(i)Analytics Process:

Initial step includes basic analysis of data, sorting the data with customer numbers and the 12 months of time period.The data given is modified with proper column names, this is done in Excel.

**Exploration of Data**

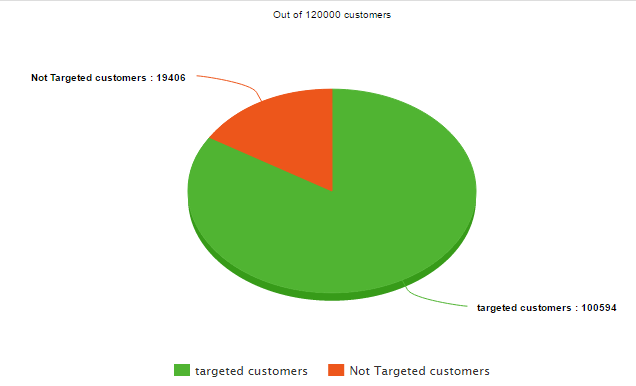
***Number of rows in data :120000***

***Number of customers: 10000***

***Number of months: 12***

Hence total data is of 10000 customers collected over a period of 12 months =120000 entries.

* **Customers Targeted transactions through channels and outreaches over 12 months :** 100594
* **Untargeted customers transactions with no outreach and no outreach through channels. i.e Customers Not reached:** 19406
  + - **Customers with no products no outreaches and no channel:** 16730
    - **Customers who purchased products in spite of no outreaches and no channels:2627**



**Targeting customers transactions through channels:** 73397

* **Customers without products 28946**
* **Customers with 44451**

**Outreaches**

**There are total 8 outreaches targeting 90859**

* **Customers with products :55215**
* **Customers without products :35644**

(ii)

New columns added indicating the change in balance, change in product presence, change in normalized balance etc

Columns added

Products Computed : sum of all the products in that particular month for the customer

Channel outreaches : Sum of channel outreaches

Outreaches\_8 : sum of outreaches other than channels

C\_prod\_purchased: identifies the month in with product C purchased or removed (indicated by 1 and -1)

D\_prod\_purchased: identifies the month in with product D purchased or removed (indicated by 1 and -1)

E\_prod\_purchased: identifies the month in with product E purchased or removed (indicated by 1 and -1)

Similarly for F and G

A\_bal\_change: change in balance of A type account wrt previous month

B\_bal\_change: change in balance of b type account wrt previous month

Norm\_bal\_change: Change in normalized balance wrt previous month

New Data generated:

products\_purchased<-subset(data4[,], data4$`Prod\_inc$Prod\_changed`>0)

products\_purchased<-products\_purchased[order(products\_purchased$cust\_num,products\_purchased$month),]

From the products purchased data set, subsets of data containing the purchases of products c,d,e,f and g have been created.

We have created datasets where A type balance increased, similarly datasets for other types of balances

(iii) Considered models of neural network, logistic regression, decision tree

(iv) Primarily, depending on selection of variables in purchased scenarios , the dependence causing successful purchases have been studied. Depending on the models metrics such as entropy, error terms etc. models have been created only with the significant variables that impacted our purchased, and caused growth in the balances we have initially considered models for each products and studied them and later based on the results, we have compiled them to create final model.

We gave also considered descriptive analytics of which demographics are covered by which channels in case of purchases and bank balances.

Also visual Analytics for the same have been done, along with the visualizations of months with purchases.

(v)Assumptions: Outreaches have an impact for one month,

Our models chosen on the channels and outreaches that impacted the most.

Outreaches and channels have no negative impact; If a customer is decreasing his balance or removing a purchased product, it could be solely due to his financial circumstances.