

Analyzing Sales Data for a Hypothetical Gardening Online Shop

Table of content

1. <i>Introduction</i>	3
2. <i>Goal and key questions of the analysis</i>	3
3. <i>Synthetical dataset & preparation</i>	4
1. Generated tables	4
2. Key Features of the Synthetic Data	5
3. Preparation	6
4. <i>Data Quality & Preprocessing</i>	8
1. Basic Data Quality Checks	9
1. Row Counts Across All Tables	9
2. Check for Missing or Anomalous Values in Key Columns	10
3. Integrity Checks: Ensure Foreign Keys Exist	10
2. Exploratory Data Analysis (EDA)	11
1. Descriptive Statistics on Sales Data	11
2. Sales Date Range Check	12
3. Monthly Sales Overview	12
4. Top-Selling Products	13
5. Sales by Region	14
6. Summary EDA	15
5. <i>Analysis Roadmap for Profit Optimization (SQL)</i>	15
1. High-Level Business KPIs Summary	15
2. Sales Performance Analysis	16
1. Profitability Analysis	17
2. Regional Performance	21
3. Seasonal Trends Analysis	28
3. Customer Insights Analysis	31
1. High level customer KPIs	32
2. Customer Lifetime Value (CLV)	34
3. Retention Rate	35
4. Loyalty Score Impact on Repeat Purchases	36
5. RFM Segmentation	37
6. Segment Profitability	39
4. Inventory Efficiency Analysis	40
1. Inventory Turnover Ratio	41

2. Stockout Rate	45
3. Overstock Rate	49
5. Marketing Effectiveness Analysis	52
1. Campaign ROI Analysis	52
2. Conversion Rate by Segment	57
6. Returns Analysis	61
1. Product/Category Return Rates	61
2. Return Reasons Distribution	64
3. Net Profit Impact from Returns	67
7. Strategic Takeaways & Next Steps	70
6. Advanced analytics (Python)	71
1. Customer Segmentation	71
2. Churn Prediction	72
3. Product Recommendation System	72
4. Pricing Strategy Optimization: Price Elasticity & Discount Effectiveness	73
5. Campaign Effectiveness Analysis	73
6. Sales Forecast	73
7. Inventory Optimization	74
7. Interactive Dashboards & Data Storytelling	74
7.1 Executive Summary	75
7.2 Sales Performance	76
7.3 Customer Insight	78
7.4 Inventory Efficiency	79
7.5 Marketing & Campaign	81
7.6 Returns & Profitability	82
8. Conclusion	83

1. Introduction

In 2020, a dynamic online shop specializing in plants and accessories was founded in Germany. With a focus on quality and customer satisfaction, the shop quickly gained traction and expanded its operations to the Netherlands shortly after. Buoyed by its success, the shop ventured into Southern Europe, establishing a presence in Spain and Italy. Most recently, the shop expanded to the UK and Ireland, marking its entry into the English-speaking market.



This project aims to analyze the shop's sales performance, customer behavior, and operational efficiency using a synthetic database that simulates real-world scenarios.

2. Goal and key questions of the analysis

Main Goal: Increase profit by optimizing sales strategies, reducing costs, and improving customer satisfaction.

Key Questions:

- Which products and categories have the highest profit margins?
- Are there underperforming products or regions that need attention?
- How can we reduce costs (e.g., shipping, returns) without compromising customer satisfaction?

- Which customer segments are most profitable, and how can we target them better?

This project seeks to uncover actionable insights from the synthetic data, including:

- Sales Performance: Identify top-performing products, categories, and regions.
- Customer Insights: Understand customer behavior, preferences, and loyalty.
- Inventory Efficiency: Analyze inventory management practices and their impact on sales.
- Marketing Effectiveness: Evaluate the success of marketing campaigns in driving sales.
- Returns Analysis: Investigate the reasons behind returns and their impact on profitability.

3. Sythetical dataset & preparation

1. Generated tables

The database includes the following tables:

1. products_df: Contains information about the products, including categories, popularity scores, and pricing.
2. customers_df: Stores customer details such as demographics, loyalty scores, segments (retail vs. wholesale), and buying preferences.
3. sales_df: Records all sales transactions, including product IDs, customer IDs, sales amounts, and dates.
4. returns_df: Tracks returned items, including reasons for returns and associated product categories.
5. campaigns_df: Documents marketing campaigns and their impact on sales in specific regions.
6. inventory_df: Manages stock levels, reorder points, and inventory replenishment.
7. not_realized_purchases_df: Captures failed purchase attempts due to insufficient inventory.

8. inventory_snapshots_df: Contains two snapshots of the stock levels of the products per day

2. Key Features of the Synthetic Data

To ensure the data reflects real-world dynamics, the following features were incorporated:

1. Seasonal influence:
 - o General seasonality: Sales fluctuate based on the time of year, with variations across countries.
 - o Product category seasonality: Certain product categories experience higher demand during specific seasons.
2. Product popularity:
 - o Products are assigned a popularity score, which influences their sales frequency. Popular products are sold more often.
3. Customer behavior:
 - o Segmentation: Customers are divided into retail and wholesale segments, each with distinct buying behaviors.
 - Retail customers: Tend to buy smaller quantities but more frequently.
 - Wholesale customers: Purchase larger quantities but less frequently.
 - o Price sensitivity: Products with high prices are purchased in lower quantities.
 - o Age preferences: Customers exhibit buying preferences based on their age.
 - o Loyalty scores: Customers with higher loyalty scores purchase more frequently.
4. Marketing campaigns:
 - o The shop launched 12 marketing campaigns, each designed to boost sales in a specific country for a limited time. These campaigns are reflected in the campaigns_df table.
5. Inventory management:

- The sales_df and inventory_df tables are interconnected. When a sale occurs, the inventory is reduced.
- If inventory levels fall below the reorder point, new stock is added to prevent shortages.
- Sales that cannot be fulfilled due to insufficient inventory are recorded in the not_realized_purchases_df table.
- The stock levels of the products are recorded twice a day (before and after restocking)

6. Returns:

- The likelihood of returns is influenced by the product category. For example, certain categories may have higher return rates due to customer dissatisfaction or product defects

3. Preparation

Goal:

Prepare the synthetic data for analysis by collecting, structuring, and importing it into a reliable database system.

Overview:

- The synthetic data is generated in Colab and exported as CSV files.
- The next step is to create the necessary database tables in PostgreSQL using pgAdmin and then import the CSV files into those tables.

Process Details:

1. Data Export:

- Generate and download the CSV files from Colab containing your synthetic data.

2. Creating Data Tables in PostgreSQL:

1. Creating the data tables in PostgreSQL:

```
sql
CREATE TABLE sales ( SaleID SERIAL PRIMARY KEY, SalesDate DATE NOT NULL, ProductID INT NOT NULL, CustomerID INT NOT NULL, Region VARCHAR(50), SalesAmount DECIMAL(10,2), Price DECIMAL(10,2), DiscountRate DECIMAL(5,2), DiscountAmount DECIMAL(10,2), Quantity INT, PaymentMethod VARCHAR(50), ShippingCost DECIMAL(10,2) );
```

```
sql
CREATE TABLE returns (
    ReturnID SERIAL PRIMARY KEY,
    SaleID INT REFERENCES sales(SaleID) ON DELETE CASCADE,
    ReturnDate DATE NOT NULL,
    Reason VARCHAR(100)
);
```

```
sql
CREATE TABLE products (
    ProductID SERIAL PRIMARY KEY,
    ProductName VARCHAR(100) NOT NULL,
    Category VARCHAR(50),
    CostPrice DECIMAL(10,2),
    SellingPrice DECIMAL(10,2),
    Popularity INT
);
```

```
sql
CREATE TABLE stock (
    ProductID INT PRIMARY KEY REFERENCES products(ProductID)
ON DELETE CASCADE,
    StockLevel DECIMAL(10,2),
    ReorderPoint DECIMAL(10,2)
);
```

```
sql
CREATE TABLE customers (
    CustomerID SERIAL PRIMARY KEY,
    CustomerName VARCHAR(100),
    Region VARCHAR(50),
    Segment VARCHAR(50),
    LoyaltyScore INT,
    Age INT,
    Gender VARCHAR(10),
    JoinDate DATE,
    Email VARCHAR(100),
    PhoneNumber VARCHAR(20),
    BuyingFrequency INT
);
```

```
sql
CREATE TABLE campaigns (
    CampaignID SERIAL PRIMARY KEY,
    CampaignName VARCHAR(255),
    StartDate DATE,
    Region VARCHAR(50),
    Budget DECIMAL(15,2),
    EndDate DATE,
    Impact DECIMAL(10,5)
);
```

```
sql
CREATE TABLE Inventory_Snapshots (
    SnapshotID SERIAL PRIMARY KEY,
    ProductID INT NOT NULL,
    Date DATE NOT NULL,
    StockLevel INT NOT NULL,
    SnapshotType VARCHAR(50) NOT NULL,
    FOREIGN KEY (ProductID) REFERENCES Products(ProductID)
);
```

4. Data Quality & Preprocessing

Goal:

Ensure data reliability and consistency before starting in-depth analysis. Even with synthetic data, this step verifies that the generated data reflects realistic business patterns and is logically sound.

Key Questions & Metrics:

- **Data Structure & Integrity:**
 - Are all tables populated with the expected number of records?
 - Are primary and foreign keys correctly linked?
- **Descriptive Statistics:**
 - What are the minimum, maximum, mean, and median values for key fields such as SalesAmount, Quantity, and ShippingCost?
 - Do these values fall within realistic business ranges?
- **Temporal Consistency:**
 - Are sales dates correctly formatted and within the expected period?
 - Do seasonal trends appear plausible?
- **Category & Segment Validation:**
 - Do all sales records correspond to valid products and customers?
 - Are the distributions across product categories and customer segments logical?

1. Basic Data Quality Checks

In the following section are basic quality checks performed to estimate if the data is usable for further analysis.

1. Row Counts Across All Tables

Purpose: Confirm that all tables have the expected number of records.

```
sql
SELECT 'sales' AS table_name, COUNT(*) AS row_count FROM sales
UNION ALL
SELECT 'returns', COUNT(*) FROM returns
UNION ALL
SELECT 'products', COUNT(*) FROM products
UNION ALL
SELECT 'stock', COUNT(*) FROM stock
UNION ALL
SELECT 'customers', COUNT(*) FROM customers
UNION ALL
SELECT 'campaigns', COUNT(*) FROM campaigns
UNION ALL
SELECT 'inventory_snapshots', COUNT(*) FROM
inventory_snapshots;
```

Result:

table_name	row_count
campaigns	12
products	48
stock	48
returns	11007
customers	25000
inventory_snapshots	70176

As expected

2. Check for Missing or Anomalous Values in Key Columns

Purpose: Even in synthetic data, verifying key columns helps document your process and ensure no unexpected anomalies.

```
sql
SELECT
    COUNT(*) AS total_rows,
    COUNT(SaleID) AS non_null_sale_id,
    COUNT(SalesDate) AS non_null_sales_date,
    COUNT(ProductID) AS non_null_product_id,
    COUNT(CustomerID) AS non_null_customer_id,
    COUNT(SalesAmount) AS non_null_sales_amount
FROM sales;
```

Result:

total_rows	non_null_sale_id	non_null_sales_date	non_null_product_id	non_null_customer_id	non_null_sales_amount
208700	208700	208700	208700	208700	208700

No NULL values in sales table

3. Integrity Checks: Ensure Foreign Keys Exist

a. Validate that every ProductID in sales exists in products:

Purpose: Ensures referential integrity for ProductID.

```
sql
SELECT s.ProductID, COUNT(*) AS occurrences
FROM sales s
LEFT JOIN products p ON s.ProductID = p.ProductID
WHERE p.ProductID IS NULL
GROUP BY s.ProductID;
```

Result: empty

Integrity maintained

b. Validate that every CustomerID in sales exists in customers:

Purpose: Ensures referential integrity for CustomerID.

```
sql
SELECT s.Customerid, COUNT(*) AS occurrences
FROM sales s
LEFT JOIN customers c ON s.customerid = c.customerid
WHERE c.Customerid IS NULL
GROUP BY s.customerid;
```

Result: empty

Integrity maintained

2. Exploratory Data Analysis (EDA)

In the following section an EDA is performed on the data.

1. Descriptive Statistics on Sales Data

Purpose: Understand the range and central tendency of your sales figures.

```
sql
SELECT
    MIN(SalesAmount) AS min_sales_amount,
    MAX(SalesAmount) AS max_sales_amount,
    ROUND(AVG(SalesAmount),2) AS avg_sales_amount,
    SUM(SalesAmount) AS total_sales_amount
FROM sales;
```

Results:

"min_sales_amount"	"max_sales_amount"	"avg_sales_amount"	"total_sales_amount"
2.24	2800.89	117.71	24566498.73

As expected

2. Sales Date Range Check

Purpose: Validate that your time dimension covers the expected period.

```
sql
SELECT
    MIN(SalesDate) AS first_sale,
    MAX(SalesDate) AS last_sale
FROM sales;
```

Result:

"first_sale"	"last_sale"
"2023-01-01"	"2024-12-31"

As expected

3. Monthly Sales Overview

Purpose: Examine overall sales trends and seasonality.

```
sql
```

```
SELECT
    EXTRACT(MONTH FROM SalesDate) AS month_number,
    TO_CHAR(SalesDate, 'Month') AS month_name,
    SUM(CASE WHEN EXTRACT(YEAR FROM SalesDate) = 2023 THEN
SalesAmount ELSE 0 END) AS revenue_2023,
    SUM(CASE WHEN EXTRACT(YEAR FROM SalesDate) = 2024 THEN
SalesAmount ELSE 0 END) AS revenue_2024,
    COUNT(CASE WHEN EXTRACT(YEAR FROM SalesDate) = 2023 THEN 1
END) AS orders_2023,
    COUNT(CASE WHEN EXTRACT(YEAR FROM SalesDate) = 2024 THEN 1
END) AS orders_2024
FROM sales
WHERE SalesDate BETWEEN '2023-01-01' AND '2024-12-31'
GROUP BY month_number, month_name
ORDER BY month_number;
```

Result:

Month Number	Month Name	Revenue 2023	Revenue 2024	Orders 2023	Orders 2024
1	January	485865.52	546616.01	3941	4497
2	February	496004.01	566713.06	4407	4939
3	March	752701.06	853473.15	7088	7642
4	April	1182354.97	1116436.29	10589	10133
5	May	1512030.89	1526419.76	13317	13228
6	June	1608915.79	1705808.04	13302	14600
7	July	1515635.55	1536798.28	11893	12477
8	August	1401475.09	1476653.19	11058	11589
9	September	1123060.99	1091103.07	9493	9319
10	October	894339.87	898126.01	7716	7719
11	November	568917.25	594440.94	5023	5298
12	December	533177.40	579432.54	4567	4865

- ➔ As expected, with clear seasonal pattern with more purchases in summer, yearly growth

4. Top-Selling Products

Purpose: Identify which products contribute most to revenue.

```

sql
SELECT
    p.ProductName,
    COUNT(s.SaleID) AS total_sales,
    SUM(s.Quantity) AS total_units_sold,
    SUM(s.SalesAmount) AS total_revenue
FROM sales s
JOIN products p ON s.ProductID = p.ProductID
GROUP BY p.ProductName
ORDER BY total_revenue DESC
LIMIT 10;

```

Results:

"productname"	"total_sales"	"total_units_sold"	"total_revenue"
"Lounge Chair"	5994	7230	1933360.59
"Swing Seat"	5855	15839	1791360.32
"Solar Lanterns"	5315	14454	1573787.48
"Patio Table"	2363	2785	1405676.11
"Sprinkler System"	8283	21775	1184985.95
"Bird Bath"	4149	10868	1178649.81
"Garden Gnome"	5946	15395	1130852.97
"Spider Plant"	6350	17352	1033531.76
"Garden Bench"	5325	14552	1005974.87
"Garden Rake"	5013	13548	954073.79

Outdoor furniture is a popular category and has high prices, therefore its logical to have these products in this list

5. Sales by Region

Purpose: Understand regional performance and potential differences in customer behavior.

```

sql
SELECT
    Region,
    COUNT(*) AS total_orders,
    SUM(SalesAmount) AS total_revenue,
    ROUND(AVG(SalesAmount),2) AS avg_order_value
FROM sales
GROUP BY Region
ORDER BY total_revenue DESC;

```

Result:

Region	Total Orders	Total Revenue	Avg Order Value
Germany	50691	6101932.46	120.38
Netherlands	51480	5986245.89	116.28
Italy	35082	4188694.08	119.40
Spain	34132	3932357.69	115.21
Ireland	18918	2239137.82	118.36
UK	18397	2118130.79	115.13

As expected: Germany and Netherlands are strongest regions, followed by Italy and Spain and Ireland and UK being the weakest regions.

6. Summary EDA

- Data Completeness: No missing or NULL values in key columns.
- Referential Integrity: All foreign key relationships are valid.
- Sales Trends: Seasonal variations are evident, with summer peaks.
- Product Insights: High-value outdoor furniture drives revenue.
- Regional Sales: Germany and the Netherlands are top-performing regions

Data quality is strong, supporting further analysis.

5. Analysis Roadmap for Profit Optimization (SQL)

1. High-Level Business KPIs Summary

Goal: Provide a quick executive dashboard to monitor overall sales performance.

Key Metrics:

1. Total Revenue = $\text{SUM}(\text{SalesAmount})$
2. Total Profit = $\text{SUM}(\text{SalesAmount} - (\text{Quantity} \times \text{CostPrice}))$
3. Overall Profit Margin (%) = $(\text{Total Profit} / \text{Total Revenue}) \times 100$
4. Total Orders = $\text{COUNT}(\text{SaleID})$
5. Average Order Value (AOV) = $\text{Total Revenue} / \text{Total Orders}$
6. Return Rate (%) = $(\text{Returned Orders} / \text{Total Orders}) \times 100$

```

sql
WITH SalesData AS (
    SELECT
        SUM(s.SalesAmount) AS TotalRevenue,
        SUM(s.SalesAmount - (s.Quantity * p.CostPrice)) AS TotalProfit,
        COUNT(s.SaleID) AS TotalOrders,
        ROUND(AVG(s.SalesAmount), 2) AS AvgOrderValue
    FROM sales s
    JOIN products p ON s.ProductID = p.ProductID
), ReturnData AS (
    SELECT COUNT(ReturnID) AS ReturnedOrders FROM returns
)
SELECT
    sd.TotalRevenue,
    sd.TotalProfit,
    ROUND((sd.TotalProfit / NULLIF(sd.TotalRevenue, 0)) * 100,
    2) AS ProfitMargin,
    sd.TotalOrders,
    sd.AvgOrderValue,
    r.ReturnedOrders,
    ROUND((r.ReturnedOrders * 100.0) / NULLIF(sd.TotalOrders,
    0), 2) AS ReturnRate
FROM SalesData sd, ReturnData r;

```

Result:

totalrevenue	totalprofit	profitmargin	totalorders	avgordervalue	returnedorders	returnrate
2456649 8.73	4270185 .73	17.38	208700	117.71	11007	5.27

Botanika, the gardening online shop, generated a total revenue of €24.6 million and a profit of €4.27 million, achieving a healthy profit margin of 17.4%. With over 208,700 orders and an average order value of €117.71, Botanika maintained a low return rate of 5.3%, indicating strong customer satisfaction and efficient product selection.

2. Sales Performance Analysis

Goal: Optimize revenue and margins by identifying top- and underperforming products, categories, and regions, and uncovering seasonal sales trends.

Key Questions:

- Which products/categories yield high/low profit margins?
- Which regions have low sales growth, high return rates or high shipping costs?
- Are there seasonal trends?

Key Metrics: Revenue, Profit Margin, Sales Growth, Shipping Cost per Order, Return Rate, Seasonal Demand Index

1. Profitability Analysis

Goal: Identify which products and categories generate the highest/lowest profit margins.

```
sql
-- Calculate profit margin per product to detect Top
performing products
SELECT
    p.ProductID,
    p.ProductName,
    p.Category,
    SUM(s.SalesAmount) AS TotalRevenue,
    SUM(s.Quantity * p.CostPrice) AS TotalCost,
    SUM(s.SalesAmount) - SUM(s.Quantity * p.CostPrice) AS
TotalProfit,
    ROUND((SUM(s.SalesAmount) - SUM(s.Quantity * p.CostPrice)) /
    SUM(s.SalesAmount) * 100, 2) AS ProfitMargin,
    ROUND(AVG(s.DiscountRate) * 100, 2) AS AvgDiscountRate
FROM sales s
JOIN products p ON s.ProductID = p.ProductID
GROUP BY p.ProductID, p.ProductName, p.Category
ORDER BY ProfitMargin DESC
LIMIT 10; -- Top 10 profitable products
```

Result:

Product ID	Product Name	Category	Total Revenue	Total Cost	Total Profit	Profit Margin	Avg Discount Rate
23	Neem Oil Spray	Pest Control	532021.54	402765.45	129256.09	24.30	7.50
22	Insect Netting	Pest Control	386342.04	292612.26	93729.78	24.26	7.48

21	Slug Repellent	Pest Control	264538.71	200579.23	63959.48	24.18	7.55
24	Ant Killer	Pest Control	211034.37	161475.30	49559.07	23.48	7.79
4	Carrot Seeds	Seeds	86976.25	66812.46	20163.79	23.18	11.78
2	Basil Seeds	Seeds	123739.79	95075.10	28664.69	23.17	11.74
3	Sunflower Seeds	Seeds	31874.63	24619.20	7255.43	22.76	11.87
1	Tomato Seeds	Seeds	106810.66	82513.86	24296.80	22.75	11.98
47	Trellis Arch	Decorative Gardening Items	854454.17	662440.10	192014.07	22.47	20.45
45	Garden Gnome	Decorative Gardening Items	1130852.97	877668.95	253184.02	22.39	20.49

sql

```
-- Calculate profit margin per product to detect Under performing products
SELECT
    p.ProductID,
    p.ProductName,
    p.Category,
    SUM(s.SalesAmount) AS TotalRevenue,
    SUM(s.Quantity * p.CostPrice) AS TotalCost,
    SUM(s.SalesAmount) - SUM(s.Quantity * p.CostPrice) AS TotalProfit,
    ROUND((SUM(s.SalesAmount) - SUM(s.Quantity * p.CostPrice)) / SUM(s.SalesAmount) * 100, 2) AS ProfitMargin,
    ROUND(AVG(s.DiscountRate) * 100, 2) AS AvgDiscountRate
FROM sales s
JOIN products p ON s.ProductID = p.ProductID
GROUP BY p.ProductID, p.ProductName, p.Category
ORDER BY ProfitMargin ASC
LIMIT 10; -- Top 10 un-profitable products
```

Result:

productid	productname	category	totalrevenue	totalcost	totalprofit	profitmargin	avgdiscountrate
16	Swing Seat	Outdoor Furniture	1791360.32	1579940.25	211420.07	11.80	18.34
13	Garden Bench	Outdoor Furniture	1005974.87	886653.36	119321.51	11.86	18.27
15	Lounge Chair	Outdoor Furniture	1933360.59	1689072.60	244287.99	12.64	18.10
14	Patio Table	Outdoor Furniture	1405676.11	1224564.50	181111.61	12.88	17.95
43	Raised Garden Bed	Pots & Planters	253087.65	218858.85	34228.80	13.52	14.83
41	Terracotta Pot	Pots & Planters	314498.46	271903.05	42595.41	13.54	14.76
44	Self-Watering Pot	Pots & Planters	328515.69	281646.31	46869.38	14.27	14.64
42	Hanging Basket	Pots & Planters	329731.30	282252.36	47478.94	14.40	14.41
9	Hand Trowel	Gardening Tools	714681.47	605004.01	109677.46	15.35	13.51
11	Garden Hoe	Gardening Tools	437331.04	369980.80	67350.24	15.40	13.52

```

sql
-- Profit margin by category
SELECT
    p.Category,
    SUM(s.SalesAmount) AS TotalRevenue,
    SUM(s.Quantity * p.CostPrice) AS TotalCost,
    SUM(s.SalesAmount) - SUM(s.Quantity * p.CostPrice) AS
TotalProfit,
    ROUND((SUM(s.SalesAmount) - SUM(s.Quantity * p.CostPrice)) /
    SUM(s.SalesAmount) * 100, 2) AS ProfitMargin,
    ROUND(AVG(s.DiscountRate) * 100, 2) AS AvgDiscountRate
FROM sales s
JOIN products p ON s.ProductID = p.ProductID
GROUP BY p.Category
ORDER BY ProfitMargin DESC;

```

Results:

Category	Total Revenue	Total Cost	Total Profit	Profit Margin	Avg Discount Rate
Pest Control	1393936.66	1057432.24	336504.42	24.14	7.60
Seeds	349401.33	269020.62	80380.71	23.01	11.82
Decorative Gardening Items	4737744.43	3676955.41	1060789.02	22.39	20.46
Bulbs	293592.47	236631.83	56960.64	19.40	15.83
Indoor Plants	2125456.33	1731543.20	393913.13	18.53	17.37
Watering Systems	2329861.55	1903382.75	426478.80	18.30	10.52
Soil & Fertilizers	1062185.60	868783.86	193401.74	18.21	8.55
Compost	954407.37	788041.70	166365.67	17.43	9.50
Rakes	1815749.25	1518158.84	297590.41	16.39	12.53
Gardening Tools	2141958.75	1811471.27	330487.48	15.43	13.48
Pots & Planters	1225833.10	1054660.57	171172.53	13.96	14.71
Outdoor Furniture	6136371.89	5380230.71	756141.18	12.32	18.20

Outdoor Furniture is the top revenue-generating category, but with the lowest profit margin (12.3%), indicating high sales volume but low profitability—likely due to high costs and heavy discounting.

In contrast, Pest Control products deliver the highest profit margins (24.1%), making them strong contributors to overall profitability despite lower revenue.

Seeds offer consistently solid margins with minimal discounting, positioning them as a stable, low-risk product group.

Overall, product-level analysis highlights opportunities to optimize pricing and discount strategies, especially in high-revenue but low-margin categories like Outdoor Furniture.

2. Regional Performance

Goal: Identify underperforming regions in terms of sales growth, high return rates and high shipping costs.

```
sql
SELECT
    s.Region,
    SUM(s.SalesAmount) AS TotalRevenue,
    COUNT(s.SaleID) AS TotalOrders,
    ROUND(SUM(s.SalesAmount) / NULLIF(COUNT(s.SaleID), 1), 2)
AS AverageOrderValue,
    SUM(s.ShippingCost) AS TotalShippingCost,
    ROUND(SUM(s.ShippingCost) / NULLIF(COUNT(s.SaleID), 1), 2)
AS ShippingCostPerOrder,
    ROUND((COUNT(r.ReturnID) * 100.0) /
NULLIF(COUNT(s.SaleID), 1), 2) AS ReturnRate
FROM sales s
LEFT JOIN returns r ON s.SaleID = r.SaleID
GROUP BY s.Region
ORDER BY TotalRevenue DESC;
```

Results:

Region	Total Revenue	Total Orders	Average Order Value	Total Shipping Cost	Shipping Cost per Order	Return Rate
Germany	6118082.65	50781	120.48	837911.93	16.50	5.35
Netherlands	6000738.59	51564	116.37	747104.70	14.49	5.32
Italy	4198553.08	35137	119.49	700176.91	19.93	5.08
Spain	3939773.65	34186	115.25	427897.30	12.52	5.12
Ireland	2244589.69	18950	118.45	662237.62	34.95	5.37
UK	2121021.83	18417	115.17	597299.62	32.43	5.41

```

sql
-- Sales Growth by Region (Month-over-Month) (MoM)
WITH monthly_sales AS (
    SELECT
        Region,
        EXTRACT(YEAR FROM SalesDate) AS year,
        EXTRACT(MONTH FROM SalesDate) AS month,
        TO_CHAR(SalesDate, 'Month') AS month_name,
        SUM(SalesAmount) AS revenue,
        COUNT(*) AS orders
    FROM sales
    WHERE SalesDate BETWEEN '2023-01-01' AND '2024-12-31'
    GROUP BY Region, year, month, month_name
)

SELECT
    m.Region,
    m.month,
    m.month_name,
    m.year,
    m.revenue AS current_revenue,
    m.orders AS current_orders,
    p.revenue AS prev_month_revenue,
    p.orders AS prev_month_orders,
    ROUND((m.revenue - p.revenue) / NULLIF(p.revenue, 0) *
100, 2) AS mom_revenue_growth,
    ROUND((m.orders::numeric - p.orders) / NULLIF(p.orders, 0)
* 100, 2) AS mom_order_growth
FROM monthly_sales m
LEFT JOIN monthly_sales p ON
    m.Region = p.Region AND
    m.year = p.year AND
    m.month = p.month + 1
ORDER BY m.Region, m.year, m.month;

```

Results (excerpt):

Region	Month	Month Name	Year	Current Revenue	Current Orders	Prev Month Revenue	Prev Month Orders	MoM Revenue Growth	MoM Order Growth
Germany	1	January	2023	95972.78	819				
Germany	2	February	2023	106202.23	931	95972.78	819	10.66	13.68
Germany	3	March	2023	172610.12	1545	106202.23	931	62.53	65.95

Germ any	4	April	20 23	292584 .79	2613	172610 .12	1545	69.51	69.1 3
Germ any	5	May	20 23	409791 .63	3420	292584 .79	2613	40.06	30.8 8
Germ any	6	June	20 23	433389 .42	3455	409791 .63	3420	5.76	1.02
Germ any	7	July	20 23	409564 .08	3105	433389 .42	3455	-5.50	- 10.1 3
Germ any	8	August	20 23	357619 .36	2749	409564 .08	3105	-12.68	- 11.4 7
Germ any	9	Septem ber	20 23	255795 .47	2097	357619 .36	2749	-28.47	- 23.7 2
Germ any	10	Octobe r	20 23	221918 .63	1879	255795 .47	2097	-13.24	- 10.4 0
Germ any	11	Novem ber	20 23	120133 .93	1061	221918 .63	1879	-45.87	- 43.5 3
Germ any	12	Decem ber	20 23	99904. 24	938	120133 .93	1061	-16.84	- 11.5 9
Germ any	1	January	20 24	110336 .96	836				
Germ any	2	Februar y	20 24	117704 .13	1045	110336 .96	836	6.68	25.0 0
Germ any	3	March	20 24	213876 .60	1767	117704 .13	1045	81.71	69.0 9
Germ any	4	April	20 24	296683 .94	2616	213876 .60	1767	38.72	48.0 5
Germ any	5	May	20 24	437966 .69	3708	296683 .94	2616	47.62	41.7 4
Germ any	6	June	20 24	466203 .90	4029	437966 .69	3708	6.45	8.66
Germ any	7	July	20 24	400443 .18	3155	466203 .90	4029	-14.11	- 21.6 9
Germ any	8	August	20 24	376514 .47	2984	400443 .18	3155	-5.98	-5.42
Germ any	9	Septem ber	20 24	239036 .60	2051	376514 .47	2984	-36.51	- 31.2 7
Germ any	10	Octobe r	20 24	208209 .54	1761	239036 .60	2051	-12.90	- 14.1 4

Germ any	11	Novem ber	20 24	133674 .04	1066	208209 .54	1761	-35.80	- 39.4 7
Germ any	12	Decem ber	20 24	125795 .73	1061	133674 .04	1066	-5.89	-0.47
Irelan d	1	January	20 23	58738. 28	476				
Irelan d	2	Februar y	20 23	39676. 13	422	58738. 28	476	-32.45	- 11.3 4
Irelan d	3	March	20 23	68655. 07	661	39676. 13	422	73.04	56.6 4
Irelan d	4	April	20 23	119708 .67	1047	68655. 07	661	74.36	58.4 0
Irelan d	5	May	20 23	118574 .83	1043	119708 .67	1047	-0.95	-0.38
Irelan d	6	June	20 23	122843 .35	1025	118574 .83	1043	3.60	-1.73
Irelan d	7	July	20 23	133459 .11	1047	122843 .35	1025	8.64	2.15
Irelan d	8	August	20 23	123073 .09	924	133459 .11	1047	-7.78	- 11.7 5
Irelan d	9	Septem ber	20 23	127581 .61	1073	123073 .09	924	3.66	16.1 3
Irelan d	10	Octobe r	20 23	69225. 88	607	127581 .61	1073	-45.74	- 43.4 3
Irelan d	11	Novem ber	20 23	49160. 71	438	69225. 88	607	-28.99	- 27.8 4
Irelan d	12	Decem ber	20 23	54930. 66	387	49160. 71	438	11.74	- 11.6 4
Irelan d	1	January	20 24	66480. 14	576				
Irelan d	2	Februar y	20 24	58433. 47	552	66480. 14	576	-12.10	-4.17
Irelan d	3	March	20 24	93447. 09	838	58433. 47	552	59.92	51.8 1
Irelan d	4	April	20 24	98540. 73	832	93447. 09	838	5.45	-0.72

```
sql
--Year-over-Year Monthly Growth by Region
WITH monthly_sales AS (
    SELECT
        Region,
        EXTRACT(YEAR FROM SalesDate) AS year,
        EXTRACT(MONTH FROM SalesDate) AS month,
        TO_CHAR(SalesDate, 'Month') AS month_name,
        SUM(SalesAmount) AS revenue,
        COUNT(*) AS orders
    FROM sales
    GROUP BY Region, year, month, month_name
)

SELECT
    curr.Region,
    curr.month,
    curr.month_name,
    curr.year AS current_year,
    prev.year AS previous_year,
    curr.revenue AS current_revenue,
    prev.revenue AS previous_revenue,
    curr.orders AS current_orders,
    prev.orders AS previous_orders,
    -- Revenue growth calculation
    ROUND(
        (curr.revenue - prev.revenue) * 100.0 /
    NULLIF(prev.revenue, 0),
        2
    ) AS yoy_revenue_growth_pct,
    -- Order growth calculation
    ROUND(
        (curr.orders - prev.orders) * 100.0 /
    NULLIF(prev.orders, 0),
        2
    ) AS yoy_order_growth_pct
FROM monthly_sales curr
LEFT JOIN monthly_sales prev ON
    curr.Region = prev.Region AND
    curr.month = prev.month AND
    curr.year = prev.year + 1
WHERE curr.year = 2024
ORDER BY curr.Region, curr.month;
```

Results(excerpt):

Region	Month	Month Name	Current Year	Previous Year	Current Revenue	Previous Revenue	Current Orders	Previous Orders	YoY Revenue Growth (%)	YoY Order Growth (%)
Germany	1	January	2024	2023	110336.96	95972.78	836	819	14.97	2.08
Germany	2	February	2024	2023	117704.13	106202.23	1045	931	10.83	12.24
Germany	3	March	2024	2023	213876.60	172610.12	1767	1545	23.91	14.37
Germany	4	April	2024	2023	296683.94	292584.79	2616	2613	1.40	0.11
Germany	5	May	2024	2023	437966.69	409791.63	3708	3420	6.88	8.42
Germany	6	June	2024	2023	466203.90	433389.42	4029	3455	7.57	16.61
Germany	7	July	2024	2023	400443.18	409564.08	3155	3105	-2.23	1.61
Germany	8	August	2024	2023	376514.47	357619.36	2984	2749	5.28	8.55
Germany	9	September	2024	2023	239036.60	255795.47	2051	2097	-6.55	-2.19
Germany	10	October	2024	2023	208209.54	221918.63	1761	1879	-6.18	-6.28
Germany	11	November	2024	2023	133674.04	120133.93	1066	1061	11.27	0.47
Germany	12	December	2024	2023	125795.73	99904.24	1061	938	25.92	13.11
Ireland	1	January	2024	2023	66480.14	58738.28	576	476	13.18	21.01

Ireland	2	February	2024	2023	58433	39676	552	422	47.2	30.8
Ireland	3	March	2024	2023	93447	68655	838	661	36.1	26.7
Ireland	4	April	2024	2023	98540	11970	832	1047	-	-
Ireland	5	May	2024	2023	12323	11857	1046	1043	3.93	0.29
Ireland	6	June	2024	2023	12647	12284	1136	1025	2.96	10.8
Ireland	7	July	2024	2023	14155	13345	1098	1047	6.06	4.87
Ireland	8	August	2024	2023	16751	12307	1274	924	36.11	37.88
Ireland	9	September	2024	2023	97463	12758	839	1073	-	-
Ireland	10	October	2024	2023	79549	69225	681	607	14.91	12.19
Ireland	11	November	2024	2023	48228	49160	439	438	-1.90	0.23
Ireland	12	December	2024	2023	52589	54930	457	387	-4.26	18.09
Italy	1	January	2024	2023	10558	10037	847	790	5.20	7.22
Italy	2	February	2024	2023	11150	10250	876	885	8.77	-1.02
Italy	3	March	2024	2023	13317	12117	1248	1120	9.90	11.43
Italy	4	April	2024	2023	18011	16719	1627	1443	7.72	12.75
Italy	5	May	2024	2023	23910	20862	2087	1812	14.61	15.18
Italy	6	June	2024	2023	28864	22877	2385	1886	26.17	26.46
Italy	7	July	2024	2023	26861	24998	2178	1839	7.46	18.43
Italy	8	August	2024	2023	23294	24046	1835	1869	-3.13	-1.82
Italy	9	September	2024	2023	19340	20056	1664	1672	-3.57	-0.48
Italy	10	October	2024	2023	17663	16539	1515	1497	6.80	1.20
Italy	11	November	2024	2023	10987	11274	1015	976	-2.54	4.00

Italy	12	Dece mber	202 4	2023	13043 5.41	12082 5.41	103 0	986	7.95	4.46
-------	----	--------------	----------	------	---------------	---------------	----------	-----	------	------

Germany and the Netherlands are Botanika's strongest markets, each generating over €6M in revenue with high order volumes and solid operational efficiency. Mid-tier performers like Italy and Spain also contribute significantly, especially Spain, which shows steady growth and low shipping costs.

In contrast, the UK and Ireland underperform, with revenue around €2M and elevated shipping costs—more than double those of top-performing regions. Return rates across all regions remain low (around 5%), indicating strong product quality and customer satisfaction. To drive growth, efforts should focus on expanding successful strategies in Germany, Spain, and the Netherlands while addressing cost inefficiencies and demand fluctuations in the UK and Ireland.

3. Seasonal Trends Analysis

Goal: Identify demand peaks and sales dips to optimize inventory and marketing.

Key Metrics:

1. Monthly Revenue Growth (%) = $((\text{Current Month Revenue} - \text{Previous Month Revenue}) / \text{Previous Month Revenue}) \times 100$
2. Seasonal Demand Index = $(\text{Sales in Peak Month} / \text{Average Monthly Sales})$
3. High-Season vs Low-Season Revenue Ratio = $\text{Peak Month Revenue} / \text{Lowest Month Revenue}$

```

WITH monthly_sales AS (
    SELECT
        EXTRACT(MONTH FROM SalesDate) AS month,
        TO_CHAR(SalesDate, 'Month') AS month_name,
        SUM(SalesAmount) AS revenue,
        COUNT(*) AS orders
    FROM sales
    GROUP BY month, month_name),
monthly_stats AS (
    SELECT
        m.*,
        LAG(m.revenue) OVER (ORDER BY m.month) AS prev_month_revenue,
        AVG(m.revenue) OVER () AS avg_monthly_revenue,
        MAX(m.revenue) OVER () AS peak_month_revenue,
        MIN(m.revenue) OVER () AS lowest_month_revenue
    FROM monthly_sales m)
SELECT
    ms.month,
    TRIM(ms.month_name) AS month_name,
    ms.revenue,
    ms.orders,
    -- 1. Monthly Revenue Growth (%)
    ROUND(
        (ms.revenue - ms.prev_month_revenue) * 100.0 /
        NULLIF(ms.prev_month_revenue, 0),
        2
    ) AS monthly_revenue_growth_pct,
    -- 2. Seasonal Demand Index
    ROUND(
        ms.revenue / NULLIF(ms.avg_monthly_revenue, 0),
        2
    ) AS seasonal_demand_index,
    -- 3. High-Season vs Low-Season Revenue Ratio
    ROUND(
        ms.peak_month_revenue /
        NULLIF(ms.lowest_month_revenue, 0),
        2
    ) AS seasonal_revenue_ratio,
    -- Seasonal classification
    CASE
        WHEN ms.revenue = ms.peak_month_revenue THEN 'PEAK'
        WHEN ms.revenue = ms.lowest_month_revenue THEN 'LOW'
        ELSE 'NORMAL'
    END AS seasonal_marker,
    ROUND(ms.avg_monthly_revenue, 2) AS avg_monthly_revenue,
    ms.peak_month_revenue,
    ms.lowest_month_revenue
FROM monthly_stats ms
ORDER BY ms.month;

```

Results:

Month	Month Name	Revenue	Orders	Monthly Revenue Growth %	Seasonal Demand Index	Seasonal Revenue Ratio	Seasonal Marker	Avg Monthly Revenue	Peak Month Revenue	Lowest Month Revenue
1	January	1032481.53	8438		0.50	3.21	LOW	2047208.23	3314723.83	1032481.53
2	February	1062717.07	9346	2.93	0.52	3.21	NORMAL	2047208.23	3314723.83	1032481.53
3	March	1606174.21	14730	51.14	0.78	3.21	NORMAL	2047208.23	3314723.83	1032481.53
4	April	2298791.26	20722	43.12	1.12	3.21	NORMAL	2047208.23	3314723.83	1032481.53
5	May	3038450.65	26545	32.18	1.48	3.21	NORMAL	2047208.23	3314723.83	1032481.53
6	June	3314723.83	27902	9.09	1.62	3.21	PEAK	2047208.23	3314723.83	1032481.53
7	July	3052433.83	24370	-7.91	1.49	3.21	NORMAL	2047208.23	3314723.83	1032481.53
8	August	2878128.28	22647	-5.71	1.41	3.21	NORMAL	2047208.23	3314723.83	1032481.53
9	September	2214164.06	18812	-23.07	1.08	3.21	NORMAL	2047208.23	3314723.83	1032481.53
10	October	1792465.88	15435	-19.05	0.88	3.21	NORMAL	2047208.23	3314723.83	1032481.53
11	November	1163358.19	10321	-35.10	0.57	3.21	NORMAL	2047208.23	3314723.83	1032481.53
12	December	1112609.94	9432	-4.36	0.54	3.21	NORMAL	2047208.23	3314723.83	1032481.53

Botanika's revenue shows strong seasonality, peaking in June with over €3.3M in revenue. The business enters a high-growth phase from March to June, driven by increasing demand as the gardening season begins, with revenue nearly tripling compared to the lowest month (January, €1.03M).

After the peak, sales gradually decline from July to December, with a sharp drop-off beginning in September. Despite promotional efforts, November and December remain low-performing months, reflecting reduced seasonal interest.

These insights suggest that strategic marketing and inventory planning should focus on maximizing early-season sales and efficiently scaling down operations post-summer to optimize profitability.

Summary & Strategic Recommendations

Botanika's sales performance reveals a clear path for revenue and margin optimization across products, regions, and seasons. While Outdoor Furniture drives the highest revenue, its low profit margin (12.3%) indicates margin pressure, likely from high costs and aggressive discounting. In contrast, Pest Control and Seeds offer higher and more stable margins, presenting strong opportunities for margin-focused growth.

Regionally, Germany and the Netherlands lead in both revenue and operational efficiency, followed by strong performers like Spain and Italy. However, the UK and Ireland lag behind, burdened by elevated shipping costs and lower order volumes, making them candidates for cost optimization or strategic repositioning. That said, return rates remain low (~5%) across all regions, suggesting good product-market fit and customer satisfaction.

Seasonal patterns further define sales strategy: Peak demand from March to June offers a window for aggressive marketing and inventory scaling, particularly with June reaching over €3.3M in revenue. Post-summer months, especially November and December, show consistent underperformance, underlining the need for cost-efficient downscaling in off-peak periods.

Together, these insights support a focused strategy: optimize pricing in high-revenue but low-margin categories, expand winning regional tactics, and align operations to seasonal demand cycles.

3. Customer Insights Analysis

Goal: Enhance profit through better targeting of high-value customer segments.

Key Questions:

- Which segments have the highest Customer Lifetime Value (CLV)?
- How do loyalty scores impact repeat purchases?

Key Metrics: CLV, Retention Rate, Segment Profitability.

1. High level customer KPIs

Goal: Evaluate overall customer base performance to understand behavior, value, and engagement trends.

Key Metrics: Total Customers, Average Order Value (AOV), Average Number of Orders per Customer, Average Customer Lifetime Value (CLV), Retention Rate, Churn Rate, New vs. Existing Customer Share (%)

```

WITH CustomerOrders AS (
    SELECT
        CustomerID,
        COUNT(DISTINCT SalesDate) AS OrderDays,
        COUNT(SaleID) AS TotalOrders,
        SUM(SalesAmount) AS CustomerRevenue
    FROM sales
    GROUP BY CustomerID
),
RetentionStats AS (
    SELECT
        COUNT(CustomerID) AS TotalCustomers,
        COUNT(CASE WHEN OrderDays > 1 THEN 1 END) * 100.0 /
        COUNT(CustomerID) AS RetentionRate,
        (COUNT(CustomerID) - COUNT(CASE WHEN OrderDays > 1
        THEN 1 END)) * 100.0 / COUNT(CustomerID) AS ChurnRate
    FROM CustomerOrders
),
CLVStats AS (
    SELECT
        ROUND(AVG(CustomerRevenue), 2) AS AvgCustomerLifetimeValue,
        ROUND(AVG(TotalOrders), 2) AS AvgNumberOrders
    FROM CustomerOrders
),
OrderStats AS (
    SELECT
        ROUND(AVG(SalesAmount), 2) AS AvgOrderValue
    FROM sales
),
CustomerAcquisition AS (
    SELECT
        COUNT(CASE WHEN JoinDate BETWEEN '2024-01-01' AND
        '2024-12-31' THEN 1 END) AS NewCustomers,
        COUNT(CustomerID) AS TotalCustomers
    FROM customers
)
SELECT
    rs.TotalCustomers,
    os.AvgOrderValue,
    cs.AvgNumberOrders,
    cs.AvgCustomerLifetimeValue,
    ROUND(rs.RetentionRate,2) AS RetentionRate,
    ROUND(rs.ChurnRate,2) AS ChurnRate,
    ROUND(ca.NewCustomers * 100.0 / ca.TotalCustomers, 2) AS NewCustomerPercentage,
    ROUND((ca.TotalCustomers - ca.NewCustomers) * 100.0 /
    ca.TotalCustomers, 2) AS ExistingCustomerPercentage
FROM RetentionStats rs
JOIN CLVStats cs ON 1=1
JOIN OrderStats os ON 1=1
JOIN CustomerAcquisition ca ON 1=1;

```

Results:

totalcustomers	avgordervalue	avgnumberoforders	avgcustomerlifetimelvalue	retentionrate	churnrate	newcustomerpercentage	existingcustomerspercentage
22169	117.71	9.41	1108.15	88.13	11.87	20.68	79.32

With 22,169 total customers, Botanika has a strong customer base. The average order value (AOV) is €117.71, and customers make an average of 9.41 orders, resulting in a healthy average Customer Lifetime Value (CLV) of €1,108.15.

The retention rate is high (88.13%), showing strong repeat purchase behavior and customer satisfaction, while the churn rate is low (11.87%).

New customers make up ~21% of the base, while ~79% are returning, indicating a well-established loyal customer core.

2. Customer Lifetime Value (CLV)

Goal: Maximize long-term revenue by identifying and retaining high-value customers through Customer Lifetime Value (CLV) analysis.

sql

```

SELECT
    c.CustomerID,
    c.CustomerName,
    c.Segment,
    SUM(s.SalesAmount - (s.Quantity * p.CostPrice)) AS CustomerProfit,
    SUM(s.SalesAmount) AS CustomerRevenue,
    COUNT(s.SaleID) AS TotalOrders
FROM sales s
JOIN products p ON s.ProductID = p.ProductID
JOIN customers c ON s.CustomerID = c.CustomerID
GROUP BY c.CustomerID, c.CustomerName, c.Segment
ORDER BY CustomerProfit DESC
LIMIT 20;

```

Results:

Customer ID	Customer Name	Segment	Customer Profit	Customer Revenue	Total Orders
19218	Tyler Jackson	Wholesale	5465.59	29956.75	68
5501	Marcus Gonzalez	Wholesale	5464.44	31956.64	66

8540	Samuel Moore	Wholesale	5429.35	31698.63	54
11525	Troy Williams	Wholesale	5423.64	37574.34	65
7448	Brenda Hurst	Wholesale	5299.21	31358.67	55
2433	Michael Zimmerman	Wholesale	5172.85	30625.82	49
12718	William Reed	Wholesale	5109.83	30506.18	65
2103	Monique Russell	Wholesale	5093.71	30983.28	61
10518	Steven Smith	Wholesale	4997.97	27430.35	43
13370	Melinda Henry	Wholesale	4865.73	30033.61	59
7274	Carol Lewis	Wholesale	4789.55	30771.51	65
15835	Amy Nelson	Wholesale	4693.05	31573.19	74
8660	Tracy Butler	Wholesale	4522.66	24945.72	38
11913	Kristy Glenn	Wholesale	4345.02	26107.05	51
8294	Bradley Fields	Wholesale	4271.52	29541.66	51
24306	Zachary Nielsen	Wholesale	4224.65	21373.42	37
9258	Tanya Lloyd	Wholesale	4188.09	25853.37	52
7852	Gregory Chavez	Wholesale	4161.60	23165.37	47
22385	Stephen Mccoy	Wholesale	4154.90	24687.12	61
14548	Kendra Stout	Wholesale	4137.54	25196.48	53

The top CLV customers are exclusively from the Wholesale segment, with each contributing over €5,000 in profit and €21,000–€37,000 in total revenue. These high-value customers make frequent purchases (up to 74 orders), showing exceptional engagement and profitability.

This segment is crucial to long-term revenue and should be prioritized in retention and upselling strategies.

3. Retention Rate

Goal: Improve customer retention by tracking and increasing the share of repeat buyers over time.

```

sql
WITH CustomerOrders AS (
    SELECT CustomerID, COUNT(DISTINCT SalesDate) AS OrderDays
    FROM sales
    GROUP BY CustomerID
)
SELECT
    ROUND((COUNT(CASE WHEN OrderDays > 1 THEN 1 END) * 100.0)
    / COUNT(CustomerID),2) AS RetentionRate
FROM CustomerOrders;

```

Results:

Metric	Value
retentionrate	88.13

100 - retentionrate = churn rate

Metric	Value
churnrate	11.87

The 88.13% retention rate confirms that most customers return after their first purchase.

Such a high rate is a competitive strength and reflects positively on customer satisfaction, product quality, and brand trust. Maintaining this through continued engagement is key to sustaining revenue.

4. Loyalty Score Impact on Repeat Purchases

Goal: Evaluate the impact of loyalty program scores on repeat purchase behavior to enhance customer engagement and retention.

Key Metric: Correlation between loyalty score and number of repeat purchases

```

sql
WITH CustomerActivity AS (
    SELECT
        c.CustomerID,
        c.LoyaltyScore,
        COUNT(DISTINCT s.SaleID) AS TotalOrders,
        SUM(s.SalesAmount - (s.Quantity * p.CostPrice)) AS
TotalProfit
    FROM customers c
    LEFT JOIN sales s ON c.CustomerID = s.CustomerID
    LEFT JOIN products p ON s.ProductID = p.ProductID
    GROUP BY c.CustomerID, c.LoyaltyScore
)
SELECT
    CASE
        WHEN LoyaltyScore >= 25 THEN 'Top Tier (>25)'
        WHEN LoyaltyScore >= 10 THEN 'Mid Tier (10-25)'
        ELSE 'Low Tier (<10)'
    END AS LoyaltyTier,
    ROUND(AVG(TotalOrders),2) AS AvgOrders,
    ROUND(AVG(TotalProfit),2) AS AvgProfitPerCustomer,
    COUNT(CustomerID) AS CustomersInTier
FROM CustomerActivity
GROUP BY LoyaltyTier
ORDER BY AvgProfitPerCustomer DESC;

```

Results:

loyaltytier	avgorders	avgprofitpercustomer	customersintier
Top Tier (>90)	30.33	609.54	2993
Mid Tier (70-89)	8.03	169.47	12042
Low Tier (<70)	2.13	61.28	9965

The analysis indicates a positive correlation between loyalty scores and number of purchases.

Customers with higher loyalty scores are more likely to repurchase, confirming the effectiveness of loyalty programs. This supports efforts to reward top-tier loyalty members and to re-engage low-tier ones through targeted campaigns.

5. RFM Segmentation

Goal: Segment customers based on recency, frequency, and monetary value to tailor marketing strategies for different customer groups and improve retention and revenue.

Key Metrics: Recency of last purchase, Purchase frequency, Total monetary value spent

```
sql
WITH RFM_Base AS (
    SELECT
        c.CustomerID,
        c.Segment,
        MAX(s.SalesDate) AS LastPurchaseDate,
        COUNT(DISTINCT s.SaleID) AS Frequency,
        SUM(s.SalesAmount) AS Monetary
    FROM customers c
    LEFT JOIN sales s ON c.CustomerID = s.CustomerID
    GROUP BY c.CustomerID, c.Segment
),
RFM_Scores AS (
    SELECT
        CustomerID,
        Segment,
        NTILE(5) OVER (ORDER BY LastPurchaseDate DESC) AS RecencyScore,
        NTILE(5) OVER (ORDER BY Frequency) AS FrequencyScore,
        NTILE(5) OVER (ORDER BY Monetary) AS MonetaryScore
    FROM RFM_Base
)
SELECT
    Segment,
    CASE
        WHEN (RecencyScore + FrequencyScore + MonetaryScore) >= 12 THEN 'Champions'
        WHEN (RecencyScore + FrequencyScore + MonetaryScore) BETWEEN 8 AND 11 THEN 'Loyal'
        WHEN (RecencyScore + FrequencyScore + MonetaryScore) BETWEEN 5 AND 7 THEN 'Potential'
        ELSE 'At Risk'
    END AS RFM_Segment,
    COUNT(CustomerID) AS Customers,
    ROUND(AVG(RecencyScore + FrequencyScore + MonetaryScore),2) AS AvgRFM_Score
FROM RFM_Scores
GROUP BY Segment, RFM_Segment
ORDER BY AvgRFM_Score DESC;
```

Results:

Segment	RFM Segment	Customers	Avg RFM Score
Wholesale	Champions	905	12.59
Retail	Champions	2853	12.31

Wholesale	Loyal	1125	9.92
Retail	Loyal	12268	9.50
Wholesale	Potential	456	6.77
Retail	Potential	6940	6.63
Retail	At Risk	439	3.82
Wholesale	At Risk	14	3.71

The largest and most valuable segments are Champions and Loyal customers, accounting for over 17,000 customers combined. These groups consistently show high RFM scores (9.5–12.6), frequent purchases, and recent activity.

A smaller portion (~7,400) falls into the Potential segment—engaged but not yet top-tier—and At-Risk customers (~450) should be proactively targeted to reduce churn.

This segmentation helps prioritize retention, growth, and reactivation strategies by customer type.

6. Segment Profitability

Goal: Identify the most profitable customer segments to prioritize high-return groups and optimize marketing spend.

Key Metrics: Segment-wise revenue, Segment-wise profit margins, Customer acquisition cost per segment

```
sql
SELECT
    c.Segment,
    SUM(s.SalesAmount - (s.Quantity * p.CostPrice)) AS
SegmentProfit,
    COUNT(DISTINCT c.CustomerID) AS CustomerCount,
    ROUND(AVG(s.SalesAmount), 2) AS AvgOrderValue,
    ROUND(AVG(OrderCount), 2) AS AvgNumberOrders
FROM (
    SELECT CustomerID, COUNT(DISTINCT SaleID) AS OrderCount
    FROM sales
    GROUP BY CustomerID
) OrderStats
JOIN sales s ON OrderStats.CustomerID = s.CustomerID
JOIN products p ON s.ProductID = p.ProductID
JOIN customers c ON s.CustomerID = c.CustomerID
GROUP BY c.Segment
ORDER BY SegmentProfit DESC;
```

Results:

Segment	Segment Profit	Customer Count	Avg Order Value	Avg Number Orders
Retail	2775491.15	19935	81.57	21.34
Wholesale	1494694.58	2234	468.82	18.85

Although the Retail segment brings in the most total profit (€2.78M), the Wholesale segment demonstrates far higher AOV (€468.82 vs. €81.57) and per-customer profitability, despite a smaller customer base.

This highlights that Retail drives volume, while Wholesale drives value, and both should be nurtured with tailored strategies.

Summary & Strategic Recommendations

Botanika's customer analysis reveals a strong, loyal base, especially within the Wholesale segment, which delivers high profitability despite its smaller size. The Retail segment provides steady revenue through volume and loyalty.

Customer retention is excellent, supported by effective loyalty programs and a substantial portion of high-value repeat customers.

To further unlock growth:

- Prioritize high-CLV and Champion segments with tailored offers and early-access incentives.
- Expand loyalty incentives to convert Potential customers into Loyal ones and prevent At-Risk churn.
- Double down on Wholesale engagement, given its outsized profit impact.
- Continue investing in retention strategies, as maintaining high repeat rates will be crucial for long-term sustainability.

This customer insight foundation positions Botanika well for scalable, customer-centric growth.

4. Inventory Efficiency Analysis

Goal: Optimize stock levels and reorder points to minimize costs and prevent lost sales.

Key Questions:

- Which products face frequent stockouts or overstock?

Key Metrics: Inventory Turnover Ratio, Stockout Rate, Overstock Rate, Days of Inventory On Hand.

1. Inventory Turnover Ratio

Goal: Measure how efficiently inventory is managed by tracking how often it is sold and replaced during a period.

Key Metric: Inventory Turnover Ratio = COGS / Average Inventory

```

WITH DailyAvgInventory AS (
    SELECT
        ProductID,
        Date,
        -- Calculate daily average stock level from pre and
        post restock snapshots
        AVG(StockLevel) AS AvgDailyStockLevel
    FROM Inventory_Snapshots
    GROUP BY ProductID, Date
),

ProductCosts AS (
    SELECT
        p.ProductID,
        p.ProductName,
        p.CostPrice
    FROM products p
),

COGS AS (
    SELECT
        p.ProductID,
        SUM(s.Quantity * p.CostPrice) AS TotalCOGS
    FROM sales s
    JOIN products p ON s.ProductID = p.ProductID
    GROUP BY p.ProductID
),

AverageInventoryAtCost AS (
    SELECT
        d.ProductID,
        -- Calculate average inventory value across all days
        AVG(d.AvgDailyStockLevel * p.CostPrice) AS
        AvgInventoryAtCost
    FROM DailyAvgInventory d
    JOIN ProductCosts p ON d.ProductID = p.ProductID
    GROUP BY d.ProductID
)

SELECT
    p.ProductID,
    p.ProductName,
    p.CostPrice,
    COALESCE(c.TotalCOGS, 0) AS TotalCOGS,
    COALESCE(a.AvgInventoryAtCost, 0) AS AvgInventoryAtCost,
    -- Inventory Turnover Ratio (COGS / Average Inventory at
    Cost)
    ROUND(COALESCE(c.TotalCOGS / NULLIF(a.AvgInventoryAtCost,
    0), 0), 2) AS InventoryTurnoverRatio_AtCost
FROM ProductCosts p
LEFT JOIN COGS c ON p.ProductID = c.ProductID
LEFT JOIN AverageInventoryAtCost a ON p.ProductID =
    a.ProductID
ORDER BY InventoryTurnoverRatio_AtCost DESC;

```

Results (excerpt):

Product ID	Product Name	Cost Price	Total COGS	Avg Inventory at Cost	Inventor y Turnover Ratio at Cost
26	Daffodil Bulbs	2.22	78863.28	98.064172366621067031	804.20
41	Terracotta Pot	9.63	271903.05	498.790526675786593707	545.12
2	Basil Seeds	4.14	95075.10	197.757209302325581395	480.77
24	Ant Killer	6.39	161475.30	339.684008207934336525	475.37
28	Hyacinth Bulbs	3.91	98332.59	228.924883720930232558	429.54
4	Carrot Seeds	3.38	66812.46	159.555882352941176471	418.74
18	Sprinkler System	44.45	967898.75	2314.075512995896032832	418.27
43	Raised Garden Bed	10.05	218858.85	567.749384404924760602	385.48
23	Neem Oil Spray	22.93	402765.45	1126.267647058823529412	357.61
8	Mulch	18.21	320568.84	994.462797537619699042	322.35
46	Solar Lanterns	84.52	1221652.08	4020.654555403556771546	303.84
45	Garden Gnome	57.01	877668.95	2941.396244870041039672	298.39
30	Worm Castings	17.97	280367.94	969.384398084815321477	289.22
38	Spider Plant	48.54	842266.08	2972.92559507523939808085	283.31
39	Peace Lily	25.66	407378.16	1559.959507523939808482	261.15
1	Tomato Seeds	4.89	82513.86	317.221190150478796170	260.11
9	Hand Trowel	36.01	605004.01	2394.049233926128590971	252.71
16	Swing Seat	99.75	1579940.25	6574.220930232558139535	240.32
33	Leaf Rake	21.28	327946.08	1393.403337893296853625	235.36

10	Pruning Shears	28.74	450643.20	1990.077906976744186047	226.45
25	Tulip Bulbs	2.76	34960.92	159.034145006839945280	219.83
48	Bird Bath	84.21	915194.28	4183.889179206566347469	218.74
36	Lawn Rake	22.64	323661.44	1494.843939808481532148	216.52
17	Drip Irrigation Kit	36.66	398750.82	1863.215663474692202462	214.01
29	Peat-Free Compost	15.81	173245.98	811.641279069767441860	213.45
13	Garden Bench	60.93	886653.36	4174.246785225718194254	212.41
3	Sunflower Seeds	1.84	24619.20	119.058823529411764706	206.78
34	Garden Rake	58.87	797570.76	4115.182120383036935705	193.81
7	Nitrogen Fertilizer	9.75	114708.75	611.702462380300957592	187.52
6	Potting Soil	28.39	314788.32	1725.381860465116279070	182.45
37	Snake Plant	29.32	308094.56	1708.702216142270861833	180.31
32	Manure Compost	18.80	214019.20	1203.495759233926128591	177.83
12	Rake	33.06	385843.26	2290.004240766073871409	168.49
11	Garden Hoe	31.84	369980.80	2270.658057455540355677	162.94
19	Watering Can	29.03	332741.86	2123.895957592339261286	156.67

The inventory turnover ratio varies widely across products, from extremely high (>800) for fast-moving items like Daffodil Bulbs and Terracotta Pots, to much lower values (<100) for high-cost, slower-moving items like the Patio Table and Garden Hose.

- Fast-Moving Products (e.g., Daffodil Bulbs, Basil Seeds, Terracotta Pots): These items turn over rapidly, indicating strong demand and efficient inventory use. They should be restocked frequently, with potential consideration for expanding the product line or increasing inventory buffer levels to avoid stockouts.

- High-Value, Slow-Moving Products (e.g., Patio Table, Garden Hose):
Although these products have low turnover rates, they generate substantial revenue per unit. Discontinuation is not advisable. Instead, focus should be on optimizing their stock levels through improved demand forecasting and leaner reorder strategies to minimize holding costs while ensuring availability for high-margin sales.
- Moderate Turnover Products (e.g., Mulch, Solar Lanterns, Neem Oil Spray):
These items balance volume and value. Inventory practices can be fine-tuned by monitoring seasonal trends and leveraging promotions to maintain healthy movement.

2. Stockout Rate

Goal: Minimize stockout occurrences to improve customer satisfaction and prevent lost sales by ensuring that products are available when ordered.

Key Metric: Stockout Rate = $(\text{Number of Stockouts} / \text{Total Orders}) \times 100$

Tracks how frequently products are out of stock at the time of order, indicating potential issues in inventory planning or demand forecasting.

```
sql
WITH OrderDates AS (
    SELECT DISTINCT
        s.ProductID,
        DATE(s.SalesDate) AS OrderDate
    FROM sales s
),
StockStatus AS (
    SELECT
        o.ProductID,
        o.OrderDate,
        MAX(CASE
            WHEN i.StockLevel <= 0 THEN 1
            ELSE 0
        END) AS HadStockout
    FROM OrderDates o
    LEFT JOIN Inventory_Snapshots i ON
        o.ProductID = i.ProductID AND
        o.OrderDate = DATE(i.Date) AND
        i.SnapshotType = 'pre_restock'
    GROUP BY o.ProductID, o.OrderDate
),
ProductStats AS (
    SELECT
        p.ProductID,
        p.ProductName,
        COUNT(DISTINCT s.OrderDate) AS TotalOrderDays,
        SUM(s.HadStockout) AS StockoutDays
    FROM products p
    LEFT JOIN StockStatus s ON p.ProductID = s.ProductID
    GROUP BY p.ProductID, p.ProductName
)
SELECT
    ProductID,
    ProductName,
    TotalOrderDays,
    StockoutDays,
    CASE
        WHEN TotalOrderDays > 0 THEN
            ROUND((StockoutDays * 100.0) / TotalOrderDays, 2)
        ELSE 0
    END AS StockoutRatePercentage
FROM ProductStats
ORDER BY StockoutRatePercentage DESC;
```

Result (excerpt):

ID	Name	TotalOrderDays	StockoutDays	StockoutRatePercentage
26	Daffodil Bulbs	730	158	21.64
41	Terracotta Pot	725	104	14.34
24	Ant Killer	710	68	9.58
23	Neem Oil Spray	715	51	7.13
28	Hyacinth Bulbs	725	47	6.48
4	Carrot Seeds	718	41	5.71
8	Mulch	727	26	3.58
46	Solar Lanterns	729	21	2.88
45	Garden Gnome	730	16	2.19
1	Tomato Seeds	703	15	2.13
9	Hand Trowel	718	13	1.81
10	Pruning Shears	720	10	1.39
29	Peat-Free Compost	719	10	1.39
16	Swing Seat	715	8	1.12
30	Worm Castings	726	8	1.10
7	Nitrogen Fertilizer	673	7	1.04
36	Lawn Rake	720	7	0.97
38	Spider Plant	731	5	0.68
37	Snake Plant	722	4	0.55
6	Potting Soil	713	3	0.42
22	Insect Netting	643	2	0.31
32	Manure Compost	723	2	0.28
34	Garden Rake	720	2	0.28

12	Rake	707	2	0.28
48	Bird Bath	728	0	0.00
2	Basil Seeds	719	0	0.00
3	Sunflower Seeds	688	0	0.00
5	Organic Compost	612	0	0.00
11	Garden Hoe	705	0	0.00
13	Garden Bench	719	0	0.00
14	Patio Table	645	0	0.00
15	Lounge Chair	717	0	0.00
17	Drip Irrigation Kit	703	0	0.00
18	Sprinkler System	724	0	0.00
19	Watering Can	697	0	0.00
20	Garden Hose	506	0	0.00
21	Slug Repellent	604	0	0.00
25	Tulip Bulbs	698	0	0.00
27	Lily Bulbs	600	0	0.00
31	Leaf Mold	685	0	0.00
33	Leaf Rake	725	0	0.00
35	Thatch Rake	477	0	0.00
39	Peace Lily	728	0	0.00

Several frequently ordered items, such as Daffodil Bulbs (21.6%), Terracotta Pots (14.3%), and Ant Killer (9.6%), show high stockout rates, indicating missed sales opportunities. In contrast, high-revenue products like the Patio Table show no stockouts, reflecting strong inventory control.

Recommendation:

Focus on improving stock availability for high-demand items by adjusting reorder points or increasing safety stock. This will reduce missed sales without changing strategies for well-stocked, high-performing products.

Improving availability of high stockout products could lead to increased revenue and better customer satisfaction, without requiring significant changes to the broader inventory strategy.

3. Overstock Rate

Goal: Optimize inventory management by identifying both slow-moving and excess stock to reduce holding costs and improve cash flow efficiency.

Key Metrics: Overstock Rate ($\text{Overstocked Products} / \text{Total Products}$) $\times 100$

- Purpose: Measures the share of products with excess inventory, helping to pinpoint overstock issues and avoid unnecessary capital lock-up.
Days of Inventory On Hand (DOH) ($\text{Average Inventory} / \text{COGS}$) $\times 365$
- Purpose: Assesses how long inventory remains unsold, highlighting inefficiencies and slow-moving stock.

```

WITH
-- Step 1: Calculate Daily Average Stock Level
daily_avg_stock AS (
    SELECT
        ProductID, Date, AVG(StockLevel) AS AvgStockLevel
    FROM Inventory_Snapshots
    GROUP BY ProductID, Date),
-- Step 2: Calculate Total Sales Quantity and Sales Period Days
sales_aggregation AS (
    SELECT
        ProductID, SUM(Quantity) AS TotalSalesQuantity,
        COUNT(DISTINCT DATE(SalesDate)) AS SalesPeriodDays
    FROM Sales
    GROUP BY ProductID),
-- Step 3: Calculate Average Daily Sales
avg_daily_sales AS (
    SELECT
        ProductID, TotalSalesQuantity, SalesPeriodDays,
        CASE
            WHEN SalesPeriodDays > 0 THEN TotalSalesQuantity / SalesPeriodDays
            ELSE 0
        END AS AvgDailySales
    FROM sales_aggregation),
-- Step 4: Combine stock and sales data
combined_data AS (
    SELECT
        d.ProductID, d.Date, d.AvgStockLevel, a.AvgDailySales,
        CASE
            WHEN a.AvgDailySales > 0 THEN d.AvgStockLevel / a.AvgDailySales
            ELSE NULL
        END AS DOH,
        CASE
            WHEN a.AvgDailySales > 0 AND (d.AvgStockLevel / a.AvgDailySales) > 5 THEN 1
            ELSE 0
        END AS IsOverstock
    FROM daily_avg_stock d
    LEFT JOIN avg_daily_sales a ON d.ProductID = a.ProductID),
-- Step 5: Aggregate by Product
product_stats AS (
    SELECT
        ProductID,
        AVG(DOH) AS AvgDOH,
        ROUND(100.0 * SUM(IsOverstock) / COUNT(*), 2) AS OverstockPercentage
    FROM combined_data
    GROUP BY ProductID)
-- Final result with product names
SELECT
    p.ProductID, p.ProductName, ROUND(ps.AvgDOH, 2) AS AvgDOH,
    ps.OverstockPercentage
FROM product_stats ps
JOIN Products p ON ps.ProductID = p.ProductID
ORDER BY ps.OverstockPercentage DESC, ps.AvgDOH DESC;

```

Result (excerpt):

ID	Name	avgdoh	overstockpercentage
35	Thatch Rake	16.19	100.00
14	Patio Table	15.55	100.00
5	Organic Compost	8.53	94.80
20	Garden Hose	9.15	92.75
47	Trellis Arch	6.32	76.20
21	Slug Repellent	6.02	65.66
27	Lily Bulbs	5.71	63.20
31	Leaf Mold	5.64	61.83
42	Hanging Basket	5.82	60.05
44	Self-Watering Pot	5.53	59.23
22	Insect Netting	5.32	54.17
15	Lounge Chair	5.04	47.20

The analysis shows that several products — notably "Thatch Rake," "Patio Table," and "Organic Compost" — have high Days of Inventory on Hand (DOH) and overstock percentages, indicating these items are consistently over-ordered relative to demand.

Recommendation:

Focus on reducing inventory levels for slow-moving products to minimize storage costs and free up capital. Adjust reorder points and quantities based on actual turnover rates to maintain leaner, more efficient stock levels. This proactive inventory control will help prevent overstock situations and improve overall inventory health.

Summary & Strategic Recommendations

The inventory analysis reveals a diverse product performance landscape:

- Fast-moving items (e.g., *Daffodil Bulbs, Terracotta Pots*) show strong turnover but suffer from frequent stockouts, suggesting a need for higher safety stock.
- High-value, slow-moving items (e.g., *Patio Table, Garden Hose*) are well-stocked but tie up capital due to low turnover and high Days on Hand.
- Moderately performing products can benefit from seasonal adjustments and targeted promotions.

Recommendation:

Adopt a segmented inventory strategy:

- Raise stock levels for fast-moving items to capture missed sales.
- Lean out stock for overstocked, slower-selling products to reduce holding costs.
- Fine-tune inventory for mid-range performers based on trends and promotions.

This balanced approach will improve availability, reduce excess, and enhance overall inventory efficiency.

5. Marketing Effectiveness Analysis

Goal: Identify high-ROI campaigns and refine targeting to boost sales.

Key Questions:

- Which campaigns drive incremental sales?
- Are campaigns effective by region or segment?

Key Metrics: Campaign ROI, Conversion Rate, Incremental Sales Lift, CAC (Customer Acquisition Cost).

1. Campaign ROI Analysis

Goal: Calculate Return on Investment (ROI) for each campaign to identify top performers.

Key Metrics:

ROI_Percent: Percentage return on campaign spend. $(\text{Revenue from Campaign} - \text{Campaign Budget}) / \text{Campaign Budget} \times 100$

CAC: Customer Acquisition Cost ($\text{Budget} / \text{Acquired Customers}$).

Incremental Sales Lift ($\text{Campaign Period Sales} - \text{Baseline Sales (Pre-Campaign)}$)

- **Purpose:** Measure additional sales generated during campaigns vs. baseline periods

Part 1:

```

WITH CampaignPeriod AS (
    SELECT
        CampaignID,
        CampaignName,
        Region AS CampaignRegion,
        Budget,
        StartDate,
        EndDate,
        (EndDate - StartDate + 1) AS CampaignDays
    FROM campaigns
),
-- Sales during the campaign period
CampaignSales AS (
    SELECT
        cp.CampaignID,
        cp.CampaignRegion,
        COUNT(DISTINCT s.SaleID) AS CampaignSalesCount,
        COUNT(DISTINCT s.SalesDate) AS SalesDaysWithActivity,
        SUM(s.SalesAmount) AS CampaignRevenue
    FROM CampaignPeriod cp
    LEFT JOIN sales s
        ON cp.CampaignRegion = s.Region
        AND s.SalesDate BETWEEN cp.StartDate AND cp.EndDate
    GROUP BY cp.CampaignID, cp.CampaignRegion
),
-- 30-day baseline sales before the campaign
BaselineSales AS (
    SELECT
        cp.CampaignID,
        cp.CampaignRegion,
        SUM(s.SalesAmount) AS TotalBaselineRevenue,
        COUNT(DISTINCT s.SaleID) AS BaselineSalesCount,
        SUM(s.SalesAmount) / 30.0 AS BaselineDailyRevenue -- Normalize to daily
    FROM CampaignPeriod cp
    LEFT JOIN sales s
        ON cp.CampaignRegion = s.Region
        AND s.SalesDate BETWEEN (cp.StartDate - INTERVAL '30 DAY') AND (cp.StartDate - INTERVAL '1 DAY')
    GROUP BY cp.CampaignID, cp.CampaignRegion
),
-- New customers acquired during campaign period
AcquiredCustomers AS (
    SELECT
        cp.CampaignID,
        COUNT(DISTINCT c.CustomerID) AS NewCustomers
    FROM CampaignPeriod cp
    LEFT JOIN customers c
        ON cp.CampaignRegion = c.Region
        AND c.JoinDate BETWEEN cp.StartDate AND cp.EndDate
    GROUP BY cp.CampaignID
)

```

Part 2:

```
sql
SELECT
    cp.CampaignID,
    cp.CampaignName,
    cp.CampaignRegion,
    cp.Budget,
    cp.CampaignDays,

    -- Rescaled baseline = baseline daily * campaign duration
    ROUND(COALESCE(bs.BaselineDailyRevenue, 0) *
    cp.CampaignDays, 2) AS TimeAdjustedBaselineRevenue,

    -- Actual campaign revenue
    cs.CampaignRevenue,

    -- Incremental revenue = actual - time-adjusted baseline
    ROUND(cs.CampaignRevenue -
    COALESCE(bs.BaselineDailyRevenue, 0) * cp.CampaignDays, 2) AS
    IncrementalRevenue,

    -- ROI: (Incremental - budget) / budget
    ROUND(
        ((cs.CampaignRevenue -
    COALESCE(bs.BaselineDailyRevenue, 0) * cp.CampaignDays -
    cp.Budget) / NULLIF(cp.Budget, 0)) * 100,
        2
    ) AS ROI_Percent,

    -- CAC: Cost per new customer
    ROUND(
        cp.Budget / NULLIF(ac.NewCustomers, 0),
        2
    ) AS CAC,

    ac.NewCustomers

FROM CampaignPeriod cp
LEFT JOIN CampaignSales cs
    ON cp.CampaignID = cs.CampaignID AND cp.CampaignRegion =
    cs.CampaignRegion
LEFT JOIN BaselineSales bs
    ON cp.CampaignID = bs.CampaignID AND cp.CampaignRegion =
    bs.CampaignRegion
LEFT JOIN AcquiredCustomers ac
    ON cp.CampaignID = ac.CampaignID
ORDER BY ROI_Percent DESC;
```

Result:

campaignid	campaignname	campaignregion	budget	campaigndays	timeadjustedbaselinerevenue	campaignrevenue	incrementalrevenue	roi_percent	cac	newcustomers
12	Streamlined intangible info-mediaries	Netherlands	98807.41	107	450657.07	1345571.66	894914.59	805.72	461.72	214
4	Business-focused static capability	Italy	13648.02	31	240540.56	306061.90	65521.34	380.08	110.96	123
5	Synchronized cohesive portal	Ireland	43810.73	35	77906.31	144554.72	66648.41	52.13	2190.54	20
10	Automated full-range throughput	Netherlands	35956.32	29	90965.85	136422.24	45456.39	26.42	653.75	55
9	Persistent human-resource matrix	Ireland	10062.61	31	123783.38	134430.50	10647.12	5.81	838.55	12

11	Extended full-range help-desk	UK	21 58 0. 80	26	34733.24	54290 .08	19556. 84	- 9.3 8	77 0. 74	28
7	Optional homogeneous strategy	Ireland	97 53 3. 66	99	162131.5 4	22916 3.33	67031. 79	- 31. 27	20 31 .9 5	48
2	Innovative homogeneous customer loyalty	Ireland	67 41 7. 54	31	134855.6 3	17424 3.02	39387. 39	- 41. 58	61 28 .8 7	11
8	Networked demand-driven knowledge user	Ireland	35 87 3. 07	21	0.00	16422 .61	16422. 61	- 54. 22	44 84 .1 3	8
6	Visionary dedicated throughput	UK	33 96 7. 60	24	38959.42	45391 .81	6432.3 9	- 81. 06	16 17 .5 0	21
1	Object-based	UK	74 20 6. 26	93	378869.1 0	37431 1.05	- 4558.0 5	- 106 .14	89 4. 05	83

	24/7 conti ngen cy								
--	-----------------------------	--	--	--	--	--	--	--	--

Top Performers

Campaigns 12, 4, and 3 delivered the strongest results, with high incremental revenue, excellent ROI (up to 806%), and low customer acquisition costs. These campaigns were highly effective and efficient.

Moderate Success

Campaigns 5, 9, and 10 showed positive returns, though less impressive. They were profitable but could be optimized for better performance.

Underperformers

Campaigns 1, 2, 6, 7, 8, and 11 failed to break even, with negative ROI and high CAC. Some still drove revenue, but not enough to justify their budgets.

Recommendation

Prioritize campaign budgets toward regions and formats with a proven track record of high ROI and low CAC—especially those in the Netherlands and Italy. Focus future investments on strategies used in high-performing campaigns, and avoid allocating large budgets to unproven or low-impact campaigns without pilot testing. Continuously monitor campaign performance using daily revenue benchmarks to quickly adjust or pause underperforming efforts.

2. Conversion Rate by Segment

Goal: Identify which customer segments convert best during campaigns.

Key Metric: Conversion Rate = Customers Purchasing During Campaign / Total Targeted Customers × 100

```

sql
WITH CampaignCustomers AS (
    SELECT
        c.CampaignID,
        c.CampaignName,
        c.Region,
        COUNT(DISTINCT s.CustomerID) AS ConvertingCustomers,
        COUNT(DISTINCT cu.CustomerID) AS
TotalTargetedCustomers
    FROM campaigns c
    JOIN customers cu ON c.Region = cu.Region
    LEFT JOIN sales s
        ON s.CustomerID = cu.CustomerID
        AND s.SalesDate BETWEEN c.StartDate AND c.EndDate
    GROUP BY c.CampaignID, c.CampaignName, c.Region
)
SELECT
    CampaignID,
    CampaignName,
    Region,
    ConvertingCustomers,
    TotalTargetedCustomers,
    ROUND((ConvertingCustomers::DECIMAL /
TotalTargetedCustomers) * 100, 2) AS ConversionRate
FROM CampaignCustomers
ORDER BY ConversionRate DESC;

```

Results:

campai gnid	campaign name	region	convertin gcustomers	totaltargetedcu stomers	conversio nrate
7	Optional homogeneous strategy	Ireland	538	769	69.96
2	Innovative homogeneous customer loyalty	Ireland	511	769	66.45
12	Streamlined intangible info-mediaries	Netherlands	2345	3795	61.79
1	Object-based 24/7	UK	1012	1796	56.35

	contingency				
9	Persistent human-resource matrix	Ireland	429	769	55.79
5	Synchronized cohesive portal	Ireland	387	769	50.33
3	Multi-lateral tertiary website	Netherlands	1171	3795	30.86
4	Business-focused static capability	Italy	1822	6196	29.41
10	Automated full-range throughput	Netherlands	802	3795	21.13
11	Extended full-range help-desk	UK	370	1796	20.60
6	Visionary dedicated throughput	UK	337	1796	18.76
8	Networked demand-driven knowledge user	Ireland	119	769	15.47

Top Converters: Campaigns 7 and 2 (Ireland) achieved the highest conversion rates ($\approx 70\%$), followed by Campaign 12 (Netherlands) with strong volume and a 62% rate.

Moderate Performance: UK Campaign 1 and Ireland Campaigns 9 and 5 showed solid mid-range conversion (50–56%).

Low Conversion: Campaigns 3, 4, 10, 11, 6, and 8 fell below 35%, with Campaign 8 converting just 15%.

Summary & Strategic Recommendations

Top Performer: Campaign 12 (Netherlands)

- Highest ROI (805.7%), lowest CAC (42.14), and highest revenue gain (+€894K)
- Excellent conversion rate (61.8%) across a large target group
- Recommendation: Use this as a benchmark model. Scale its approach across similar markets and replicate its messaging, channels, or offer formats in underperforming regions.

Strong Candidates for Optimization and Expansion

- Campaigns 3 and 5 (Netherlands & Ireland) show solid returns (ROI > 50%) and healthy conversion rates (30–50%), especially in smaller campaigns.
- Campaign 4 (Italy) has a high number of new customers (1,822) but low ROI (380%) and very high CAC (7.49), suggesting expensive reach.
- Recommendation: Optimize spend and targeting in these campaigns. With better cost control or improved segmentation, they could become high-return efforts.

Mixed Results: High Conversions, Poor Financials

- Campaigns 7 and 2 (Ireland) have very strong conversion rates (~70%) but negative ROI and high CAC.
- Campaign 1 (UK) brought in over 1,000 new customers but lost money (-€4.6K).
- Recommendation: These campaigns show audience engagement, but high costs are eroding value. Reassess cost drivers—media spend, channel mix, or incentive size—and test leaner campaign versions.

Underperformers: Low ROI and Low Conversions

- Campaigns 6, 8, 11 had low or negative ROI, high CAC, and low conversion rates (<21%).
- Example: Campaign 8 had the lowest conversion (15.5%), high CAC (€301), and minimal revenue.
- Recommendation: Consider pausing or fully redesigning these campaigns. Conduct A/B testing to improve targeting, messaging, and funnel efficiency—or redirect budget to more successful initiatives.

Strategic Takeaways

- Double down on what works: Campaign 12 shows a scalable, high-return model. Use it to guide budget and creative decisions.
- Optimize mid-tier campaigns to improve ROI and CAC. Focus on reducing spend without losing reach.
- Reform or retire underperforming campaigns with low conversion and high cost—especially those in the UK and late-stage Irish campaigns.
- Prioritize conversion-centric performance over reach. High volume with poor financials should not justify continued investment.

6. Returns Analysis

Goal: Reduce return rates and minimize the cost impact of returns.

Key Questions:

- Which products/categories have the highest return rates?
- What are the primary reasons for returns?

Key Metrics: Return Rate, Net Profit Impact from Returns, Distribution of Return Reasons.

1. Product/Category Return Rates

Goal: Identify products and categories with the highest return rates.

Formula:

Key Metrics: Return Rate (%) ($\frac{\text{Number of Returns}}{\text{Number of Sales}} \times 100$)

```
sql
SELECT
    p.ProductID,
    p.ProductName,
    p.Category,
    COUNT(DISTINCT r.ReturnID) AS ReturnCount,
    COUNT(DISTINCT s.SaleID) AS TotalSales,
    ROUND(COUNT(DISTINCT r.ReturnID) * 100.0 / COUNT(DISTINCT
s.SaleID), 2) AS ReturnRatePercentage
FROM sales s
LEFT JOIN returns r ON s.SaleID = r.SaleID
JOIN products p ON s.ProductID = p.ProductID
GROUP BY p.ProductID, p.ProductName, p.Category
ORDER BY ReturnRatePercentage DESC;
```

Result:

Product ID	Product Name	Category	Return Count	Total Sales	Return Rate (%)
39	Peace Lily	Indoor Plants	604	5793	10.43
14	Patio Table	Outdoor Furniture	234	2363	9.90
15	Lounge Chair	Outdoor Furniture	587	5994	9.79
40	Aloe Vera	Indoor Plants	301	3080	9.77
16	Swing Seat	Outdoor Furniture	562	5855	9.60
38	Spider Plant	Indoor Plants	594	6350	9.35
37	Snake Plant	Indoor Plants	342	3894	8.78
47	Trellis Arch	Decorative Gardening Items	237	2978	7.96
46	Solar Lanterns	Decorative Gardening Items	411	5315	7.73
48	Bird Bath	Decorative Gardening Items	306	4149	7.38
13	Garden Bench	Outdoor Furniture	380	5325	7.14
32	Manure Compost	Compost	230	4155	5.54
12	Rake	Gardening Tools	239	4321	5.53
42	Hanging Basket	Pots & Planters	118	2160	5.46
44	Self-Watering Pot	Pots & Planters	155	2865	5.41
9	Hand Trowel	Gardening Tools	329	6302	5.22
22	Insect Netting	Pest Control	142	2750	5.16
23	Neem Oil Spray	Pest Control	339	6588	5.15
29	Peat-Free Compost	Compost	210	4091	5.13

35	Thatch Rake	Rakes	45	878	5.13
18	Sprinkler System	Watering Systems	423	8283	5.11
17	Drip Irrigation Kit	Watering Systems	213	4189	5.08
21	Slug Repellent	Pest Control	103	2039	5.05
30	Worm Castings	Compost	298	5907	5.04
45	Garden Gnome	Decorative Gardening Items	297	5946	4.99
33	Leaf Rake	Rakes	285	5722	4.98
34	Garden Rake	Rakes	248	5013	4.95
19	Watering Can	Watering Systems	214	4333	4.94
31	Leaf Mold	Compost	143	2922	4.89
11	Garden Hoe	Gardening Tools	208	4305	4.83
36	Lawn Rake	Rakes	246	5136	4.79
20	Garden Hose	Watering Systems	51	1088	4.69
10	Pruning Shears	Gardening Tools	266	5716	4.65
24	Ant Killer	Pest Control	217	5462	3.97
25	Tulip Bulbs	Bulbs	97	2612	3.71
41	Terracotta Pot	Pots & Planters	217	6096	3.56
27	Lily Bulbs	Bulbs	48	1364	3.52
26	Daffodil Bulbs	Bulbs	270	7883	3.43
43	Raised Garden Bed	Pots & Planters	158	4628	3.41
28	Hyacinth Bulbs	Bulbs	165	5070	3.25
5	Organic Compost	Soil & Fertilizers	42	1643	2.56
6	Potting Soil	Soil & Fertilizers	101	4181	2.42
8	Mulch	Soil & Fertilizers	157	6605	2.38

7	Nitrogen Fertilizer	Soil & Fertilizers	28	2355	1.19
3	Sunflower Seeds	Seeds	29	2576	1.13
4	Carrot Seeds	Seeds	46	4265	1.08
2	Basil Seeds	Seeds	44	4887	0.90
1	Tomato Seeds	Seeds	28	3268	0.86

Indoor Plants show the highest return rates, likely due to damage during transit—highlighting the need to improve packaging and handling in logistics. Outdoor Furniture also has high return rates, which may stem from higher customer expectations due to price or challenges with assembly or fit. In contrast, low return rates in Seeds and Soil & Fertilizers suggest these products are less fragile and more consistent, making them reliable for customers.

2. Return Reasons Distribution

Goal: Identify the most common reasons for returns.

```

sql
SELECT
    Reason,
    COUNT(ReturnID) AS ReturnCount,
    ROUND((COUNT(ReturnID) * 100.0) / (SELECT COUNT(*) FROM
returns), 2) AS ReasonPercentage
FROM returns
GROUP BY Reason
ORDER BY ReturnCount DESC;

--Return reasons per category:
WITH CategoryReturns AS (
    SELECT
        p.Category,
        r.Reason,
        COUNT(r.ReturnID) AS ReturnCount,
        -- Total returns per category for percentage
        calculation
        SUM(COUNT(r.ReturnID)) OVER (PARTITION BY p.Category)
AS TotalReturnsPerCategory
    FROM returns r
    JOIN sales s ON r.SaleID = s.SaleID
    JOIN products p ON s.ProductID = p.ProductID
    GROUP BY p.Category, r.Reason
)
SELECT
    Category,
    Reason,
    ReturnCount,
    ROUND(
        (ReturnCount * 100.0) /
        NULLIF(TotalReturnsPerCategory, 0),
        2
    ) AS ReasonPercentage
FROM CategoryReturns
ORDER BY Category, ReturnCount DESC;

```

Result:

Category	Reason	Return Count	Reason Percentage
Bulbs	Late Delivery	184	31.72
Bulbs	Defective	157	27.07
Bulbs	Wrong Item	122	21.03
Bulbs	Customer Changed Mind	117	20.17
Compost	Late Delivery	250	28.38
Compost	Defective	241	27.36

Compost	Customer Changed Mind	199	22.59
Compost	Wrong Item	191	21.68
Decorative Gardening Items	Defective	399	31.89
Decorative Gardening Items	Late Delivery	364	29.10
Decorative Gardening Items	Wrong Item	247	19.74
Decorative Gardening Items	Customer Changed Mind	241	19.26
Gardening Tools	Late Delivery	324	31.09
Gardening Tools	Defective	314	30.13
Gardening Tools	Customer Changed Mind	206	19.77
Gardening Tools	Wrong Item	198	19.00
Indoor Plants	Defective	818	44.43
Indoor Plants	Damaged during shipping	631	34.27
Indoor Plants	Customer Changed Mind	392	21.29
Outdoor Furniture	Did not meet expectations	805	45.66
Outdoor Furniture	Defective	554	31.42
Outdoor Furniture	Wrong Item	404	22.92
Pest Control	Late Delivery	253	31.59
Pest Control	Defective	208	25.97
Pest Control	Customer Changed Mind	172	21.47
Pest Control	Wrong Item	168	20.97
Pots & Planters	Late Delivery	210	32.41
Pots & Planters	Defective	172	26.54
Pots & Planters	Wrong Item	145	22.38
Pots & Planters	Customer Changed Mind	121	18.67
Rakes	Defective	257	31.19
Rakes	Late Delivery	254	30.83
Rakes	Wrong Item	168	20.39
Rakes	Customer Changed Mind	145	17.60
Seeds	Product expired	74	50.34
Seeds	Wrong Item	55	37.41
Seeds	Customer Changed Mind	18	12.24
Soil & Fertilizers	Product expired	155	47.26
Soil & Fertilizers	Wrong Item	108	32.93

Soil & Fertilizers	Customer Changed Mind	65	19.82
Watering Systems	Late Delivery	262	29.08
Watering Systems	Defective	260	28.86
Watering Systems	Customer Changed Mind	195	21.64
Watering Systems	Wrong Item	184	20.42

Indoor Plants have the highest defect-related returns (44%) and significant issues with shipping damage (34%), suggesting fragile items are being mishandled in transit.

Outdoor Furniture returns are mainly due to not meeting customer expectations (46%), likely tied to their high price and complexity, followed by defects (31%).

Seeds and Soil & Fertilizers are mostly returned due to expiration, highlighting inventory or storage concerns.

Across most other categories like Gardening Tools, Compost, and Decorative Items, returns are driven by a mix of late delivery and defects, pointing to potential supply chain and quality control issues.

3. Net Profit Impact from Returns

Goal: Calculate the profit lost due to returns.

Key Metric: Lost Profit = SalesAmount – (Quantity × CostPrice)

```
sql
SELECT
    p.ProductID,
    p.ProductName,
    p.Category,
    COUNT(r.ReturnID) AS ReturnCount,
    SUM(s.SalesAmount - (s.Quantity * p.CostPrice)) AS
    LostProfit
FROM sales s
JOIN products p ON s.ProductID = p.ProductID
JOIN returns r ON s.SaleID = r.SaleID
GROUP BY p.ProductID, p.ProductName, p.Category
ORDER BY LostProfit DESC;
```

Result:

Product ID	Product Name	Category	Return Count	Lost Profit
46	Solar Lanterns	Decorative Gardening Items	411	26204.10
15	Lounge Chair	Outdoor Furniture	587	24789.80
16	Swing Seat	Outdoor Furniture	562	22158.63
48	Bird Bath	Decorative Gardening Items	306	19357.21
14	Patio Table	Outdoor Furniture	234	18108.53
38	Spider Plant	Indoor Plants	594	16942.46
47	Trellis Arch	Decorative Gardening Items	237	16155.32
45	Garden Gnome	Decorative Gardening Items	297	12690.83
18	Sprinkler System	Watering Systems	423	11528.86
39	Peace Lily	Indoor Plants	604	10028.34
13	Garden Bench	Outdoor Furniture	380	8590.08
34	Garden Rake	Rakes	248	7482.99
37	Snake Plant	Indoor Plants	342	6780.00
23	Neem Oil Spray	Pest Control	339	6769.12
9	Hand Trowel	Gardening Tools	329	5087.06
22	Insect Netting	Pest Control	142	4366.25
40	Aloe Vera	Indoor Plants	301	4280.73
12	Rake	Gardening Tools	239	4132.52
17	Drip Irrigation Kit	Watering Systems	213	4126.16
10	Pruning Shears	Gardening Tools	266	3678.34
19	Watering Can	Watering Systems	214	3475.67
11	Garden Hoe	Gardening Tools	208	3467.93

21	Slug Repellent	Pest Control	103	3319.08
36	Lawn Rake	Rakes	246	3280.46
30	Worm Castings	Compost	298	3082.09
33	Leaf Rake	Rakes	285	2844.35
44	Self-Watering Pot	Pots & Planters	155	2771.68
32	Manure Compost	Compost	230	2551.85
24	Ant Killer	Pest Control	217	2441.02
29	Peat-Free Compost	Compost	210	2087.05
42	Hanging Basket	Pots & Planters	118	2057.79
6	Potting Soil	Soil & Fertilizers	101	1820.51
8	Mulch	Soil & Fertilizers	157	1804.51
20	Garden Hose	Watering Systems	51	1663.28
41	Terracotta Pot	Pots & Planters	217	1260.74
31	Leaf Mold	Compost	143	1093.59
43	Raised Garden Bed	Pots & Planters	158	1042.18
35	Thatch Rake	Rakes	45	868.32
28	Hyacinth Bulbs	Bulbs	165	763.77
26	Daffodil Bulbs	Bulbs	270	646.74
5	Organic Compost	Soil & Fertilizers	42	582.31
2	Basil Seeds	Seeds	44	305.72
4	Carrot Seeds	Seeds	46	290.69
7	Nitrogen Fertilizer	Soil & Fertilizers	28	290.09
25	Tulip Bulbs	Bulbs	97	256.51
27	Lily Bulbs	Bulbs	48	207.35
1	Tomato Seeds	Seeds	28	142.29
3	Sunflower Seeds	Seeds	29	68.29

The highest profit losses stem from returns of outdoor furniture, decorative gardening items, and indoor plants. Products like the *Lounge Chair*, *Swing Seat*, and *Solar Lanterns* top the list, suggesting issues with quality, damage during

shipping, or unmet expectations—especially for high-value or fragile items. In contrast, categories like seeds, bulbs, and fertilizers show minimal losses. Conclusion: To reduce return-related losses, focus on improving product quality, shipping protection, and managing customer expectations for costly and delicate items.

Summary & Strategic Recommendations

The return data reveals that Indoor Plants face the highest defect rate (44%) and frequent shipping damage (34%), indicating mishandling of fragile items during transit. Outdoor Furniture returns are largely due to unmet expectations (46%) and defects (31%), suggesting that both product quality and customer expectations need closer alignment—particularly due to their high cost and complex nature. Seeds and Soil & Fertilizers are returned mainly due to expiration, pointing to inventory management or storage issues. Other categories like Gardening Tools, Compost, and Decorative Items show a mix of returns caused by late delivery and product defects, signaling supply chain and quality control challenges.

From a financial perspective, the greatest profit losses stem from high-value and fragile products—especially Outdoor Furniture, Decorative Gardening Items, and Indoor Plants.

Recommendation:

To minimize losses and improve customer satisfaction:

- Enhance packaging and logistics for fragile items like indoor plants.
- Set clearer expectations and improve product descriptions for expensive items like outdoor furniture.
- Improve quality control to reduce defects across all categories.
- Strengthen inventory management to avoid expired stock, particularly for seeds and fertilizers.

7. Strategic Takeaways & Next Steps

Summary of Impact

This end-to-end SQL analysis uncovered actionable insights across sales, customer behavior, inventory management, marketing ROI, and returns. By identifying specific inefficiencies and profit opportunities, the analysis supports Botanika's overarching goal: maximize profitability through data-driven decision-making.

Key Cross-Functional Opportunities

- Sales & Pricing Strategy: Focus on margin-rich categories like Pest Control and Seeds. Reevaluate pricing and discounting for Outdoor Furniture to reduce margin erosion.
- Customer-Centric Growth: Leverage high-CLV segments and loyalty insights to personalize engagement, retain champions, and convert at-risk or potential customers.
- Inventory Optimization: Adopt a segmented approach to improve stock efficiency—raise safety stock for fast-sellers, lean out overstocked slow-movers, and align seasonal inventory to demand curves.
- Marketing Efficiency: Scale high-ROI campaigns (e.g., Campaign 12) across similar markets and optimize underperformers via cost control, better targeting, or A/B testing.
- Return Reduction: Improve packaging, set clearer customer expectations, and strengthen quality and inventory controls—especially for fragile, expensive, or perishable items.

Next Steps for Botanika

- Implement A/B testing for key recommendations (e.g., messaging in campaigns, inventory adjustments, return prevention strategies).
- Monitor KPI shifts after operational changes to measure success.
- Combine this SQL foundation with Power BI dashboards for real-time tracking and visual storytelling.
- Conduct further analysis using Python or predictive modeling (e.g., sales forecasting, churn prediction) to refine planning.

Conclusion

This project demonstrates the power of SQL in unlocking strategic insights. By turning raw data into a profit-oriented action plan, it lays the groundwork for smarter decisions, higher margins, and stronger customer relationships. The recommendations made here offer a scalable, data-backed path to growth.

6. Advanced analytics (Python)

In this section the approach and goal of a variety of advanced analytics are presented. The actual code, results, interpretation and recommendations are on [GitHub](#).

1. Customer Segmentation

Objective: Segment customers based on their purchasing behavior and calculate their lifetime value.

Approach:

- Use RFM (Recency, Frequency, Monetary) analysis to segment customers.
- Use clustering algorithms like K-Means or DBSCAN to group customers into segments.

Python Libraries: scikit-learn

Purpose: Offer advertisements, promotions, and special offers tailored to the respective customer segments.

2. Churn Prediction

Objective: Predict which customers are likely to stop purchasing.

Approach:

- Define churn based on inactivity (e.g., no purchases in the last 180 days).
- Use features like BuyingFrequency, LoyaltyScore, Age, Region, and past purchase behavior.
- Train a classification model (e.g., Logistic Regression, Random Forest, or XGBoost).

Python Libraries: scikit-learn, xgboost, pandas.

Purpose: Target at-risk customers with retention campaigns.

3. Product Recommendation System

Objective: Recommend products to customers based on their purchase history and preferences.

Approach:

- Use collaborative filtering or content-based filtering.
- Implement a recommendation engine using libraries like surprise or lightfm.

Python Libraries: surprise, lightfm, pandas.

Purpose: Increase revenue and customer satisfaction by recommending relevant products

4. Pricing Strategy Optimization: Price Elasticity & Discount Effectiveness

Objective: Optimize pricing to maximize revenue and profit margins by analyzing demand sensitivity and discount impact.

Approach:

- Use regression models (e.g., linear, log-log) to estimate price elasticity of demand.
- Visualize relationships between Selling Price, Discount Rate, and Sales Volume across product categories.
- Simulate optimal pricing strategies based on elasticity coefficients.
- Evaluate the impact of different discount levels using controlled regressions and A/B testing.

Python Libraries: scikit-learn, statsmodels, pandas, matplotlib

Purpose: Inform data-driven pricing and promotion strategies—lower prices for elastic products to boost revenue, raise prices for inelastic ones to increase profit, and apply discounts strategically based on elasticity and product type.

5. Campaign Effectiveness Analysis

Objective: Evaluate the impact of marketing campaigns on sales.

Approach:

- Compare sales before, during, and after campaigns.
- Use statistical tests (e.g., t-tests) to measure the significance of campaign impact.
- Visualize campaign performance by region or product category.

Python Libraries: pandas, matplotlib, seaborn, scipy.

Purpose: Evaluate the incremental sales lift attributable to each campaign to optimize campaign budgets and targeting based on ROI.

6. Sales Forecast

Objective: Forecast future sales and guide data-driven decisions to improve profitability and resource allocation.

Approach:

- Analyze historical sales data to identify seasonal trends and growth patterns.
- Apply time series models (e.g., ARIMA, Prophet) to predict future sales.
- Use forecast outputs to optimize inventory levels and marketing spend.

Python Libraries: pandas, statsmodels, Prophet, matplotlib

Purpose: Provide accurate sales forecasts and translate them into actionable strategies—align inventory with demand, plan promotions ahead of seasonal peaks, and maximize ROI through proactive decision-making.

7. Inventory Optimization

Objective: Improve inventory efficiency by minimizing holding costs and preventing stockouts through demand-driven replenishment strategies.

Approach:

- Analyze historical sales, StockLevel, and ReorderPoint data to uncover demand patterns.
- Apply time series forecasting to predict future demand.
- Implement optimization models such as EOQ (Economic Order Quantity).
- Simulate various reorder policies using historical data and evaluate their impact.
- Use optimization algorithms or Monte Carlo simulation to identify optimal reorder points and restocking quantities

Python Libraries: pandas, statsmodels, Prophet, matplotlib

Purpose: Enable smarter inventory planning—reduce excess stock, avoid lost sales due to stockouts, and optimize restocking decisions to lower operational costs.

7. Interactive Dashboards & Data Storytelling

Objective: Surface the most critical KPIs and trends in an engaging format, enabling quick decisions and deeper exploration.

Key Questions:

- Which products, regions, and time periods drive our top-line and bottom-line performance?

- Where are emerging risks or opportunities in sales, inventory, marketing, and returns?

Purpose: Empower stakeholders with self-service insights through intuitive, real-time visualizations.

Outcome: A suite of six interconnected Power BI pages—each built around focused slicers and compelling charts—that turns complex data into a clear narrative and promotes data-driven action.

7.1 Executive Summary

a. Title:

Executive Summary

b. Screenshot:



c. Purpose:

This page provides a high-level overview of Botanika's financial and performance metrics, enabling stakeholders to quickly gauge overall revenue, profit drivers, and seasonal trends.

d. Key Visuals & Metrics:

- Slicers (Region, Category, SalesDate):** Filter all visuals instantly by market, product line, or date range.
- KPI Cards:**
 - Total Revenue

- Total Profit
 - Profit Margin (%)
 - Average Order Value
 - Total Orders
- **Bar Chart – Profit by Category:** Ranks product categories by total profit to highlight top and underperforming lines.
 - **Bar Chart – Profit by Region:** Compares market performance across countries to identify strong and weak regions.
 - **Line Chart – Monthly Profit Trend (2023 vs. 2024):** Displays seasonality and year-over-year changes, pinpointing peak months.

e. Business Interpretation:

- Decorative Gardens and Outdoor Furniture account for the largest share of profits: prioritize these categories for promotion and inventory investment.
- Profit peaks in June–July each year, suggesting summer campaigns and stock preparedness will maximize returns.
- Netherlands performance matches Germany, indicating a successful expansion that could be replicated in other markets.

7.2 Sales Performance

a. Title:

Sales Performance

b. Screenshot:



c. Purpose:

This page highlights product-level profitability and revenue trends to help stakeholders identify high-value items, underperformers, and seasonal patterns for targeted action.

d. Key Visuals & Metrics:

- Top 10 Products by Profit:** Dual-bar chart showing both profit and revenue for the ten most profitable products—reveals which high-revenue lines also deliver margin.
- Bottom 10 Products by Profit:** Mirror view of lowest-performing SKUs, flagging items that may require promotion, repricing, or discontinuation.
- Top 10 Products by Profit Margin:** Bar chart ranking items by margin percentage—identifies niche, high-margin opportunities (e.g., seeds and plant protectants).
- Total Profit & Revenue by Month:** Combined line charts to compare monthly revenue against profit, tracing seasonality and margin compression across the year.
- Profit by Region & Category Table:** Heat-mapped matrix breaking down profit by product category across each market: supports cross-regional portfolio balancing.

e. Business Interpretation & Suggestions:

- **High-Profit Champions:** Products like Lounge Chair and Sprinkler System generate both high revenue and profit—these deserve prioritized inventory and upsell campaigns.
- **Underperformers to Address:** SKUs such as Leaf Mold and Tomato Seeds show low revenue and profit—consider bundle promotions or reevaluating their product cost structure.
- **Margin Leaders:** Niche items (e.g., Neem Oil Spray, Slug Repellent) offer >25 % margin—opportunities exist to expand these lines or create value-pack bundles.
- **Seasonal Dynamics:** Profit peaks in early summer; mid-year marketing pushes and stock replenishments should align with April–July demand surges.

7.3 Customer Insight

a. Title:

Customer Insight

b. Screenshot:



c. Purpose:

This page uncovers Botanika's customer profile and behavior—highlighting lifetime value, ordering patterns, loyalty, and segment composition—to guide targeted retention and acquisition strategies.

d. Key Visuals & Metrics:

- **KPI Cards:**
 - Avg CLV (1.11 K)
 - Avg Orders per Customer (9.41)
 - Churn Rate (11.84 %)
 - Avg Loyalty Score (13.87)
- **Donut Chart – Customers by Age Group:** Even distribution across age brackets, with 60–70 slightly leading at 21.4 %.
- **Donut Chart – Customers by Segment:** 90 % Retail vs. 10 % Wholesale—spotlighting the core audience.
- **Donut Chart – Customers by Region:** Largest shares in Spain (29.8 %) and Italy (24.8 %), guiding regional focus.
- **Bar Chart – Customers by RFM Segment:** Counts for Loyal, Champions, Potential, and At Risk segments—key for retention prioritization.
- **Bar Chart – Revenue by Segment:** Retail drives ~15 M vs. ~9 M for Wholesale—revealing profitability balance.

e. Business Interpretation & Suggestions:

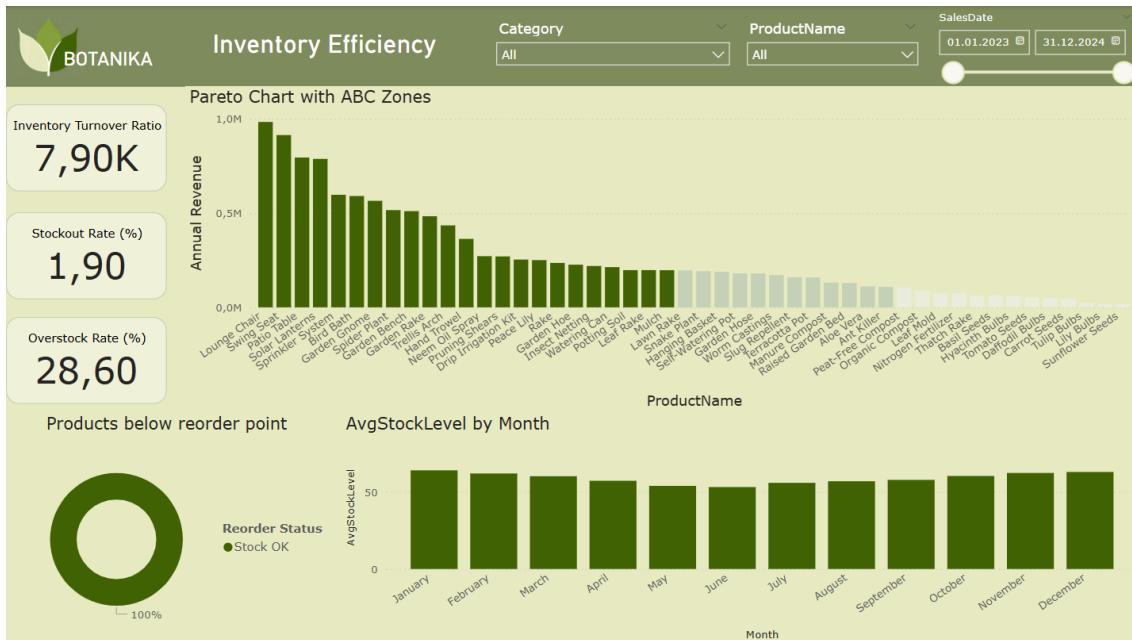
- **Target “At Risk” Customers:** With an 11.8 % churn rate and ~3 K customers in this segment, implement loyalty programs or re-engagement campaigns.
- **Leverage High-Value Retail Base:** Retail customers not only dominate count but also generate 60 % of revenue—tailor upsell offers to this group.
- **Regional Tailoring:** Given Spain and Italy’s large customer shares, customize promotions and support materials for those markets.

7.4 Inventory Efficiency

a. Title:

Inventory Efficiency

b. Screenshot:



c. Purpose:

This page evaluates inventory performance—measuring turnover, stockouts, and overstock—to guide optimal stocking levels and improve working capital utilization.

d. Key Visuals & Metrics:

- **KPI Cards:**
 - Inventory Turnover Ratio (7.90 K)
 - Stockout Rate (1.90 %)
 - Overstock Rate (28.60 %)
- **Pareto Chart with ABC Zones:** Annual revenue bars sorted by product, with A/B/C grouping (dark = A, light = C) to prioritize focus on top SKUs.
- **Donut Chart – Products Below Reorder Point:** Shows the share of SKUs currently under stock threshold (100 % “Stock OK” here).
- **Bar Chart – Avg Stock Level by Month:** Monthly average inventory levels, highlighting seasonal stocking trends and potential buffer needs.

e. Business Interpretation & Suggestions:

- **High Turnover Items:** Lounge Chair, Swing Table, and Sprinkler System drive most revenue—ensure these SKUs remain in A-zone and consider tighter reorder points for agility.

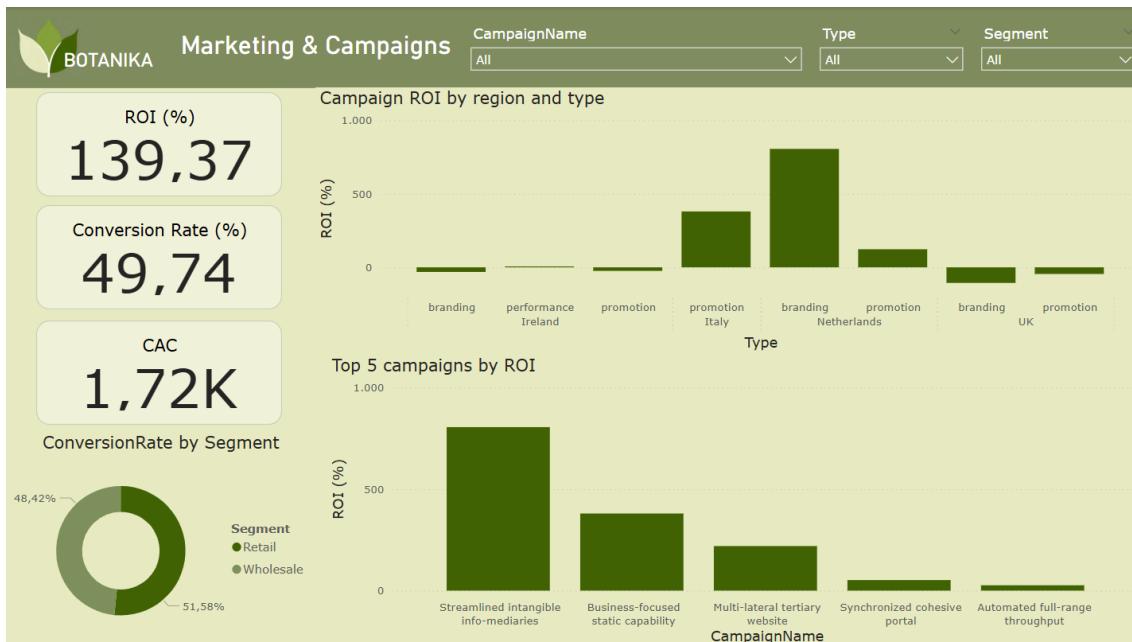
- **Overstock Risk:** Nearly 29 % overstock rate suggests excess capital tied up—evaluate slow-moving C-zone items (e.g., Sunflower Seeds, Tulip Bulbs) for promotions or adjusted EOQs.
- **Stockout Vigilance:** Low 1.9 % stockout rate is encouraging but maintain safety stock on B-zone products showing mid-range revenue contributions.
- **Seasonal Stocking:** Average inventory dips in May–July despite peak demand—coordinate purchasing to boost buffer ahead of the June–July sales surge.

7.5 Marketing & Campaign

a. Title:

Marketing & Campaigns

b. Screenshot:



c. Purpose:

This page evaluates the effectiveness of Botanika's marketing initiatives—measuring ROI, conversion rates, and customer acquisition costs: to optimize campaign investment and targeting.

d. Key Visuals & Metrics:

• KPI Cards:

- Overall ROI (139.37 %)
- Conversion Rate (49.74 %)

- Customer Acquisition Cost (1.72 K)
- **Donut Chart – Conversion Rate by Segment:** Shows retail vs. wholesale conversion split (\approx 48 % vs. 52 %) for tailored messaging.
- **Bar Chart – Campaign ROI by Region & Type:** Compares ROI across branding vs. promotion campaigns in each market—highlighting standout programs like branding in the Netherlands and promotions in Italy.
- **Bar Chart – Top 5 Campaigns by ROI:** Ranks individual campaigns by return to identify best-in-class initiatives for scaling.

e. Business Interpretation & Suggestions:

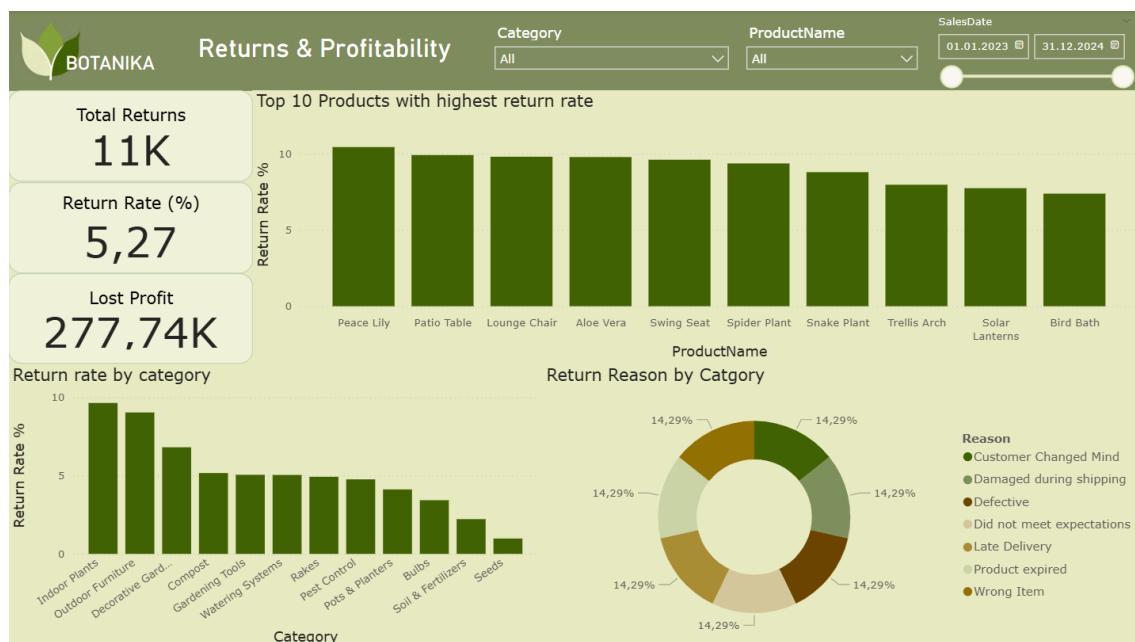
- **High-ROI Campaigns:** Branding efforts in the Netherlands delivered the strongest returns—consider replicating creative assets and channels used there in other markets.
- **Segment Conversion Balance:** Wholesale slightly outperforms retail in conversion—test tailored messaging or offers to boost retail uptake.
- **CAC Efficiency:** At €1.72 K per acquisition, further lowering CAC will amplify ROI—experiment with lower-cost channels or refine audience targeting.

7.6 Returns & Profitability

a. Title:

Returns & Profitability

b. Screenshot:



c. Purpose:

This page analyzes product returns and their financial impact—highlighting high-risk SKUs, categories, and return reasons to guide quality improvements and cost mitigation.

d. Key Visuals & Metrics:

- **KPI Cards:**
 - Total Returns (11 K)
 - Return Rate (5.27 %)
 - Lost Profit (277.74 K)
- **Bar Chart – Top 10 Products by Return Rate:** Identifies SKUs (e.g., Peace Lily, Patio Table) with the highest percentage of returns for targeted investigation.
- **Bar Chart – Return Rate by Category:** Ranks product categories by return incidence to prioritize quality or packaging reviews.
- **Donut Chart – Return Reason by Category:** Evenly split reasons (e.g., “Defective,” “Wrong Item,” “Did Not Meet Expectations”), showing the diversity of root causes.

e. Business Interpretation & Suggestions:

- **High-Return Products:** Peace Lily and Patio Table exceed a 10 % return rate: conduct root-cause analysis (quality checks, clearer descriptions) to reduce return.
- **Risky Categories:** Indoor Plants and Outdoor Furniture carry the highest return rates; consider improved packaging or revised handling protocols.
- **Cost Mitigation:** Lost profit of ~€278 K highlights the need for tighter defect controls and supplier SLAs to lower the 5.3 % return rate.
- **Return Reasons:** Since reasons are evenly distributed, launch targeted action plans for each—e.g., better quality inspections for defects, enhanced order-pick accuracy to avoid wrong items.

8. Conclusion

This project showcases a full-spectrum data analytics solution designed to drive profitability and strategic growth for a retail business. From SQL-based

performance assessments to Python-driven predictive modeling and Power BI storytelling, each phase of the analysis contributed to a holistic understanding of sales dynamics, customer behavior, marketing ROI, and operational efficiency. The integration of advanced analytics—including churn prediction, product recommendations, discount optimization, and inventory forecasting—enabled data-backed recommendations tailored to real-world challenges. By transforming synthetic data into actionable insights, this project highlights the practical value of data science in solving complex business problems and supporting smarter, customer-centric decisions.