Dear XX YY ZZ

Associate Director (AD)

BCG

As you know that we are currently working on PowerCo's project. The client is a major gas and electricity utility which supplies power to corporate, SME and residential customers. The power-liberalization of the energy market in Europe has led to significant customer churn, especially in the SME segment. Hence, they came to us to get help with the SME segment.

There are mainly 2 hypotheses to test:

1. The customer churn is driven by customer price sensitivity.
2. Offering customers at high propensity to churn a 20% discount might be effective.

In order to test the two hypotheses, we would need to model the 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 following models:

1. The SME customer data which contains the SME customer's characteristics of each client
2. Churn data which indicate if the SME customer has churned or not.
3. Price data which indicates the prices of different SME customers at different times.

When we get the data and after initial basic descriptive statistics, we need to do the Exploratory Data Analysis (EDA) to confirm if the churn is driven by customer price sensitivity. If it is driven by customer price sensitively, then we can build a binary model (e.g. Logistic Regression, Random Forest, Gradient Boosted Machines to name a few), to predict customers likely to churn. We can find the most appropriate model that fits best. Once we get the model, based on the model, we would be able to understand the impact of price on churn rates and we can size the business impact of the second hypothesis.

Regards,

Akshay Pandurang Paunikar

Intern – Data Science & Analytics.