**Subject: Approach for Testing Hypothesis on Customer Churn and Discount Strategy**

Respected AD’s,

I hope this email finds you well. After reviewing the client's situation, I've outlined an approach to address their problem and test the specific hypothesis regarding the impact of discounts on customer churn. Below are the steps and considerations for this data science task.

Situation Overview

The client is experiencing a churn rate of approximately 9.7%. To mitigate this, we hypothesize that offering a 20% discount to specific customer segments can effectively reduce churn rates, especially among highly profitable customers. The goal is to determine whether targeted discount strategies can reduce churn without significantly impacting overall revenue.

**Hypothesis:**

**Offering a 20% discount to highly profitable customers will result in a significant reduction in churn rates compared to not offering any discount, leading to higher overall profitability.**

Data Requirements

To test this hypothesis, we will need the following data from the client:

1. Customer Data:

* Customer segmentation (highly profitable, average to low-profit customers)
* Historical purchase data (at least 12 months)
* Churn status (binary indicator)

2. Transaction Data and Dates:

* Date history (Dates of all the activities performed by the customer)
* Power or Gas Consumption history
* Revenue generated per customer over the last 12 months

3. Discount Data:

- Acceptance of discount offers (Assuming all the customers agreed to stay for the discount)

Approach and Analytical Models

1. Data Preparation:

* Clean and preprocess the data to ensure consistency and accuracy.
* Segment customers into groups based on profitability and churn status.
* Create a dataset that includes customers who received a discount and those who did not.

2. Exploratory Data Analysis (EDA):

* Analyze churn rates across different customer segments.
* Examine the distribution of revenue and churn before and after applying discounts.
* Visualize the relationship between discounts and churn rates.

3. A/B Testing Framework:

* Control Group: Customers who did not receive any discount.
* Treatment Group: Customers who received a 20% discount.
* Randomly assign customers to control and treatment groups to ensure unbiased results.

4. Predictive Modeling:

* Build a predictive **Random Forest model** (as mentioned) to identify potential churners.
* Evaluate model performance using metrics such as accuracy, precision, recall, and F1-score.
* Use the model to simulate different discount strategies and predict their impact on churn rates.

5. Evaluation and Recommendations:

* Compare the financial impact of offering discounts to highly profitable customers versus all customers.
* Provide recommendations on the most cost-effective discount strategy based on the analysis.
* Suggest next steps for implementation and monitoring.

Summary

By following this approach, we aim to validate whether targeted discounts can effectively reduce churn rates and enhance overall profitability. The analysis will provide insights into the optimal discount strategy and help the client make data-driven decisions.

Please let me know if you need any further details or have any questions about the proposed approach.

Best regards,