**Problem Statement**

There is need to identify customers that are likely to churn so as to offer relevant incentives to prevent this proactively.

**Hypothesis Generation – Factors that affect customer churn.**

1. Price Sensitivity

Customer churn is related to changes in pricing.

1. Customer Service

The perceived level of customer satisfaction is related to customer churn.

1. Product/ Service

Ability to meet the customer specific product requirement is related to customer churn.

**Data Type and Sources**

Historical data of customers

1. Customer Details – Name, Age, Education, gender (Highest Education Level), Address, customer type (individual or corporate)

There is the assumption that the demography features of the customer might influence how sensitive they are to the factors affecting churn. For example, a more educated person might be more intolerant to poor service quality.

1. Service Details – price at sign-up, date of Sign-up, date of last recharge, price at last recharge, days since last recharge, status, product type(s)
2. Customer Satisfaction Surveys

**Modelling Approach**

The target label that will be predicted is whether a customer will churn or not. Thus we will need to build a binary classification model with two labels – Churn or No Churn. The data field stated above from historical records will be used by labelling the appropriately the churned customers and those that did not churn with within the period.

We would also be able to carry out exploratory data analysis on the price trend and amount of churn witnessed to test any correlation.