Telecom churn analysis - Requirements

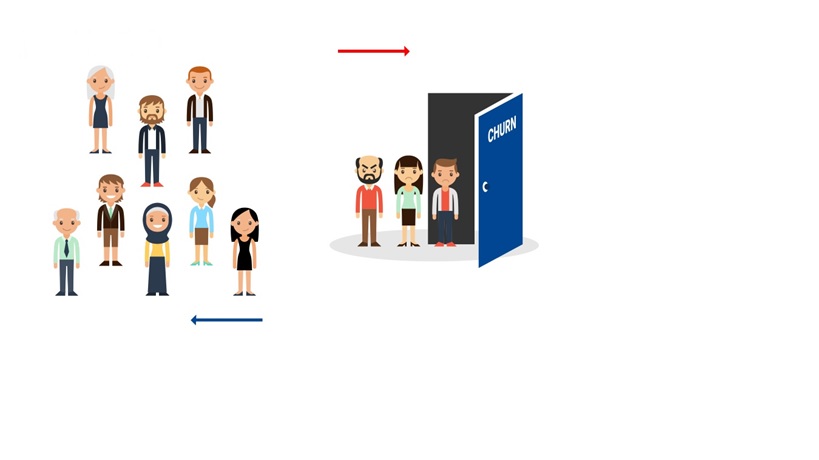
### **BUSNIESS OVERVEW**

For any business to run in a profitable way the most important factor is their customers. As the number of customers increase the revenue generated by the business will also increase.

As our economy is growing at such a fast pace, for any business there will be multiple competitors available in the market. This will eventually led the customers to jump from one company to another based on product quality, price, discounts, customer services etc. This is what is called as churn in any business.

Telecom industry experiences an average churn rate of 15-21%. Therefore to reduce churn the main aim in telecom companies is retention of high value customers. A churn can be defined in two ways:

* **Revenue based churn** – Customers who have not used any revenue generating facilities like Internet, SMS, calls, etc. are considered likely to churn customers.
* **Usage based churn** – Customers who have not used calls as in incoming or outgoing, internet services and other facilities are considered likely to churn.



Customers unusually don’t decide to choose other operators instantly instead this happens over a period of time. For prediction of churn three main phases of customer lifecycle is to be considered:

* **Good phase**: In this phase the customer is happy with the services provided by the operator and it as usual.
* **Action phase**: In this phase the customer starts to experience some difficulties with the operator i.e. offers from other competitors, poor service, unjust charges etc. This is a crucial phase, to take some preventive measures by the operator to retain the customers.
* **Churn phase**: In this phase the customer is said to have churned

Our provided dataset is of four months, the first two months are “Good phase”, third month is “Action Phase” and fourth month is “Churn phase”.

### **OBJECTIVE**

The main goal of this project is to do **Exploratory Data Analysis** on the given real time dataset and come up with insights which would result in reduction in churn of customers in the month of September.

### **DATASET**

* The dataset provided for analysis is of four months – June, July, August & September
* It consists of 226 features and 99,9999 rows
* Size of csv file is 76 MB