Executive Summary: The online national clothing chain is aiming to revitalize its sales by implementing a targeted marketing campaign to win back lost customers. To achieve this, they plan to advertise specific products to specific customers in particular locations. However, they lack the necessary information on whom to target. Three products have been identified for the campaign: a shirt priced at $25, a sweater at $100, and a leather bag at $1,000.

To gain insights into their customer database and inform their marketing strategy, several analysis questions were addressed and the results have been summarized as in the following:

- The correlation (R2 value) between sales and income was found to be 0.78, indicating a strong positive relationship. This suggests that as customer income increases, so does their likelihood of making purchases and spending more on average based on data dating back to 6 months prior. **(See Page 2 for more info)**
- The correlation (R2 value) between customer ratings and product return rate was determined to be 0.69, demonstrating a strong positive relationship. This implies that higher customer ratings are associated with lower product return rates, indicating greater customer satisfaction towards company's products. Further analysis can be done to determine what makes some products better than others and the reasoning behind having high or low ratings for company's products. **(See Page 3 for more info)**
- A Linear Regression Model was developed to construct a formula to predict customer sales and incomes. The formula for predicting sales based on income is Sales = 0.0107 * Income - 722.1416. Conversely, the formula for predicting income based on sales is Income = (Sales + 722.1416)/0.0107. These formulas enable estimation of sales and incomes based on available data. **(See Page 2 for more info)**
- It was determined that customers who are willing to spend or have spent over 300$ and above for the last 6 months in total purchases are predicted to have a yearly income over 100,000$ and above according to predictions made by the Regression Model (the regression model confirms that higher income levels correspond to increased spending). In addition, it was determined that customers with an average yearly income of 100k and above are responsible for **over 60%** of company's total sales **(See Page 6 for more info and Summary Table for further detailed info)**
- The product that will be advertised the most was predicted to be the **25$ shirt**, while the **100$ sweater** comes in second and the **1000$ leather bag** comes in third. Based on previous spending behaviours dating back to 6 months, it was found that a significant portion of the company's customer database falls within an **income range of $70,000-100,000** accounting **for 63% of total customers** and contributing up to **35% to total sales**; as a result, the regression model predicts that most of former mentioned income customers segment are willing to spend $25 on a shirt and perhaps a little more. Consequently, the shirt is expected to receive the highest advertising focus.

A decision tree in addition to several other variables that were analysed was provided to facilitate product offering decisions based on predicted customer incomes derived from average spending patterns. This additional tool aids in tailoring the marketing campaign to specific customer segments. **(See Pages 1, 4, 5 and 7 for more info)**.

In summary, the clothing chain's targeted marketing campaign should prioritize advertising the shirt, given the customer income profiles and spending behaviour. By leveraging the insights from data analysis, the company can effectively target and engage customers, leading to improved sales and customer retention.