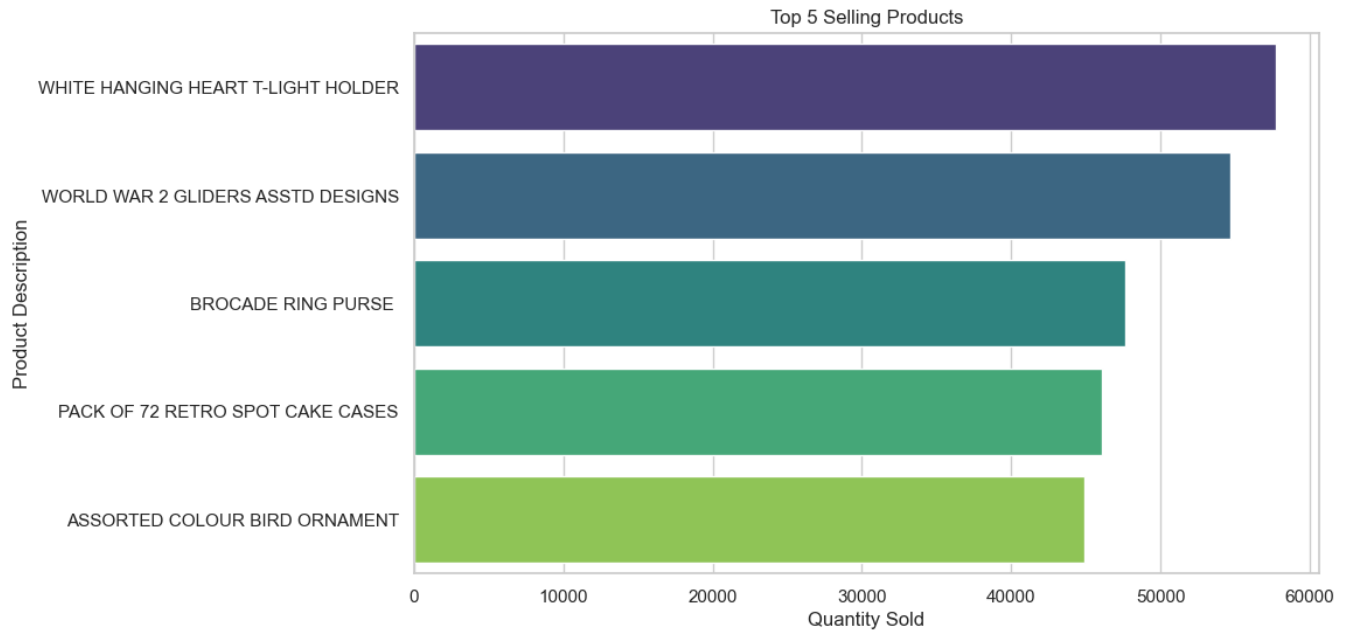


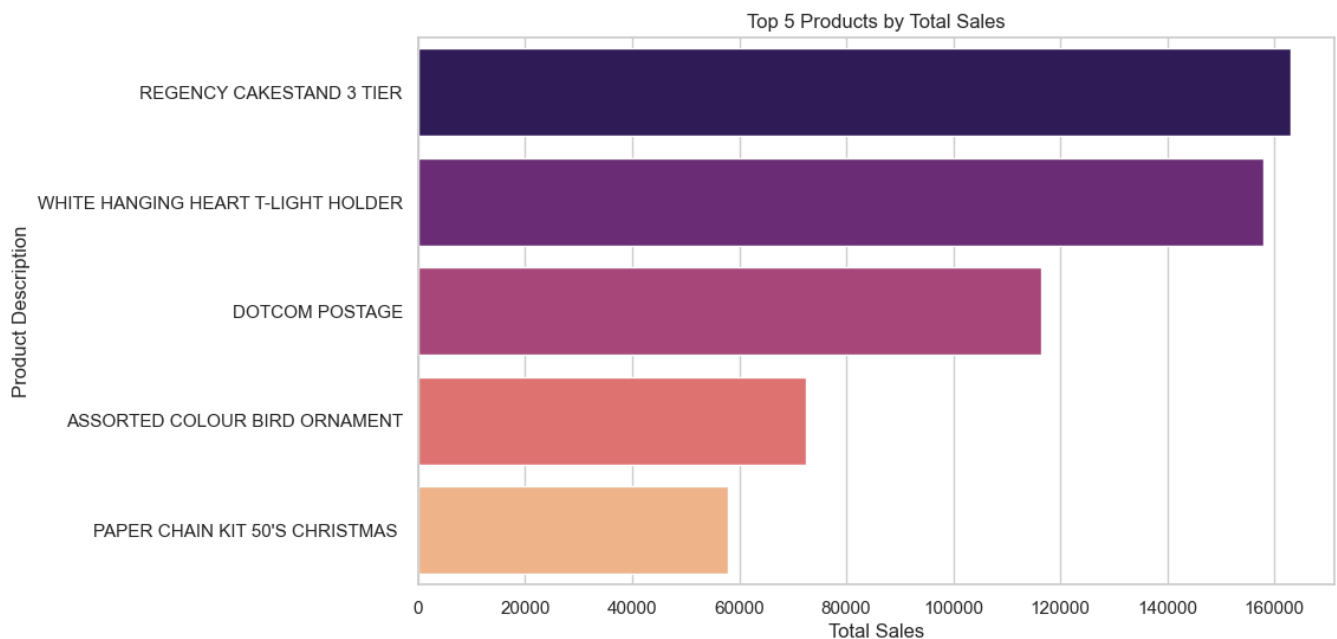
ANALYSIS OF ONLINE RETAIL STORE II DATA

Throughout the analysis of the dataset, I have taken a question-answer based approach to understanding the data. Below are the insights I gained from the analysis:

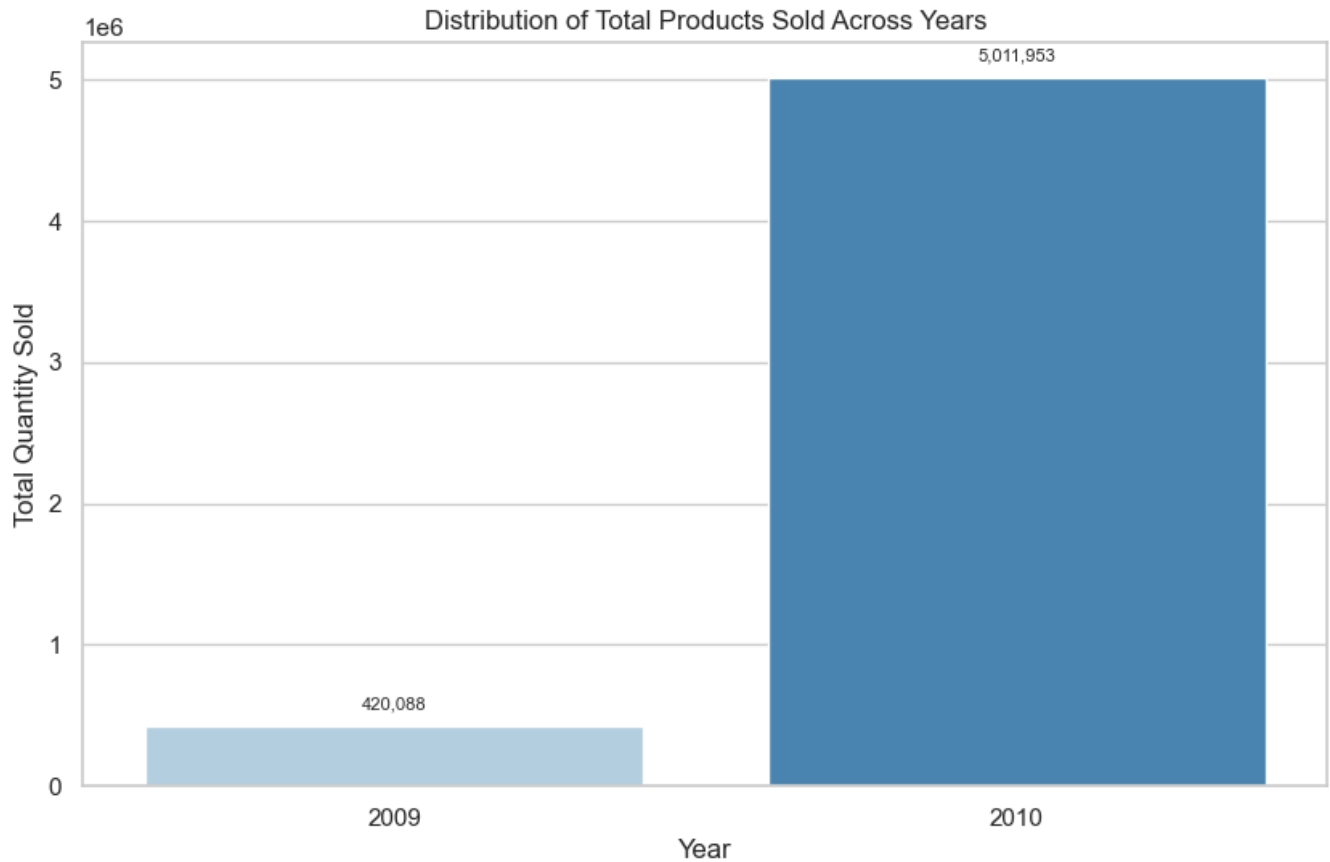
- The dataset contains approximately **526 thousand** purchase records with **WHITE HANGING HEART T-LIGHT HOLDER** being the most purchased product category (**57733** times).



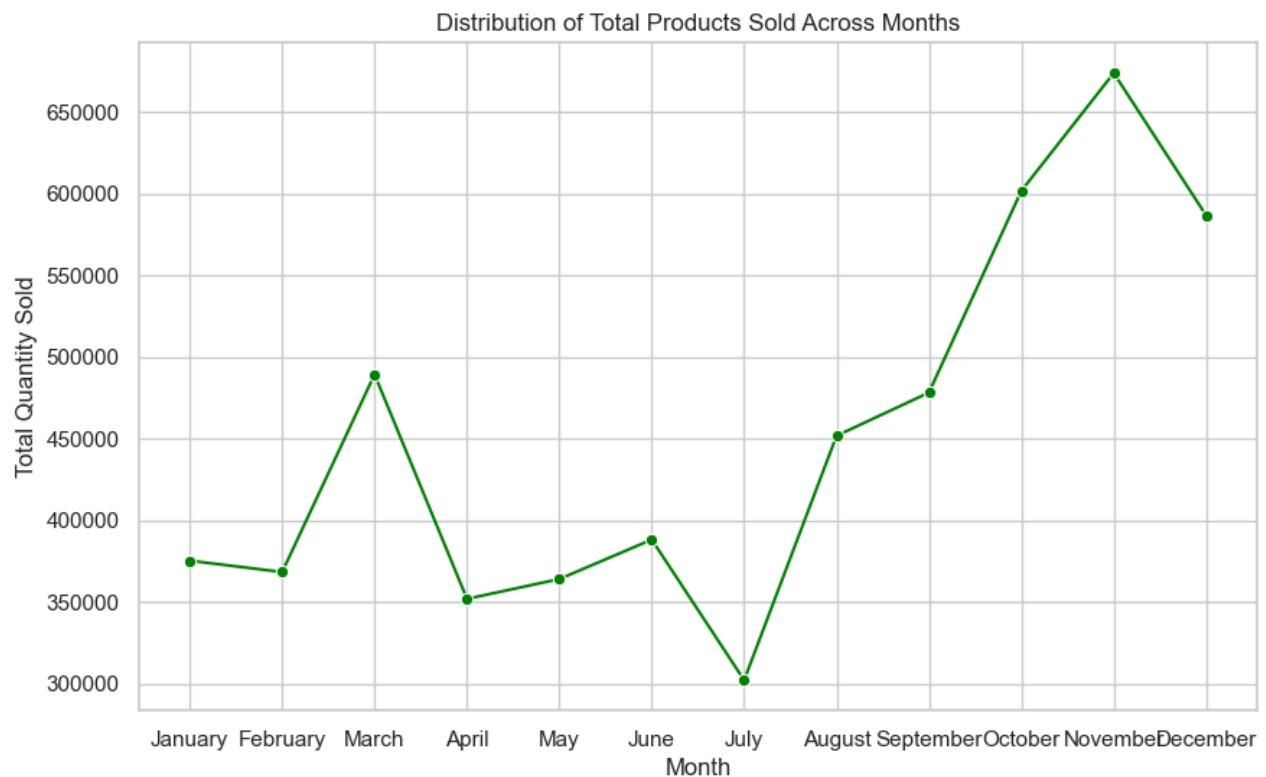
- REGENCY CAKESTAND 3 TIER** and **WHITE HANGING HEART T-LIGHT HOLDER** are the top purchased products by total sales



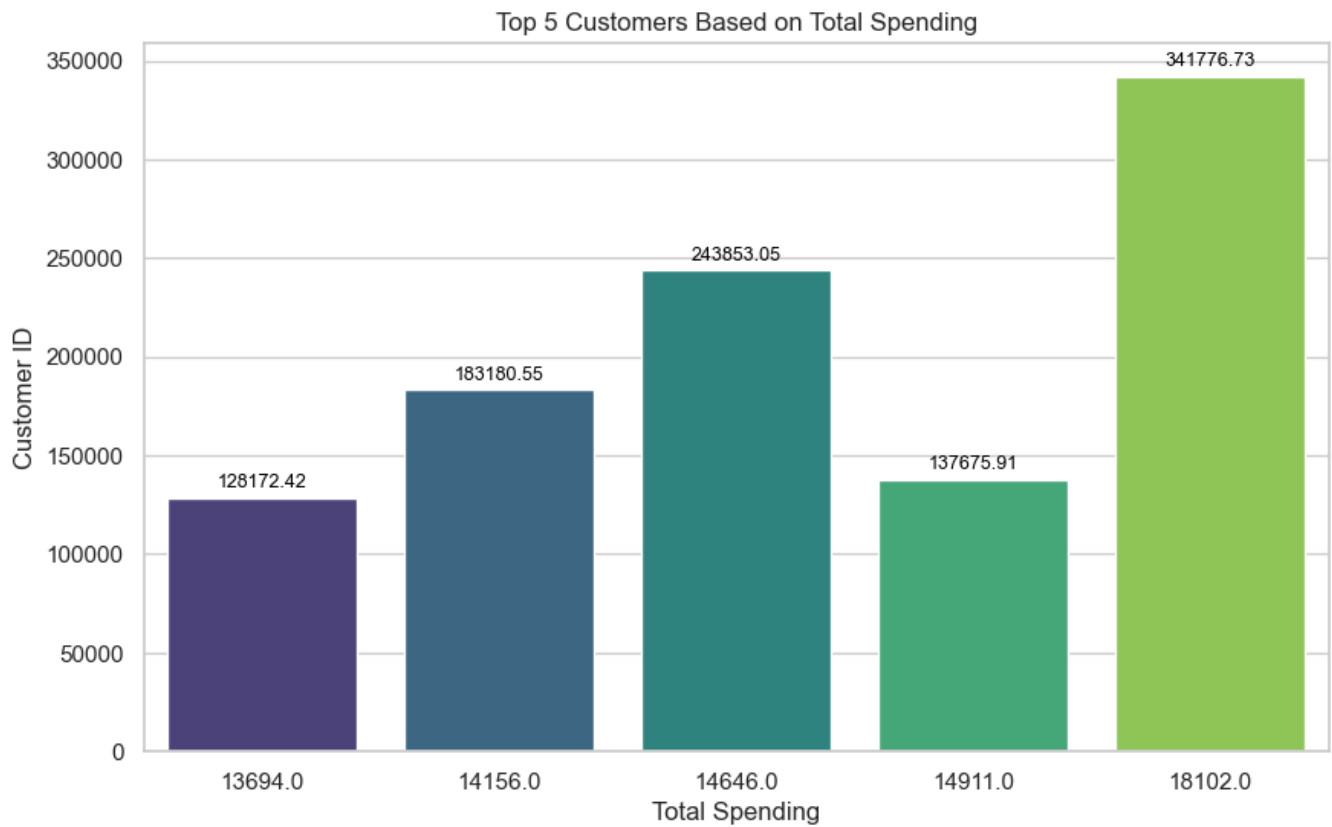
- The total number of products sold in 2010 was significantly higher.



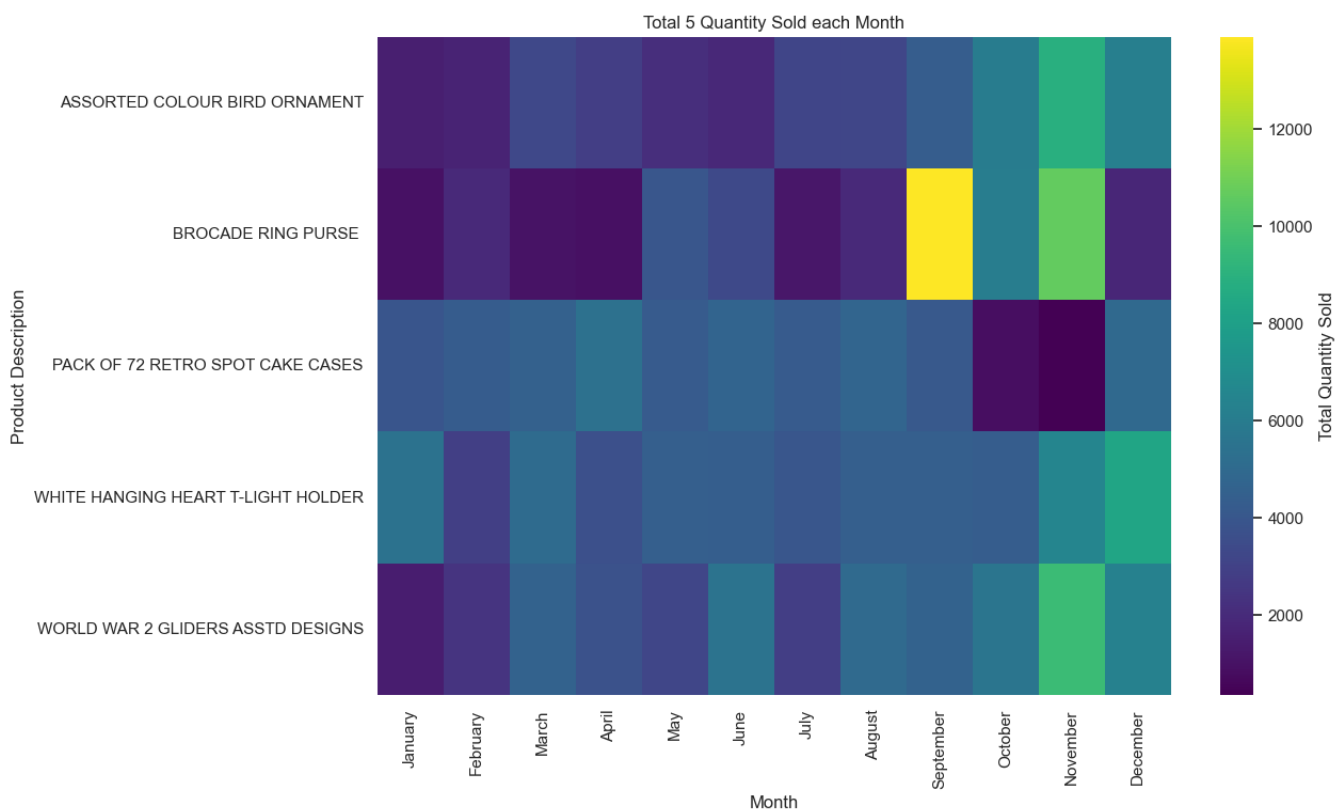
- Across 2009 and 2010 combined, **the highest number of product appear to be brought in the month of November.**



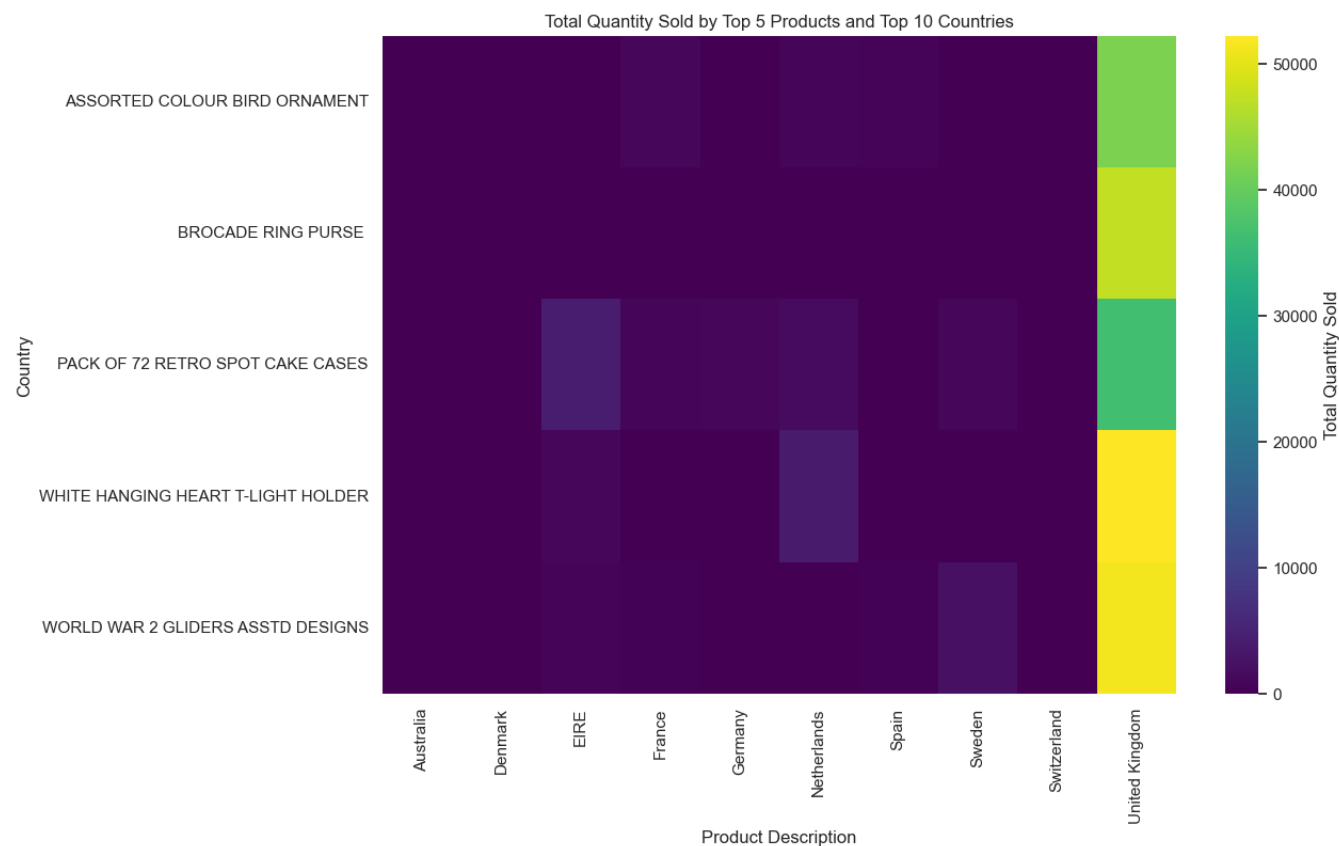
- Customer with id 18102 appears to be the highest spender



- The heatmap below answers the question “do certain products sell better during specific months?”. **BROCADE RING PURSE** seems to be the best-selling product in September while WORLD WAR 2 GLIDERS and BROCADE RING PURSE the best-selling in November and so on.



- There appears to be very less product sold in other countries. Majority of the total sales was completed in United Kingdom (92%). In addition, WHITE HANGING HEART T-LIGHT HOLDER are the most sold products.



- Customer Clustering using Elbow method

