

Dear Associate Director,

I hope this email finds you well. After our team meeting regarding the hypothesis that customer churn in the SME segment is driven by price sensitivities, I have been pondering the best approach to test this hypothesis. Based on our discussion, I have outlined a plan encompassing the necessary steps and analytical models to investigate this matter further. I have used the **ask, prepare, process, analyze, share, and act** process to solve this data science problem.

Ask:

1. Hypothesis: Churn in the SME segment is driven by price sensitivities.
2. Objective: Develop a predictive model to identify customers at risk of churning and determine the effectiveness of a 20% discount in retaining these customers.

Prepare:

- Historical customer data, including churn information, pricing details, customer demographics, usage patterns, and data on customer interactions, such as complaints, inquiries, and feedback, to capture their engagement levels.
- Collect data from PowerCo's CRM systems and other relevant sources, ensuring that we have a comprehensive dataset covering the relevant time period.

Process:

- Clean and validate the data, addressing missing values, outliers, and any data quality issues that may affect the analysis.
- Create additional variables that might be indicative of price sensitivity, such as price changes over time, price relative to competitors, or the proportion of a customer's budget spent on energy bills.

Analyze:

- Analyze the distribution of churners and non-churners in different price segments.
- Investigate correlations between churn and other relevant factors, such as contract duration, payment history, or customer satisfaction ratings.
- Employ techniques like correlation analysis, feature importance, or domain knowledge to identify the most relevant features for predicting churn.
- Focus on variables related to price sensitivity, such as historical price changes, pricing plans, or discounts received.
- Conduct statistical tests, such as t-tests or chi-square tests, to examine the relationship between churn and pricing variables, comparing churned and retained customers.
- Split the dataset into training and testing sets, ensuring that churned and retained customers are well-represented in both sets.

- Select appropriate machine learning algorithms for churn prediction, such as logistic regression, decision trees, random forests, or gradient boosting models.
- Train the predictive model on the training set and evaluate its performance using suitable metrics like accuracy, precision, recall, and F1 score.
- Validate the model using the testing set to assess its generalizability and reliability.

Share:

- Develop a user-friendly interface or integrate the predictive model into PowerCo's existing systems for easy monthly usage.

Act:

- Generate churn predictions for each SME customer on the 1st working day of every month.
- Offer a 20% discount to those customers who are predicted to churn and whose price sensitivity is considered influential, based on the model's interpretability.

By following this proposed approach, we can effectively test the hypothesis that customer churn in the SME segment is driven by price sensitivities.

Regards,
Rahul Gupta