

A Comprehensive Data Analysis on Medical Expenses



Retail Sales Presentation

Vinit Solanki

[Click here for watch video](#)

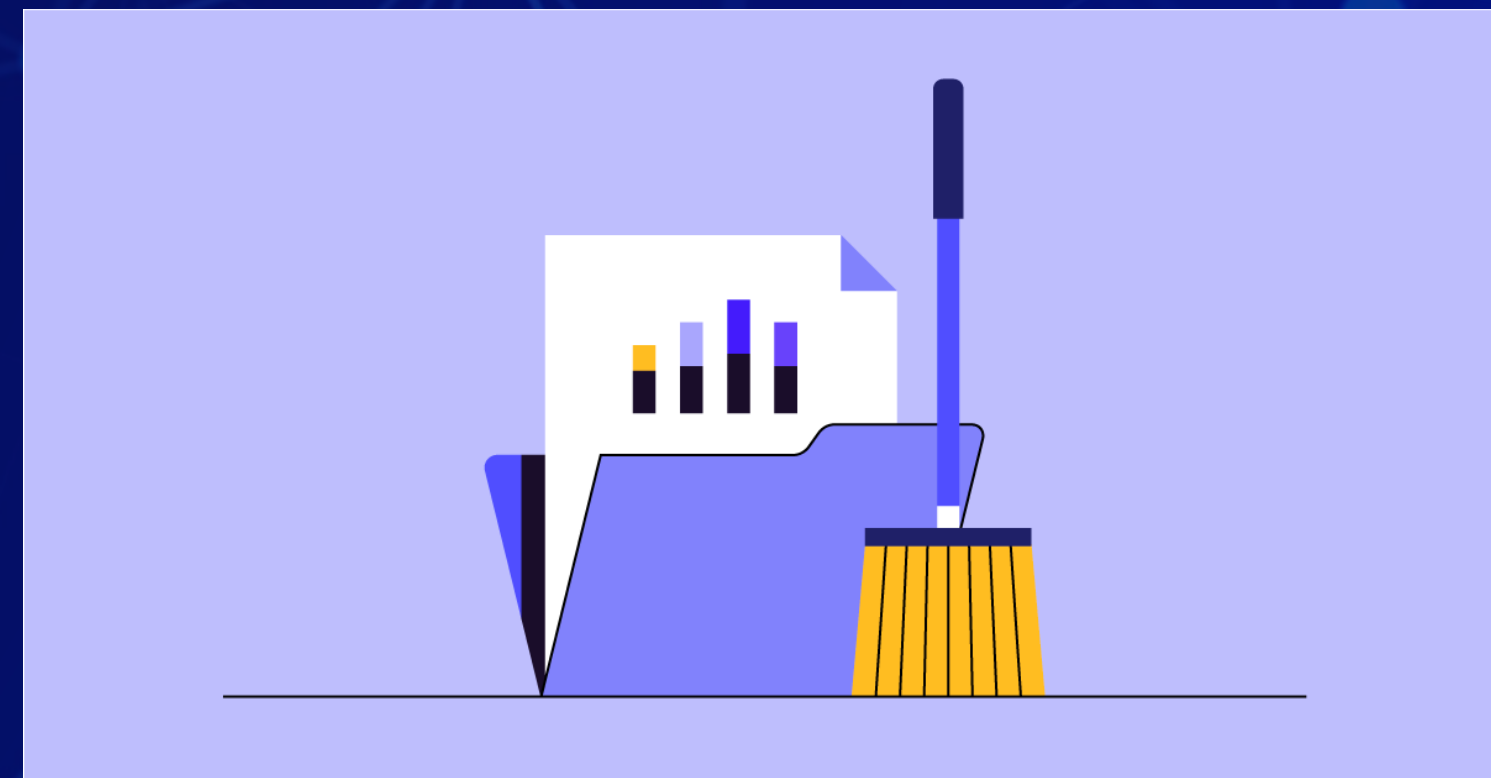
Data Cleaning and Preparation



- **Steps Taken:**
- **Identified Duplicates:** Used **COUNTIF** to check for duplicate InvoiceNo entries.
- **Removed Duplicates:** Applied **Remove Duplicates** feature in Excel, selecting key columns (InvoiceNo, StockCode, CustomerID, etc.).
- **Ensured Uniqueness:** Verified using **COUNTIF** to confirm that no duplicates remain.
- **Outcome:**
Dataset now contains only unique transactions, ensuring accurate analysis.

Handling Cancellations

- **Handling Cancellations**
- **Approach:**
- Transactions with invoice numbers starting with 'C' indicate cancellations.
- A new column was created to **flag these transactions** and calculate the total value of cancellations.
- **Steps Taken:**
- **Flagged Cancellations** using:
 - excel
 - `=IF(LEFT(InvoiceNo,1)="C","Cancelled","Normal")`
- **Calculated Total Cancelled Value** using:
 - excel
 - `=SUMIFS(TotalValueColumn, CancellationFlagColumn, "Cancelled")`
- **Key Insights:**
- Identified and categorized cancelled transactions.
- Calculated their impact on overall sales.
- Helps in understanding **return trends and potential revenue loss.**



Top Selling Products

- **Top Products by Revenue:**
- **Leading in 2009-10:**
- *WHITE HANGING HEART T-LIGHT HOLDER* generated the highest revenue with **£157,580.28**.
- Other top contributors were *ASSORTED COLOUR BIRD ORNAMENT* and *PACK OF 72 RETRO SPOT CAKE CASES*.
- **Leading in 2010-11:**
- *ASSORTED COLOUR BIRD ORNAMENT* topped the revenue chart with **£58,792.42**, followed by *PACK OF 72 RETROSPOT CAKE CASES* and *POPCORN HOLDER*.
- **Consistent High Revenue:**
- *ASSORTED COLOUR BIRD ORNAMENT* and *PACK OF 72 RETROSPOT CAKE CASES* remained in the top 5 for both years.
- **Top Products by Quantity:**
- **2009-10 High Volume Sellers:**
 - *WHITE HANGING HEART T-LIGHT HOLDER* sold **57,632 units**, making it the top-selling product by quantity.
 - It was followed by *WORLD WAR 2 GLIDERS ASSTD DESIGNS* and *BROCADE RING PURSE*.
- **2010-11 Leading Products:**
- *WORLD WAR 2 GLIDERS ASSTD DESIGNS* led with **53,751 units**, while *JUMBO BAG RED RETROSPOT* and *POPCORN HOLDER* entered the top ranks.
- *WHITE HANGING HEART T-LIGHT HOLDER* still held a strong position.
- **Consistent Sellers:**
- *ASSORTED COLOUR BIRD ORNAMENT* and *PACK OF 72 RETROSPOT CAKE CASES* appeared in both years' top 5 by quantity.
- **Key Insights:**
- *WHITE HANGING HEART T-LIGHT HOLDER* was a top performer in both quantity and revenue, especially dominant in 2009-10.
- *WORLD WAR 2 GLIDERS ASSTD DESIGNS* consistently achieved high sales

2009-10

Row Labels	Sum of Quantity	Sum of LineTotal
WHITE HANGING HEART T-LIGHT HOLDER	57632	157580.28
WORLD WAR 2 GLIDERS ASSTD DESIGNS	54554	11276.21
BROCADE RING PURSE	47647	8879.82
PACK OF 72 RETRO SPOT CAKE CASES	46079	23744.41
ASSORTED COLOUR BIRD ORNAMENT	44805	72251.32
Grand Total	250717	273732.04

2010-11

Row Labels	Sum of Quantity	Sum of Revenue
WORLD WAR 2 GLIDERS ASSTD DESIGNS	53751	13560.09
JUMBO BAG RED RETROSPOT	47260	92175.79
POPCORN HOLDER	36322	33959.26
ASSORTED COLOUR BIRD ORNAMENT	36282	58792.42
PACK OF 72 RETROSPOT CAKE CASES	36016	21047.07
Grand Total	209631	219534.63

Pricing Strategy Analysis



➤ Pricing Strategy Analysis

- Objective: Identify products with unit prices above the 90th percentile to support pricing decisions.

➤ Key Metrics

- 90th Percentile Unit Price: £7.95
- Average Unit Price:
 - 2009–10: £9.0194
 - 2010–11: £3.2055

➤ Methodology

- Used PERCENTILE.INC on all unit prices to determine the 90th percentile threshold
- Applied the IF function to filter products with unit prices greater than £7.95
- Used AVERAGEIF to calculate the average unit price per item

➤ Top Products Above 90th Percentile

- Selected the actual top 10 high-priced products for each year based on the threshold
- Examples of frequently high-priced products include:
 - REGENCY CAKESTAND 3 TIER
 - POSTAGE / DOTCOM POSTAGE
 - IVORY KITCHEN SCALES

➤ Business Insight

- Identifying high-priced products enables better pricing strategies
- Supports premium inventory planning and targeted marketing efforts

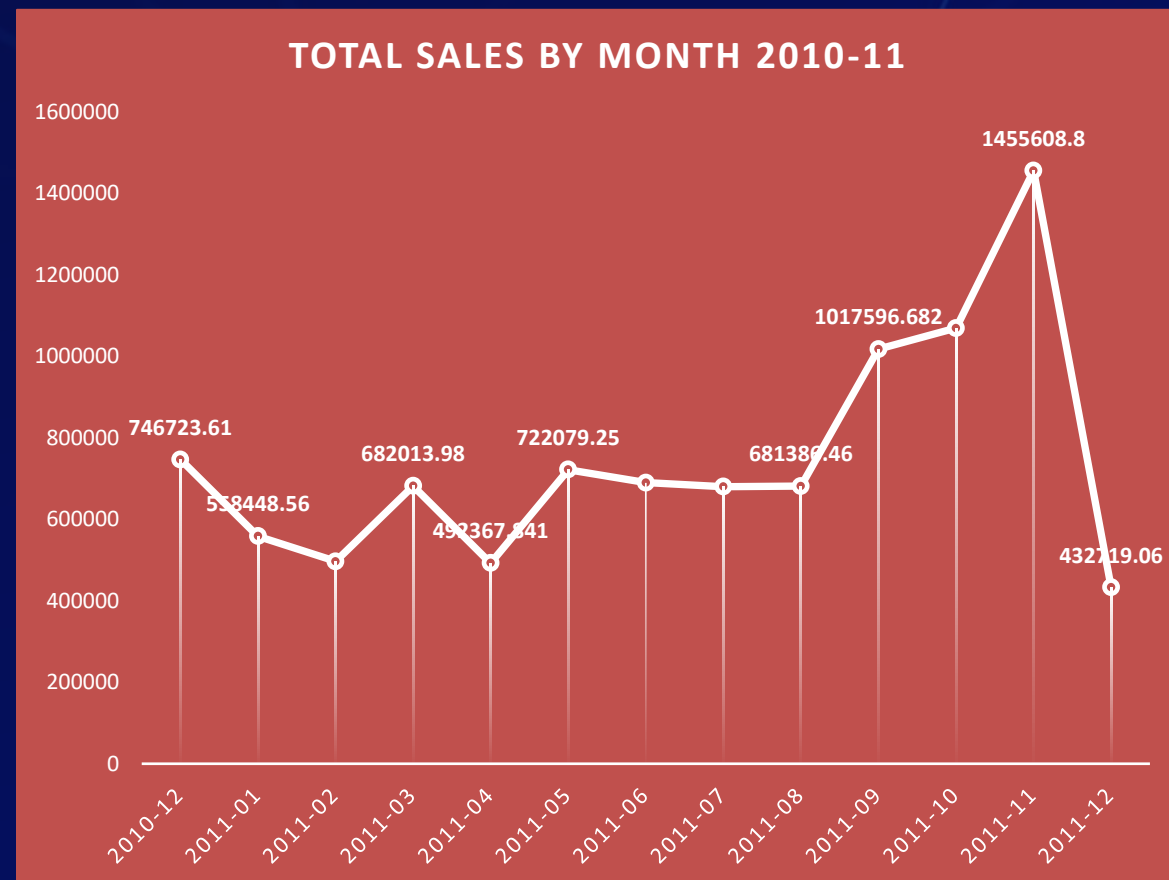
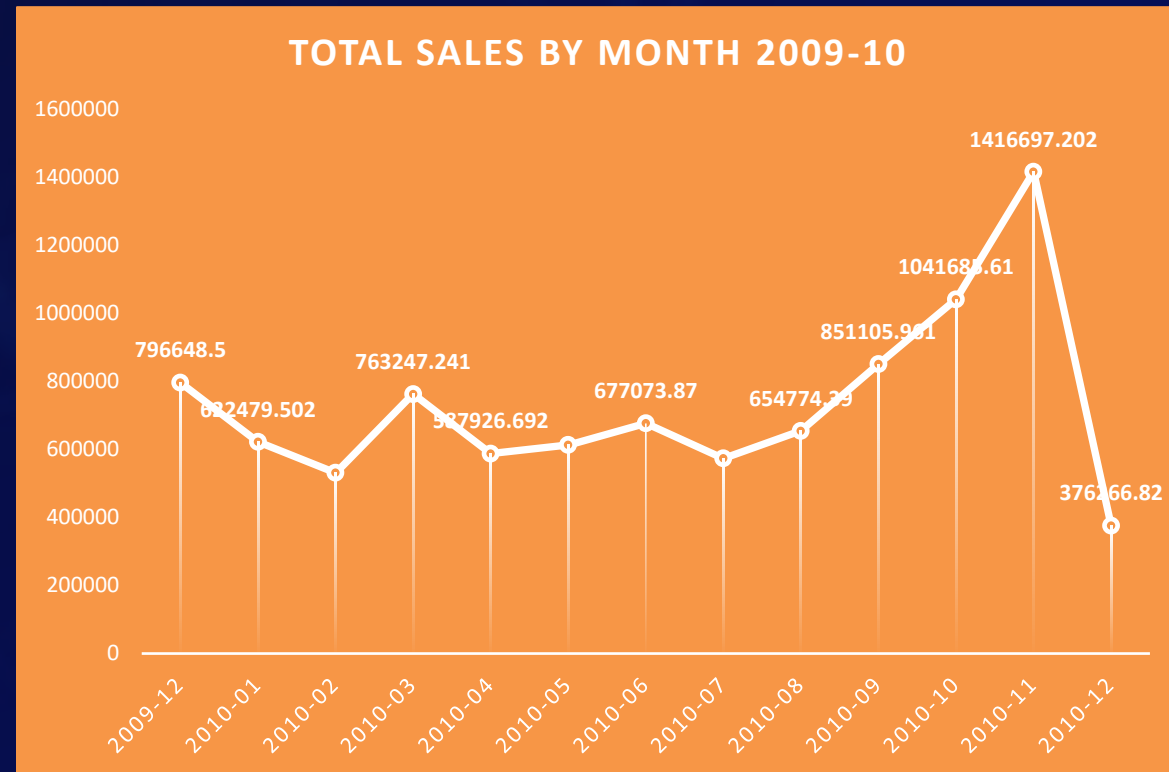
Customer Analysis

- • Objective: Identify the top 10 customers by total revenue contribution across 2009–10 and 2010–11.
- **Key Metrics**
 - Total Revenue (2009–10): £9,502,341.84
 - Total Revenue (2010–11): £2,756,843.98
 - Combined Revenue from Top 10 Customers: Significant share of yearly totals
- **Methodology**
 - Used SUMIF to aggregate revenue by CustomerID from transaction-level data
 - Applied RANK.EQ to rank customers based on their total spending
 - Leveraged INDEX + MATCH to extract corresponding customer details
 - Excluded blank CustomerIDs from actionable insights due to missing identifiers
- **Insights Observed**
 - Customer 18102 appeared in both years' top spenders list, indicating strong recurring value
 - 2010–11 shows more concentrated spending among fewer high-value customers compared to 2009–10
 - New top customers emerged in 2010–11, suggesting potential market shifts or successful acquisition strategies
- **Business Implications**
 - Loyal and high-value customers like 18102 should be prioritized for retention initiatives
 - High-value customers can be offered tailored promotions or exclusive benefits
 - Understanding year-over-year changes in top customers helps shape sales and retention strategies

2009-10	
Row Labels	Sum of Revenue
(blank)	9495311.674
18102	2286.24
13758	996.1
13085	651.1
13078	630.33
17865	589.33
15413	570.24
16167	485.79
12682	426.3
18087	394.74
Grand Total	9502341.844

2010-11	
Row Labels	Sum of Revenue
(blank)	1447487.53
14646	279489.02
18102	256438.49
17450	187322.17
14911	132458.73
12415	123725.45
14156	113214.59
17511	88125.38
16684	65892.08
13694	62690.54
Grand Total	2756843.98

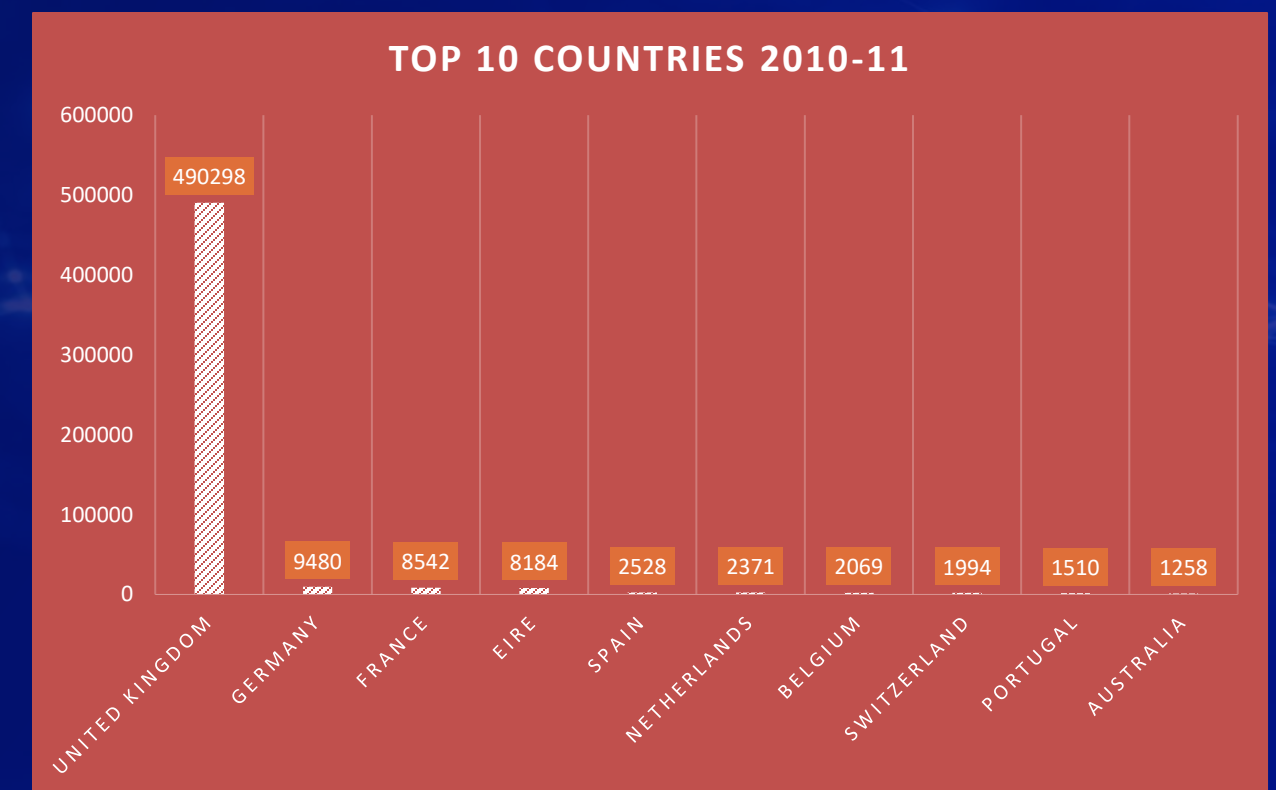
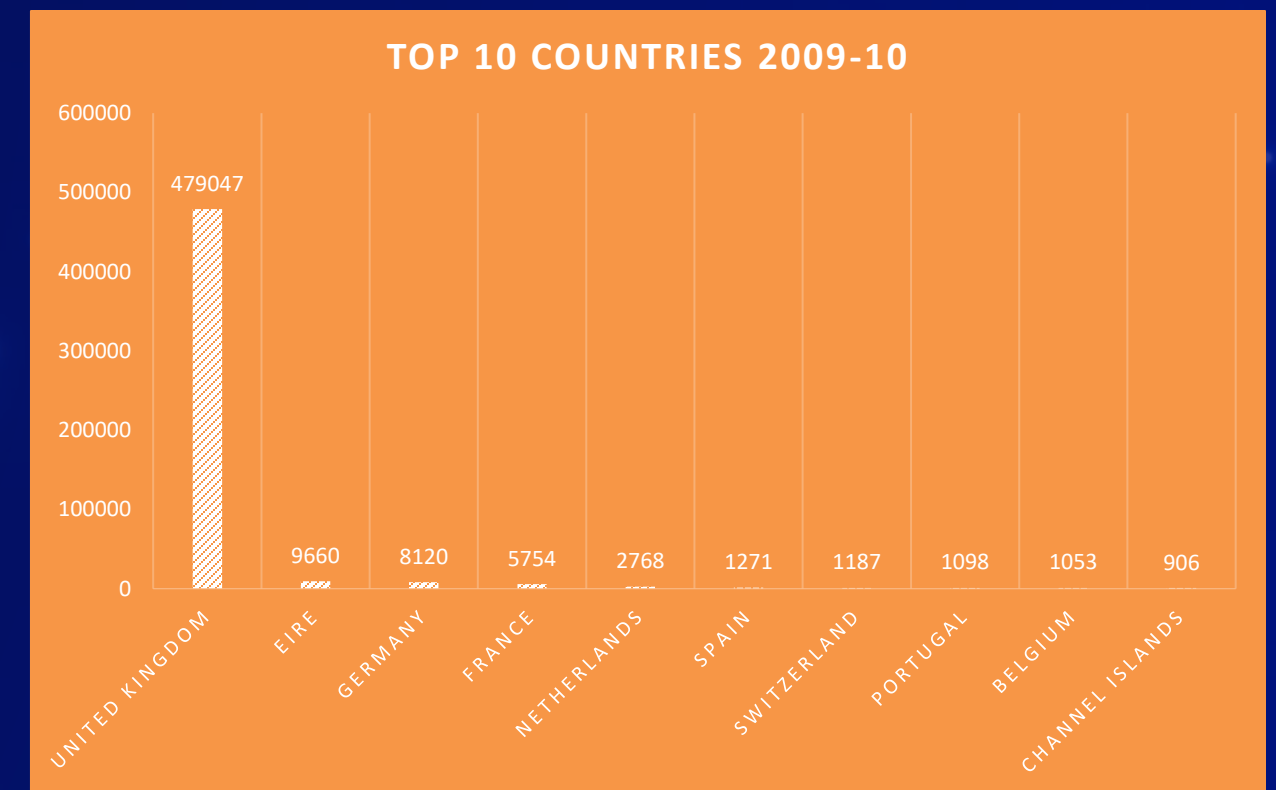
Temporal Sales Trends



- Objective: Analyze monthly sales fluctuations to identify seasonality or demand spikes in 2009–10 and 2010–11.
- **Key Observations**
 - 2009–10:
 - Sales remained steady through most months with slight dips and rises
 - Major spike in November 2010 (£1,416,697), followed by a sharp drop in December
 - Seasonal peak likely due to holiday-related demand
 - 2010–11:
 - Sales showed a gradual rise mid-year with a notable jump in November 2011 (£1,455,608)
 - December again saw a significant drop, indicating similar seasonal patterns
- **Methodology**
 - Used TEXT to convert invoice dates to Year-Month format
 - Applied SUMIFS to calculate monthly sales totals
 - Plotted using a line chart to visualize monthly trends and highlight peaks
- **Business Insights**
 - Strong November performance in both years suggests effectiveness of pre-holiday campaigns
 - Consistent December drop could indicate end-of-season clearance or lower post-peak buying behavior
 - Useful for planning inventory, staffing, and marketing spend during peak periods

Geographical Insights

- **Objective:** Identify countries with the highest number of transactions to uncover major markets.
- **Key Findings**
- The **United Kingdom** dominated both years:
- 2009–10: 479,047 transactions
- 2010–11: 490,298 transactions
- This reflects the business's strong domestic presence.
- Other top countries include:
- **Germany, France, and Ireland** consistently ranked high across both years.
- Emerging contributors include **Australia** and **Switzerland** in 2010–11.
- **Methodology**
- Created a Pivot Table using COUNTIF on InvoiceNo grouped by Country.
- Filtered to display only the top 10 countries based on transaction count.
- Visualized through a bar chart for easy comparison across years.
- **Business Insights**
- Focused marketing and logistics efforts in the UK can be reinforced due to its overwhelming volume.
- The presence of multiple European countries indicates regional growth potential and market loyalty.
- Exploring why countries like **Australia** entered the top 10 in 2010–11 can guide future international expansion strategy.



○ Product Return Rate



- **Overview of Key Metrics**
- To assess product return rates, we compared the **total quantity ordered** with the **total quantity returned** for each year. This analysis helps evaluate product performance, customer satisfaction, and potential areas for improvement.
- **Average Values and Trends**
- The overall **return rate was 4% in 2009-10**, which improved to **3% in 2010-11**, showing a positive trend in product retention. A **1% decrease** in returns suggests improved product quality, better customer experience, or enhanced order accuracy.
- **Central Tendency and Dispersion**
- The **median return rate per product** was lower than the overall average, indicating that most products had minimal returns, while a few items contributed significantly to the total return rate. The **standard deviation of return rates** suggests variation across different product categories, with some experiencing higher return frequencies.
- **Key Observations**
- **Declining Return Rates:** The reduction from **4% to 3%** suggests **better inventory management and customer satisfaction**.
- **High-Return Products:** Certain items may have **consistently higher return rates**, requiring further analysis to identify quality or fulfillment issues.
- **Impact on Revenue:** Returns contribute to revenue loss, and minimizing returns can significantly **improve profit margins**.
- **Possible Causes:** Return trends may be influenced by **product defects, incorrect orders, or shifting customer preferences**.

Data Segmentation for Marketing

2009-10

Quartile	Value
Q1 (25th Percentile)	291.88
Q2 (50th Percentile)	644.24
Q3 (75th Percentile)	1610.15

2010-11

Quartile	Value
Q1 (25th Percentile)	220.91
Q2 (50th Percentile)	310.75
Q3 (75th Percentile)	579.785

- **Objective:**
Classify customers into High, Medium, and Low tiers based on total spending.
- **Segmentation Logic (Based on Quartiles):**
Customers were segmented using the total expenditure and applying the QUARTILE.INC function.
- **Thresholds (2009–10):**
- **Low Tier:** $\leq \text{£}291.88$
- **Medium Tier:** $> \text{£}291.88$ and $\leq \text{£}1610.15$
- **High Tier:** $> \text{£}1610.15$
- **Thresholds (2010–11):**
- **Low Tier:** $\leq \text{£}220.91$
- **Medium Tier:** $> \text{£}220.91$ and $\leq \text{£}579.79$
- **High Tier:** $> \text{£}579.79$
- **Excel Functions Used:**
- QUARTILE.INC – for calculating Q1, Q2, and Q3
- IF – to classify customers into tiers
- VLOOKUP – for mapping and lookup if needed
- **Business Use:**
This segmentation helps target marketing campaigns more effectively:
- Focus high-value offers on High Tier
- Use retention strategies for Medium Tier
- Apply engagement campaigns for Low Tier

Dashboard for Stakeholders

Retails Sales

Total Revenue
2009-10
9505775.564

Total Revenue
2010-11
9725473.104

Average Sales Per
Transaction 2010-11
375.5008

Average Sales Per
Transaction 2009-10
329.8789

Total Cancellation
Value 2010-11
9251

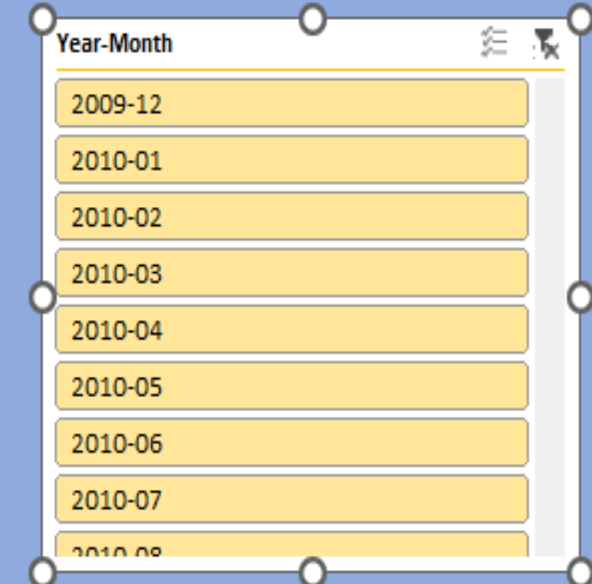
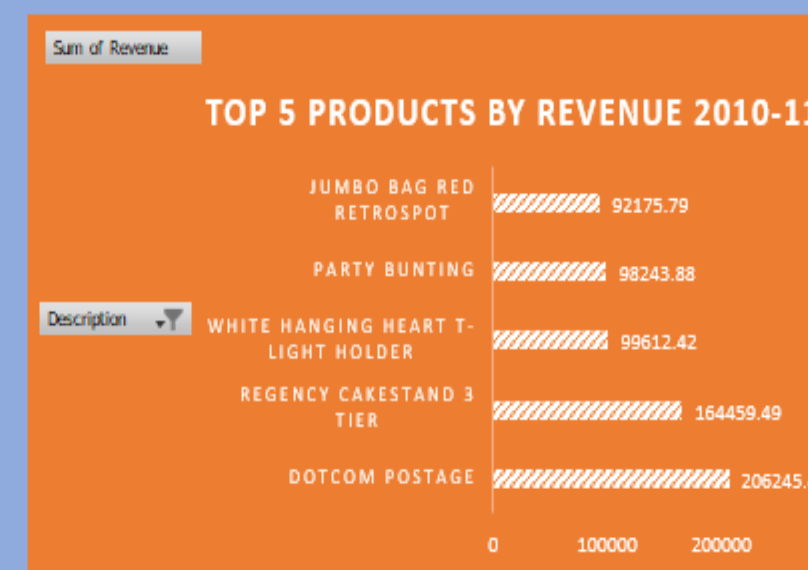
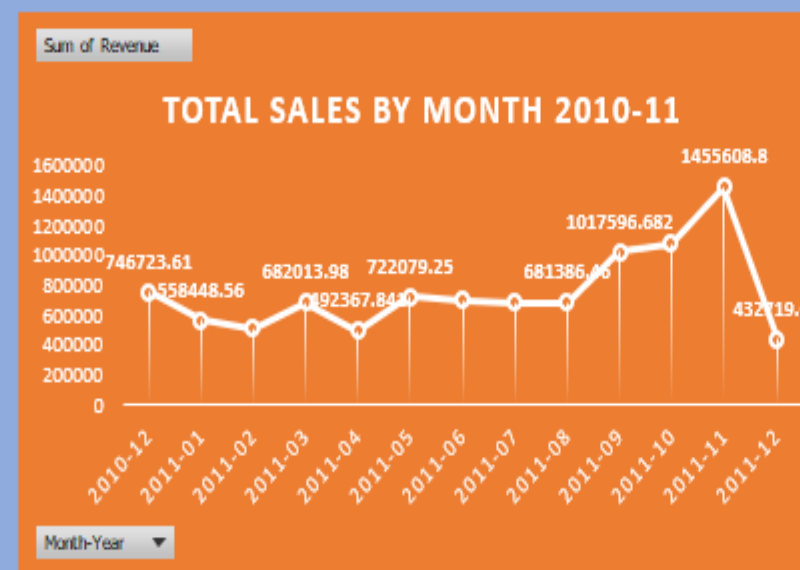
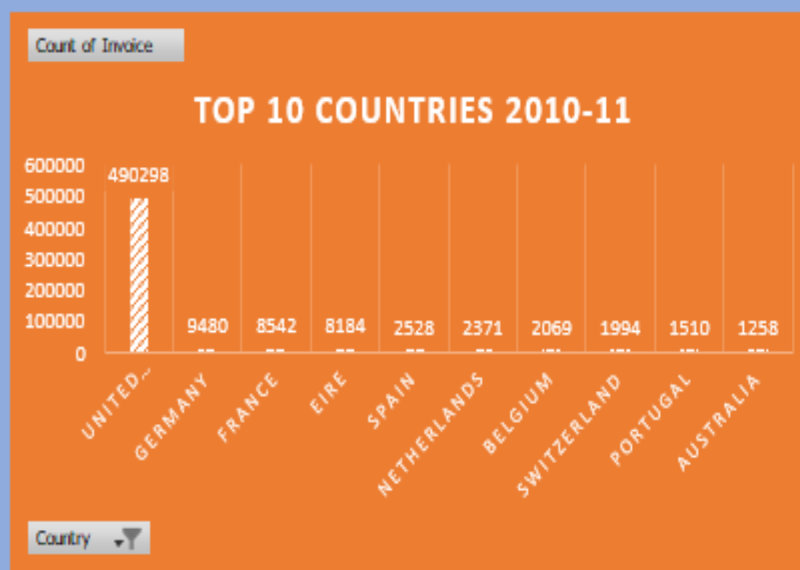
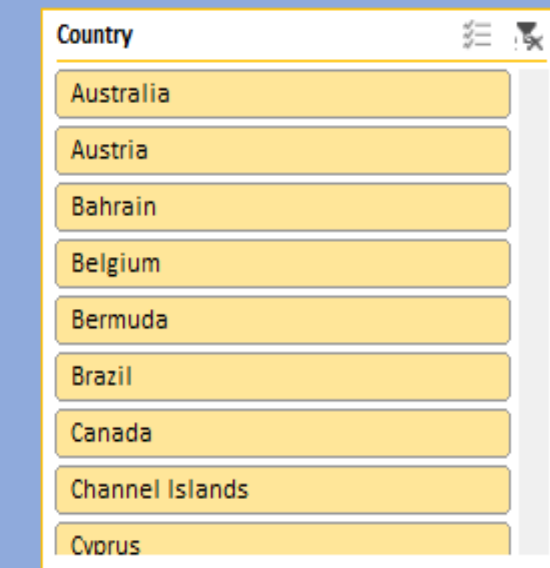
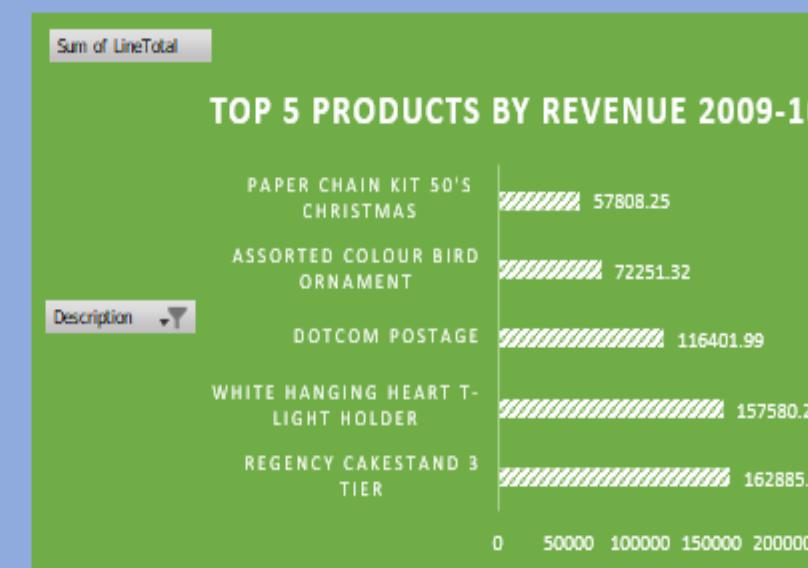
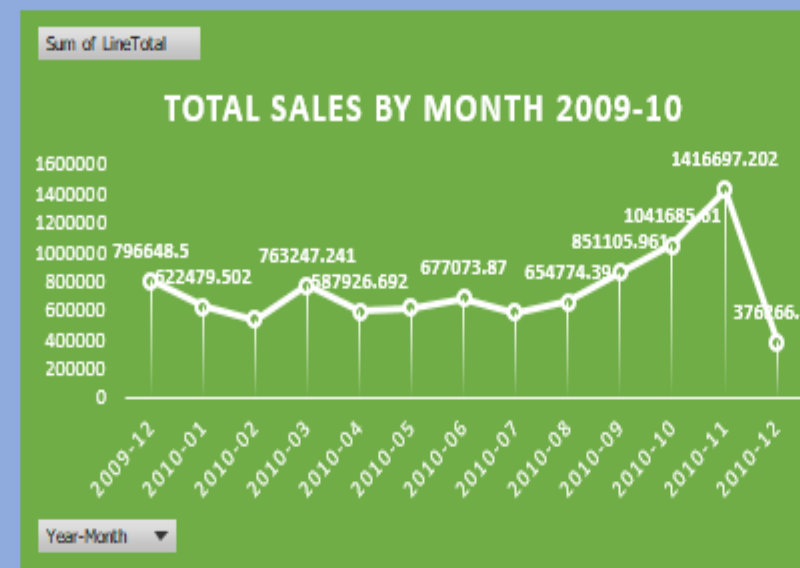
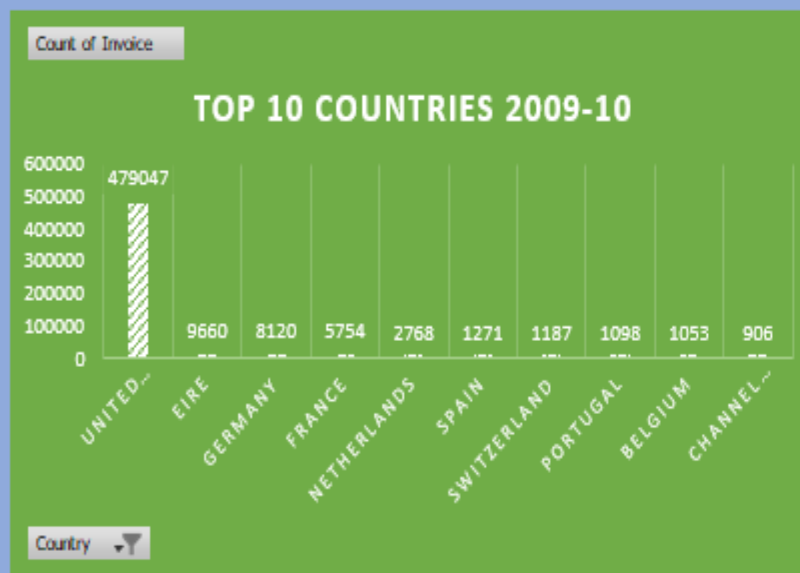
Total Cancellation
Value 2009-10
10182

Total Order
Quantity 2009-10
5992600

Total Order
Quantity 2010-11
5644717

Total Retured
Quantity 2009-10
213289

Total Retured
Quantity 2010-11
275560





Thank you

Data Speaks. Insights Lead.

