data2 <- read.csv("C:/Users/Pavan/Desktop/wellsfargo2/data2.csv", header=TRUE)

sqldf("SELECT count(\*) from data2")

count(\*)

1 120000

install.packages("sqldf")

library(sqldf)

products\_computed <- (data2$C\_product\_presence + data2$D\_product\_presence + data2$E\_product\_presence + data2$F\_product\_presence + data2$G\_product\_presence)

outreaches\_8\_computed <- (data3$Type\_A\_customer\_outreaches + data3$Type\_B\_customer\_outreaches + data3$Type\_C\_customer\_outreaches + data3$Type\_D\_customer\_outreaches + data3$Type\_E\_customer\_outreaches + data3$Type\_F\_customer\_outreaches + data3$Type\_G\_customer\_outreaches + data3$Typ\_H\_customer\_outreaches )

channel\_outreaches\_computed <- (data3$Channel\_1\_outreach + data3$service\_contact\_using\_channel\_2 + data3$X3\_market\_outreach\_using\_channe + data3$X4\_market\_outreach\_using\_channel )

data3 <- cbind(data2, products\_computed)

data3 <- cbind(data3, channel\_outreaches\_computed)

data3 <- cbind(data3, outreaches\_8\_computed)

> sqldf("SELECT count(\*) FROM data3 ")

count(\*)

1 120000

**Number of customers:**

sqldf("SELECT count(distinct(cust\_num)) FROM data4 ")

count(distinct(cust\_num))

1 10000

**The database contains 10000 customers data for a period of 12 months**

**Customers Targeted through channels and outreaches**

> sqldf("SELECT count(\*) FROM data3 WHERE channel\_outreaches\_computed > '0' or outreaches\_8\_computed >'0' ")

count(\*)

1. 100594

Customers Not reached

> sqldf("SELECT count(\*) FROM data3 WHERE channel\_outreaches\_computed = '0' and outreaches\_8\_computed ='0' ")

count(\*)

1 19406

**Total:120000**

**Untargeted customers with no products, no outreach and no outreach through channels**

**Unexplored customers:**

sqldf("SELECT count(\*) FROM data3 WHERE products\_computed ='0'AND channel\_outreaches\_computed = '0' and outreaches\_8\_computed ='0' ")

count(\*)

1 16730

**Untargeted customers with products**

sqldf("SELECT count(\*) FROM data3 WHERE products\_computed >'0'AND channel\_outreaches\_computed = '0' and outreaches\_8\_computed ='0' ")

count(\*)

1 2676

**Total 19406 untargeted**

**Targeted customers 100594**

Targeted\_customers<-sqldf("SELECT \* FROM data3 WHERE channel\_outreaches\_computed > '0' or outreaches\_8\_computed >'0' ")

> sqldf("SELECT count(\*) FROM Targeted\_customers WHERE channel\_outreaches\_computed > '0' or outreaches\_8\_computed >'0' ")

count(\*)

1 100594

**Responding**

> sqldf("SELECT count(\*) FROM Targeted\_customers WHERE products\_computed =0 ")

count(\*)

1 42858

Unresponsive

> sqldf("SELECT count(\*) FROM Targeted\_customers WHERE products\_computed !='0' ")

count(\*)

1 57736

**Channels**

**Targeting customers**

> sqldf("SELECT count(\*) FROM data3 WHERE channel\_outreaches\_computed > '0' ")

count(\*)

1 73397

**Unresponsive**

sqldf("SELECT count(\*) FROM data3 WHERE products\_computed ='0' AND channel\_outreaches\_computed > '0' ")

count(\*)

1 28946

**Responsive**

> sqldf("SELECT count(\*) FROM data3 WHERE products\_computed !='0' AND channel\_outreaches\_computed > '0' ")

count(\*)

1 44451

**Outreaches**

**Targeting customers**

> sqldf("SELECT count(\*) FROM data3 WHERE outreaches\_8\_computed > '0' ")

count(\*)

1. 90859

Responding

> sqldf("SELECT count(\*) FROM data3 WHERE products\_computed !='0' AND outreaches\_8\_computed > '0' ")

count(\*)

1 55215

**Not responding**

> sqldf("SELECT count(\*) FROM data3 WHERE products\_computed ='0' AND outreaches\_8\_computed > '0' ")

count(\*)

1 35644

**cust\_1 <-subset(data4[,c("cust\_num","month","A\_type\_account\_balance","B\_type\_account\_balance","C\_type\_account\_balance","D\_type\_account\_balance","E\_type\_account\_balance", "products\_computed","outreaches\_8\_computed","channel\_outreaches\_computed")], cust\_num==434)**

which(diff(cust\_1$E\_type\_account\_balance)!=0,arr.ind = T)

[1] 6

subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==436)

[1] 4 4 4 3 3 3 4 4 4 4 3 3

> which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==436)) >0)

[1] 6

> which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==436)) <0)

[1] 3 10

**for (i in 435:436) print(which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==i)) >0))**

for (i in 434:436) { x<- subset(data4, data4$cust\_num==i)[print(which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==i)) >0)),] }

**customer with increased b type account**

**subset(data4, data4$cust\_num==436)[which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==436)) >0),]**

**for (i in 434:436) { print(subset(data4, data4$cust\_num==i)[c(print(which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==i)) >0))),]) }**

**for (i in 434:436) { x[i] <-print(subset(data4, data4$cust\_num==i)[c((which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==i)) >0)),(which(diff(subset(data4[,c("B\_type\_account\_balance")], data4$cust\_num==i)) >0))-1),]) }**

**Each customers demographic data**

**for (i in 1:10000) print(unique(subset(data4[,c("cust\_num","cust\_demographics\_ai","cust\_demographics\_aii")], cust\_num==i)))**

> nrow((subset(final2[,],final2$Prod\_inc.Prod\_changed ==0 & final2$a\_bal.A\_bal\_change >0 | final2$b\_bal.B\_bal\_change >0)))

[1] 17176

> nrow((subset(final2[,],final2$Prod\_inc.Prod\_changed ==0 & final2$a\_bal.A\_bal\_change >0 & final2$b\_bal.B\_bal\_change >0)))

[1] 700

> nrow((subset(final2[,], final2$a\_bal.A\_bal\_change >0 )))

[1] 13374

> nrow((subset(final2[,], final2$b\_bal.B\_bal\_change >0 )))

[1] 4753

|  |
| --- |
| > b\_bal\_inc<-(final2[,], final2$b\_bal.B\_bal\_change >0 )  Error: unexpected ',' in "b\_bal\_inc<-(final2[,],"  > b\_bal\_inc<-subset(final2[,], final2$b\_bal.B\_bal\_change >0 )  > View(b\_bal\_inc)  > View(b\_bal\_inc)  >  >  > a\_bal\_inc<-(final2[,], final2$a\_bal.a\_bal\_change >0 )  Error: unexpected ',' in "a\_bal\_inc<-(final2[,],"  > a\_bal\_inc<-subset(final2[,], final2$a\_bal.a\_bal\_change >0 )  >  >  >  >  >  >  >  >  >  >  >  > write.csv(a\_bal\_inc,"a\_bal\_inc.csv")  > write.csv(b\_bal\_inc,"b\_bal\_inc.csv")  >  >  > write.csv(prod\_purchased,"products\_purchased.csv") |
|  |
| |  | | --- | |  | |
| unique(data4[order(data4$cust\_demographics\_ai,data4$cust\_demographics\_aii),c(3:4)])  cust\_demographics\_ai cust\_demographics\_aii  2317 0 1  1489 0 2  1801 0 3  1609 0 4  6049 0 5  1 1 1  37 1 2  145 1 3  985 1 4  32653 1 5  85 2 1  109 2 2  253 2 3  49 2 4  820 2 5  133 3 1  195 3 2  236 3 3  212 3 4  433 3 5  241 4 1  313 4 2  169 4 3  517 4 4  157 4 5  181 5 1  73 5 2  534 5 3  25 5 4  13 5 5 |
|  |
| |  | | --- | | > | |

**Customers in demography 0,1**

**data4[ which( data4$cust\_demographics\_ai ==0 & data4$cust\_demographics\_aii ==1 ) , c(1,3,4) ]**

|  |
| --- |
| for (i in 0:5) for(j in 1:5) print(count(data4[ which( data4$cust\_demographics\_ai ==i & data4$cust\_demographics\_aii ==j ) , c(3,4) ] ))  cust\_demographics\_ai cust\_demographics\_aii freq  1 0 1 2425  cust\_demographics\_ai cust\_demographics\_aii freq  1 0 2 923  cust\_demographics\_ai cust\_demographics\_aii freq  1 0 3 431  cust\_demographics\_ai cust\_demographics\_aii freq  1 0 4 191  cust\_demographics\_ai cust\_demographics\_aii freq  1 0 5 410  cust\_demographics\_ai cust\_demographics\_aii freq  1 1 1 7715  cust\_demographics\_ai cust\_demographics\_aii freq  1 1 2 6351  cust\_demographics\_ai cust\_demographics\_aii freq  1 1 3 2837  cust\_demographics\_ai cust\_demographics\_aii freq  1 1 4 1260  cust\_demographics\_ai cust\_demographics\_aii freq  1 1 5 72  cust\_demographics\_ai cust\_demographics\_aii freq  1 2 1 5238  cust\_demographics\_ai cust\_demographics\_aii freq  1 2 2 7061  cust\_demographics\_ai cust\_demographics\_aii freq  1 2 3 7022  cust\_demographics\_ai cust\_demographics\_aii freq  1 2 4 3631  cust\_demographics\_ai cust\_demographics\_aii freq  1 2 5 742  cust\_demographics\_ai cust\_demographics\_aii freq  1 3 1 3667  cust\_demographics\_ai cust\_demographics\_aii freq  1 3 2 3934  cust\_demographics\_ai cust\_demographics\_aii freq  1 3 3 5874  cust\_demographics\_ai cust\_demographics\_aii freq  1 3 4 6923  cust\_demographics\_ai cust\_demographics\_aii freq  1 3 5 3200  cust\_demographics\_ai cust\_demographics\_aii freq  1 4 1 2955  cust\_demographics\_ai cust\_demographics\_aii freq  1 4 2 3395  cust\_demographics\_ai cust\_demographics\_aii freq  1 4 3 4280  cust\_demographics\_ai cust\_demographics\_aii freq  1 4 4 6576  cust\_demographics\_ai cust\_demographics\_aii freq  1 4 5 8118  cust\_demographics\_ai cust\_demographics\_aii freq  1 5 1 1945  cust\_demographics\_ai cust\_demographics\_aii freq  1 5 2 2356  cust\_demographics\_ai cust\_demographics\_aii freq  1 5 3 3564  cust\_demographics\_ai cust\_demographics\_aii freq  1 5 4 5437  cust\_demographics\_ai cust\_demographics\_aii freq  1 5 5 11467 |
|  |
| |  | | --- | |  | |