Wells Fargo Competition

Identify which outreach methods & channels yield the best product portfolio

Shishir Shirbhate

June 07, 2017

Contents

1. Load Libraries

```
suppressPackageStartupMessages({
library(data.table)
library(ggplot2)
library(dplyr)
library(Hmisc)
library(rpart)
library(caret)
library(e1071)
library(pscl)
library(reshape)
library(plm)
library(gplots)
library(fields)
library(arules)
library(arulesViz)
library(gridExtra)
library(MASS)
library(reshape2)
library(coefplot)
})
```

2. Data Insertion & summarization

```
file_path <- "C:/MS Material/Competitions/Wells Fargo/Campus Challenge.csv"

bank_data = read.table(file=file_path, header=TRUE, sep = ",")

head(bank_data)
```

```
##   cust_num month cust_demographics_ai cust_demographics_aii typeA_ct
## 1        1     1                    1                     1        1
## 2        1     2                    1                     1        1
## 3        1     3                    1                     1        1
## 4        1     4                    1                     1        1
## 5        1     5                    1                     1        1
## 6        1     6                    1                     1        1
##   typeB_ct typeC_flag typeD_flag typeE_flag typeF_flag typeG_flag
## 1        2          0          0          0          0          0
## 2        2          0          0          0          0          0
## 3        2          0          0          0          0          0
## 4        2          0          0          0          0          0
## 5        2          0          0          0          0          0
## 6        2          0          0          0          0          0
##   typeA_bal_cat typeB_bal_cat typeC_bal_cat typeD_bal_cat typeE_bal_cat
## 1             3             1             0             0             0
## 2             2             1             0             0             0
## 3             2             1             0             0             0
```

```
## 4             1             1             0             0             0
## 5             1             1             0             0             0
## 6             1             1             0             0             0
##   normal_tot_bal cust_outreach_ai cust_outreach_aii cust_outreach_aiii
## 1     -0.3739689                0                 2                  1
## 2     -0.3843220                0                 0                  0
## 3     -0.3910417                0                 0                  0
## 4     -0.3943622                0                 0                  0
## 5     -0.3976828                0                 0                  0
## 6     -0.3986766                0                 0                  0
##   cust_outreach_aiv cust_outreach_av cust_outreach_avi cust_outreach_avii
## 1                 0                0                 4                  0
## 2                 0                0                 4                  0
## 3                 0                0                 3                  0
## 4                 0                0                 1                  0
## 5                 0                0                 2                  0
## 6                 0                0                 1                  0
##   cust_outreach_aviii wf_outreach_flag_chan_i wf_outreach_flag_chan_ii
## 1                   0                       0                        0
## 2                   0                       0                        0
## 3                   0                       0                        1
## 4                   0                       0                        1
## 5                   0                       0                        0
## 6                   0                       0                        0
##   wf_outreach_flag_chan_iii wf_outreach_flag_chan_iv
## 1                         0                       0
## 2                         0                       0
## 3                         0                       0
## 4                         0                       0
## 5                         0                       0
## 6                         0                       0
```

```
summary(bank_data)
```

```
##     cust_num         month       cust_demographics_ai
##  Min.   :    1   Min.   : 1.00   Min.   :0.000
##  1st Qu.: 2501   1st Qu.: 3.75   1st Qu.:2.000
##  Median : 5000   Median : 6.50   Median :3.000
##  Mean   : 5000   Mean   : 6.50   Mean   :3.013
##  3rd Qu.: 7500   3rd Qu.: 9.25   3rd Qu.:4.000
##  Max.   :10000   Max.   :12.00   Max.   :5.000
##  cust_demographics_aii    typeA_ct         typeB_ct        typeC_flag
##  Min.   :1.000         Min.   : 0.000   Min.   : 1.000   Min.   :0.0000
##  1st Qu.:2.000         1st Qu.: 1.000   1st Qu.: 1.000   1st Qu.:0.0000
##  Median :3.000         Median : 1.000   Median : 1.000   Median :0.0000
##  Mean   :3.001         Mean   : 1.366   Mean   : 1.628   Mean   :0.1375
##  3rd Qu.:4.000         3rd Qu.: 2.000   3rd Qu.: 2.000   3rd Qu.:0.0000
##  Max.   :5.000         Max.   :10.000   Max.   :15.000   Max.   :1.0000
##    typeD_flag        typeE_flag        typeF_flag       typeG_flag
##  Min.   :0.00000   Min.   :0.00000   Min.   :0.0000   Min.   :0.00000
##  1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.00000
##  Median :0.00000   Median :0.00000   Median :0.0000   Median :0.00000
##  Mean   :0.05641   Mean   :0.03971   Mean   :0.4045   Mean   :0.02186
##  3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:0.00000
```

```
##  Max.   :1.00000   Max.   :1.00000   Max.   :1.0000   Max.   :1.00000
##  typeA_bal_cat   typeB_bal_cat typeC_bal_cat    typeD_bal_cat
##  Min.   :0.000   Min.   :1     Min.   :0.0000   Min.   :0.0000
##  1st Qu.:1.000   1st Qu.:2     1st Qu.:0.0000   1st Qu.:0.0000
##  Median :3.000   Median :3     Median :0.0000   Median :0.0000
##  Mean   :2.632   Mean   :3     Mean   :0.4125   Mean   :0.1403
##  3rd Qu.:4.000   3rd Qu.:4     3rd Qu.:0.0000   3rd Qu.:0.0000
##  Max.   :5.000   Max.   :5     Max.   :5.0000   Max.   :5.0000
##  typeE_bal_cat    normal_tot_bal     cust_outreach_ai cust_outreach_aii
##  Min.   :0.0000   Min.   :-0.41062   Min.   : 0.000   Min.   :  0.000
##  1st Qu.:0.0000   1st Qu.:-0.36029   1st Qu.: 0.000   1st Qu.:  0.000
##  Median :0.0000   Median :-0.27181   Median : 0.000   Median :  0.000
##  Mean   :0.1191   Mean   : 0.00000   Mean   : 1.262   Mean   :  1.076
##  3rd Qu.:0.0000   3rd Qu.:-0.00913   3rd Qu.: 2.000   3rd Qu.:  1.000
##  Max.   :5.0000   Max.   :55.49929   Max.   :34.000   Max.   :103.000
##  cust_outreach_aiii cust_outreach_aiv cust_outreach_av  cust_outreach_avi
##  Min.   : 0.000     Min.   : 0.0000   Min.   :  0.000   Min.   :  0.000
##  1st Qu.: 0.000     1st Qu.: 0.0000   1st Qu.:  0.000   1st Qu.:  0.000
##  Median : 0.000     Median : 0.0000   Median :  0.000   Median :  1.000
##  Mean   : 1.054     Mean   : 0.1552   Mean   :  2.257   Mean   :  7.672
##  3rd Qu.: 1.000     3rd Qu.: 0.0000   3rd Qu.:  1.000   3rd Qu.:  9.000
##  Max.   :45.000     Max.   :25.0000   Max.   :213.000   Max.   :394.000
##  cust_outreach_avii cust_outreach_aviii wf_outreach_flag_chan_i
##  Min.   : 0.0000    Min.   : 0.0000     Min.   :0.0000
##  1st Qu.: 0.0000    1st Qu.: 0.0000     1st Qu.:0.0000
##  Median : 0.0000    Median : 0.0000     Median :0.0000
##  Mean   : 0.6282    Mean   : 0.2046     Mean   :0.2606
##  3rd Qu.: 0.0000    3rd Qu.: 0.0000     3rd Qu.:1.0000
##  Max.   :58.0000    Max.   :51.0000     Max.   :1.0000
##  wf_outreach_flag_chan_ii wf_outreach_flag_chan_iii
##  Min.   :0.0000           Min.   :0
##  1st Qu.:0.0000           1st Qu.:0
##  Median :1.0000           Median :0
##  Mean   :0.5469           Mean   :0
##  3rd Qu.:1.0000           3rd Qu.:0
##  Max.   :1.0000           Max.   :0
##  wf_outreach_flag_chan_iv
##  Min.   :0.0000
##  1st Qu.:0.0000
##  Median :0.0000
##  Mean   :0.1477
##  3rd Qu.:0.0000
##  Max.   :1.0000
```

```
describe(bank_data)
```

```
##                                                                    25%
##  3.480000e+06  1.737323e+02  1.058032e+03 -4.106217e-01  0.000000e+00
##           50%           75%
##  0.000000e+00  2.000000e+00  1.000000e+04 -3.479971e+06
```

3. Data Wrangling & Feature Engineering

```
bank_data$portfolio <- bank_data$typeA_ct + bank_data$typeB_ct +
  bank_data$typeC_flag + bank_data$typeD_flag + bank_data$typeE_flag + bank_data$typeF_flag + bank_data$

# Find change in balance from previous month
setDT(bank_data)[,balance_change:=bank_data$normal_tot_bal-shift(bank_data$normal_tot_bal,1,type="lag")]

bank_data$balance_change[which(bank_data$month == 1)] <- 0

# Find change in portfolio from previous month
setDT(bank_data)[,portfolio_change:=bank_data$portfolio-shift(bank_data$portfolio,1,type="lag")]

bank_data$portfolio_change[which(bank_data$month == 1)] <- 0

# Flag increase / decrease in balance
bank_data$balance_change_flag <- NA
bank_data$balance_change_flag[which(bank_data$balance_change < 0)] <- 0
bank_data$balance_change_flag[which(bank_data$balance_change > 0)] <- 1
bank_data$balance_change_flag[which(bank_data$balance_change == 0)] <- 0

# Flag increase / decrease in portfolio
bank_data$portfolio_change_flag <- NA
bank_data$portfolio_change_flag[which(bank_data$portfolio_change < 0)] <- 0
bank_data$portfolio_change_flag[which(bank_data$portfolio_change > 0)] <- 1
bank_data$portfolio_change_flag[which(bank_data$portfolio_change == 0)] <- 0

# Customer Outreaches changes and flags

setDT(bank_data)[,cust_outreach_ai_change:=bank_data$cust_outreach_ai -
                   shift(bank_data$cust_outreach_ai,1,type="lag")]
bank_data$cust_outreach_ai_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_ai_flag <- NA
bank_data$cust_outreach_ai_flag[which(bank_data$cust_outreach_ai_change < 0)] <- 0
bank_data$cust_outreach_ai_flag[which(bank_data$cust_outreach_ai_change > 0)] <- 1
bank_data$cust_outreach_ai_flag[which(bank_data$cust_outreach_ai_change == 0)] <- 0

setDT(bank_data)[,cust_outreach_aii_change:=bank_data$cust_outreach_aii -
                   shift(bank_data$cust_outreach_aii,1,type="lag")]
bank_data$cust_outreach_aii_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_aii_flag <- NA
bank_data$cust_outreach_aii_flag[which(bank_data$cust_outreach_aii_change < 0)] <- 0
bank_data$cust_outreach_aii_flag[which(bank_data$cust_outreach_aii_change > 0)] <- 1
bank_data$cust_outreach_aii_flag[which(bank_data$cust_outreach_aii_change == 0)] <- 0

setDT(bank_data)[,cust_outreach_aiii_change:=bank_data$cust_outreach_aiii -
                   shift(bank_data$cust_outreach_aiii,1,type="lag")]
bank_data$cust_outreach_aiii_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_aiii_flag <- NA
bank_data$cust_outreach_aiii_flag[which(bank_data$cust_outreach_aiii_change < 0)] <- 0
bank_data$cust_outreach_aiii_flag[which(bank_data$cust_outreach_aiii_change > 0)] <- 1
bank_data$cust_outreach_aiii_flag[which(bank_data$cust_outreach_aiii_change == 0)] <- 0
```

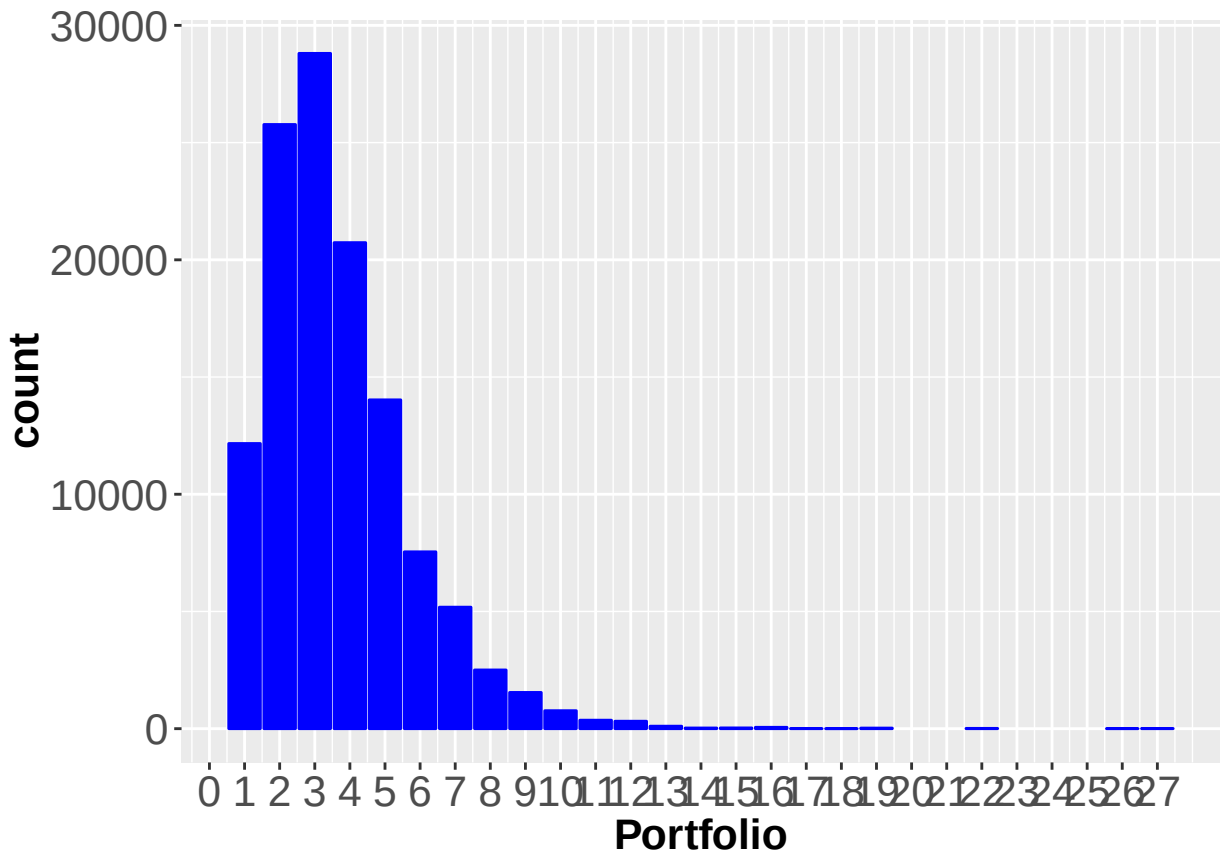
```
setDT(bank_data)[,cust_outreach_aiv_change:=bank_data$cust_outreach_aiv -
                   shift(bank_data$cust_outreach_aiv,1,type="lag")]
bank_data$cust_outreach_aiv_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_aiv_flag <- NA
bank_data$cust_outreach_aiv_flag[which(bank_data$cust_outreach_aiv_change < 0)] <- 0
bank_data$cust_outreach_aiv_flag[which(bank_data$cust_outreach_aiv_change > 0)] <- 1
bank_data$cust_outreach_aiv_flag[which(bank_data$cust_outreach_aiv_change == 0)] <- 0

setDT(bank_data)[,cust_outreach_av_change:=bank_data$cust_outreach_av -
                   shift(bank_data$cust_outreach_av,1,type="lag")]
bank_data$cust_outreach_av_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_av_flag <- NA
bank_data$cust_outreach_av_flag[which(bank_data$cust_outreach_av_change < 0)] <- 0
bank_data$cust_outreach_av_flag[which(bank_data$cust_outreach_av_change > 0)] <- 1
bank_data$cust_outreach_av_flag[which(bank_data$cust_outreach_av_change == 0)] <- 0

setDT(bank_data)[,cust_outreach_avi_change:=bank_data$cust_outreach_avi -
                   shift(bank_data$cust_outreach_avi,1,type="lag")]
bank_data$cust_outreach_avi_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_avi_flag <- NA
bank_data$cust_outreach_avi_flag[which(bank_data$cust_outreach_avi_change < 0)] <- 0
bank_data$cust_outreach_avi_flag[which(bank_data$cust_outreach_avi_change > 0)] <- 1
bank_data$cust_outreach_avi_flag[which(bank_data$cust_outreach_avi_change == 0)] <- 0

setDT(bank_data)[,cust_outreach_avii_change:=bank_data$cust_outreach_avii -
                   shift(bank_data$cust_outreach_avii,1,type="lag")]
bank_data$cust_outreach_avii_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_avii_flag <- NA
bank_data$cust_outreach_avii_flag[which(bank_data$cust_outreach_avii_change < 0)] <- 0
bank_data$cust_outreach_avii_flag[which(bank_data$cust_outreach_avii_change > 0)] <- 1
bank_data$cust_outreach_avii_flag[which(bank_data$cust_outreach_avii_change == 0)] <- 0

setDT(bank_data)[,cust_outreach_aviii_change:=bank_data$cust_outreach_aviii -
                   shift(bank_data$cust_outreach_aviii,1,type="lag")]
bank_data$cust_outreach_aviii_change[which(bank_data$month == 1)] <- 0
bank_data$cust_outreach_aviii_flag <- NA
bank_data$cust_outreach_aviii_flag[which(bank_data$cust_outreach_aviii_change < 0)] <- 0
bank_data$cust_outreach_aviii_flag[which(bank_data$cust_outreach_aviii_change > 0)] <- 1
bank_data$cust_outreach_aviii_flag[which(bank_data$cust_outreach_aviii_change == 0)] <- 0

bank_data$binary_channels <- paste(bank_data$wf_outreach_flag_chan_i,
                                   bank_data$wf_outreach_flag_chan_ii,
                                   bank_data$wf_outreach_flag_chan_iv,sep="")
bank_data$demo_combo <- paste(bank_data$cust_demographics_ai,bank_data$cust_demographics_aii,sep="")


bank_data$channel_combination <- strtoi(bank_data$binary_channels,base=2)
```

3. Change data types

```
bank_data$cust_demographics_ai <- as.factor(bank_data$cust_demographics_ai)
bank_data$cust_demographics_aii <- as.factor(bank_data$cust_demographics_aii)

bank_data$channel_combinationF <- as.factor(bank_data$channel_combination)

bank_data$monthF <- as.factor(bank_data$month)

bank_data$typeA_bal_cat <- as.factor(bank_data$typeA_bal_cat)
bank_data$typeB_bal_cat <- as.factor(bank_data$typeB_bal_cat)
bank_data$typeC_bal_cat <- as.factor(bank_data$typeC_bal_cat)
bank_data$typeD_bal_cat <- as.factor(bank_data$typeD_bal_cat)
bank_data$typeE_bal_cat <- as.factor(bank_data$typeE_bal_cat)
```

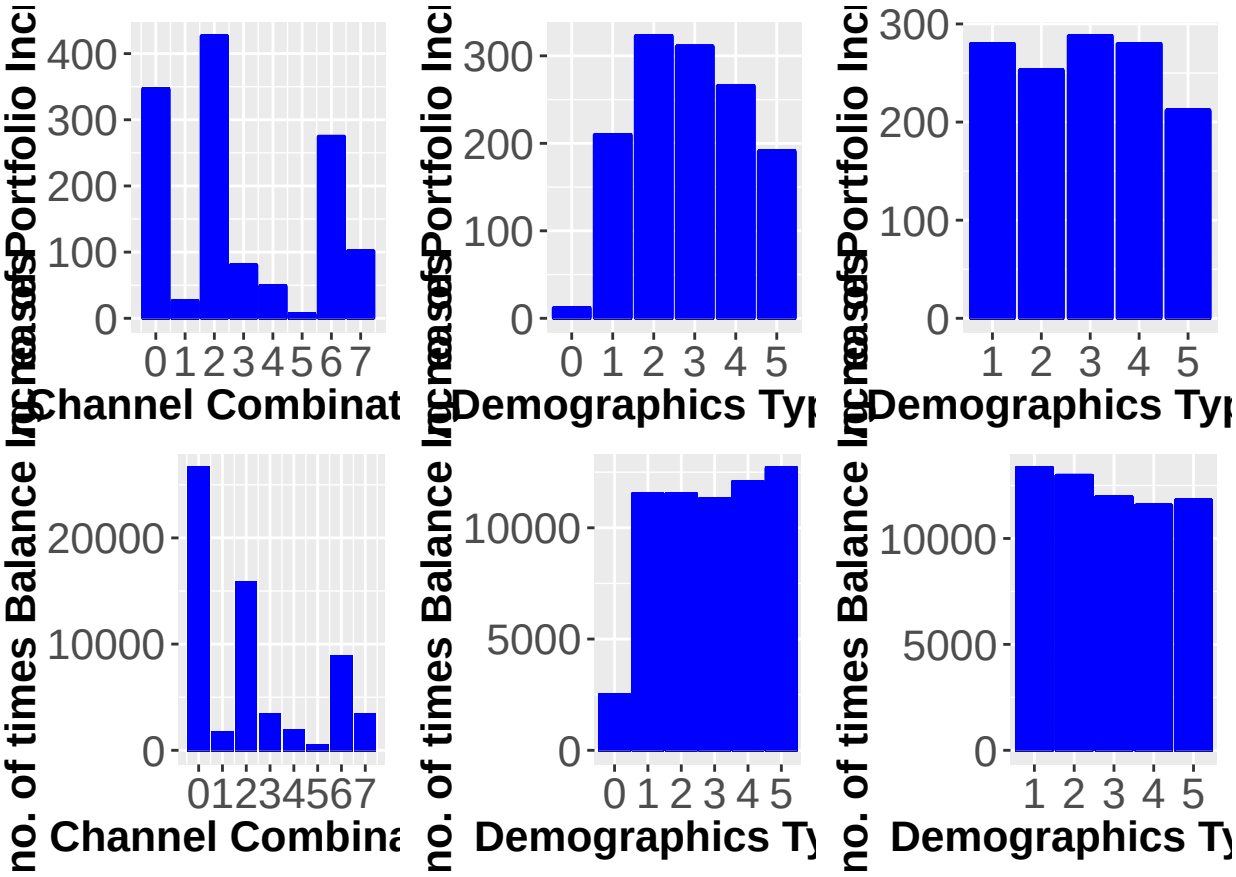
4. Data Exploration

```
ggplot(bank_data, aes(x=portfolio)) +
  geom_bar(color="blue",fill="blue") +
  xlab("Portfolio")+
  scale_x_continuous(breaks = seq(0,max(bank_data$portfolio),1)) +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5
```



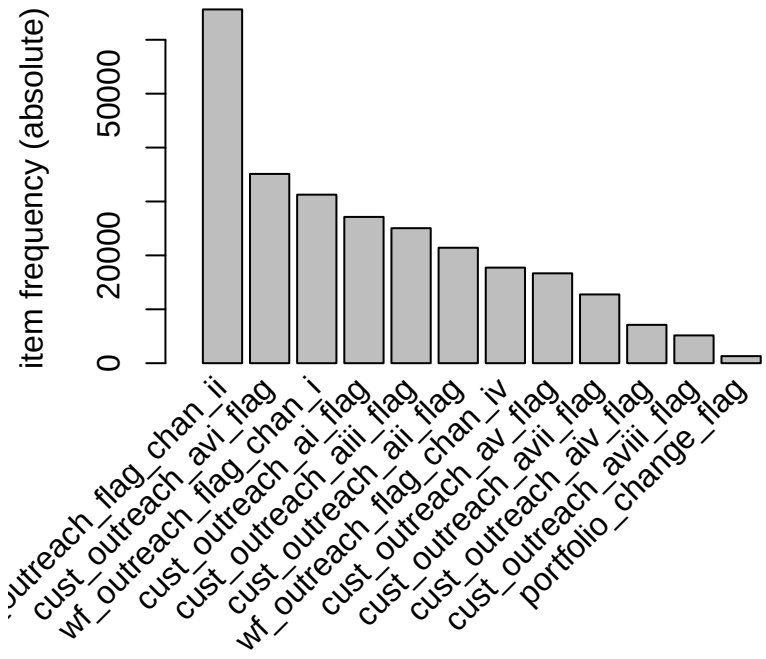

```
# Check if first bar can be made of different color
portfolio_channel <- ggplot(bank_data, aes(x=channel_combination,y=portfolio_change_flag)) +
  geom_bar(stat="identity",color="blue",fill="blue") +
  stat_summary(fun.y="mean", geom="bar") +
  xlab("Channel Combinations") +
  ylab("Avg no. of Portfolio Increases") +
  scale_x_continuous(breaks = seq(0,7,1)) +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5

balance_channel <- ggplot(bank_data, aes(x=channel_combination,y=balance_change_flag)) +
  geom_bar(stat="identity",color="blue",fill="white") +
  stat_summary(fun.y="mean", geom="bar") +
  xlab("Channel Combinations") +
  ylab("Avg no. of times Balance Increases") +
  scale_x_continuous(breaks = seq(0,7,1)) +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5

portfolio_demo_A <- ggplot(bank_data, aes(x=cust_demographics_ai,y=portfolio_change_flag)) +
  geom_bar(stat="identity",color="blue",fill="blue") +
  stat_summary(fun.y="mean", geom="bar") +
  xlab("Demographics Type A") +
  ylab("Avg no. of Portfolio Increases") +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5
```

```
balance_demo_A <- ggplot(bank_data, aes(x=cust_demographics_ai,y=balance_change_flag)) +
  geom_bar(stat="identity",color="blue",fill="white") +
  stat_summary(fun.y="mean", geom="bar") +
  xlab("Demographics Type A") +
  ylab("Avg no. of times Balance Increases") +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5

portfolio_demo_B <- ggplot(bank_data, aes(x=cust_demographics_aii,y=portfolio_change_flag)) +
  geom_bar(stat="identity",color="blue",fill="blue") +
  stat_summary(fun.y="mean", geom="bar") +
  xlab("Demographics Type B") +
  ylab("Avg no. of Portfolio Increases") +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5

balance_demo_B <- ggplot(bank_data, aes(x=cust_demographics_aii,y=balance_change_flag)) +
  geom_bar(stat="identity",color="blue",fill="white") +
  stat_summary(fun.y="mean", geom="bar") +
  xlab("Demographics Type B") +
  ylab("Avg no. of times Balance Increases") +
  theme(axis.title.x = element_text(face="bold", size=16),axis.text.x  = element_text(angle=0, vjust=0.5
  theme(axis.title.y = element_text(face="bold", size=16),axis.text.y  = element_text(angle=0, vjust=0.5

#multiple graphs on one panel
grid.arrange(portfolio_channel,portfolio_demo_A,portfolio_demo_B,balance_channel,balance_demo_A,balance_
```



5. Check Association between portfolio change and customer outreaches change

```
association_data <- subset(bank_data,select=c("portfolio_change_flag",
                                "cust_outreach_ai_flag",
                                "cust_outreach_aii_flag",
                                "cust_outreach_aiii_flag",
                                "cust_outreach_aiv_flag",
                                "cust_outreach_av_flag",
                                "cust_outreach_avi_flag",
                                "cust_outreach_avii_flag",
                                "cust_outreach_aviii_flag",
                                "wf_outreach_flag_chan_i",
                                "wf_outreach_flag_chan_ii",
                                "wf_outreach_flag_chan_iv"))

transaction <- as(data.matrix(association_data), "transactions")

itemFrequencyPlot(transaction,topN=20,type="absolute")
```



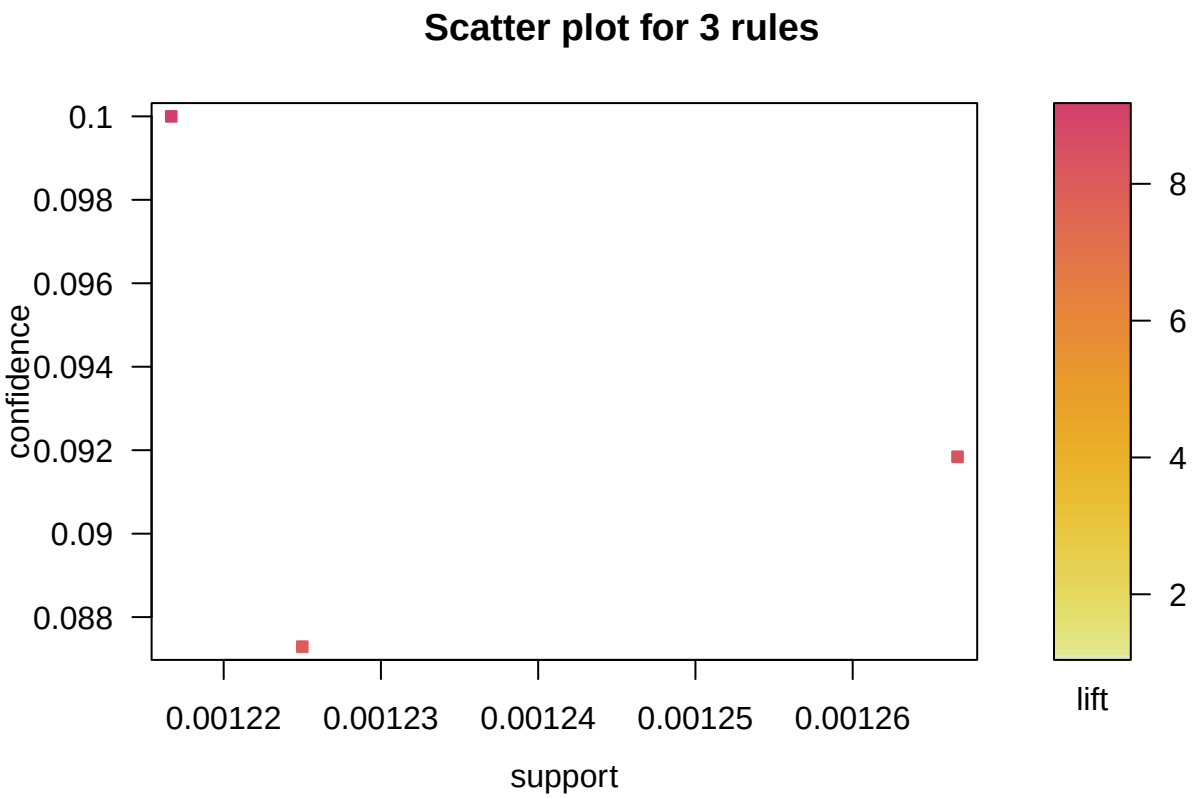

```
rules <- apriori(data = transaction, parameter = list(support = 0.001, confidence = 0.005))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport support minlen maxlen
##       0.005    0.1    1 none FALSE            TRUE   0.001      1     10
##  target   ext
##   rules FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 120
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[12 item(s), 120000 transaction(s)] done [0.02s].
## sorting and recoding items ... [12 item(s)] done [0.00s].
## creating transaction tree ... done [0.03s].
## checking subsets of size 1 2 3 4 5 6 7 8 done [0.00s].
## writing ... [4678 rule(s)] done [0.00s].
## creating S4 object  ... done [0.01s].
```

```r
rules <- sort(rules,by="confidence",decreasing=TRUE)

#inspect(rules[1:3])

rules_subset <- subset(rules, subset = rhs %pin% "portfolio_change_flag" & lift > 0)[1:3]

inspect(rules_subset)
```

```
##   lhs                            rhs                          support confidence     lift
## 1 {cust_outreach_ai_flag,
##    cust_outreach_aiii_flag,
##    cust_outreach_aiv_flag,
##    cust_outreach_avi_flag}   => {portfolio_change_flag} 0.001216667 0.10000000 9.139375
## 2 {cust_outreach_ai_flag,
##    cust_outreach_aiv_flag,
##    cust_outreach_avi_flag}   => {portfolio_change_flag} 0.001266667 0.09184290 8.393868
## 3 {cust_outreach_ai_flag,
##    cust_outreach_aiii_flag,
##    cust_outreach_av_flag,
##    cust_outreach_avi_flag,
##    wf_outreach_flag_chan_ii} => {portfolio_change_flag} 0.001225000 0.08729216 7.977958
```

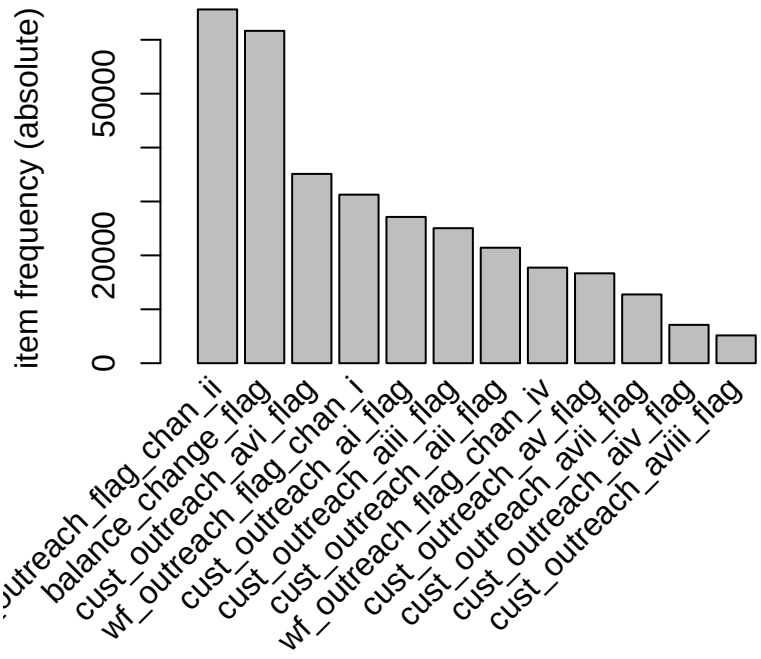
```r
plot(rules_subset)
```



```
# Confidence - This is the primary mode of sorting since we want to know out of all balance increases
# which combinations of outreaches & channels occur the most frequently

# Support - We have put a cut-off of 10% of actual proportion of portfolio changes

# Lift - Combinations of <> outreaches and channels are <>% more likely to have balance increases
```

6. Check Association between balance change and customer outreaches change

```
association_data <- subset(bank_data,select=c("balance_change_flag",
                                              "cust_outreach_ai_flag",
                                              "cust_outreach_aii_flag",
                                              "cust_outreach_aiii_flag",
                                              "cust_outreach_aiv_flag",
                                              "cust_outreach_av_flag",
                                              "cust_outreach_avi_flag",
                                              "cust_outreach_avii_flag",
                                              "cust_outreach_aviii_flag",
                                              "wf_outreach_flag_chan_i",
                                              "wf_outreach_flag_chan_ii",
                                              "wf_outreach_flag_chan_iv"))

transaction <- as(data.matrix(association_data), "transactions")

itemFrequencyPlot(transaction,topN=20,type="absolute")
```



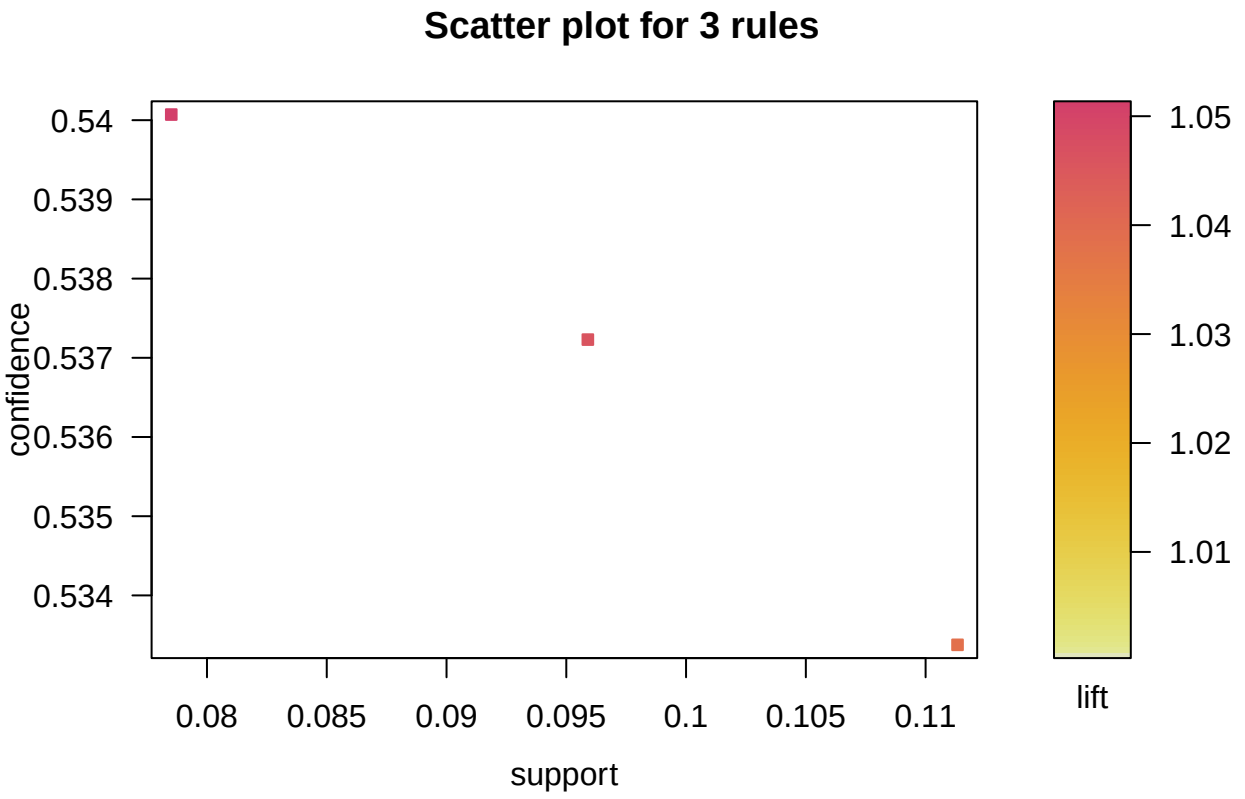

```
rules <- apriori(data = transaction, parameter = list(support = 0.05, confidence = 0.005))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport support minlen maxlen
##       0.005    0.1    1 none FALSE            TRUE    0.05      1     10
##  target   ext
##   rules FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 6000
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[12 item(s), 120000 transaction(s)] done [0.02s].
## sorting and recoding items ... [11 item(s)] done [0.00s].
## creating transaction tree ... done [0.04s].
## checking subsets of size 1 2 3 4 done [0.00s].
## writing ... [150 rule(s)] done [0.00s].
## creating S4 object  ... done [0.01s].
```

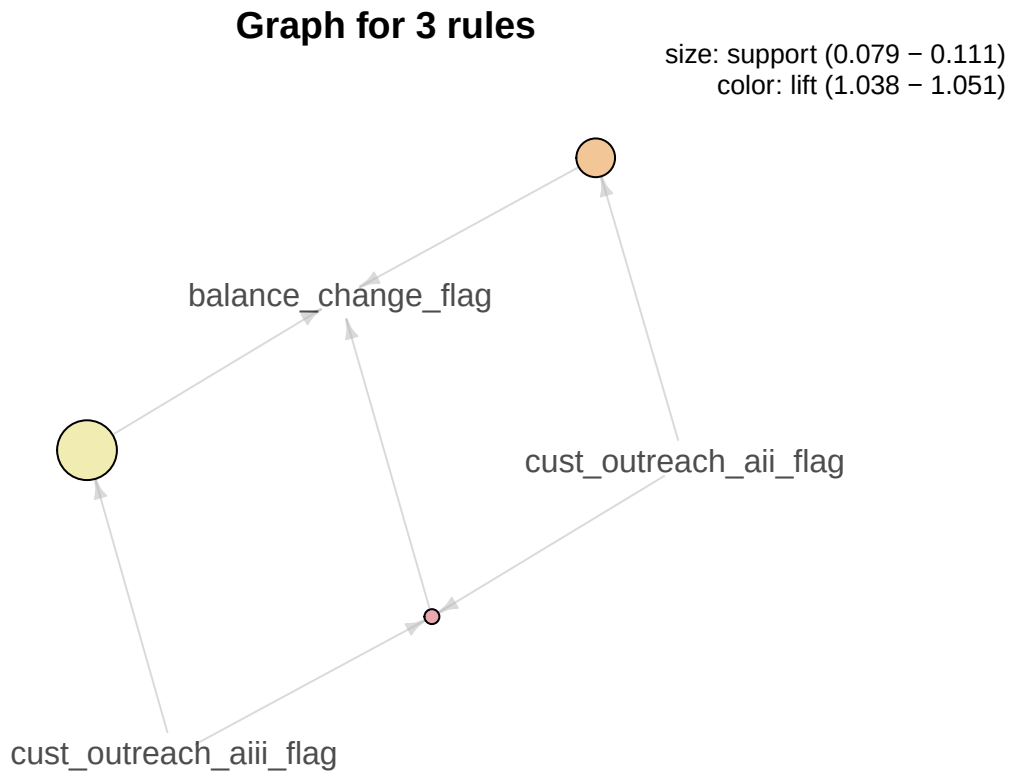
```
rules <- sort(rules,by="confidence",decreasing=TRUE)

rules_subset <- subset(rules, subset = rhs %pin% "balance_change_flag" & lift > 0)[1:3]

inspect(rules_subset)
```

```
##   lhs                           rhs                       support confidence     lift
## 1 {cust_outreach_aii_flag,
##    cust_outreach_aiii_flag} => {balance_change_flag} 0.07850833  0.5400711 1.051114
## 2 {cust_outreach_aii_flag}  => {balance_change_flag} 0.09590000  0.5372298 1.045584
## 3 {cust_outreach_aiii_flag} => {balance_change_flag} 0.11133333  0.5333759 1.038083
```

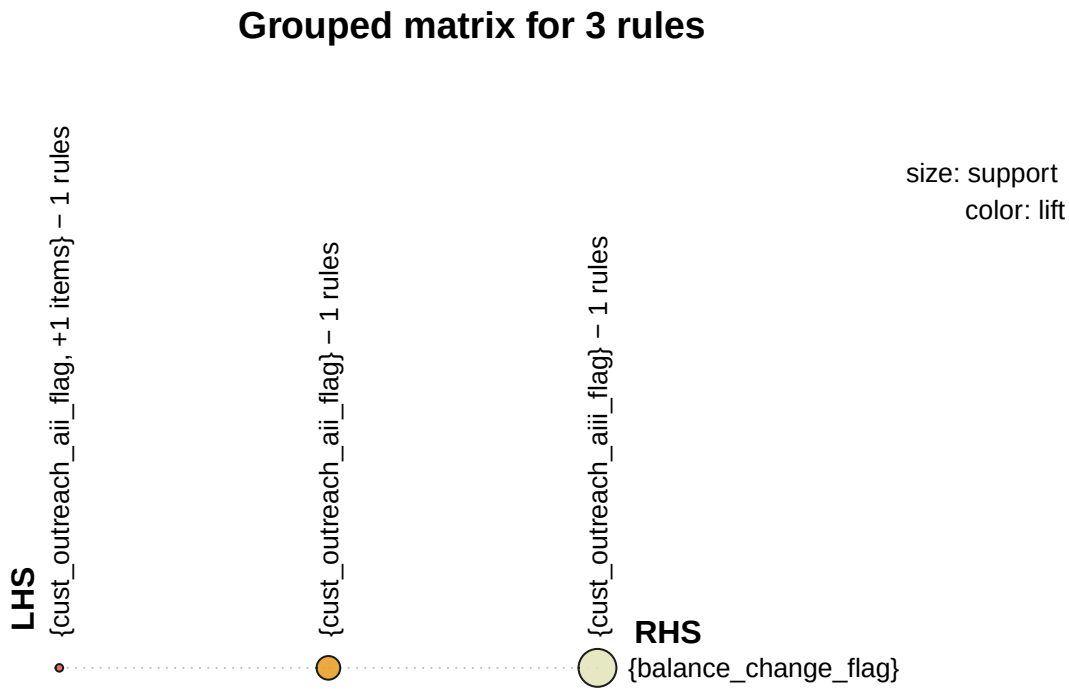
```
plot(rules_subset)
```




```
plot(rules_subset,method="graph")
```




```
plot(rules_subset,method="grouped")
```




```
# Confidence - This is the primary mode of sorting since we want to know out of all balance increases
# which combinations of outreaches & channels occur most frequently

# Support - We have put a cut-off of 10% of actual proportion of balance changes

# Lift - Combinations of <> outreaches and channels are <>% more likely to have balance increases
```

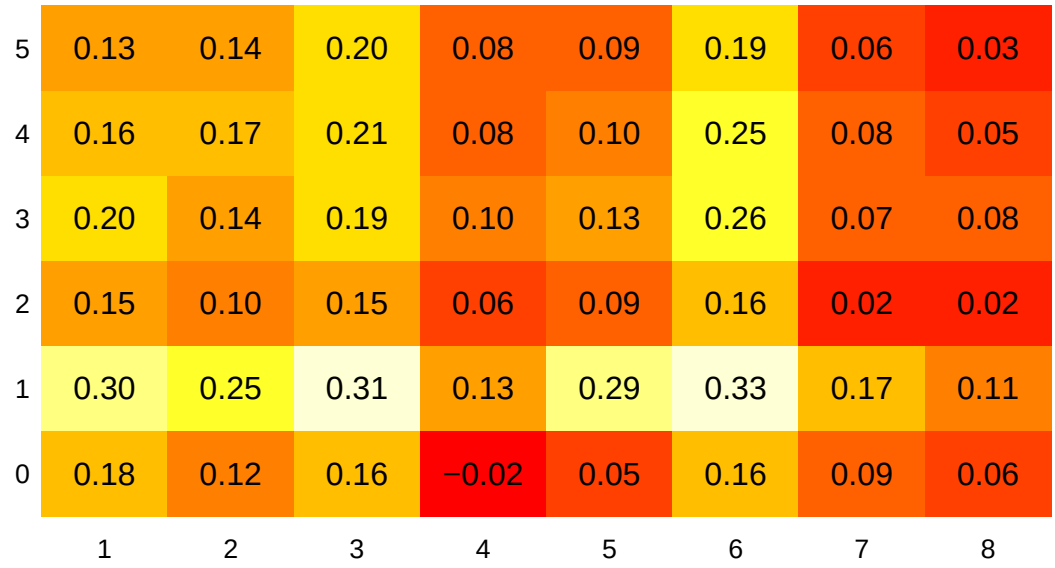
7. Finding correlations of various outreach programs with portfolio

```
demo1_portfolio <- data.frame(matrix(nrow=8,ncol=6))
colnames(demo1_portfolio) <- c("0","1","2","3","4","5")

for (i in 1:length(unique(bank_data$cust_demographics_ai)))
{
  bank_data1 <- bank_data[bank_data$cust_demographics_ai == i-1]
  for (j in 1:nrow(demo1_portfolio))
  {
    if (j == 1)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_ai),2)
    }else if(j==2)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aii),2)
```

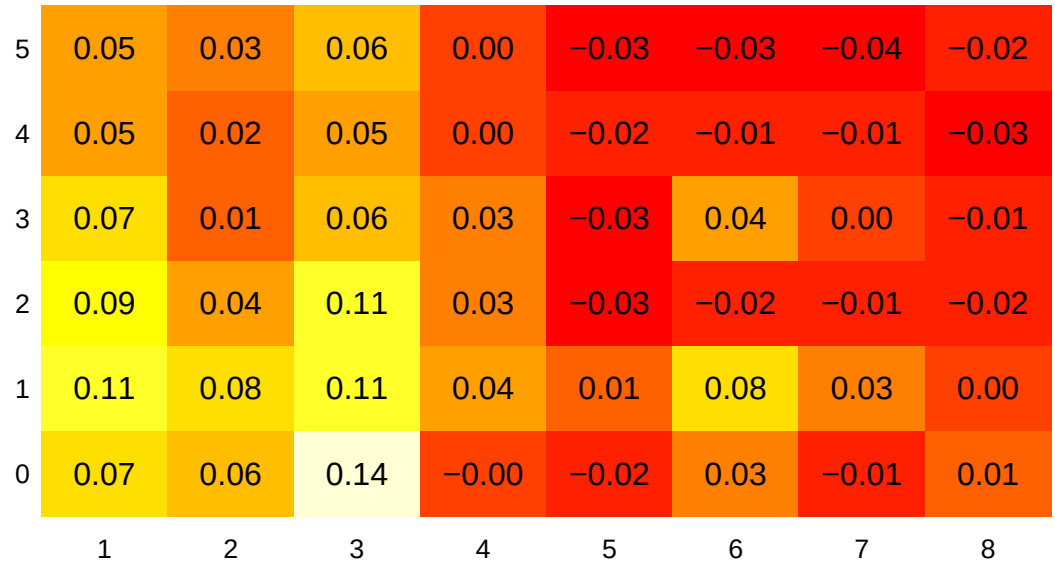
```
    }else if(j==3)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aiii),2)
    }else if(j==4)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aiv),2)
    }else if(j==5)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_av),2)
    }else if(j==6)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_avi),2)
    }else if(j==7)
    {
      demo1_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_avii),2)
    }else
    {
      demo1_portfolio[[i]][j]<- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aviii),2)
    }
  }
}

demo2_portfolio <- data.frame(matrix(nrow=8,ncol=5))
colnames(demo2_portfolio) <- c("1","2","3","4","5")

for (i in 1:length(unique(bank_data$cust_demographics_aii)))
{
  bank_data1 <- bank_data[bank_data$cust_demographics_aii == i]
  for (j in 1:nrow(demo2_portfolio))
  {
    if (j == 1)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_ai),2)
    }else if(j==2)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aii),2)
    }else if(j==3)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aiii),2)
    }else if(j==4)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aiv),2)
    }else if(j==5)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_av),2)
    }else if(j==6)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_avi),2)
    }else if(j==7)
    {
      demo2_portfolio[[i]][j] <- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_avii),2)
    }else
    {
```

```
      demo2_portfolio[[i]][j]<- round(cor(bank_data1$portfolio,bank_data1$cust_outreach_aviii),2)
    }
  }
}

demo1_bal <- data.frame(matrix(nrow=8,ncol=6))
colnames(demo1_bal) <- c("0","1","2","3","4","5")

for (i in 1:length(unique(bank_data$cust_demographics_ai)))
{
  bank_data1 <- bank_data[bank_data$cust_demographics_ai == i-1]
  for (j in 1:nrow(demo1_bal))
  {
    if (j == 1)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_ai),2)
    }else if(j==2)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aii),2)
    }else if(j==3)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aiii),2)
    }else if(j==4)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aiv),2)
    }else if(j==5)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_av),2)
    }else if(j==6)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_avi),2)
    }else if(j==7)
    {
      demo1_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_avii),2)
    }else
    {
      demo1_bal[[i]][j]<- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aviii),2)
    }
  }
}

demo2_bal <- data.frame(matrix(nrow=8,ncol=5))
colnames(demo2_bal) <- c("1","2","3","4","5")

for (i in 1:length(unique(bank_data$cust_demographics_aii)))
{
  bank_data1 <- bank_data[bank_data$cust_demographics_aii == i]
  for (j in 1:nrow(demo2_bal))
  {
    if (j == 1)
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_ai),2)
    }else if(j==2)
```

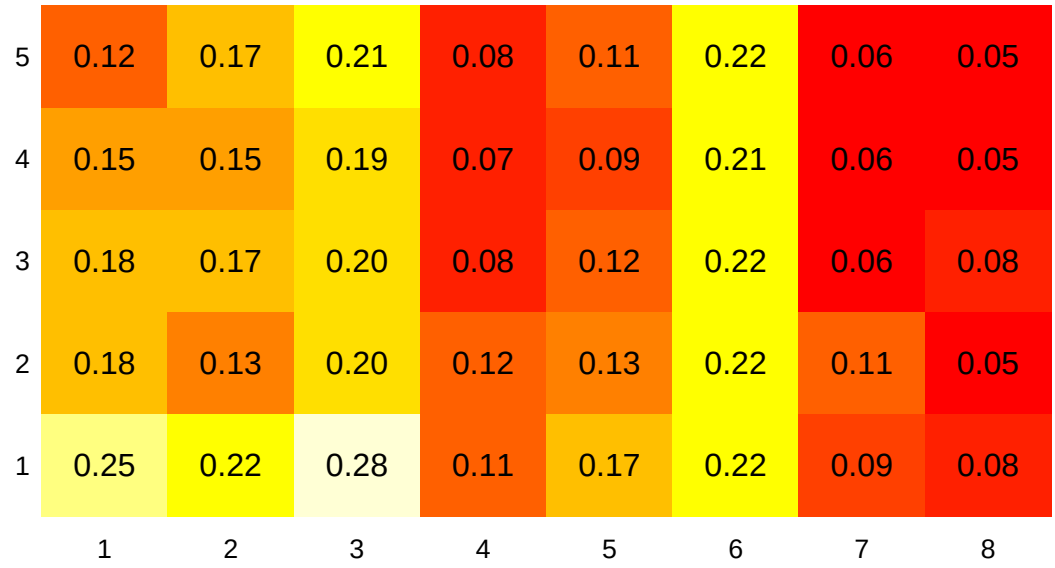
```
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aii),2)
    }else if(j==3)
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aiii),2)
    }else if(j==4)
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aiv),2)
    }else if(j==5)
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_av),2)
    }else if(j==6)
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_avi),2)
    }else if(j==7)
    {
      demo2_bal[[i]][j] <- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_avii),2)
    }else
    {
      demo2_bal[[i]][j]<- round(cor(bank_data1$normal_tot_bal,bank_data1$cust_outreach_aviii),2)
    }
  }
}

# Demographics A - Portfolio
image(as.matrix(demo1_portfolio),axes=F)
mtext(text=c(1:nrow(demo1_portfolio)), side=1, line=0.3, at=seq(0,1,1/(nrow(demo1_portfolio)-1)), las=1
mtext(text=c(0:(ncol(demo1_portfolio)-1)), side=2, line=0.3, at=seq(0,1,1/(ncol(demo1_portfolio)-1)), la
for (x in 1:nrow(demo1_portfolio))
  for (y in 1:ncol(demo1_portfolio))
    text((x-1)/(nrow(demo1_portfolio)-1), (y-1)/(ncol(demo1_portfolio)-1),
         sprintf("%0.2f", demo1_portfolio[x,y]))
```



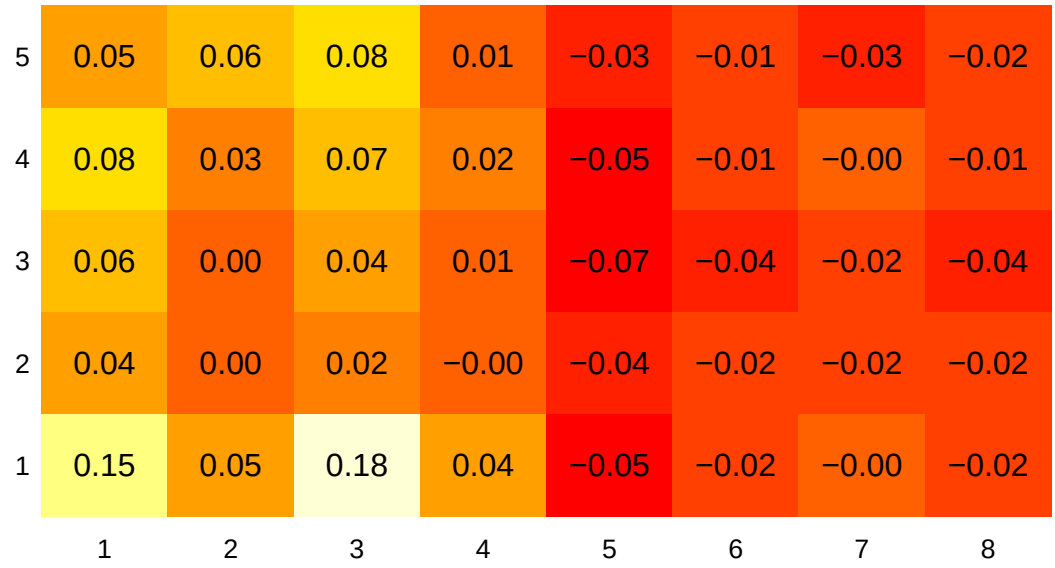

```
# Demographics A - Balance
image(as.matrix(demo1_bal),axes=F)
mtext(text=c(1:nrow(demo1_bal)), side=1, line=0.3, at=seq(0,1,1/(nrow(demo1_bal)-1)), las=1, cex=0.8)
mtext(text=c(0:(ncol(demo1_bal)-1)), side=2, line=0.3, at=seq(0,1,1/(ncol(demo1_bal)-1)), las=1, cex=0.8
for (x in 1:nrow(demo1_bal))
  for (y in 1:ncol(demo1_bal))
    text((x-1)/(nrow(demo1_bal)-1), (y-1)/(ncol(demo1_bal)-1),
         sprintf("%0.2f", demo1_bal[x,y]))
```




```
# Demographics B - Portfolio
image(as.matrix(demo2_portfolio),axes=F)
mtext(text=c(1:nrow(demo2_portfolio)), side=1, line=0.3, at=seq(0,1,1/(nrow(demo2_portfolio)-1)), las=1
mtext(text=c(1:ncol(demo2_portfolio)), side=2, line=0.3, at=seq(0,1,1/(ncol(demo2_portfolio)-1)), las=1
for (x in 1:nrow(demo2_portfolio))
  for (y in 1:ncol(demo2_portfolio))
    text((x-1)/(nrow(demo2_portfolio)-1), (y-1)/(ncol(demo2_portfolio)-1),
         sprintf("%0.2f", demo2_portfolio[x,y]))
```




```r
# Demographics B - Balance
image(as.matrix(demo2_bal),axes=F)
mtext(text=c(1:nrow(demo2_bal)), side=1, line=0.3, at=seq(0,1,1/(nrow(demo2_bal)-1)), las=1, cex=0.8)
mtext(text=c(1:ncol(demo2_bal)), side=2, line=0.3, at=seq(0,1,1/(ncol(demo2_bal)-1)), las=1, cex=0.8)
for (x in 1:nrow(demo2_bal))
  for (y in 1:ncol(demo2_bal))
    text((x-1)/(nrow(demo2_bal)-1), (y-1)/(ncol(demo2_bal)-1),
         sprintf("%0.2f", demo2_bal[x,y]))
```




```
# image.plot(as.matrix(demo2_bal),axes=F)
```

8. Modeling

```
#Create model to predict change in total balance -- find which methods increase balance!
balance_model<-lm(balance_change~month+cust_outreach_ai+
                    cust_outreach_aiii, data=bank_data)

#Create model to predict change in portfolio
portfolio_model<-lm(portfolio_change~ month + cust_demographics_ai +
                      cust_demographics_aii + typeA_ct + typeB_ct + typeA_bal_cat +
                      typeC_bal_cat + typeE_bal_cat +
                      typeF_flag + cust_outreach_ai + cust_outreach_aii +
                      cust_outreach_aiii + cust_outreach_aiv +
                      cust_outreach_avi + cust_outreach_avii +
                      wf_outreach_flag_chan_i + wf_outreach_flag_chan_ii,
                    data = bank_data)


##What drives growth in account balance?

bal_growth_model<-glm(balance_change_flag ~ cust_outreach_ai+cust_outreach_aiii+cust_outreach_aiv+
                        cust_outreach_avi+cust_outreach_avii+cust_outreach_aviii
```

```
                      +channel_combination, data=bank_data,
                      family="binomial")

sort(exp(coef(bal_growth_model)), decreasing = TRUE)
```

```
##         (Intercept) cust_outreach_aviii   cust_outreach_avi
##           1.3256371           1.0651777           0.9955342
##    cust_outreach_ai  cust_outreach_avii  cust_outreach_aiii
##           0.9917196           0.9883660           0.9691434
##   cust_outreach_aiv channel_combination
##           0.9685078           0.9367480
```

```
#growth in number of accounts?

portfolio_model<-glm(portfolio_change_flag~cust_outreach_ai+cust_outreach_aii+
                       cust_outreach_aiii+cust_outreach_aiv+cust_outreach_av+
                       cust_outreach_avi+cust_outreach_avii
                     +channel_combination+cust_demographics_ai, data=bank_data,
                     family="binomial")

sort(exp(coef(portfolio_model)), decreasing = TRUE)
```

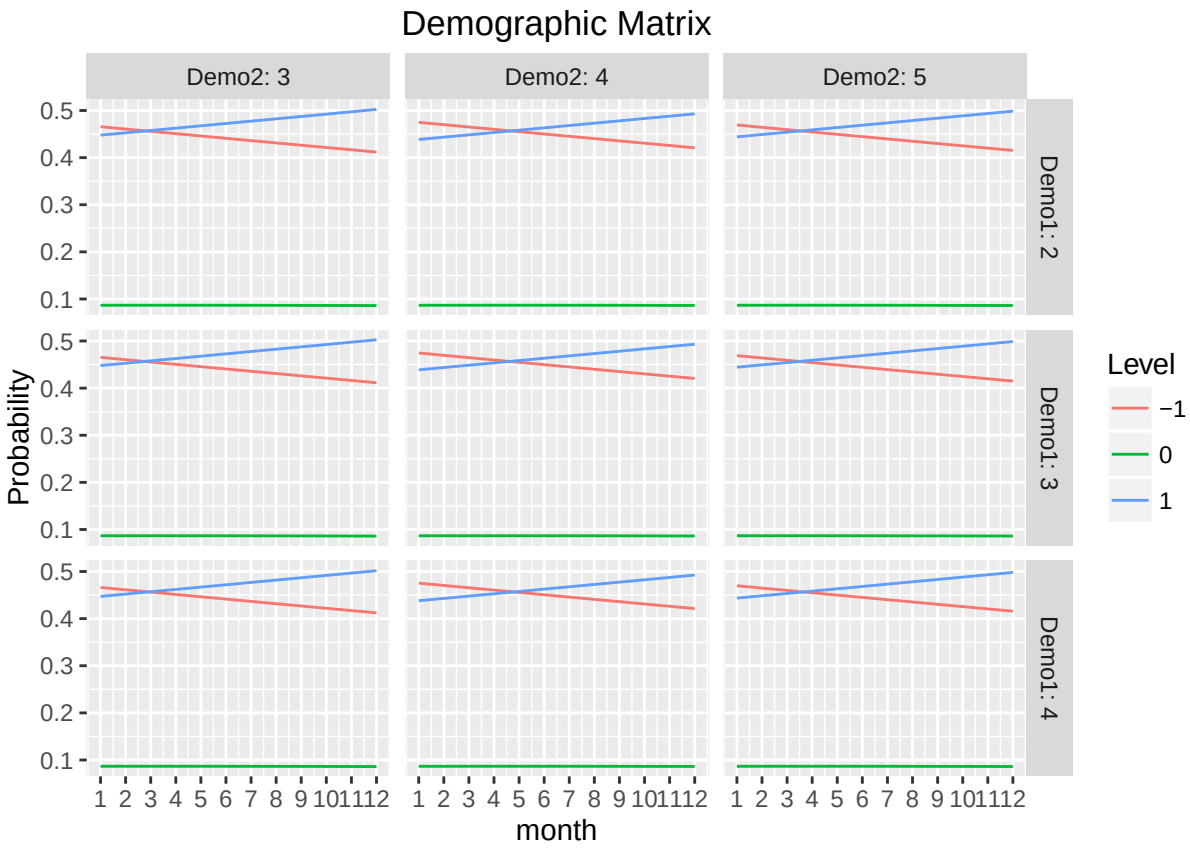
```
## cust_demographics_ai1 cust_demographics_ai2 cust_demographics_ai3
##           4.190320220           3.476470577           3.218382808
## cust_demographics_ai4 cust_demographics_ai5    cust_outreach_aiii
##           2.666504743           1.985192087           1.408439649
##      cust_outreach_ai   channel_combination     cust_outreach_avi
##           1.127367256           1.075904776           1.007861805
##      cust_outreach_av    cust_outreach_avii     cust_outreach_aiv
##           1.004718582           0.971276636           0.915585376
##     cust_outreach_aii           (Intercept)
##           0.811596382           0.001634574
```

```
## demographic regressions

bank_data$balance_change_flag_new <- NA
bank_data$balance_change_flag_new[which(bank_data$balance_change < 0)] <- -1
bank_data$balance_change_flag_new[which(bank_data$balance_change > 0)] <- 1
bank_data$balance_change_flag_new[which(bank_data$balance_change == 0)] <- 0

bank_data$portfolio_change_flag_new <- NA
bank_data$portfolio_change_flag_new[which(bank_data$portfolio_change < 0)] <- -1
bank_data$portfolio_change_flag_new[which(bank_data$portfolio_change > 0)] <- 1
bank_data$portfolio_change_flag_new[which(bank_data$portfolio_change == 0)] <- 0

bank_data$balance_change_flag_new<-as.factor(bank_data$balance_change_flag_new)
bank_data$portfolio_change_flag_new<-as.factor(bank_data$balance_change_flag_new)


bank_data$month<-as.integer(bank_data$month)

portfolio_order_model<-polr(portfolio_change_flag_new~cust_demographics_ai+cust_demographics_aii+month,
```

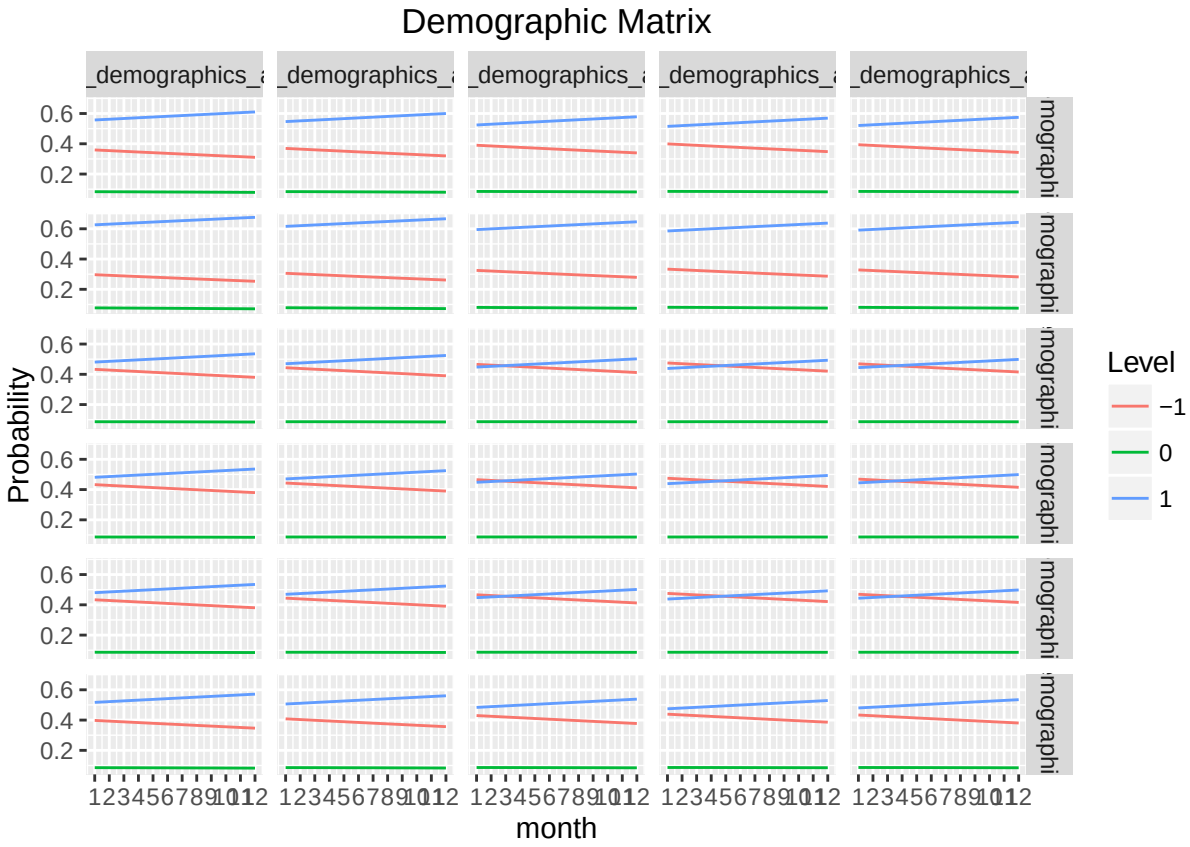
```
ctable<-coef(summary(portfolio_order_model))

p<-pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
ctable<-cbind(ctable, "p value" = p)
sort(exp(coef(portfolio_order_model)), decreasing = TRUE)
```

```
##  cust_demographics_ai1                 month cust_demographics_aii2
##              1.3298882             1.0200962              0.9575404
## cust_demographics_aii3 cust_demographics_aii5  cust_demographics_ai5
##              0.8756605             0.8630276              0.8497922
## cust_demographics_aii4  cust_demographics_ai3  cust_demographics_ai2
##              0.8436911             0.7360504              0.7351507
##  cust_demographics_ai4
##              0.7336007
```

```
newdat<-data.frame(cust_demographics_ai = as.factor(rep(0:5, 200)),
  cust_demographics_aii = as.factor(rep(1:5, 240)),
  month = rep(1:12, each = 100))

prediction<- cbind(newdat, predict(portfolio_order_model, newdat, type = "probs"))

lnewdat<-melt(prediction, id.vars = c("cust_demographics_ai", "cust_demographics_aii", "month"),
              variable.name = "Level", value.name = "Probability")

colnames(lnewdat) <- c("cust_demographics_ai","cust_demographics_aii","month","Level","Probability")

labels<-lnewdat

labels$Demo1<-labels$cust_demographics_ai
labels$Demo2<-labels$cust_demographics_aii

ggplot(subset(labels,Demo1 %in% c("2","3","4")& Demo2 %in% c("3","4","5")), aes(x=month, y=Probability,
  geom_line() + facet_grid(Demo1~Demo2, labeller = "label_both")+
  ggtitle("Demographic Matrix") + scale_x_continuous(breaks=seq(1, 12, 1))
```




```
ggplot(lnewdat, aes(x=month, y=Probability, color = Level))+
  geom_line() + facet_grid(cust_demographics_ai~cust_demographics_aii, labeller = "label_both")+
  ggtitle("Demographic Matrix") + scale_x_continuous(breaks=seq(1, 12, 1))
```




```
#Ordinal Logistic Regression
#Account A
A<-polr(typeA_bal_cat~month+cust_demographics_ai+cust_demographics_aii+
          cust_outreach_ai+cust_outreach_aii+cust_outreach_aiii+cust_outreach_aiv+
          cust_outreach_av+cust_outreach_avi+cust_outreach_avii+
          channel_combination, data=bank_data, Hess=TRUE)

ctableA<-coef(summary(A))

pc<-pnorm(abs(ctableA[, "t value"]), lower.tail=FALSE *2)
(ctableA <- cbind(ctableA, "p value" = pc))
```

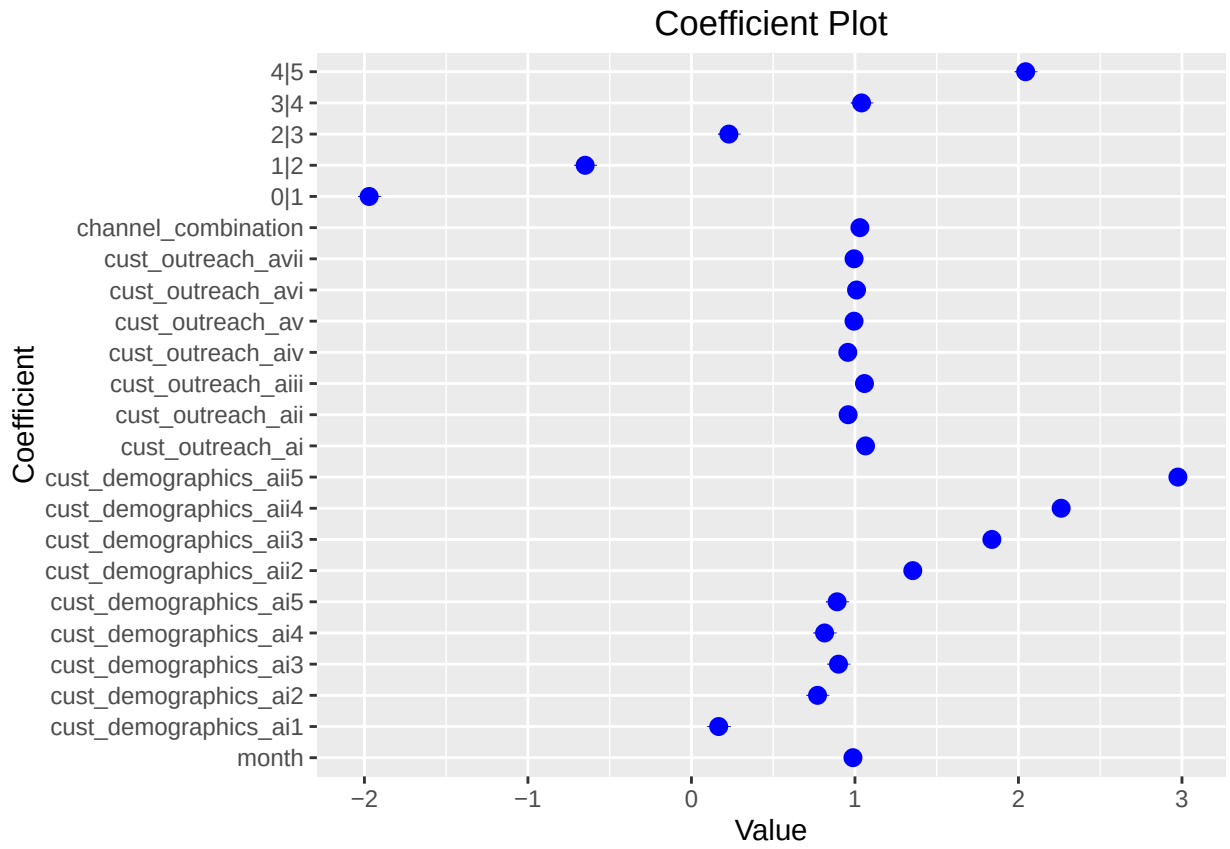
```
##                            Value   Std. Error    t value       p value
## month                 -0.012298382 0.0016580686  -7.417294  5.976882e-14
## cust_demographics_ai1 -1.791398367 0.0340907277 -52.547965  0.000000e+00
## cust_demographics_ai2 -0.259968270 0.0335747345  -7.742973  4.855925e-15
## cust_demographics_ai3 -0.106025169 0.0337879352  -3.137959  8.506422e-04
## cust_demographics_ai4 -0.205630034 0.0337956588  -6.084510  5.842394e-10
## cust_demographics_ai5 -0.115728309 0.0339213509  -3.411666  3.228362e-04
## cust_demographics_aii2  0.302842884 0.0166928412  18.142081  7.415331e-74
## cust_demographics_aii3  0.608396207 0.0169587335  35.875097 3.735749e-282
## cust_demographics_aii4  0.815822660 0.0175665717  46.441769  0.000000e+00
## cust_demographics_aii5  1.090281759 0.0189165946  57.636260  0.000000e+00
## cust_outreach_ai        0.062656131 0.0028769642  21.778557 1.852794e-105
## cust_outreach_aii      -0.042647700 0.0033700292 -12.654994  5.249184e-37
## cust_outreach_aiii      0.056485672 0.0053378110  10.582179  1.802377e-26
```

```
## cust_outreach_aiv      -0.044598126 0.0082376377  -5.413946  3.082537e-08
## cust_outreach_av       -0.005766822 0.0007525733  -7.662805  9.095783e-15
## cust_outreach_avi       0.009676780 0.0003733859  25.916300 2.181488e-148
## cust_outreach_avii     -0.005524252 0.0025730614  -2.146957  1.589836e-02
## channel_combination     0.029887629 0.0022166503  13.483241  9.814616e-42
## 0|1                    -1.971019519 0.0335732156 -58.708095  0.000000e+00
## 1|2                    -0.650031913 0.0329226987 -19.744187  4.500106e-87
## 2|3                     0.228901361 0.0328326689   6.971756  1.565044e-12
## 3|4                     1.040462251 0.0328913223  31.633336 6.427583e-220
## 4|5                     2.043801053 0.0332395972  61.486938  0.000000e+00
```

```
sort(exp(coef(A)), decreasing = TRUE)
```

```
## cust_demographics_aii5 cust_demographics_aii4 cust_demographics_aii3
##              2.9751122              2.2610350              1.8374821
## cust_demographics_aii2       cust_outreach_ai     cust_outreach_aiii
##              1.3537018              1.0646607              1.0581115
##    channel_combination      cust_outreach_avi     cust_outreach_avii
##              1.0303387              1.0097238              0.9944910
##       cust_outreach_av                  month      cust_outreach_aii
##              0.9942498              0.9877769              0.9582489
##      cust_outreach_aiv  cust_demographics_ai3  cust_demographics_ai5
##              0.9563817              0.8994020              0.8907172
##  cust_demographics_ai4  cust_demographics_ai2  cust_demographics_ai1
##              0.8141342              0.7710761              0.1667269
```

```
A$coefficients<-exp(coef(A))

coefplot(A, zeroType = 0)
```




```
#Account B
B<-polr(typeB_bal_cat~+cust_demographics_ai+cust_demographics_aii+
          cust_outreach_ai+cust_outreach_aii+cust_outreach_aiii+cust_outreach_aiv+
          cust_outreach_avi+cust_outreach_avii+cust_outreach_aviii+
          channel_combination, data=bank_data, Hess=TRUE)

ctableB<-coef(summary(B))

pcB<-pnorm(abs(ctableB[, "t value"]), lower.tail=FALSE *2)
(ctableB <- cbind(ctableB, "p value" = pcB))
```

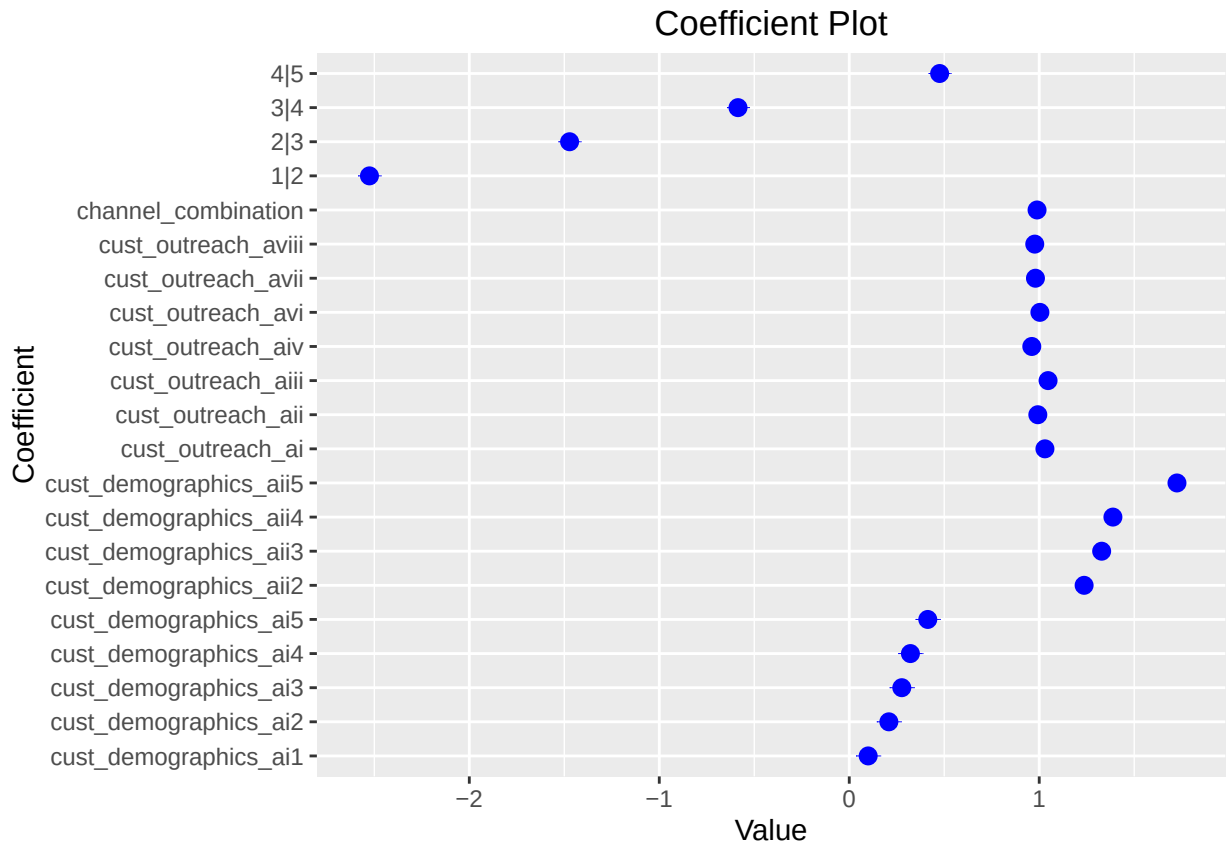
```
##                            Value   Std. Error    t value       p value
## cust_demographics_ai1  -2.307377031 0.0317415879 -72.692552  0.000000e+00
## cust_demographics_ai2  -1.567586990 0.0316747288 -49.490147  0.000000e+00
## cust_demographics_ai3  -1.285241358 0.0318509612 -40.351729  0.000000e+00
## cust_demographics_ai4  -1.133450491 0.0318185286 -35.622341 3.159303e-278
## cust_demographics_ai5  -0.883612612 0.0319223382 -27.680072 6.066265e-169
## cust_demographics_aii2  0.212119160 0.0163450490  12.977579  8.199147e-39
## cust_demographics_aii3  0.284169634 0.0168262380  16.888483  2.734559e-64
## cust_demographics_aii4  0.327750582 0.0173878522  18.849400  1.485969e-79
## cust_demographics_aii5  0.545206298 0.0187550263  29.069876 4.315215e-186
## cust_outreach_ai        0.029539861 0.0028226770  10.465193  6.241064e-26
## cust_outreach_aii      -0.007482683 0.0034138895  -2.191835  1.419570e-02
## cust_outreach_aiii      0.045417857 0.0053941967   8.419763  1.886237e-17
## cust_outreach_aiv      -0.039966195 0.0084667980  -4.720343  1.177234e-06
## cust_outreach_avi       0.003468738 0.0003463614  10.014794  6.561759e-24
```

```
## cust_outreach_avii     -0.020070062 0.0025728197  -7.800804  3.075699e-15
## cust_outreach_aviii    -0.023874513 0.0045165789  -5.285973  6.251924e-08
## channel_combination    -0.011541563 0.0022276747  -5.180991  1.103549e-07
## 1|2                    -2.525587218 0.0296541735 -85.168019  0.000000e+00
## 2|3                    -1.472105170 0.0291900286 -50.431782  0.000000e+00
## 3|4                    -0.585633663 0.0289578481 -20.223660  3.030773e-91
## 4|5                     0.475765174 0.0288881333  16.469225  3.052361e-61
```

```
sort(exp(coef(B)), decreasing = TRUE)
```

```
## cust_demographics_aii5 cust_demographics_aii4 cust_demographics_aii3
##             1.72496420             1.38784278             1.32865830
## cust_demographics_aii2    cust_outreach_aiii       cust_outreach_ai
##             1.23629519             1.04646504             1.02998049
##      cust_outreach_avi     cust_outreach_aii    channel_combination
##             1.00347476             0.99254524             0.98852479
##     cust_outreach_avii    cust_outreach_aviii      cust_outreach_aiv
##             0.98013000             0.97640823             0.96082192
##  cust_demographics_ai5  cust_demographics_ai4  cust_demographics_ai3
##             0.41328717             0.32192055             0.27658382
##  cust_demographics_ai2  cust_demographics_ai1
##             0.20854780             0.09952195
```

```
B$coefficients<-exp(coef(B))
coefplot(B, zeroType = 0)
```



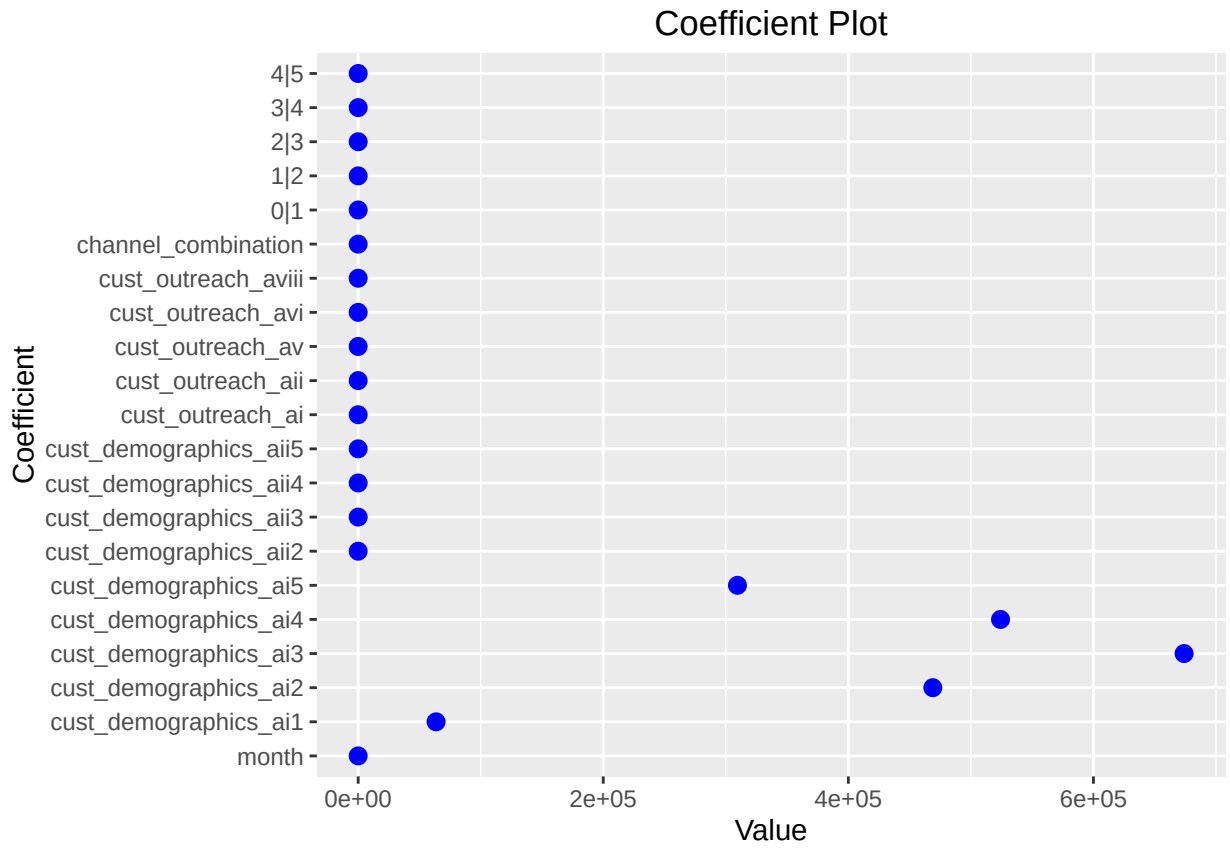
```
#Account C
C<-polr(typeC_bal_cat~month+cust_demographics_ai+cust_demographics_aii+
          cust_outreach_ai+cust_outreach_aii+
          cust_outreach_av+cust_outreach_avi+cust_outreach_aviii+
          channel_combination, data=bank_data, Hess=TRUE)

ctableC<-coef(summary(C))

pcC<-pnorm(abs(ctableC[, "t value"]), lower.tail=FALSE *2)
(ctableC <- cbind(ctableC, "p value" = pcC))
```

```
##                              Value  Std. Error    t value       p value
## month                  -0.016652371 0.002920155  -5.702565  5.900900e-09
## cust_demographics_ai1  11.060104609 0.051483942 214.826296  0.000000e+00
## cust_demographics_ai2  13.058215025 0.021425684 609.465496  0.000000e+00
## cust_demographics_ai3  13.420986146 0.020130510 666.698767  0.000000e+00
## cust_demographics_ai4  13.169638709 0.021087863 624.512707  0.000000e+00
## cust_demographics_ai5  12.642773418 0.025421627 497.323538  0.000000e+00
## cust_demographics_aii2  0.686397190 0.038053933  18.037484  4.947411e-73
## cust_demographics_aii3  1.024854741 0.036182318  28.324740 8.569479e-177
## cust_demographics_aii4  0.914682815 0.036658865  24.951204 1.036038e-137
## cust_demographics_aii5  0.781693057 0.039176076  19.953327  7.011952e-89
## cust_outreach_ai        0.022202911 0.004409402   5.035356  2.384803e-07
## cust_outreach_aii      -0.010091452 0.004267377  -2.364791  9.020135e-03
## cust_outreach_av       -0.005386310 0.001365300  -3.945148  3.987540e-05
## cust_outreach_avi       0.008422811 0.000589298  14.292956  1.210561e-46
## cust_outreach_aviii     0.024573576 0.007710276   3.187120  7.184849e-04
## channel_combination     0.529371204 0.004285455 123.527433  0.000000e+00
## 0|1                    17.368451921 0.035230677 492.992284  0.000000e+00
## 1|2                    17.707628822 0.035874781 493.595458  0.000000e+00
## 2|3                    18.103282046 0.036591969 494.733754  0.000000e+00
## 3|4                    18.609857349 0.037589374 495.082931  0.000000e+00
## 4|5                    19.396168897 0.039816243 487.142112  0.000000e+00
```

```
sort(format(exp(coef(C)), scientific = FALSE),decreasing = TRUE)
```

```
##  cust_demographics_ai3  cust_demographics_ai4  cust_demographics_ai2
##       "674000.5094113"       "524205.3151078"       "468932.9274169"
##  cust_demographics_ai5  cust_demographics_ai1 cust_demographics_aii3
##       "309518.5851506"       " 63583.2028681"       "     2.7866906"
## cust_demographics_aii4 cust_demographics_aii5 cust_demographics_aii2
##       "     2.4959834"       "     2.1851687"       "     1.9865455"
##    channel_combination    cust_outreach_aviii       cust_outreach_ai
##       "     1.6978644"       "     1.0248780"       "     1.0224512"
##      cust_outreach_avi       cust_outreach_av      cust_outreach_aii
##       "     1.0084584"       "     0.9946282"       "     0.9899593"
##                 month
##       "     0.9834855"
```

```
C$coefficients<-exp(coef(C))
coefplot(C, zeroType = 0)
```




```
#Account D
D<-polr(typeD_bal_cat~month+cust_demographics_ai+cust_demographics_aii+
          cust_outreach_ai+cust_outreach_aiii+
          cust_outreach_av+cust_outreach_avi+cust_outreach_avii+cust_outreach_aviii+
          channel_combination, data=bank_data, Hess=TRUE)

ctableD<-coef(summary(D))

pcD<-pnorm(abs(ctableD[, "t value"]), lower.tail=FALSE *2)
(ctableD <- cbind(ctableD, "p value" = pcD))
```

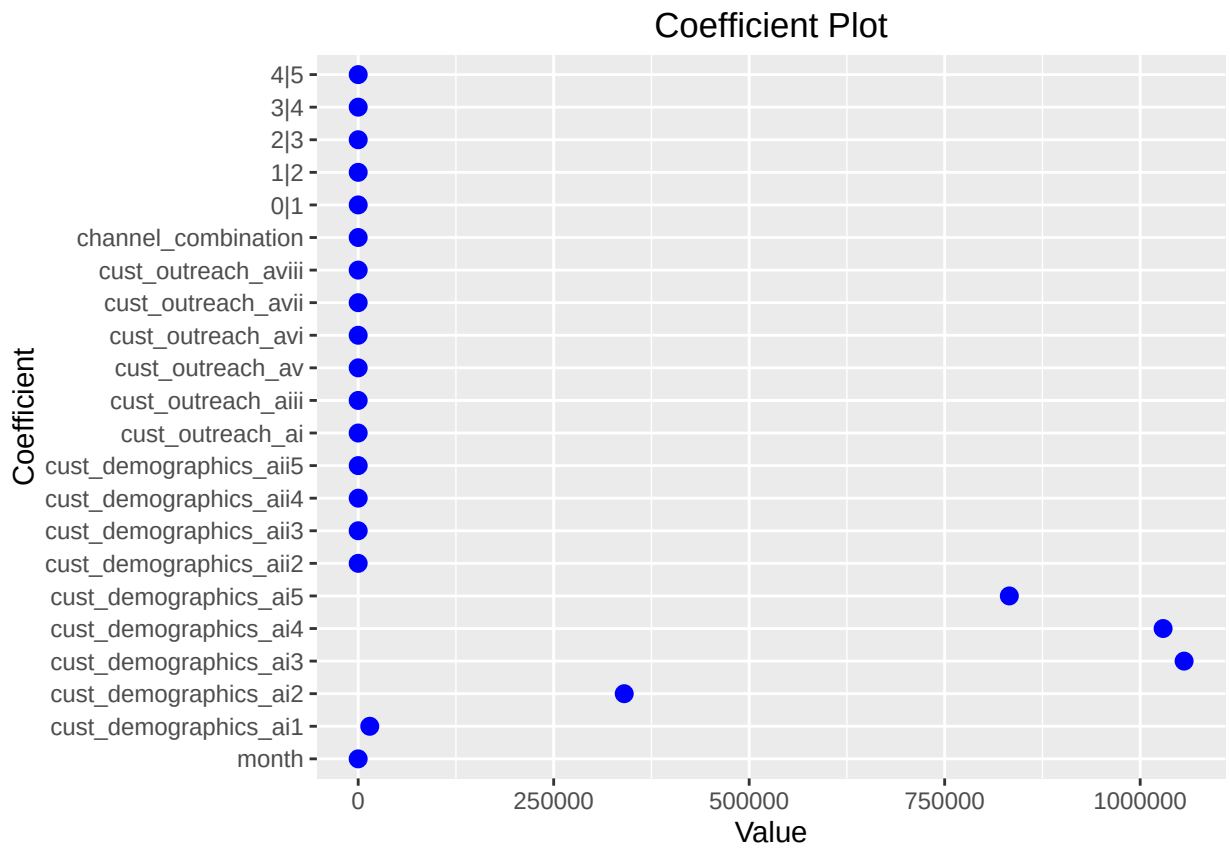
```
##                            Value   Std. Error    t value       p value
## month                 -0.02282928 0.0042189784  -5.411092  3.132084e-08
## cust_demographics_ai1  9.60293043 0.2642884240  36.335040 2.263820e-289
## cust_demographics_ai2 12.73732762 0.0671107434 189.795657  0.000000e+00
## cust_demographics_ai3 13.87014912 0.0597214451 232.247379  0.000000e+00
## cust_demographics_ai4 13.84441376 0.0595699006 232.406192  0.000000e+00
## cust_demographics_ai5 13.63239061 0.0607672757 224.337696  0.000000e+00
## cust_demographics_aii2  1.11349882 0.0970319363  11.475591  8.748893e-31
## cust_demographics_aii3  1.80149279 0.0905517005  19.894632  2.264583e-88
## cust_demographics_aii4  2.05933787 0.0891928127  23.088608 3.013168e-118
## cust_demographics_aii5  2.34359520 0.0891811321  26.279047 1.664650e-152
## cust_outreach_ai        0.03264173 0.0057280212   5.698605  6.039576e-09
## cust_outreach_aiii      0.03809989 0.0067238602   5.666372  7.292638e-09
## cust_outreach_av       -0.01052434 0.0023508347  -4.476853  3.787564e-06
## cust_outreach_avi       0.01195464 0.0007435985  16.076743  1.856989e-58
```

```
## cust_outreach_avii      0.01981694 0.0054814031   3.615305  1.499974e-04
## cust_outreach_aviii    -0.03882271 0.0129360621  -3.001123  1.344930e-03
## channel_combination     0.10750622 0.0052670706  20.411009  6.675011e-93
## 0|1                    18.41476565 0.0908336722 202.730609  0.000000e+00
## 1|2                    18.98153235 0.0914912844 207.468203  0.000000e+00
## 2|3                    19.28685311 0.0919817334 209.681340  0.000000e+00
## 3|4                    19.70971238 0.0929166427 212.122520  0.000000e+00
## 4|5                    20.42058897 0.0955927896 213.620599  0.000000e+00
```

```
sort(format(exp(coef(D)), scientific = FALSE), decreasing = TRUE)
```

```
##  cust_demographics_ai3  cust_demographics_ai4  cust_demographics_ai5
##      "1056158.8132834"      "1029324.9488112"       " 832668.2715305"
##  cust_demographics_ai2  cust_demographics_ai1 cust_demographics_aii5
##       " 340213.1529963"       "  14808.1121171"       "     10.4186263"
## cust_demographics_aii4 cust_demographics_aii3 cust_demographics_aii2
##       "      7.8407765"       "      6.0586851"       "      3.0449937"
##    channel_combination     cust_outreach_aiii       cust_outreach_ai
##       "      1.1134978"       "      1.0388350"       "      1.0331803"
##     cust_outreach_avii      cust_outreach_avi       cust_outreach_av
##       "      1.0200146"       "      1.0120264"       "      0.9895308"
##                  month    cust_outreach_aviii
##       "      0.9774293"       "      0.9619212"
```

```
D$coefficients<-exp(coef(D))
coefplot(D, zeroType = 0)
```



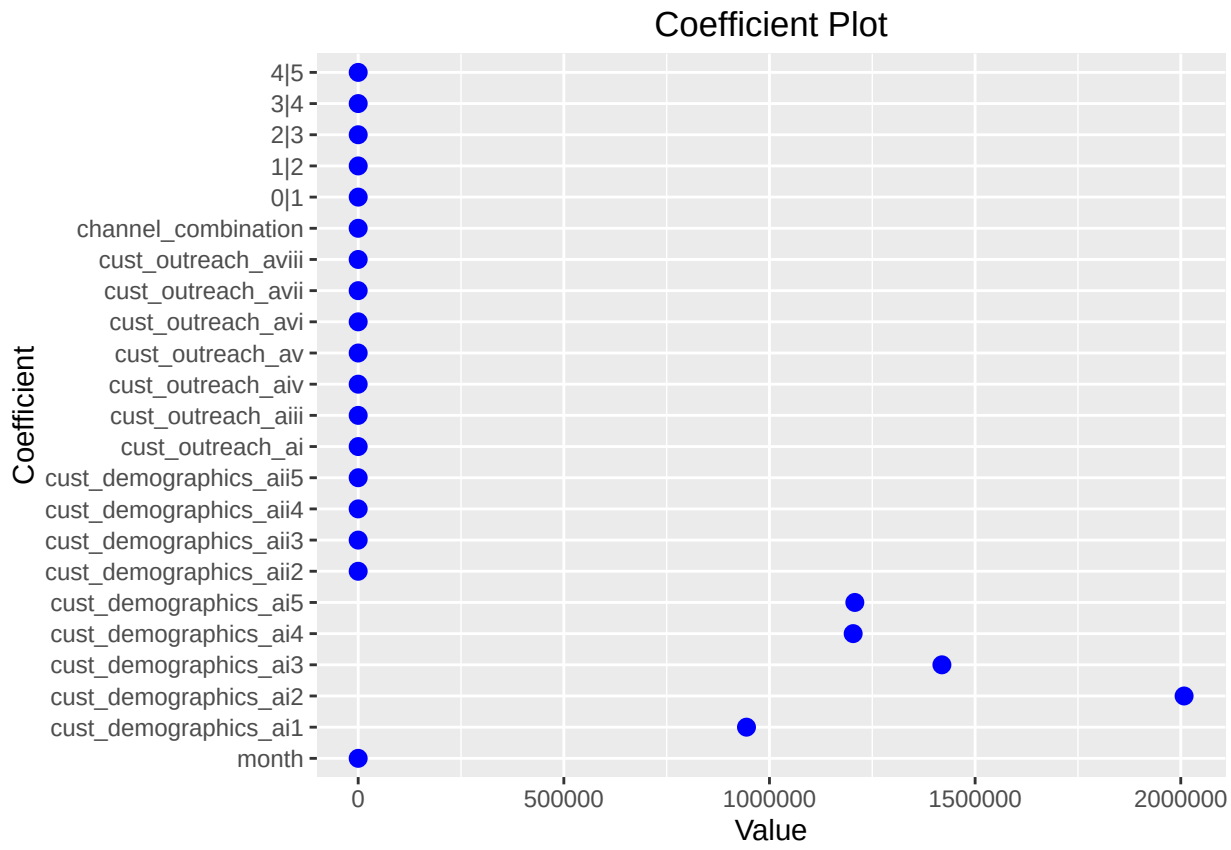
```r
#Account E
E<-polr(typeE_bal_cat~month+cust_demographics_ai+cust_demographics_aii+
          cust_outreach_ai + cust_outreach_aiii+ cust_outreach_aiv+
          cust_outreach_av+cust_outreach_avi+cust_outreach_avii+cust_outreach_aviii+
          channel_combination, data=bank_data, Hess=TRUE)

ctableE<-coef(summary(E))

pcE<-pnorm(abs(ctableE[, "t value"]), lower.tail=FALSE *2)
(ctableE <- cbind(ctableE, "p value" = pcE))
```

```
##                              Value   Std. Error    t value      p value
## month                  -0.023604174 0.0048166384  -4.900549 4.778453e-07
## cust_demographics_ai1  13.758020546 0.0422748878 325.441917 0.000000e+00
## cust_demographics_ai2  14.512685344 0.0279027768 520.116167 0.000000e+00
## cust_demographics_ai3  14.165623972 0.0312055764 453.945276 0.000000e+00
## cust_demographics_ai4  14.000697121 0.0339251475 412.693773 0.000000e+00
## cust_demographics_ai5  14.004087871 0.0376995916 371.465241 0.000000e+00
## cust_demographics_aii2  0.195972605 0.0499077515   3.926697 4.306023e-05
## cust_demographics_aii3  0.364270094 0.0485770080   7.498817 3.219817e-14
## cust_demographics_aii4  0.234821436 0.0511745987   4.588633 2.230793e-06
## cust_demographics_aii5 -0.325658655 0.0623728742  -5.221158 8.890369e-08
## cust_outreach_ai        0.028768442 0.0075857040   3.792455 7.458259e-05
## cust_outreach_aiii      0.030807830 0.0085655931   3.596696 1.611427e-04
## cust_outreach_aiv       0.055290904 0.0169053779   3.270610 5.365784e-04
## cust_outreach_av        0.012778785 0.0014775580   8.648585 2.607102e-18
## cust_outreach_avi       0.007456013 0.0007370392  10.116170 2.341946e-24
## cust_outreach_avii      0.024262281 0.0060868299   3.986029 3.359413e-05
## cust_outreach_aviii     0.018572250 0.0105058685   1.767798 3.854737e-02
## channel_combination     0.104327558 0.0059755355  17.459115 1.467251e-68
## 0|1                    17.780668558 0.0430322780 413.193756 0.000000e+00
## 1|2                    18.015771318 0.0436803795 412.445394 0.000000e+00
## 2|3                    18.315935270 0.0447226137 409.545278 0.000000e+00
## 3|4                    18.734087774 0.0466865753 401.273550 0.000000e+00
## 4|5                    19.439423878 0.0520422884 373.531304 0.000000e+00
```

```r
sort(format(exp(coef(E)), scientific = FALSE), decreasing = TRUE)
```

```
##  cust_demographics_ai2  cust_demographics_ai3  cust_demographics_ai5
##      "2008071.4546786"      "1419228.4992774"      "1207530.4370652"
##  cust_demographics_ai4  cust_demographics_ai1 cust_demographics_aii3
##      "1203442.9368969"      " 944131.3208280"      "      1.4394630"
## cust_demographics_aii4 cust_demographics_aii2    channel_combination
##      "      1.2646829"      "      1.2164936"      "      1.1099640"
##      cust_outreach_aiv     cust_outreach_aiii       cust_outreach_ai
##      "      1.0568480"      "      1.0312873"      "      1.0291863"
##     cust_outreach_avii    cust_outreach_aviii       cust_outreach_av
##      "      1.0245590"      "      1.0187458"      "      1.0128608"
##      cust_outreach_avi                   month cust_demographics_aii5
##      "      1.0074839"      "      0.9766722"      "      0.7220516"
```

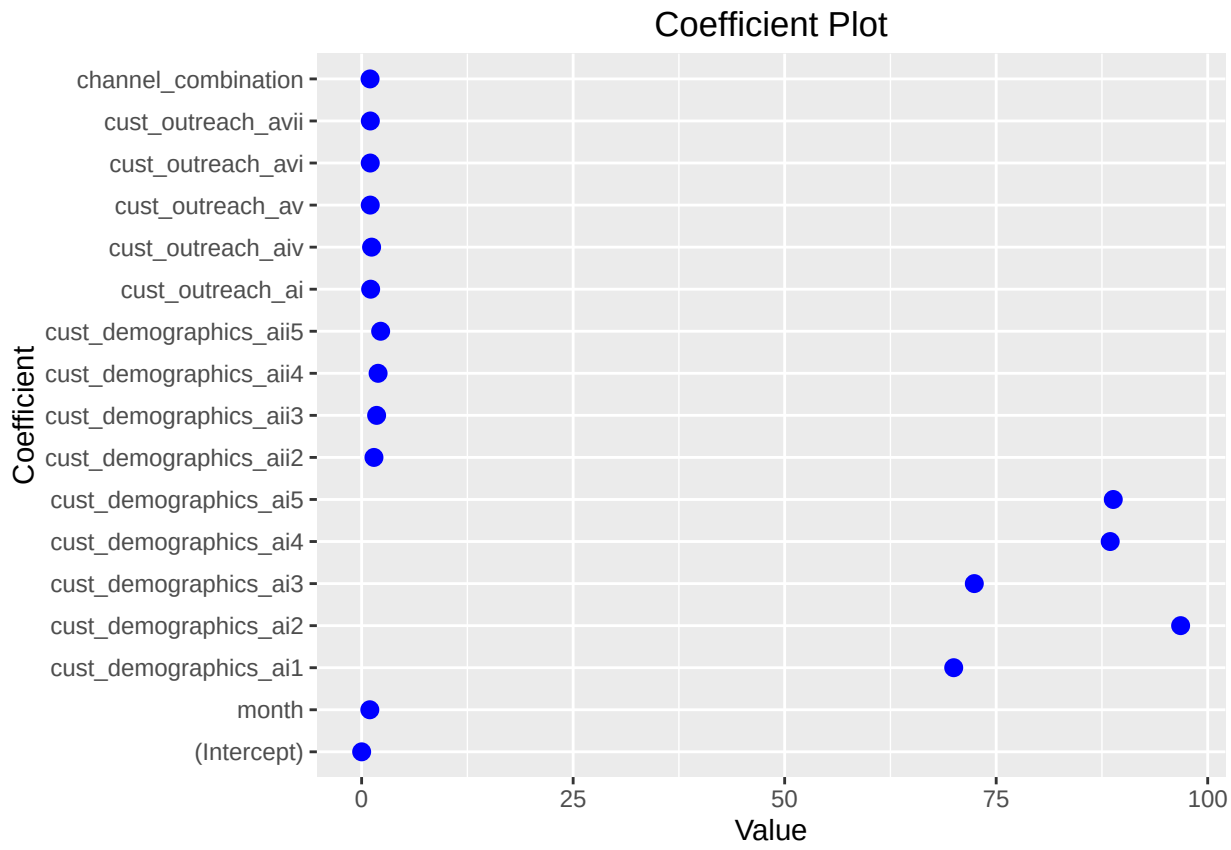
```
E$coefficients<-exp(coef(E))
coefplot(E, zeroType = 0)
```




```
#Account F
F_account<-glm(typeF_flag~month+cust_demographics_ai+cust_demographics_aii +
                 cust_outreach_ai+cust_outreach_aiv+cust_outreach_av+
                 cust_outreach_avi+cust_outreach_avii
               +channel_combination, data=bank_data,
               family="binomial")

sort(exp(coef(F_account)), decreasing=TRUE)
```

```
##  cust_demographics_ai2  cust_demographics_ai5  cust_demographics_ai4
##           96.798357215           88.837073181           88.476581846
##  cust_demographics_ai3  cust_demographics_ai1 cust_demographics_aii5
##           72.415228683           69.982519405            2.251104094
## cust_demographics_aii4 cust_demographics_aii3 cust_demographics_aii2
##            1.947676838            1.775478108            1.459901537
##      cust_outreach_aiv       cust_outreach_ai     cust_outreach_avii
##            1.180051219            1.068911741            1.031512357
##      cust_outreach_avi       cust_outreach_av    channel_combination
##            1.020983935            1.020235245            0.998053062
##                  month            (Intercept)
##            0.969844804            0.004544448
```

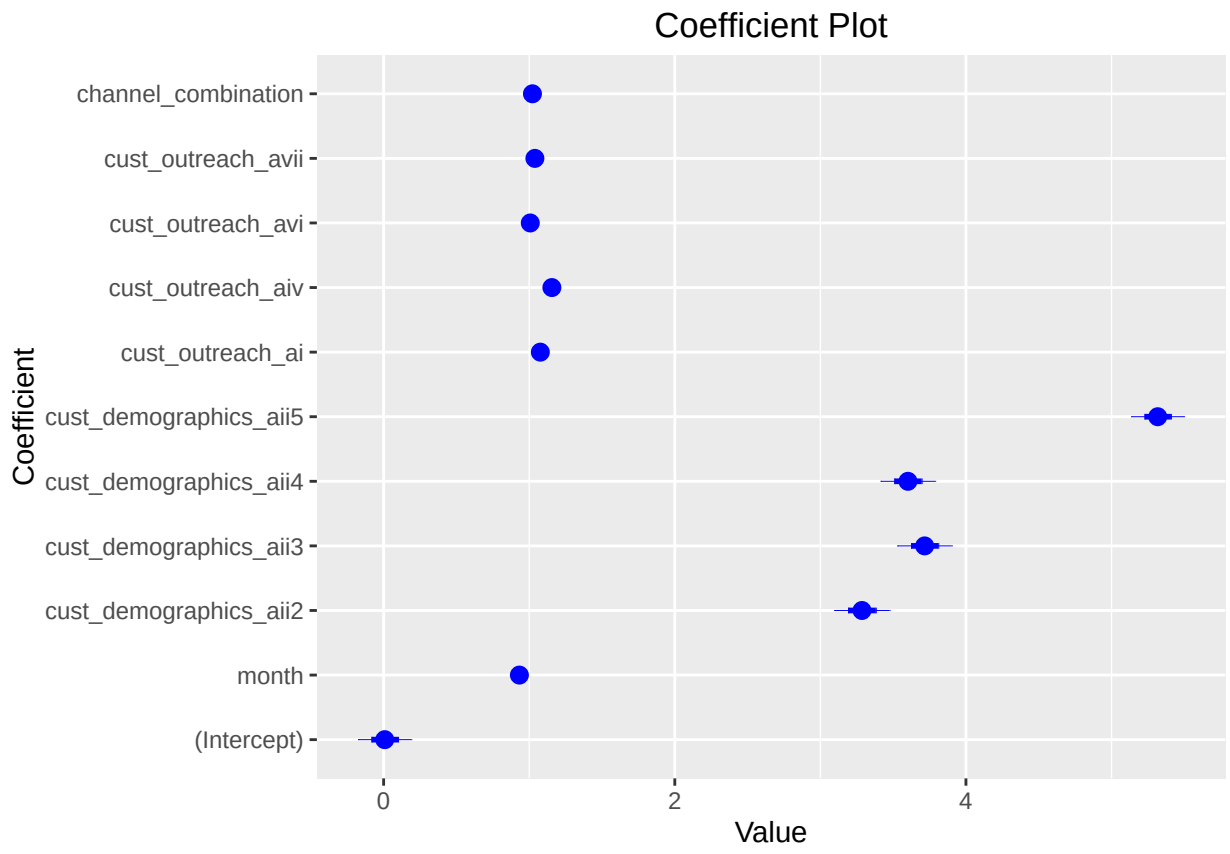
```
F_account$coefficients<-exp(coef(F_account))
coefplot(F_account, zeroType = 0)
```



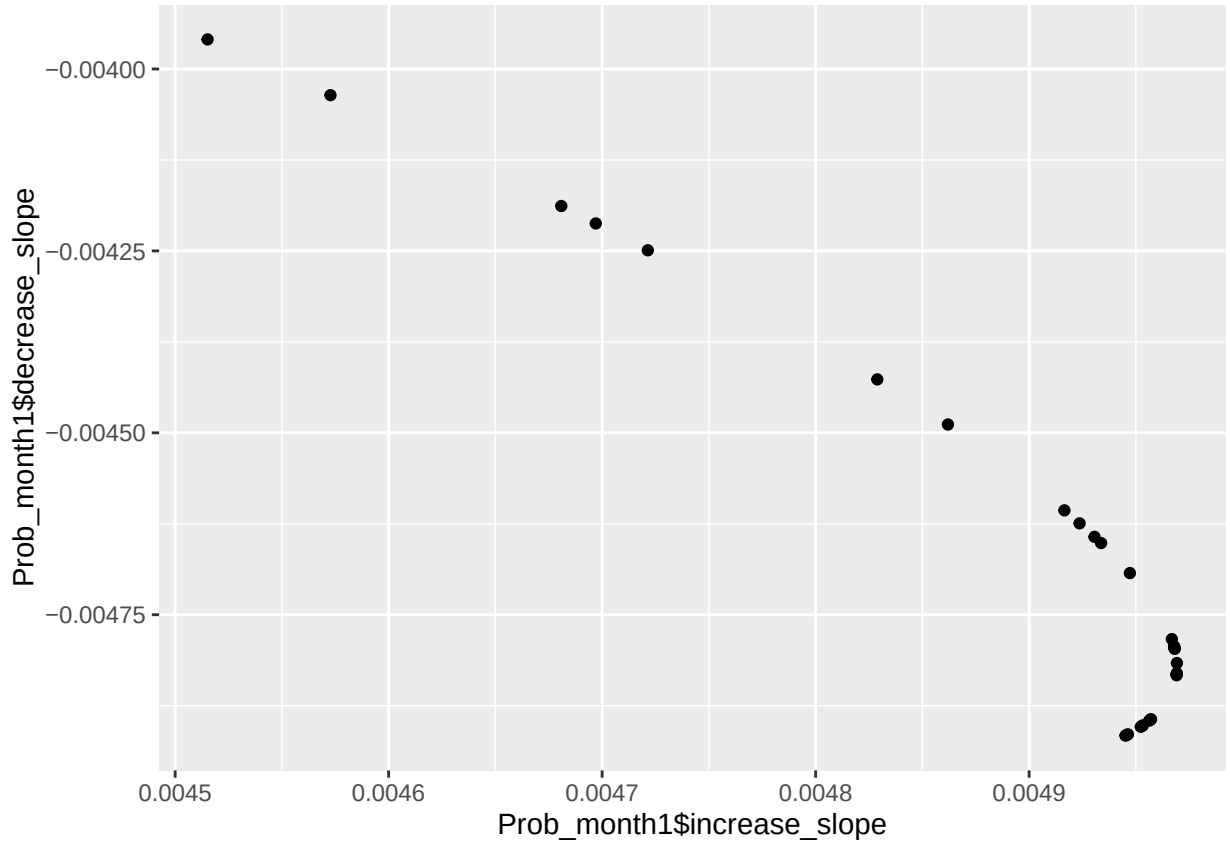

```
#Account G
G_account<-glm(typeG_flag~month+cust_demographics_aii+cust_outreach_ai+cust_outreach_aiv+
                cust_outreach_avi+cust_outreach_avii+channel_combination, data=bank_data,family="binom

sort(exp(coef(G_account)), decreasing=TRUE)
```

```
## cust_demographics_aii5 cust_demographics_aii3 cust_demographics_aii4
##            5.317196950            3.716206031            3.601648872
## cust_demographics_aii2      cust_outreach_aiv       cust_outreach_ai
##            3.285788350            1.155693813            1.076386375
##     cust_outreach_avii    channel_combination      cust_outreach_avi
##            1.038977044            1.022213512            1.007306130
##                  month            (Intercept)
##            0.932536974            0.007611338
```

```
G_account$coefficients<-exp(coef(G_account))
coefplot(G_account, zeroType = 0)
```




```
Prob_month1 <- subset(prediction, month == 1)
Prob_month12 <- subset(prediction, month == 12)

Prob_month1 <- Prob_month1[order(Prob_month1$cust_demographics_ai,Prob_month1$cust_demographics_aii),]
Prob_month12 <- Prob_month12[order(Prob_month12$cust_demographics_ai,Prob_month12$cust_demographics_aii

Prob_month1 <- Prob_month1[!duplicated(Prob_month1),]
Prob_month12 <- Prob_month12[!duplicated(Prob_month12),]

Prob_month1$decrease_slope <- (Prob_month12$`-1` - Prob_month1$`-1`) / 11
Prob_month1$increase_slope <- (Prob_month12$`1` - Prob_month1$`1`) / 11

ggplot(Prob_month1) + geom_point(aes(Prob_month1$increase_slope, Prob_month1$decrease_slope))
```




```
write.table(Prob_month1,file="output.csv",sep=",",row.names = FALSE)

# Use this output in tableau to create a heatmap
```