

Dear Associate Director,

I'm writing to you to outline my approach to understanding PowerCo's churn problem and testing the hypothesis that churn is driven by customer price sensitivity.

Understanding the Problem:

1. **Client Context:** PowerCo is a major gas and electricity utility that supplies to corporate, SME, and residential customers. They are experiencing a significant churn problem (especially in the SME segment), driven by power-liberalization of the energy market in Europe.
2. **Client Hypothesis:** The client believes that it is possible to predict customers likely to churn using a predictive model and that churn is driven by customer price sensitivity. They suggest giving a 20% discount to customers who are likely to churn.

Testing the Hypothesis:

Formulating the Data Science Problem: We can formulate this problem as a binary classification problem, where the target variable is whether a customer churned or not. We can then use various machine learning models to predict churn and evaluate the importance of price sensitivity in the churn decision.

Data Needed from the Client:

Customer level data including:

- Contractual information (contract start date, end date, type of contract)
- Customer demographics (industry, company size, location)
- Consumption data (historical energy consumption, meter readings)
- Billing data (monthly bills, payment history)
- Churn information (date of churn, reason for churn if available)

Pricing data:

- Historical pricing information for different customer segments
- Information on competitor pricing

Analytical Models:

We can use various machine learning models for churn prediction such as:

- Logistic regression
- Decision trees
- Random forests

We can then compare the performance of these models and select the best performing one for further analysis.

To evaluate the impact of price, we'll analyze the weights assigned to different features by the model. These weights indicate how much emphasis the model places on each feature, including price.

Next Steps:

- Once we have the necessary data from the client, we can start cleaning and preprocessing the data.
- We can then train and evaluate the machine learning models for churn prediction.
- We will analyze the weights to assess the impact of price sensitivity on churn.
- Finally, we will communicate our findings and recommendations to the client.

I believe that this approach will allow us to effectively test the client's hypothesis and provide valuable insights.

Sincerely,