Task 1

Dear Sir/Madam,

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. We would need the following data to be able to build the models.

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 a 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. We would need to define what price sensitivity is and calculate it
2. 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),
3. The best model would be picked based on the tradeoff between the complexity, the explainability, and the accuracy of the models.
4. We would subsequently dive deeper into why and how price changes impact churn.
5. Last but not least, the model would allow us to size the business impact of the client's proposed discounting strategy.

Regards,

Tuan Nguyen

Task 4

Churn rate is 9.7% across 14606 customers

Predictive model is able to predict churn but the main driver is not customer price sensitivity. Yearly consumption, forecasted consumption and net margin are the 3 largest drivers.

Discount strategy of 20% is effective but only if targeted appropriately. Offer discount to only to high-value customers with high churn probability