Dear LDS,

The hypothesis provided by the customers states that churn is driven by customer price sensitivity.

To verify such hypothesis: We need following data

Data needed:

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 have to build the corresponding data pipeline and need to engineer features based on the data we had. Then, we would further need to build classification model picking the most appropriate model based on the tradeoff between the complexity, the explanability, and the accuracy of the models. Based on the model picked, we would be able to understand the direction and magnitude of the impact of prices on churn rates, as well as the relative importance of prices compared to other factors. Furthermore, the model would allow us to size the business impact of the client’s proposed discounting strategy.

Best,

Brian