Hi Estelle,

In order to test the hypothesis that churn is driven by the customers’ price sensitivity, we 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. Define what price sensitivity is and calculate it

2. Prepare the data and engineer features

3. Test our hypothesis using a binary classification model (e.g. Logistic Regression, Random Forest) to predict likelihood of churn

4. Select the best model based on accuracy metrics and other considerations such as explainability

5. With the trained model, extrapolate the extent to which price sensitivity influences churn

Regards,

Sushil Suhang