Dear LDS

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.
And the second hypotheses to test is offering customers at a high propensity to churn a 20% discount might be effective or not.

In order to test the two hypotheses, we would need to model the 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 following 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, we would need to engineer features based on the data that we obtain, and build a binary classification model (e.g. Logistic Regression, Random Forest, Gradient Boosted Machines to name a few), picking the most appropriate model based on the trade-off between the complexity, the explainability, and the accuracy of the models.

Once we get the model, based on the model, we would able to understand the impact of price on churn rates and we can size the business impact of the second hypothesis.

Regards,
Aryan Singh