**Hypothesis Formulation and Testing**

We can formulate our hypothesis around the price . We can check whether giving discount on price there is impact on the churn or not

So our hypothesis is:-

H0 = there is no impact of discount on customer churning or price is irrelevant of churning.

H1 = price is oppositely related to churn or null hypothesis is false

Steps we are going to use for the testing of these hypothesis are:

Step 1

We will take the sample size of 10000

Step 2

We will take the window size of 8 months and from these 8 months we will take 5 months to monitor our clients

Step 3

We will monitor 5000 samples from the sample of 10000 for 8 month without giving any discount and monitor 5000 sample for 8 months by giving discount. On these two sample set we will run our t-test independent test, if P-value is less than .05 then we will reject null hypothesis else we fail to reject the null hypothesis

**Data needed**

1. We need data consumption history of the customer , how consumer is charged by client throughout different period for both electricity and gas
2. Whether client is churned or not
3. Several other features like customer profile, industry client is related, whether it’s a domestic client or commercial client, date from which he initiated the service etc.

Based on above feature we will be doing feature engineering and we will be drawing meaning full insight through EDA. After that we will be applying binary class classifier for building the model (Logistic Regression, Random Forest, gradient boost, SVM, XGB classifier). On the basis of model building we will be able to predict whether client is going to churn or not