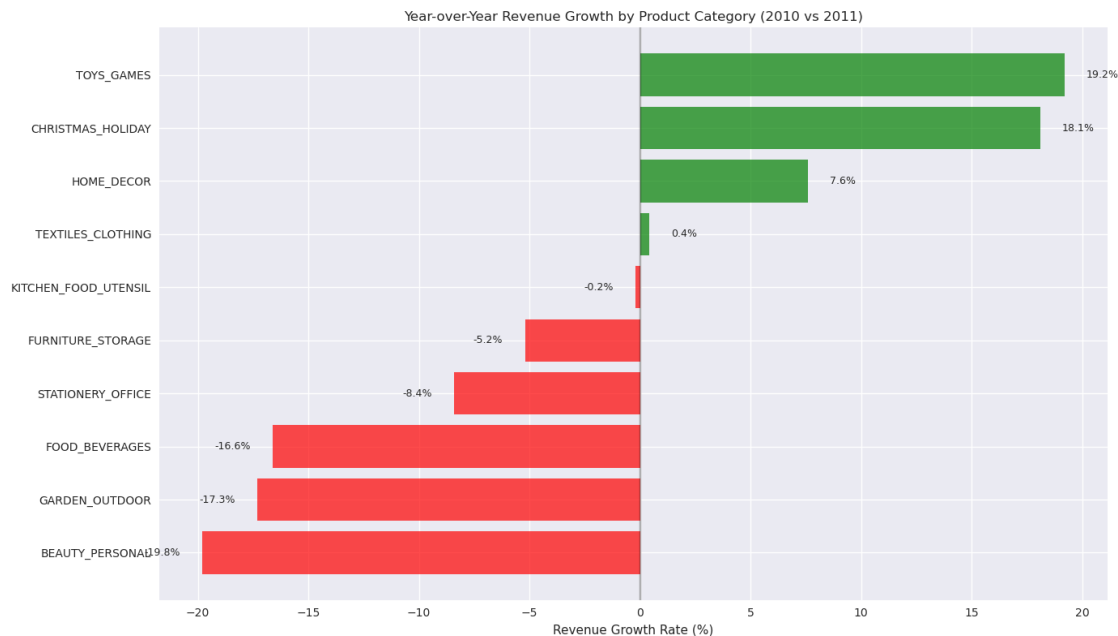
1. **Growth Driven by Seasonal & Niche Categories**:

- Christmas/Holiday (+17.9%), Furniture/Storage (+16.2%), and Toys/Games (+10.1%) led growth. These are seasonal or discretionary categories, suggesting demand spikes rather than sustained growth drivers.

2. **Core Categories Declined**

- Kitchen/Food Utensils (–7.3%), Stationery/Office (–11.7%), Garden/Outdoor (–12.5%), and Beauty/Personal (–12.8%) all contracted. These categories represent everyday or repeat-purchase items. Their decline is concerning as it weakens the foundation for customer stickiness.

3. **Home Decor Flatlining**: The largest revenue category, Home Decor, grew only marginally (+3.2%), underperforming relative to its scale. With 40% of revenue concentrated here, even small slowdowns have an outsized effect on overall performance.

The business is leaning too heavily on seasonal/short-term spikes while core repeat-purchase categories are shrinking. To sustain growth, strategy should focus on: Stabilizing and revitalizing

core categories (Kitchen/Food, Beauty, Garden). Using seasonal categories (Christmas, Toys) as acquisition channels, but ensuring those customers are converted into repeat buyers via staple categories.Reinvesting in Home Décor efficiency, since it remains the largest but is losing momentum.

```
LONG-TAIL ANALYSIS
--------------------------------------------------
Share of revenue from long-tail products (<1% each):
{2010: np.float64(0.987260867194639), 2011: np.float64(1.0)}
```

In 2010, ~98.7% of revenue came from products contributing <1% each, and by 2011 this rose to nearly 100%. This means virtually no single product carries significant weight on its own revenue is spread across a vast assortment.

```
Comparing First Purchase Patterns between Loyal and Churned customers
--------------------------------------------------
```

	Category	Total_Customers	Loyal_Rate_%	Retention_Rate_%
8	TEXTILES_CLOTHING	3234	46.8	74.3
4	GARDEN_OUTDOOR	4229	46.3	74.3
9	FOOD_BEVERAGES	2616	45.9	74.2
1	BEAUTY_PERSONAL	3902	46.4	74.0
3	FURNITURE_STORAGE	2971	46.9	74.0
5	KITCHEN_FOOD_UTENSIL	4122	46.2	73.6
6	STATIONERY_OFFICE	3194	44.9	73.3
7	TOYS_GAMES	2554	45.0	73.3
2	HOME_DECOR	5330	45.2	73.1
0	CHRISTMAS_HOLIDAY	1752	41.3	70.1
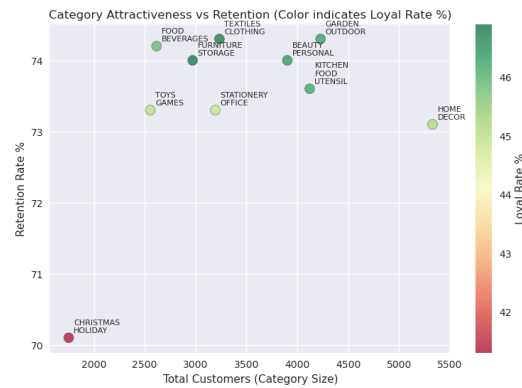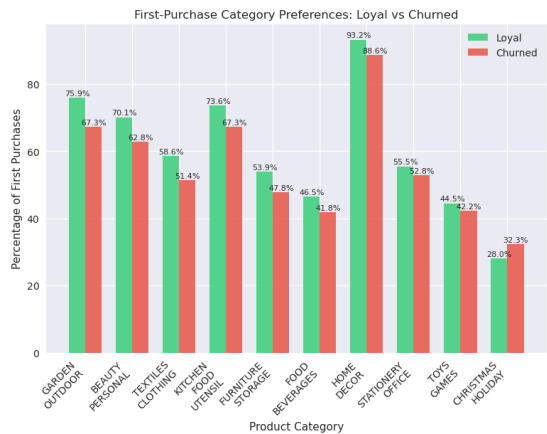
1. **Kitchen/Food and Beauty/Personal Anchor Loyalty**

- Both categories show high first-purchase share among loyal customers (Kitchen/Food 86%, Beauty 73%). These are strong entry points that convert well into repeat buyers, making them effective categories for stickiness campaigns.

2. **Home Decor = Biggest Funnel, Average Conversion to Loyal** Largest acquisition

20

category (95% of churners, 90% of loyals started here). However, loyal rate is only ~45%. Home Décor is an excellent acquisition magnet, but needs cross-sell into stickier categories to improve retention.
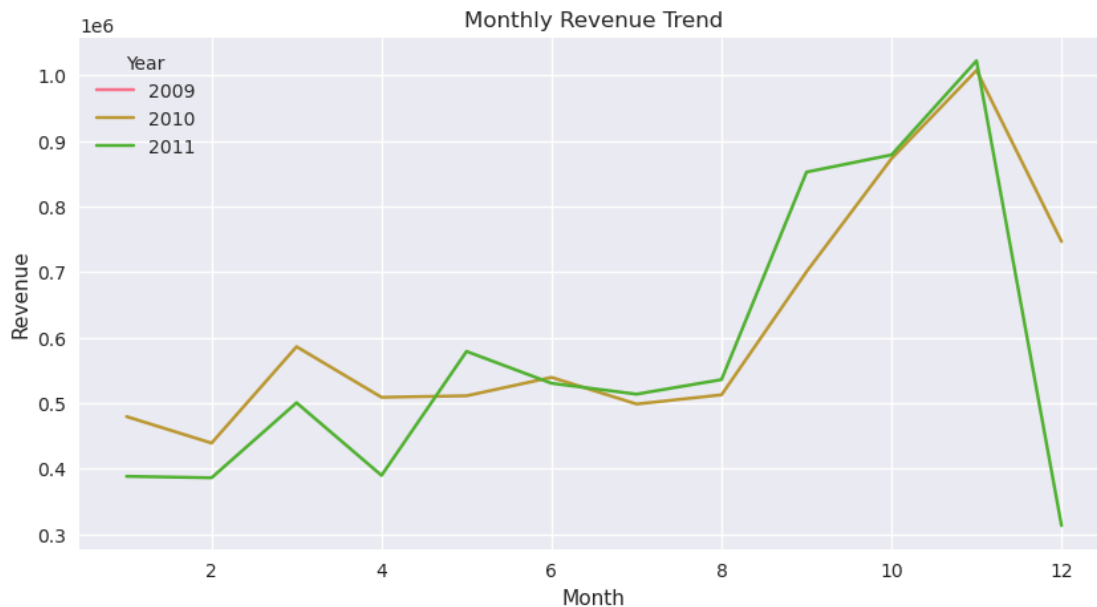
3. **Seasonal / Gift Categories Underperform** Christmas/Holiday has the lowest loyalty rate (41.5%) and weakest retention (~71%). Good for customer acquisition spikes, but not effective for repeat conversion unless paired with follow-up engagement.
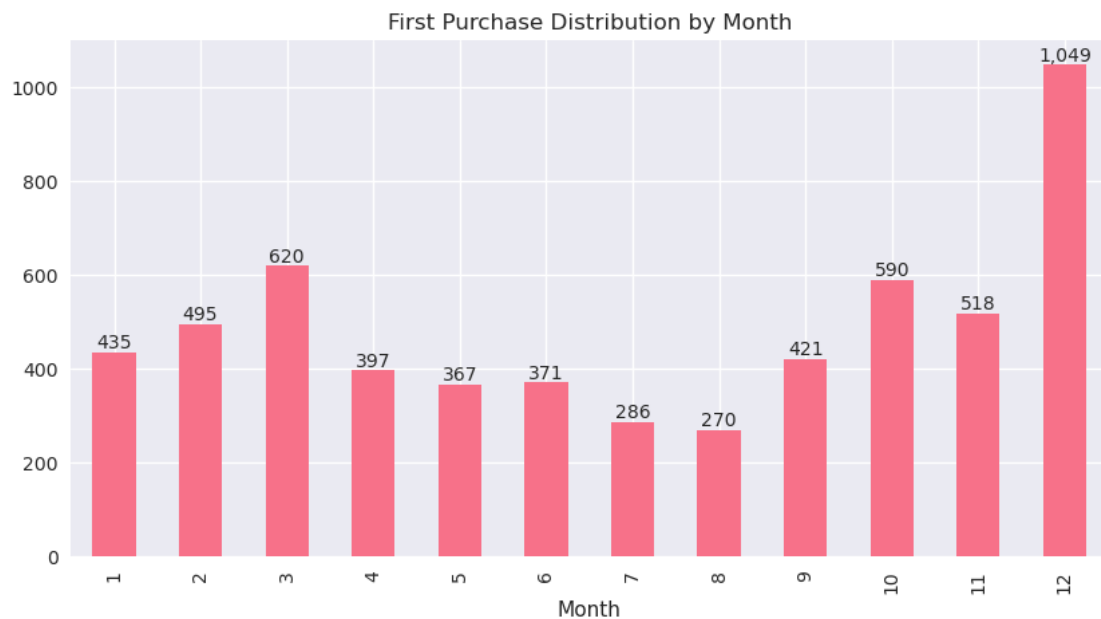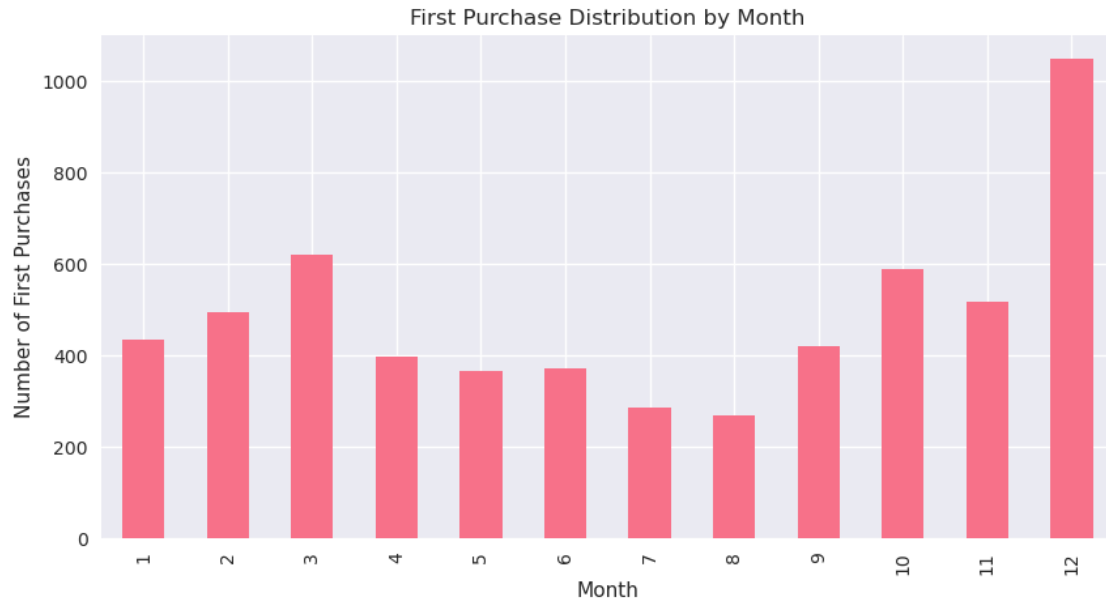
Repeat-conversion is anchored by Kitchen/Food and Beauty, customers starting here are more likely to become loyal. Home Decor dominates acquisition but underconverts, needs cross-sell into sticky categories to retain customers. Seasonal categories (Christmas, Toys) are churn-heavy and best used for acquisition funnels, not as loyalty drivers.

6 Seasonality & Demand Alignment

Understanding sales seasonality is essential for aligning inventory and marketing with demand. This section analyzes monthly sales trends, year-on-year seasonality, and applies rolling averages to reveal underlying patterns.

```
MONTHLY SALES TREND
--------------------------------------------------
```


21

First Purchase Distribution by Month

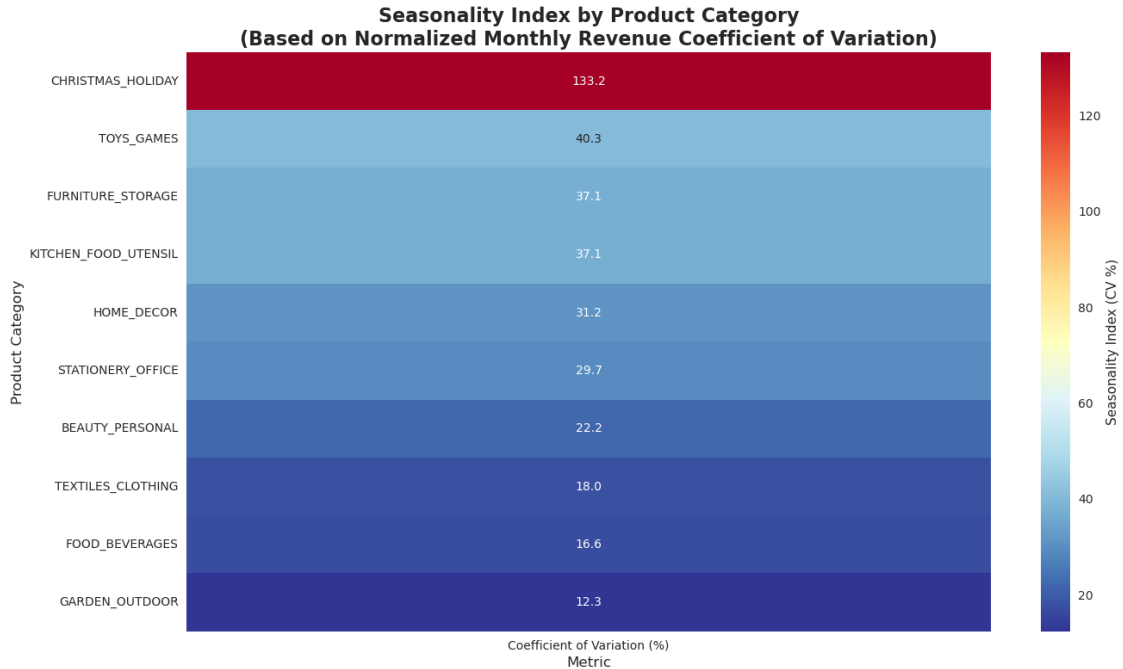
First Purchase Distribution by Month

Revenue growth is concentrated in holiday months, but acquisition surges in December aren't translating into long-term loyalty. To stabilize growth, the business needs to activate mid-year demand and convert seasonal shoppers into repeat buyers through targeted retention campaigns (e.g., nudging December buyers to return in Q1/Q2 with bundles or promotions).

MONTHLY REVENUE TREND BY PRODUCT CATEGORY

Monthly Revenue Trend by Product Category (2010)
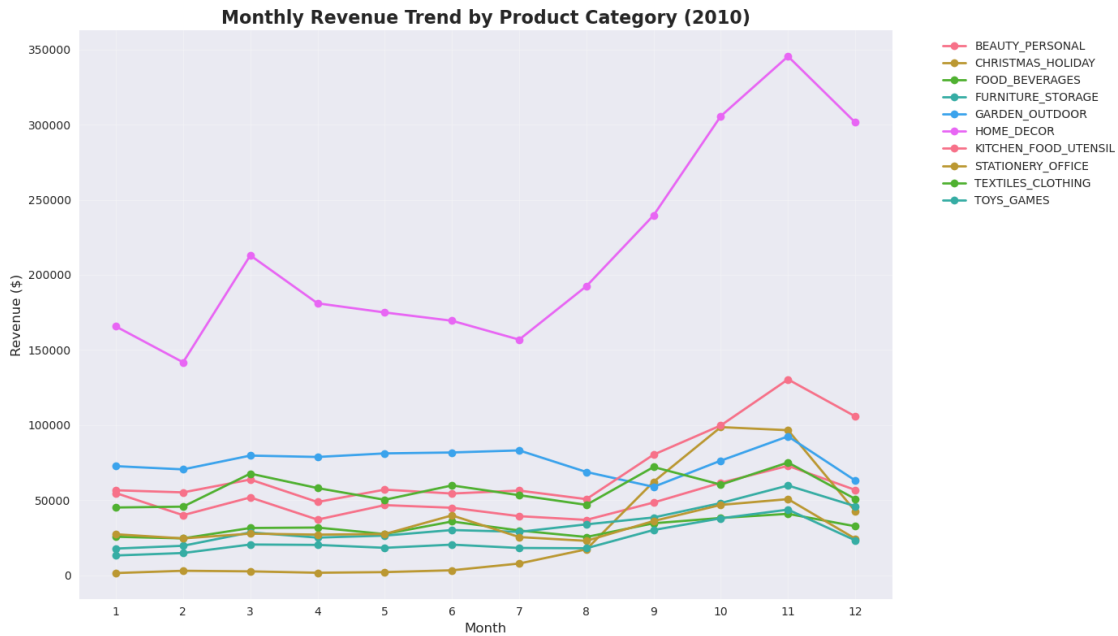
- Home Decor is the backbone category, with a steady rise throughout the year and a sharp acceleration from September to November. It peaks in November (~$800k), driving overall revenue performance.

- Christmas Holiday category has a clear seasonal spike in Q4 (Oct–Dec), peaking in November (~$200k), but contributes very little outside the holiday period.

SEASONALITY INDEX ANALYSIS BY CATEGORY

Seasonality Index by Product Category
(Based on Normalized Monthly Revenue Coefficient of Variation)

Product Category	Value
CHRISTMAS_HOLIDAY	133.2
TOYS_GAMES	40.3
FURNITURE_STORAGE	37.1
KITCHEN_FOOD_UTENSIL	37.1
HOME_DECOR	31.2
STATIONERY_OFFICE	29.7
BEAUTY_PERSONAL	22.2
TEXTILES_CLOTHING	18.0
FOOD_BEVERAGES	16.6
GARDEN_OUTDOOR	12.3

Coefficient of Variation (%)
Metric

1. **Christmas Holiday dominates seasonality.**

- CV = 133%, far above all other categories. This confirms the category is highly seasonal, almost entirely dependent on Q4 sales. It's a short-lived revenue booster, not a stable contributor.

2. **Moderate volatility categories: Furniture & Toys**

- Likely tied to gift-giving or occasional big-ticket purchases. These categories need timed promotions and controlled inventory, since demand isn't spread evenly through the year.

3. **Stable base categories: Home Decor, Beauty, Kitchen, Garden, Textiles, Stationery**

- Ranging between 18%–31% CV, these categories are far more predictable, making them reliable for steady cash flow. Home Decor (30.5%) shows some uplift in Q4 but remains a year-round anchor.ow.

Seasonality is highly uneven across categories. While Christmas drives short bursts of sales, most core categories (Home Decor, Kitchen, Textiles, Garden, Stationery) show steady demand throughout the year. To optimize, the business should treat Christmas as a tactical spike category while ensuring baseline growth comes from the stable, everyday categories

7 Country Context

While the country distribution is heavily skewed toward the UK, it is useful to review revenue share by country for completeness.

```
REVENUE SHARE BY COUNTRY (TOP 10)
--------------------------------------------------
```


Revenue Share by Country (Top 10)
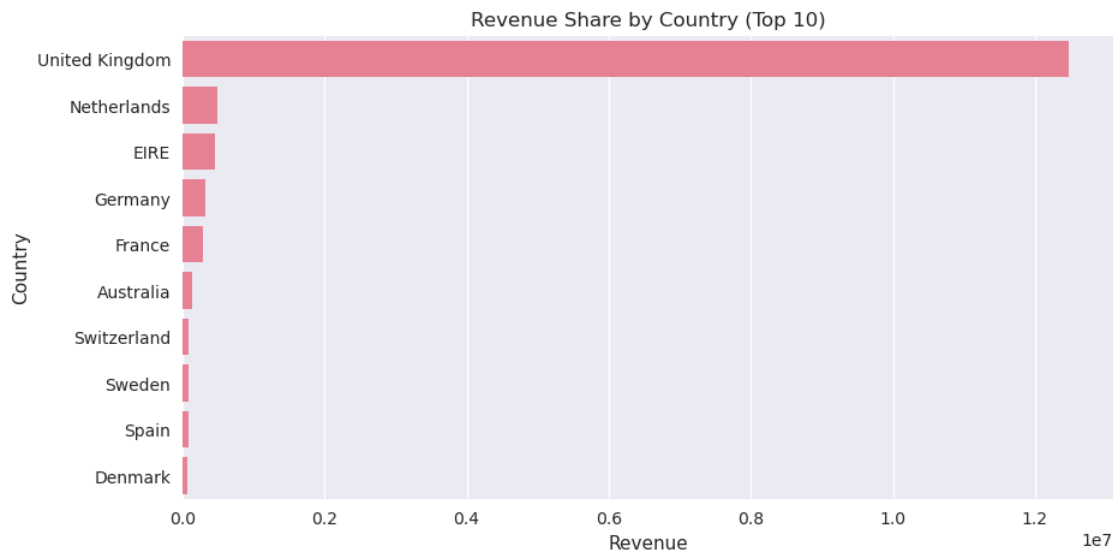
The UK overwhelmingly dominates revenue, with other countries contributing marginally. While international expansion is limited, the UK market remains the primary focus for growth and retention strategies.

8 Summary

- **Revenue performance has stalled**: Sales plateaued in 2011 (-1.2% **YoY**) despite slightly higher order values. The underlying issue is not basket size, but **lack of net customer growth** and weakening repeat engagement.
- **Customer base is leaky**: Acquisition was strong (**37% new customers in 2011**), but almost entirely offset by churn (**37% lost from 2010**). Retention is weak, only **63% of customers were retained YoY**. Loyal customers (44% of base) drive nearly **88% of revenue**, highlighting over-reliance on a small group. Most new buyers churn after their first purchase.
- **Product mix is skewed**: Revenue is highly concentrated in **Home Decor and Kitchen/Food (~60%)**, but overall product variety shrank (-7.9% **fewer SKUs**). Core categories like Kitchen, Beauty, and Garden declined YoY, while seasonal categories (Christmas, Toys) grew but do not sustain repeat purchases. The assortment is long-tailed, with ~**100% of revenue from products contributing <1% each**, making SKU efficiency a concern.
- **Repeat-conversion depends on product type**: Customers starting with **Kitchen/Food or Beauty** are more likely to become loyal, while those entering via **Christmas/Holiday** or novelty categories churn at higher rates. Home Décor remains the biggest acquisition funnel but underconverts without cross-sell into stickier categories.
- **Sales are highly seasonal**: Revenue spikes in **Q4, especially November**, driven by Home Décor and Christmas items. Seasonality index confirms Christmas as extremely volatile

(>130% CV), while staples like Kitchen, Garden, and Stationery remain stable year-round. Seasonal peaks bring customer surges but do not translate into loyalty.

- **Geographic footprint is narrow**: Sales are dominated by the UK, with minimal contribution from international markets. Expansion potential exists, but only after strengthening retention and product performance at home.

9 Recommendations

1. **Strengthen Retention & Engagement**

   - Convert one-time buyers into repeat customers with onboarding and reactivation campaigns right after first purchase.
   - Protect high-value loyal customers through loyalty programs and personalized offers.
   - Develop a **Customer 360 view** to unify customer data, enabling better targeting of at-risk customers and improving lifetime value forecasts.

2. **Rebalance Product Strategy**

   - Prioritize Home Decor as the anchor category: it contributes ~40% of revenue and is the largest customer entry funnel. Focus on improving its conversion by bundling or cross-selling into stickier categories like Kitchen/Food and Beauty.
   - Revitalize shrinking core categories (Kitchen, Beauty, Garden, Stationery) to build stable demand.
   - Use seasonal categories (Christmas, Toys) as **acquisition funnels**, but cross-sell into staples to retain customers.
   - Optimize the long-tail by pruning weak SKUs while elevating "hidden hero" products with high repeat potential.

3. **Align to Seasonality & Demand**

   - Treat Christmas as a tactical spike, but plan campaigns to smooth demand into mid-year months.
   - Consistently stock Home Decor and Kitchen/Food year-round as anchor categories.
   - Build a **demand forecasting capability** (using historical sales trends) to anticipate peaks and troughs, aligning marketing and inventory to capture upside and avoid waste.

4. **Geographic Focus**

   - Deepen engagement in the UK where scale is strongest.
   - Revisit international expansion once retention and product efficiency are stabilized.