Assignment 2

This is an individual assignment. Please submit by due date on Canvas

For this assignment, you will use RFM_data.csv file

Sephora: Using RFM Analysis to Target Email Campaigns

Sephora is a global leader in beauty retail, offering a wide range of products, including cosmetics, skincare, fragrances, and more. With millions of customers worldwide, Sephora relies heavily on digital marketing to engage its audience and drive sales. Email marketing, in particular, has been one of the successful channels for promoting exclusive deals, new product launches, and loyalty rewards.

However, Sephora faces a common challenge in email marketing: how to maximize engagement while minimizing the risk of customer dissatisfaction. In recent years, Sephora has noticed that sending promotional emails to all customers often leads to higher unsubscribe rates and even account deletions. Customers who feel overwhelmed by frequent or irrelevant emails may disengage entirely, which reduces Sephora's ability to communicate with them in the future.

To address this issue, Sephora's marketing team wants to use **targeted email campaigns** to send coupon offers only to the customers most likely to respond positively. By identifying and targeting the right customers, Sephora aims to:

1. Increase the likelihood that recipients will use the coupon code.
2. Reduce the risk of unsubscribes or account deletions.
3. Optimize the overall effectiveness of its email marketing campaigns.

The Coupon Campaign

Sephora is preparing a campaign to offer a **20% off coupon** on a customer's next purchase. The coupon will be valid for 30 days and will apply to all items in the customer's online cart. The goal is to target customers who are most likely to:

1. Use the coupon to make a new purchase within 30 days, who otherwise might not purchase.
2. Continue engaging with Sephora's email marketing in the future.

At the same time, Sephora wants to minimize the cost of the campaign. Although sending emails seems that it does not incur a cost, marketing team at Sephora knows that sending emails to customers who are unlikely to use the coupon—or who might unsubscribe due to annoying promotional emails—represents a wasted cost. To estimate the cost of annoyance, Sephora's team considers:

- The likelihood that a customer will unsubscribe after receiving the email.

- The average loss in revenue from an unsubscribed customer over their lifetime.

Using RFM Analysis to Target Customers

To make the campaign successful, Sephora plans to use RFM analysis to identify and target customers who are most likely to respond positively to the coupon. RFM (Recency, Frequency, Monetary) is a simple but powerful method for segmenting customers based on their past behavior:

- **Recency (R)**: How recently a customer has made a purchase.
- **Frequency (F)**: How often a customer has made purchases in the past year.
- **Monetary (M)**: How much a customer has spent in the past year.

The logic behind RFM is straightforward: customers who have purchased recently, frequently, and spent more are more likely to respond positively to promotions.

Sephora's marketing team had long relied on email campaigns to engage customers, promote new products, and drive sales. However, excessive or poorly targeted emails could annoy customers, leading them to unsubscribe from Sephora's email marketing list. This posed a significant challenge: without email as a channel, Sephora would lose a critical way to reach its customers, making future promotions less effective.

To better manage this risk, Sephora's Chief Marketing Officer decided to quantify the marginal cost of sending a promotional email. To do so, the team needed to calculate two key metrics:

1. The likelihood of a customer unsubscribing after receiving an additional email.
2. The financial difference in behavior between customers who remained subscribed to emails and those who unsubscribed.

By multiplying these two metrics, Sephora could estimate the marginal cost of sending one additional promotional email.

Sephora had sent millions of promotional emails in recent years for various reasons: announcing new product launches, sharing discount codes, highlighting limited-time offers, and sending loyalty program updates. By analyzing historical data, the marketing team found that approximately 1.8% of customers unsubscribed from Sephora's email list after receiving an additional promotional email. This unsubscribe rate varied slightly depending on the content of the email and the customer's engagement level, but 1.8% was a reliable average.

Next, the team needed to determine the financial impact of a customer unsubscribing from email marketing. Sephora knew that customers who remained subscribed were significantly more engaged: they shopped more frequently, spent more per purchase, and were more likely to respond to future email campaigns. By analyzing past purchase data, the team estimated that subscribed customers spent $60 more per year compared to unsubscribed customers. This difference represented the financial value of keeping a customer subscribed to email marketing.

Using these two metrics, Sephora calculated the marginal cost of sending an additional promotional email to be $60 \times 0.018 = \$1.08$. This meant that every additional promotional email Sephora sent resulted in an average loss of $1.08 per recipient due to unsubscribes. While this cost might seem small, it added up quickly when scaled across Sephora's massive customer base.

Implications for the Coupon Campaign

The insights from this analysis underscored the importance of targeting email campaigns effectively. For Sephora's upcoming 20% coupon campaign, sending emails to disengaged customers could result in significant losses if those customers unsubscribed without using the coupon. On the other hand, carefully targeting the most engaged and responsive customers could maximize the campaign's effectiveness while minimizing the marginal cost.

Sephora's marketing team decided to use RFM analysis to identify the best customers for the campaign, ensuring that the emails reached those most likely to redeem the coupon and remain engaged with the brand.

Running a Test Campaign: Sephora's 20% Coupon Offer

To identify which customers would be most likely to respond positively to a promotional email, Sephora decided to run a test campaign. The campaign targeted 35,000 randomly selected customers from three regions in the United States: New York City, Chicago, and Los Angeles. These customers received an email with a 20% off coupon for their next Sephora purchase. The email informed recipients that the discount would automatically apply to any eligible online purchase during the promotion period. Sephora chose this approach to simplify customer experience. There was no need for recipients to enter a code; they simply had to make a purchase within the time window to take advantage of the discount.

Sephora already had RFM data for all of its customers, including both the 30,000 customers in the test campaign and millions of others who were not included. By combining the test results with the RFM data, Sephora planned to:

- Measure Engagement: Determine how many customers from each RFM segment responded positively by using the coupon.

- Assess Profitability: Calculate the return on marketing investment (ROI) for each RFM segment by comparing the revenue generated with the campaign costs.
- Identify Target Segments: Pinpoint the RFM segments that were profitable to target, based on their engagement and ROI.

A week after the test campaign ended, Sephora's marketing team reviewed the data. Of the 35,000 customers who received the email, a portion made purchases using the coupon. The team found clear differences in response rates between RFM segments. For example:

Customers with high Recency, Frequency, and Monetary scores were much more likely to use the coupon.

Customers with low Recency scores (i.e., those who hadn't shopped recently) were far less likely to respond, making them unprofitable to target.

Sephora also calculated the profitability of the campaign for each RFM segment. This included: The total revenue generated from each segment during the campaign and the total cost of sending emails to each segment (including the marginal cost of unsubscribes).

By comparing the revenue and costs, the team was able to calculate the return on marketing expenditure for each segment.

In this assignment, you are tasked with using RFM analysis to determine segments of the customers to be targeted and to calculate the economic analysis of this segmentation. To calculate profit and return on marketing expenditures, assume that Sephora's cost is 76% of the revenue (normal price paid by the customer). This implies that the usual gross margin for the company is 24%.

Q1) What percentage of customers in the data have used the offer? Among those who have used the offer, how much they spent in the data?

Q2) Assuming a 5-quantile method for R, F, and M (i.e., a total of 125 buckets), calculate the RFM index for each customer. Subsequently, plot all the bar chart for the case of R, F, and M separately as we discussed in the class. Explain your findings by looking at each of these graphs.

Q3) Using the same indexes you calculated in the previous question, plot R, F, M graphs with respect to the money they spent on the offer only for customers who have responded yes to the offers. What do you understand from these graphs compared to the previous question? Explain

Q4) Calculate the average revenue, cost and profit the company can make from each customer who has used the offer.

Q5) Based on the previous question, calculate the break-even response rate. Hint: Use the cost of annoying the customer in this question.

Q6) For this question, imagine that the customer does not use selective targeting and instead sends offer to everyone. What are the profit and return-on-marketing cost if the company gives the offer to everyone in the customer base? Assume the customer base for Sephora is 39,968,762 customers in the US.

Q7) Now consider the case of RFM where we only target people in the customer base if they are above the break-even response rate. How many customers to target? Calculate the profit and return-on-marketing cost. Compare the results to the previous question and explain any noticeable differences in the results.

Q8) Using a 5-quantile for the response probability (as described previously in the class), calculate Lift and Gain for RFM. Plot both Lift and Gain graphs. What do you understand? You are allowed to use Excel for the last part to calculate Lift and Gain calculations and to produce graphs.

Q9) What are some shortcomings, limitations, or caveats for the profitability estimates that you conducted? Were there any inherent assumptions in this analysis that you find implausible?

Good luck 😊

Data disclaimer: Cases are not intended to serve as endorsements, sources of primary data, or illustrations of effective or ineffective management. Many elements of this case and future cases have been disguised or fictionalized by the instructor for pedagogical reasons and/or to assure confidentiality.

Instructions for the assignment:

You are required to do every part of the numerical analysis in R (unless specified in the question) and to attach the code at the end of each question or combine them and put them at the end of the document. It is always good practice to write all the middle-steps so that a wrong answer can get partial credit based on the resemblance of the approach taken to the correct approach and the efforts made to answer the questions.