Hi LDS,

In order to test the hypothesis of whether churn is driven by the customer's price sensitivity, we would need to model churn probabilities of customers, and derive the effects of price on churn rates. For this we would need the following data to build the models.

Needed Data such as :

1. Customer Data - which should include characteristics of each customer, for example, industry type, historical electricity consumption, date joined as customer, electricity consumption purpose, etc.
2. Churn Data - which should indicate if a customer has churned or not.
3. Historical Price Data - which should indicate the prices the client charged to customers for both electricity and gas at regular time intervals.

Once we have this data, we would need to engineer features based on the data that we obtained, and build a Binary Classification Model (e.g. Logistic Regression, Random Forest, Gradient Boosting Machine learning algorithms to name of few), by picking up the most appropriate model based on the trade-off between the complexity, the explainability and the accuracy of models.

Based on the models picked, we would be able to understand the direction and magnitude of the impact of price on churn rates, as well as the relative importance of prices compared to other factors. Furthermore, the model would allow us to resize the business impact of the client's proposed discounting strategy.

Regards,
Shrikant Biyani