Trendify E-Commerce:

Sales Analysis & Customer Segmentation Dashboard

By: Daffa Kaisha Pratama Chandra

Overview

Introduction

Trendify is a dynamic and growing online retailer dedicated to bringing a diverse range of quality products to customers worldwide.

With an expanding digital footprint, Trendify aims to understand its sales performance and customer base more deeply to enhance shopping experiences and drive strategic growth.

Project Overview

Project Goals:

To analyze Trendify's e-commerce sales data to identify key trends, understand customer behavior, and provide actionable insights through an interactive dashboard.

Key Questioned Addressed

1. What are the overall sales trends on a monthly and quarterly basis?
2. Which are the top 10 best-selling products by revenue and quantity?
3. What is the geographical distribution of sales?
4. How can we segment our customers based on their purchasing behavior?

Dataset

The dataset source is from Kaggle.com, the name of the dataset is "Online Retail II UCI".

This Online Retail II data set contains all the transactions occurring for a UK-based and registered, non-store online retail between 01/12/2009 and 09/12/2011.The company mainly sells unique all-occasion gift-ware. Many customers of the company are wholesalers.

Brief information of the dataset:
- The dataset contained 8 features and 1,067,371 records.
- There are 243.007 missing values of Customer ID and 4.382 missing values of product Description.
- There are 34.335 duplicate records in the dataset.

These inconsistencies were addressed during the data cleaning phase to ensure the analysis was accurate.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1067371 entries, 0 to 1067370
Data columns (total 8 columns):
 #   Column       Non-Null Count    Dtype
---  ------       --------------    -----
 0   Invoice      1067371 non-null  object
 1   StockCode    1067371 non-null  object
 2   Description  1062989 non-null  object
 3   Quantity     1067371 non-null  int64
 4   InvoiceDate  1067371 non-null  datetime64[s]
 5   Price        1067371 non-null  float64
 6   Customer ID  824364 non-null   float64
 7   Country      1067371 non-null  object
dtypes: datetime64[s](1), float64(2), int64(1), object(4)
memory usage: 65.1+ MB
None
```

Methodology

Methodology used to analyze the "Online Retail II UCI" dataset.

01 Data Cleaning & Preprocessing
Handled missing values, removed duplicates, created TotalPrice column (for further analysis).

02 Exploratory Data Analysis
Analyzed sales trends, product performance, geographical distribution.

03 Customer Segmentation
Implemented RFM Analysis (Recency, Frequency, Monetary) and customer clustering using K-Means.

04 Dashboard Development
Designed and built an interactive dashboard using Microsoft Power BI.

Dashboard Overview

Home Page

Trendify E-Commerce Dashboard

| Home | Sales Trend | Top Products | Sales By Country | Customer Segmentation |

Select Date

All

Buyer's Country

Total Customers

797.89K

Order's Made

44.88K

NORTH AMERICA

Atlantic Ocean

SOUTH AMERICA

ASIA

EUROPE

AFRICA

Indian Ocean

AUSTRALIA

Pacific O...

Microsoft Bing

© 2025 Microsoft Corporation   Terms

Key Findings:

- **Significant Market Reach:** Acquired a significant customer base of nearly 800,000.

- **Proven Sales Volume:** Processed over 44,000 orders, proving consistent sales volume.

- **Strong Global Footprint:** Established a global footprint with a primary customer base in Europe.
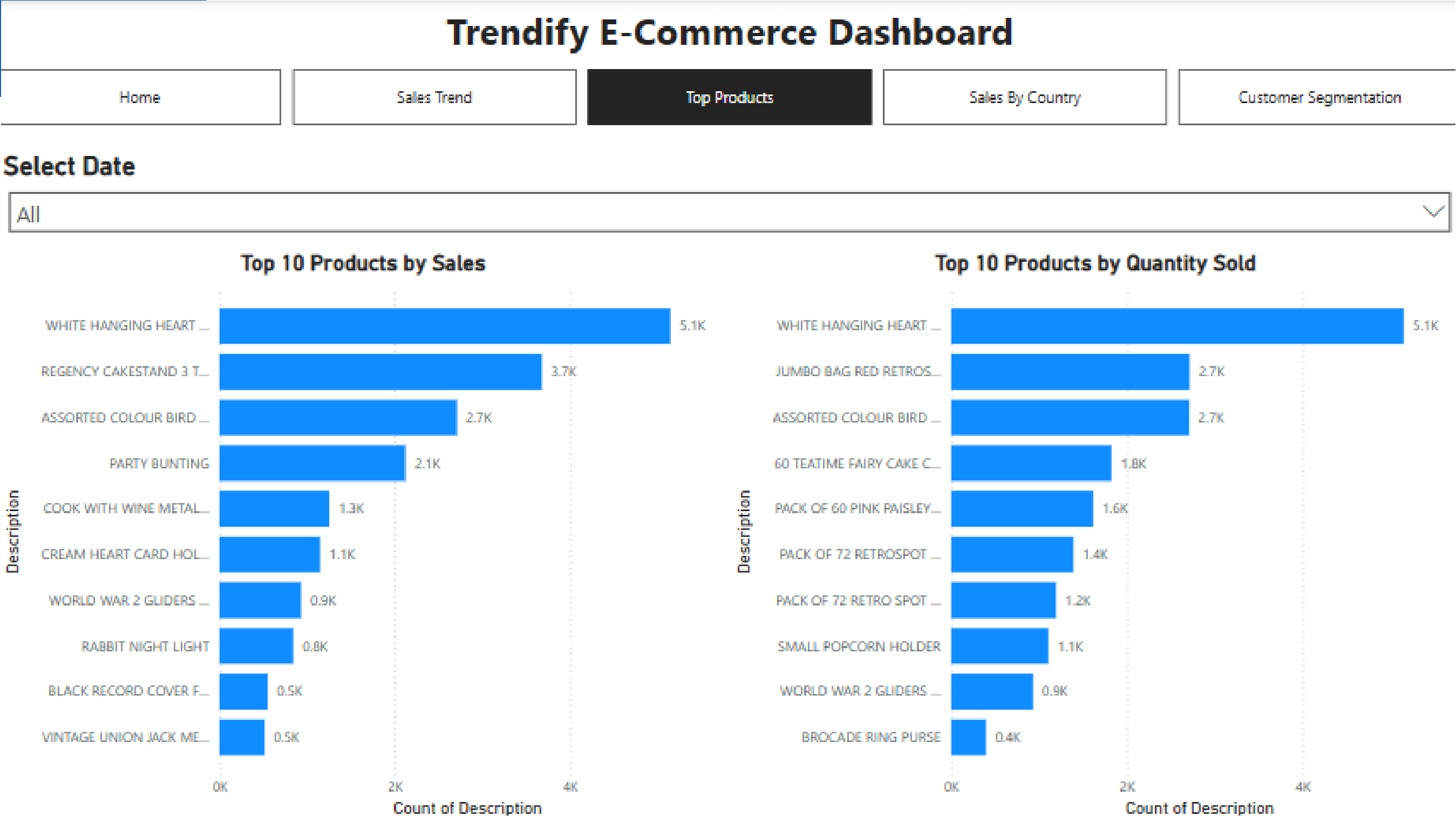
Sales Performance Analysis
Sales Trend

Trendify E-Commerce Dashboard

| Home | Sales Trend | Top Products | Sales By Country | Customer Segmentation |

Select Date

All ▼

Sum of Total Sales by Year and Month
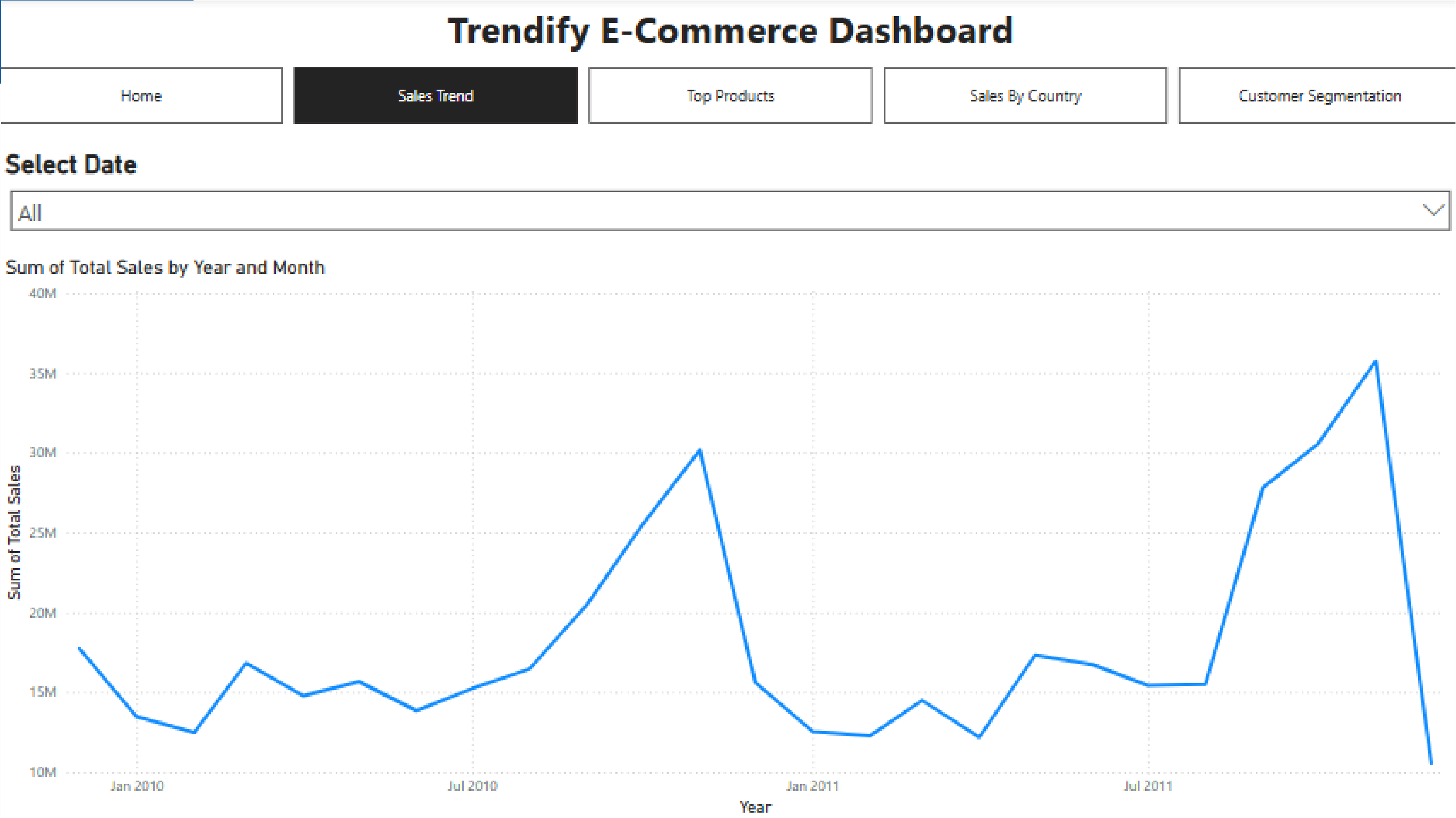
Key Findings:

- **Identified Strong Seasonality:** Strong seasonality identified, with sales peaking every November due to holiday shopping.

- **Positive Year-Over-Year Growth:** Clear Year-over-Year (YoY) growth observed, with 2011 sales outperforming 2010.

- **Note:** The dataset for Dec 2011 is incomplete, impacting final Q4 analysis.
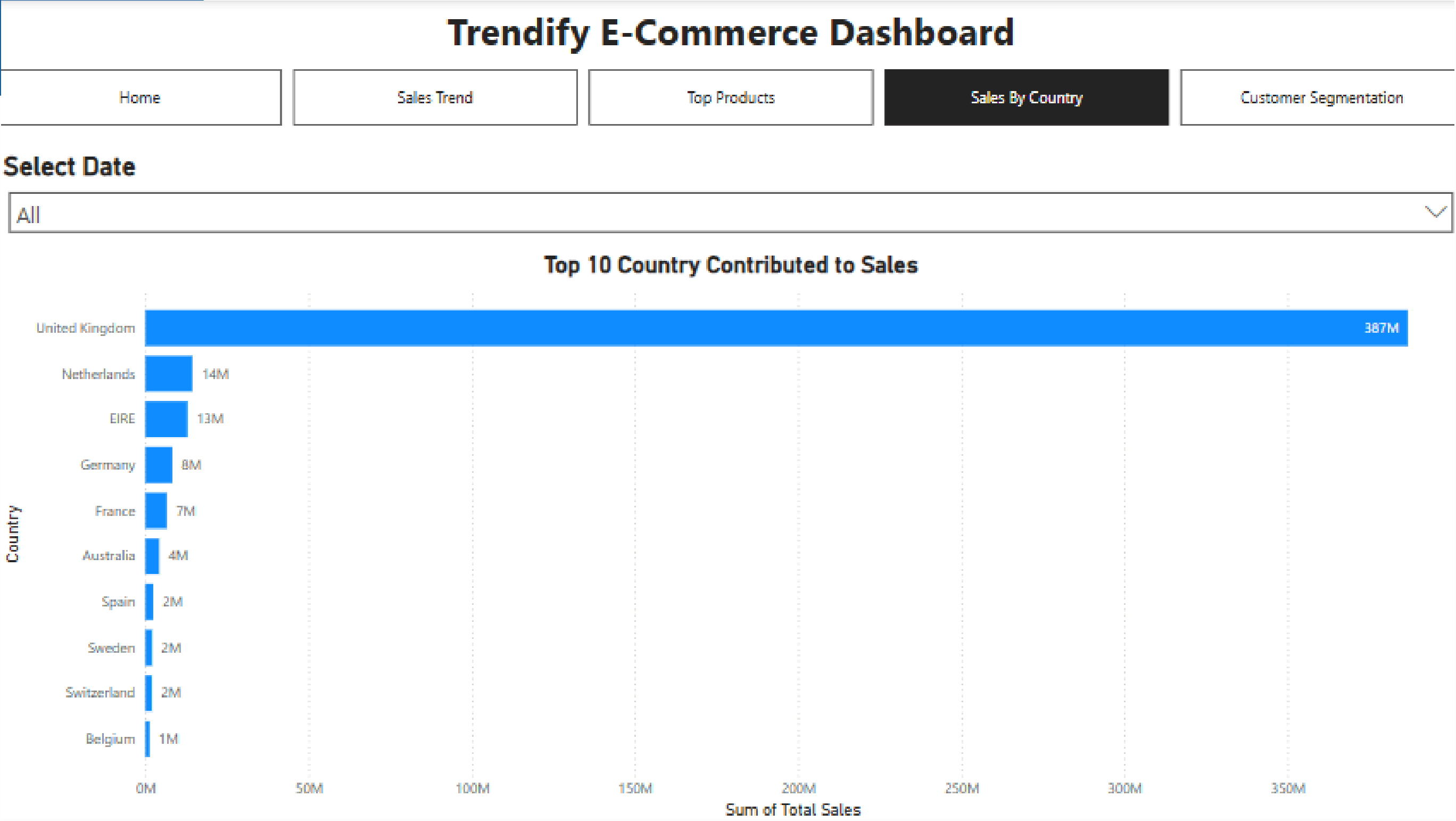
Product Sales Analysis
Top Products

Trendify E-Commerce Dashboard

| Home | Sales Trend | Top Products | Sales By Country | Customer Segmentation |

Select Date

All

Top 10 Products by Sales

WHITE HANGING HEART ...	5.1K
REGENCY CAKESTAND 3 T...	3.7K
ASSORTED COLOUR BIRD ...	2.7K
PARTY BUNTING	2.1K
COOK WITH WINE METAL...	1.3K
CREAM HEART CARD HOL...	1.1K
WORLD WAR 2 GLIDERS ...	0.9K
RABBIT NIGHT LIGHT	0.8K
BLACK RECORD COVER F...	0.5K
VINTAGE UNION JACK ME...	0.5K

Count of Description

Top 10 Products by Quantity Sold

WHITE HANGING HEART ...	5.1K
JUMBO BAG RED RETROS...	2.7K
ASSORTED COLOUR BIRD ...	2.7K
60 TEATIME FAIRY CAKE C...	1.8K
PACK OF 60 PINK PAISLEY...	1.6K
PACK OF 72 RETROSPOT ...	1.4K
PACK OF 72 RETRO SPOT ...	1.2K
SMALL POPCORN HOLDER	1.1K
WORLD WAR 2 GLIDERS ...	0.9K
BROCADE RING PURSE	0.4K

Count of Description

Key Findings:
- **Star Product:** 'WHITE HANGING HEART...' is #1 by both revenue and quantity sold.

- **High-Value Item:** 'REGENCY CAKESTAND' drives high revenue from low sales volume.

- **High-Volume Item:** 'JUMBO BAG...' is popular by quantity but has a lower revenue impact.

Sales Performance by Country
Sales by Country

Trendify E-Commerce Dashboard

| Home | Sales Trend | Top Products | Sales By Country | Customer Segmentation |

Select Date

All

Top 10 Country Contributed to Sales

Country	Sum of Total Sales
United Kingdom	387M
Netherlands	14M
EIRE	13M
Germany	8M
France	7M
Australia	4M
Spain	2M
Sweden	2M
Switzerland	2M
Belgium	1M

Key Findings:

- **UK Market Dominance:** The UK is the primary market, driving significantly more revenue than all other countries combined.

- **Key European Presence:** The Netherlands and Ireland lead as the key secondary European markets.

- **Expansion Opportunity:** Australia is the top-performing non-European market, signaling potential for growth.

Customer Segmentation
RFM Segmentation Analysis

Trendify E-Commerce Dashboard

| Home | Sales Trend | Top Products | Sales By Country | Customer Segmentation |

Select Date

All

Cluster Distribution

0.04K (0.61%)
2.04K (34.28%)
3.87K (65.05%)
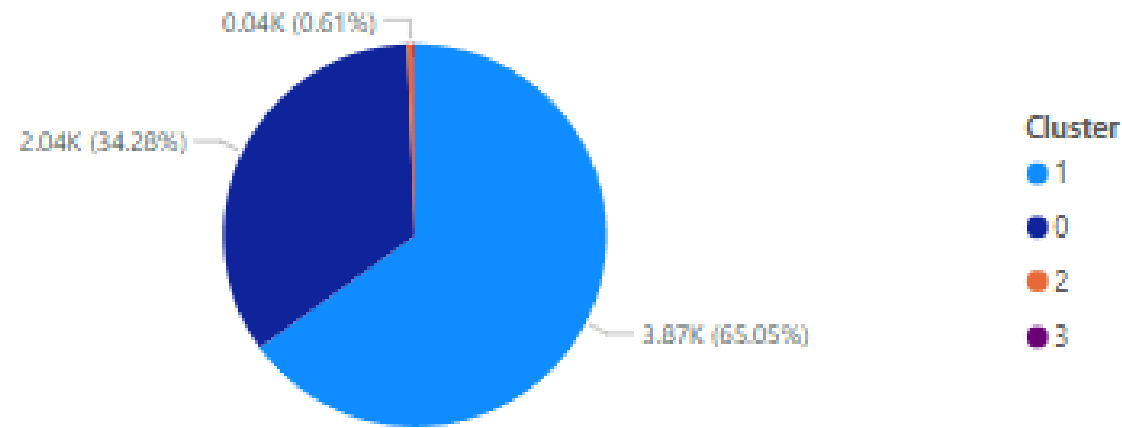
Cluster
- 1
- 0
- 2
- 3

Average of Monetary by Cluster

3 — 0.41M
2 — 0.07M
1 — 0.00M
0 — 0.00M

0.0M   0.1M   0.2M   0.3M   0.4M
Average of Monetary
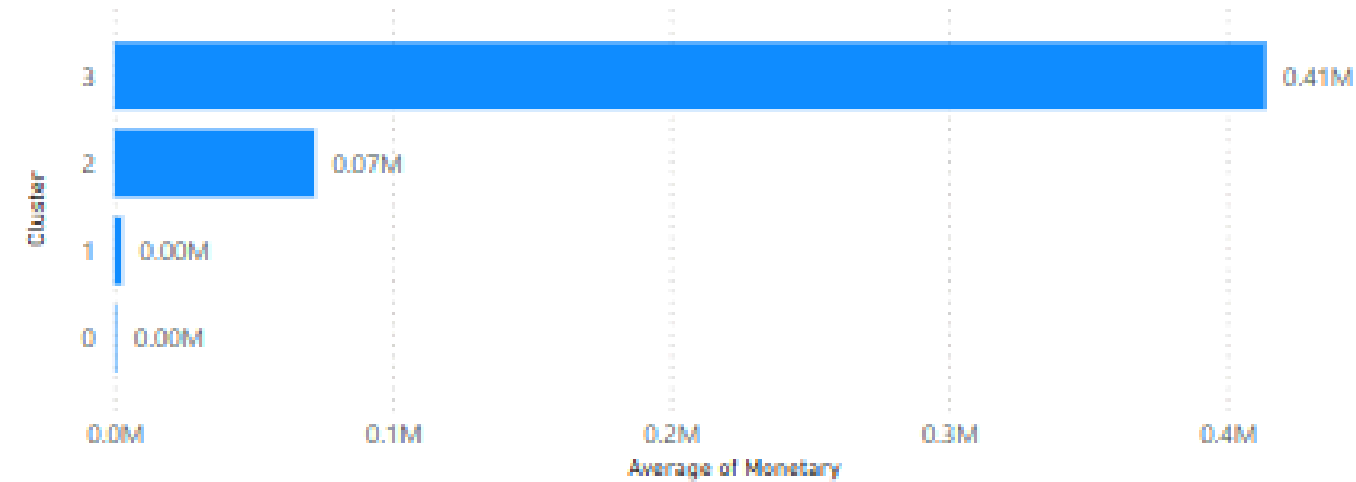
Average of Recency by Cluster

0 — 466.26
1 — 66.00
2 — 22.81
3 — 3.50

0   100   200   300   400   500
Average of Recency
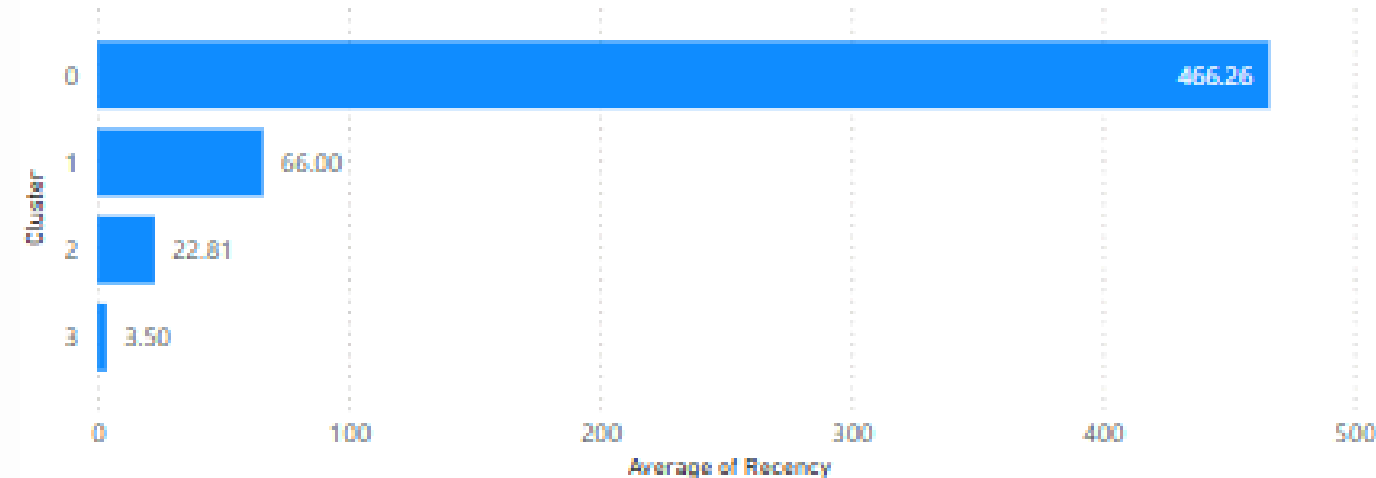
Average of Frequency by Cluster

3 — 257.25
2 — 128.97
1 — 8.78
0 — 2.58

0   50   100   150   200   250
Average of Frequency
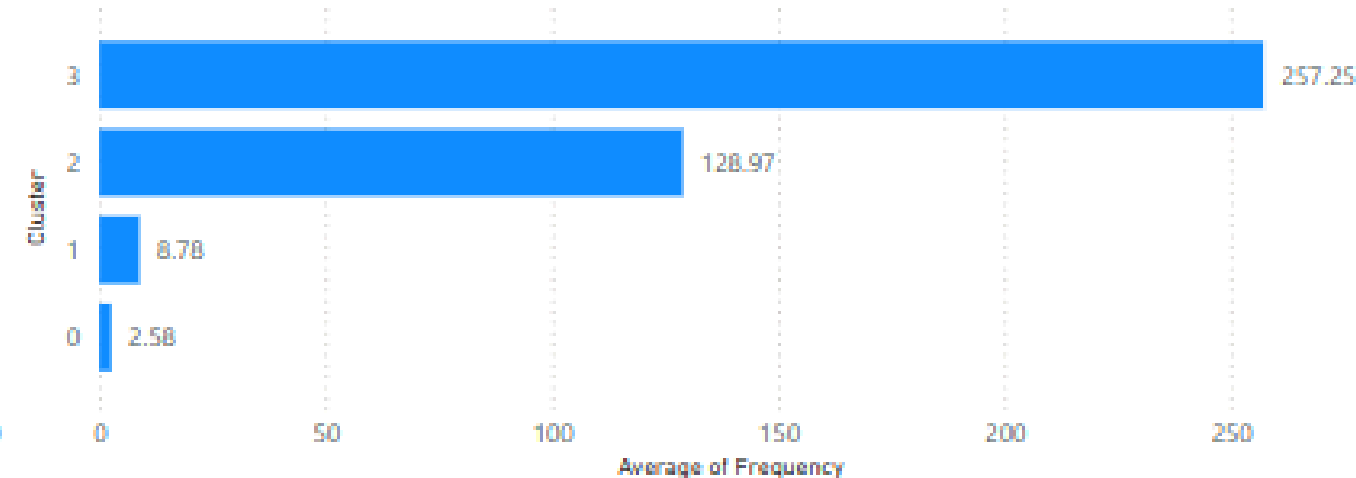
Key Findings:

- **Customers were segmented into 4 key personas**: At-Risk, Regular, Champions, and VIPs.

- **A large "At-Risk" group was identified**, highlighting a key opportunity for re-engagement campaigns.

- **Elite "VIPs" & "Champions" (<1% of all customers)** drive a massive share of revenue, making their retention critical.

Actionable Insights
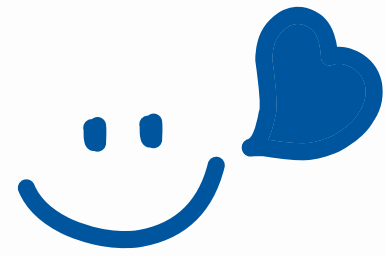
Re-engage At-Risk Customers

Analysis revealed a large "At-Risk" customer segment. Action: Launch a targeted email campaign with a special discount to win back these customers and prevent churn.
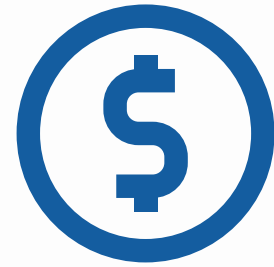
Retain High-Value Customers

An elite group of "VIP" & "Champion" customers (<1%) generates a massive share of revenue. Action: Implement an exclusive loyalty program with special perks to ensure retention of these critical customers.
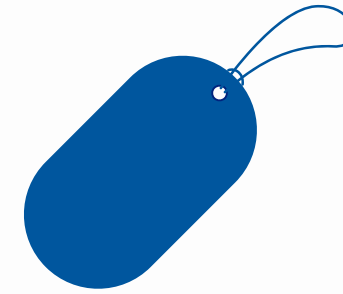
Optimize Product Strategy

Analysis identified distinct product groups (high-value vs. high-volume). Action: Promote high-value items to VIPs and use popular, low-cost items as "add-ons" to increase average order value.

Drive Strategic Expansion

Geographical analysis shows Australia is the top-performing non-European market. Action: Launch targeted digital advertising for Australia to accelerate growth and capitalize on this proven product-market fit.

Tools
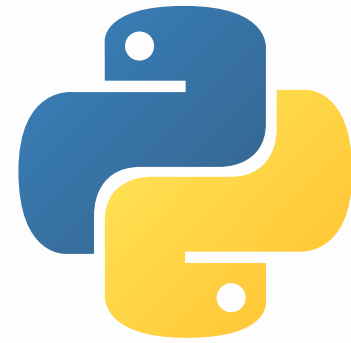
Python

Leveraged Python's data science ecosystem, primarily using the Pandas library for data cleaning and manipulation, and Scikit-learn for performing the RFM analysis and customer clustering.

Jupyter Notebook

Served as the interactive development environment for the entire analytical process, allowing for step-by-step data exploration, code execution, and documentation of the findings.

Microsoft Power BI

The primary business intelligence tool used to connect to the cleaned dataset, create all interactive charts and KPIs, and build the final, user-friendly dashboard for presenting the key insights.

Challenges

Challenge	How I Overcame It
Implementing RFM Analysis for the first time	• Broke the problem into smaller stages (calculating R, F, & M individually). • Researched best practices to ensure a robust and accurate model.
Applying K-Means Clustering effectively	• Implemented the 'Elbow Method' to find the optimal number of clusters. • Used this data-driven approach to select 4 meaningful segments.

THANK YOU!

Daffa Kaisha Pratama Chandra

daffakpc21@gmail.com

linkedin.com/daffakaisha