**Telco Customer Churn**

**Planning:**

1. The dataset consists of 21 columns which contain customers’ account information, demographic details, services the customers have signed up for and customers who have churned in last month.

**Churn** is the most important column, and all the analysis will be based on this column. It tells us about churn status of the customer. We will focus on the customers who have churned so that we can analyze their data and predict who else is going to churn and prevent it from happening.

Most relevant columns are –

* gender: Gender of customer, i.e., male or female.
* SeniorCitizen: Is customer a senior citizen or not (1 for yes, 0 for no)?
* Partner: Does the customer have any partner or not?
* Dependents: Does the customer have any dependents or not?
* tenure: How long the person has been a customer for (in months)?
* PhoneService: Has the customer opted for phone service?
* MultipleLines: Does the customer have multiple telephone lines?
* InternetService: Type of internet service opted by customer.
* OnlineSecurity: Has the customer opted for online security?
* OnlineBackup: Has the customer opted for online backup?
* DeviceProtection: Has the customer opted for device protection?
* TechSupport: Has the customer opted for technical support?
* StreamingTV: Whether the customer streams TV using internet service?
* StreamingMovies: Whether the customer streams movies using internet service?
* Contract: Type of contract the customer has purchased.
* PaymentMethod: Type of payment method used by customer.

All analysis will be done around these variables.

1. Tenure is in months and, MonthlyCharge and TotalCharge are in dollars($). Else have no units.
2. Assumptions:
   1. Whole dataset can be grouped based on following criteria:
      1. Types of services purchased.
      2. Contracts: Which type of contract is purchased by the customer – month to month, one year or two years.
      3. Payment method.

After grouping we can analyze and identify which group is contributing the most to churn rate and can try to upsell the better performing service or contract.

* 1. We can also categorize customers based on senior citizenship, partners and dependents and can try to sell category-specific plans.
  2. Tenure can be used as an indicator of loyalty of a customer. We should focus on loyal customers and can offer exciting deals to those who are likely to churn.

1. TotalCharge column has some missing entries but that is because the customers haven’t paid any bills yet. So, we can replace the missing values with 0.
2. Yes. The dataset consists of only one file which is in excel sheet format.
3. To begin with:
   1. First, the data needs to be loaded into the environment.
   2. Then it is examined to get a sense of variables and type of data present.
   3. Then, check for null or missing values is done. And it is to be decided whether the row needs to be deleted or the value needs to be replaced.