

Hi [AD],

In order to test the hypothesis of whether churn is driven by the customers' price sensitivity, we would need to model churn probabilities of customers, and derive the effect of prices on churn rates.

We would need the following data to be able to build the models.

1. Customer data - which should include characteristics of each client, for example, industry, historical electricity consumption, date joined as customer etc.
2. Churn data - which should indicate if customer has churned
3. Historical price data – which should indicate the prices the client charges to each customer for both electricity and gas at granular time intervals

Once we have the data, the work plan would be:

1. We need to define what price sensitivity is and calculate it
2. We need to prepare the data and engineer features
3. Then, we can test our hypothesis using a binary classification model (e.g. Logistic Regression, Random Forest, Gradient Boosted Machines to name a few)
4. We would choose a model from one of the tested algorithms based on the model complexity, the explainability, and the accuracy of the models.
5. With the trained model, we would be able to extrapolate the extent to which price sensitivity influences churn

Regards,

Chandra