Dear Associate Director,

To test the hypothesis of whether the churn is driven by the customers’ price sensitivity, we would need to model churn probability of customer and derive the effect of prices on churn rate. For this purpose, we need following data sets:

* Customer data – should include characteristics of each client, for eg, industry, historical electric consumption, date joined as customer etc.
* Churn data – should include the details of churned customer.
* Historic price data – should include the prices of the client changes to each customer for both electricity and gas at granular time intervals.

Once we have these datasets, the work plan would be:

* We would need to define what price sensitivity is and calculate it
* We would need to engineer features based on the data that we obtain, and build a binary classification model (eg, Logistic Regression, Random Forest, Gradient Boosted Machine)
* Based on the complexity, explanability, and the accuracy of the models, we can pick the correct model.
* Once we pick our model, we would be able to dive deeper and understand the direction and magnitude of the impact of prices on churn rates, as well as relative importance of prices compared to other factors.
* Finally, the model would allow us to size the business impact of the client’s proposed discounting strategy.

Regards,

Sudeep Dhakal