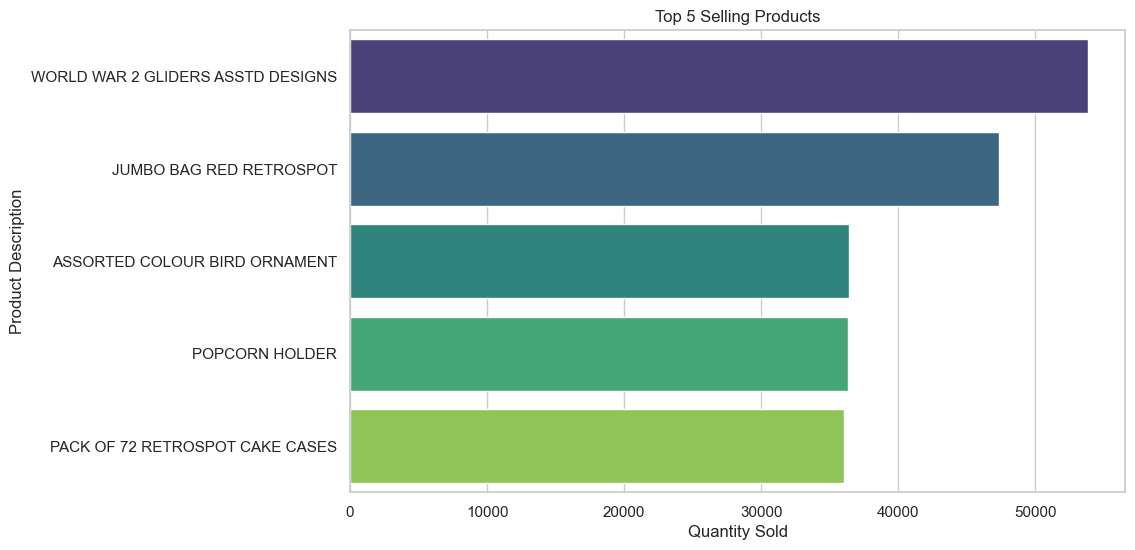
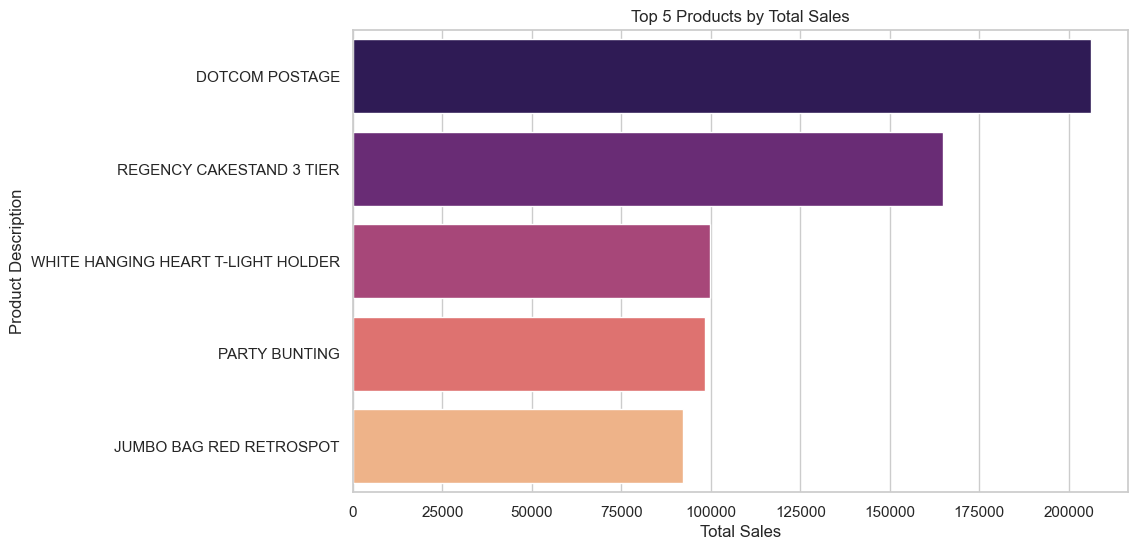
**ANALYSIS OF ONLINE RETAIL STORE DATA**

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:

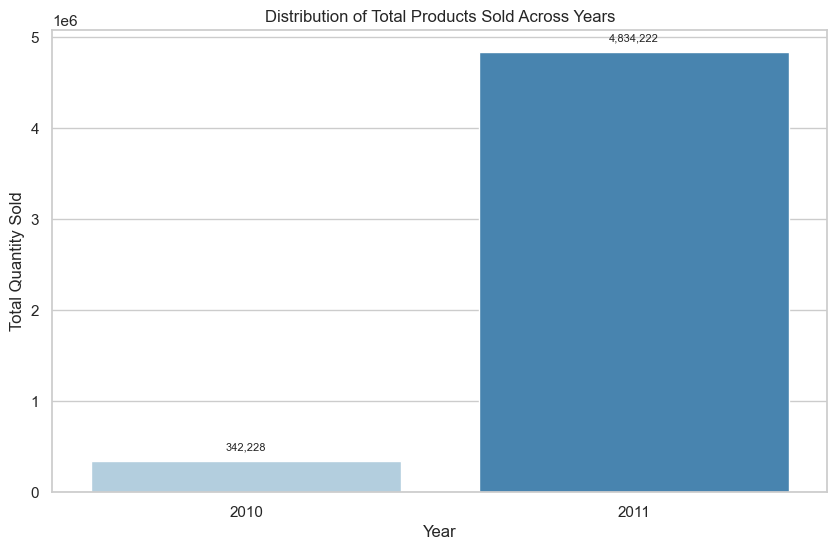
* The dataset contains approximately **542 thousand** purchase records with **World War 2 Glider designs being the most purchased product category (53847 times).**



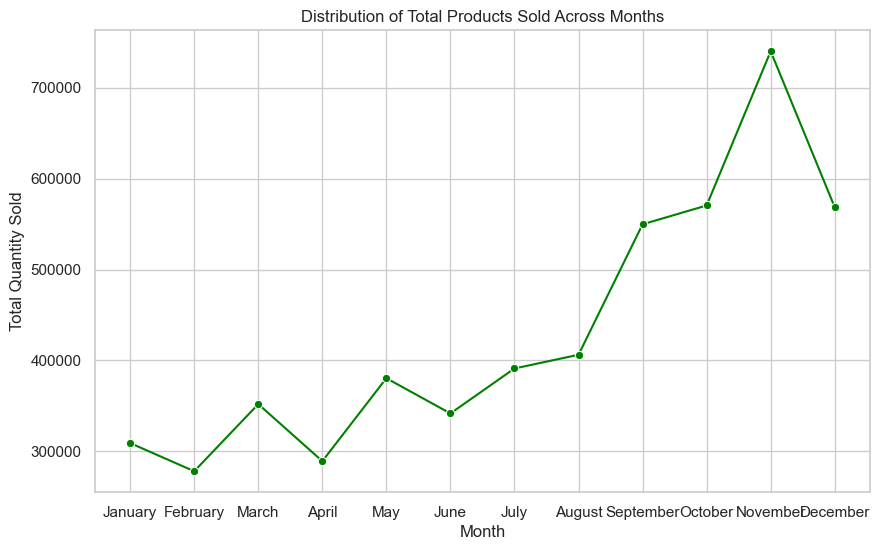
* **DOTCOM POSTAGE** and **REGENCY CAKESTAND 3 TIER** are the top purchased products by total sales



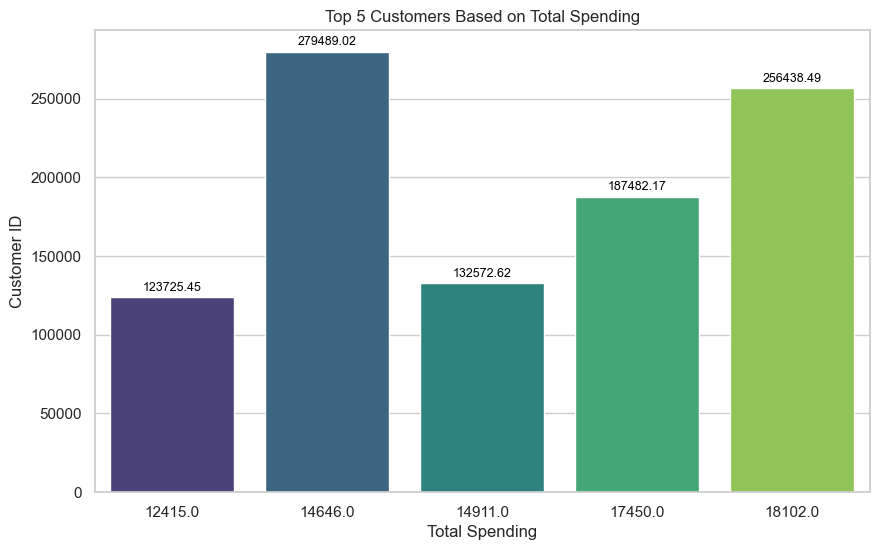
* The total number of products sold in 2011 was significantly higher.



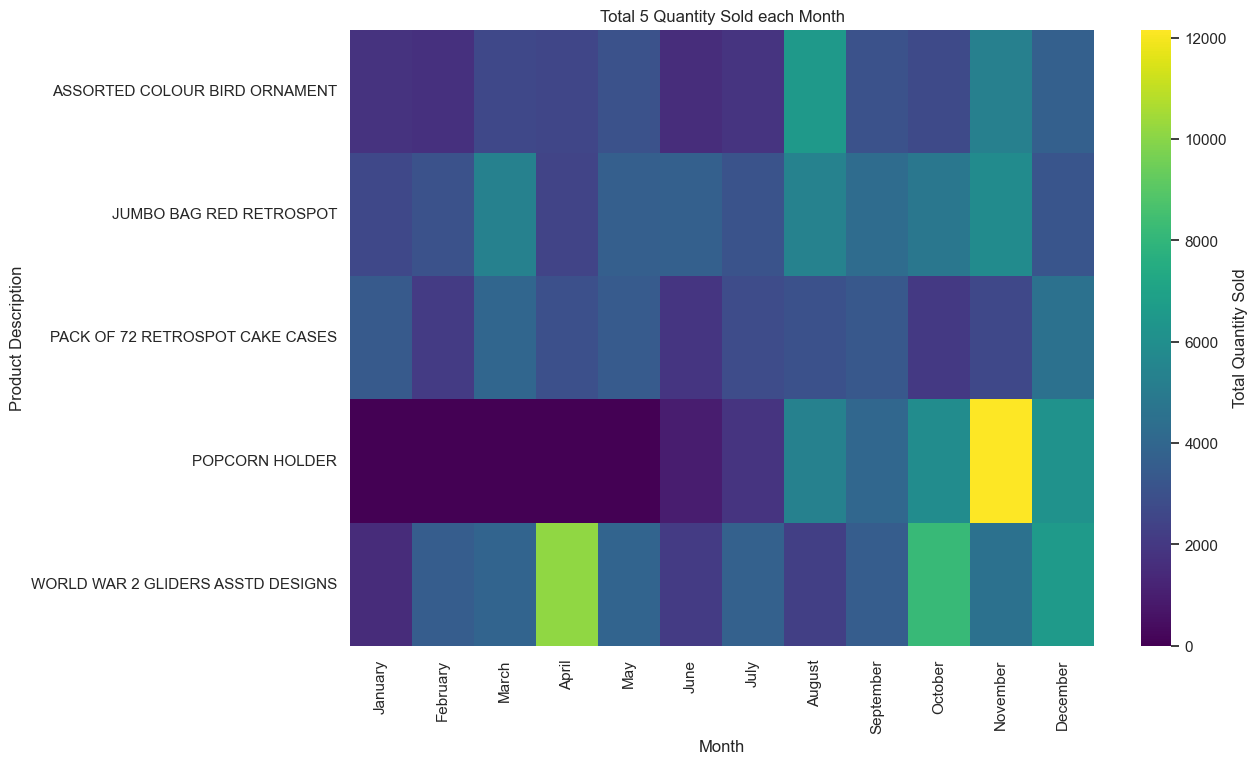
* Across 2010 and 2011 combined**, the highest number of product appear to be brought in the month of November.**



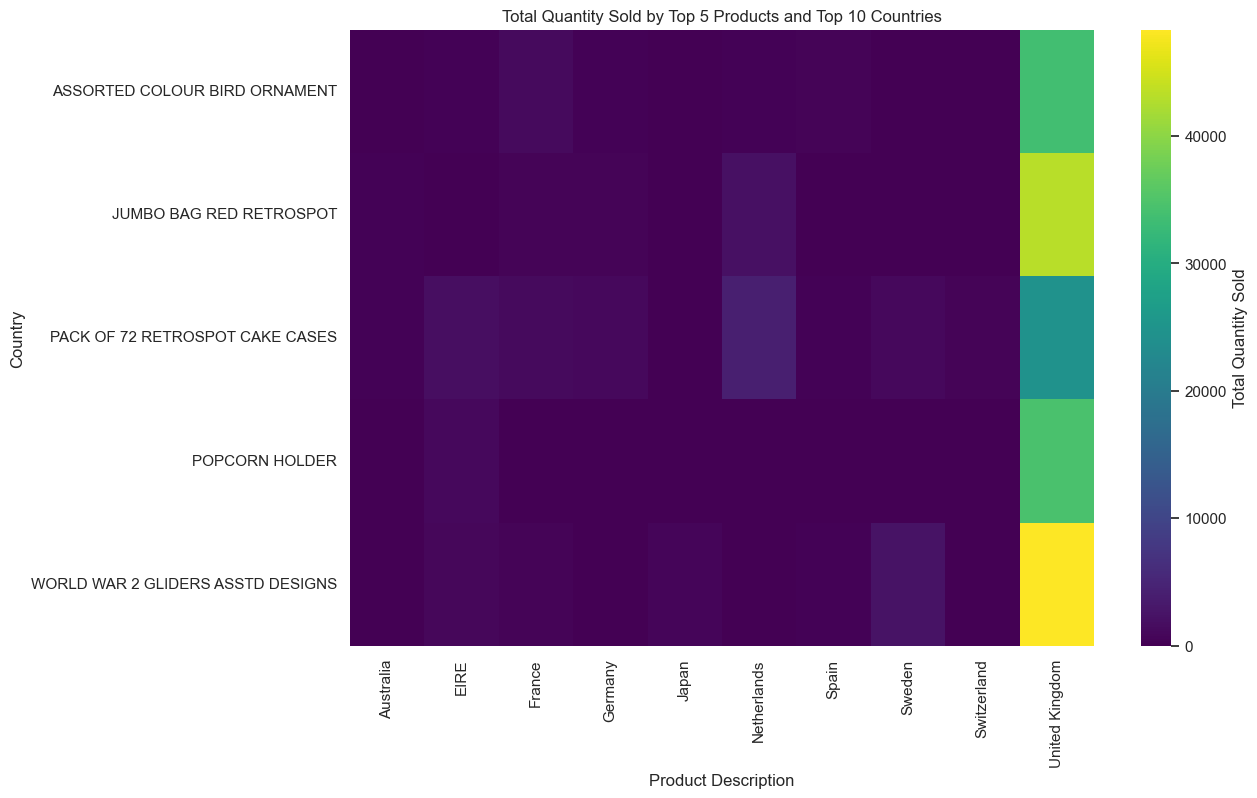
* Customer with id 14646 appears to be the highest spender



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



* There appears to be very less product sold in other countries. Majority of the total sales was completed in United Kingdom (91%). In addition, World War 2 Gliders are the most sold products.



* Customer Clustering using Elbow method

