

MV Project Team 7

Team 7

2020/3/18

```
library(caret)
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
library(leaps)
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(glmulti)
```

```
## Loading required package: rJava
```

```
library(sjPlot)
```

```
library(sjmisc)
```

```
## Learn more about sjmisc with 'browseVignettes("sjmisc")'.
```

```
library(ggplot2)
```

```
library(corrplot)
```

```
## corrplot 0.84 loaded
```

Visualizations

```

setwd("E:/W6400 MV/project")
churn <- read.csv("Churn_Modelling.csv")
# 3 variables not in use
churn <- subset(churn, select = -c(LineNumber, CustomerId, Surname))
churn$Exited <- as.factor(churn$Exited)

# distribution plot
# Personal information variables
churn$Geography <- as.factor(churn$Geography)
#ggplot(data=churn, aes(x=Geography)) +
# geom_bar() +
# geom_text(stat='count', aes(label=..count..), vjust=-1)
summary(churn$Geography)

```

```

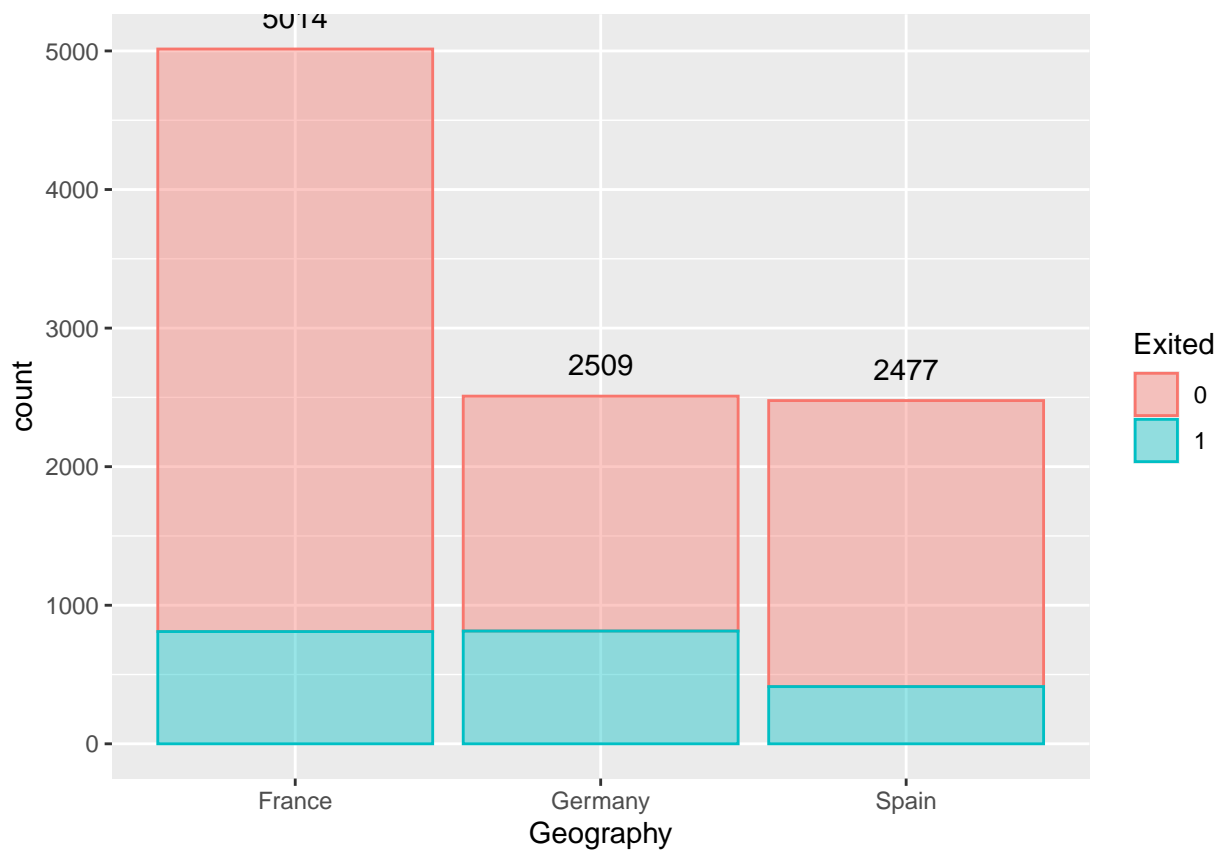
## France Germany Spain
## 5014 2509 2477

```

```

ggplot(churn, aes(x = Geography)) +
  geom_bar(aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  geom_text(stat='count', aes(label=..count..), vjust=-1)

```



```

summary(churn[churn$Geography=='France',]$Exited)

```

```

## 0 1

```

```
## 4204 810
```

```
summary(churn[churn$Geography=='Germany',]$Exited)
```

```
##    0    1  
## 1695 814
```

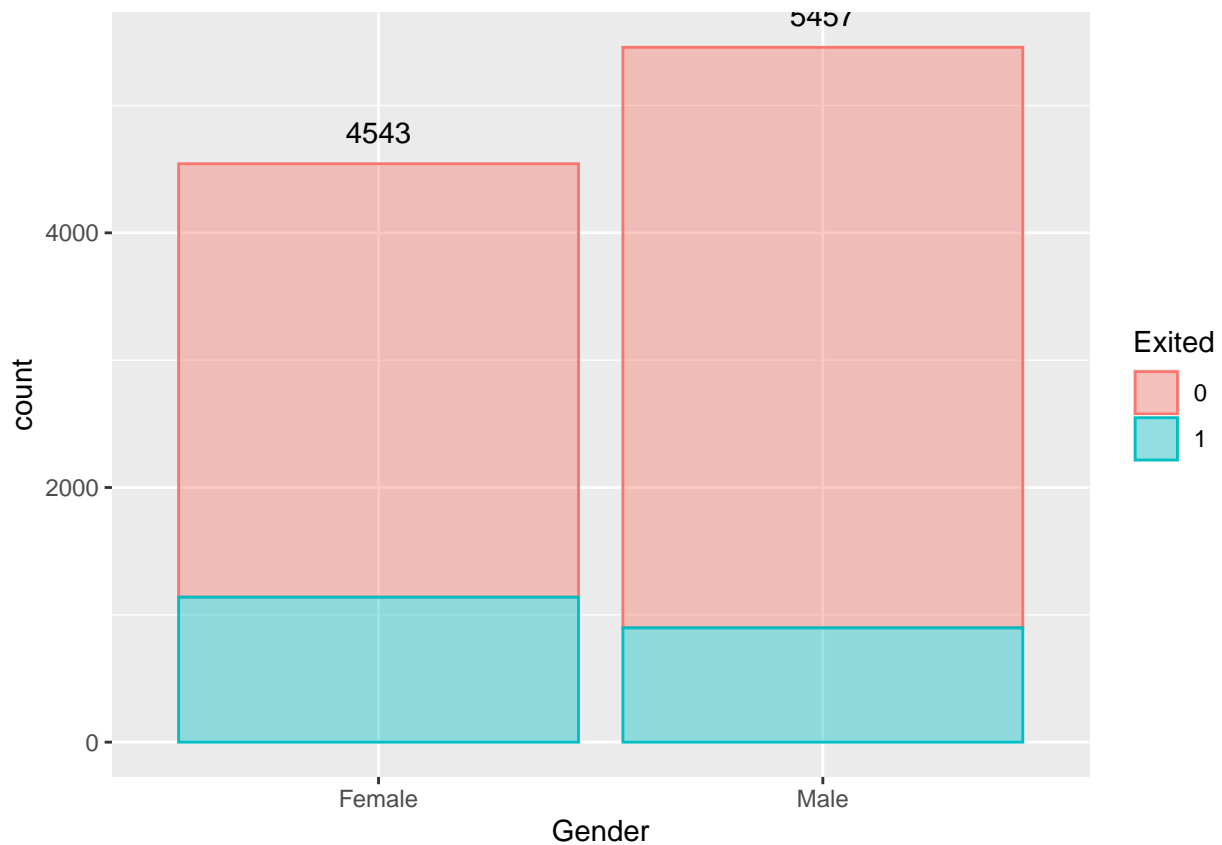
```
summary(churn[churn$Geography=='Spain',]$Exited)
```

```
##    0    1  
## 2064 413
```

```
churn$Gender <- as.factor(churn$Gender)  
#ggplot(data=churn, aes(x=Gender)) +  
#  geom_bar() +  
#  geom_text(stat='count', aes(label=..count..), vjust=-1)  
summary(churn$Gender)
```

```
## Female  Male  
##  4543   5457
```

```
ggplot(churn, aes(x = Gender)) +  
  geom_bar(aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4) +  
  geom_text(stat='count', aes(label=..count..), vjust=-1)
```



```
summary(churn[churn$Gender=='Male',]$Exited)
```

```
##      0      1
## 4559  898
```

```
summary(churn[churn$Gender=='Female',]$Exited)
```

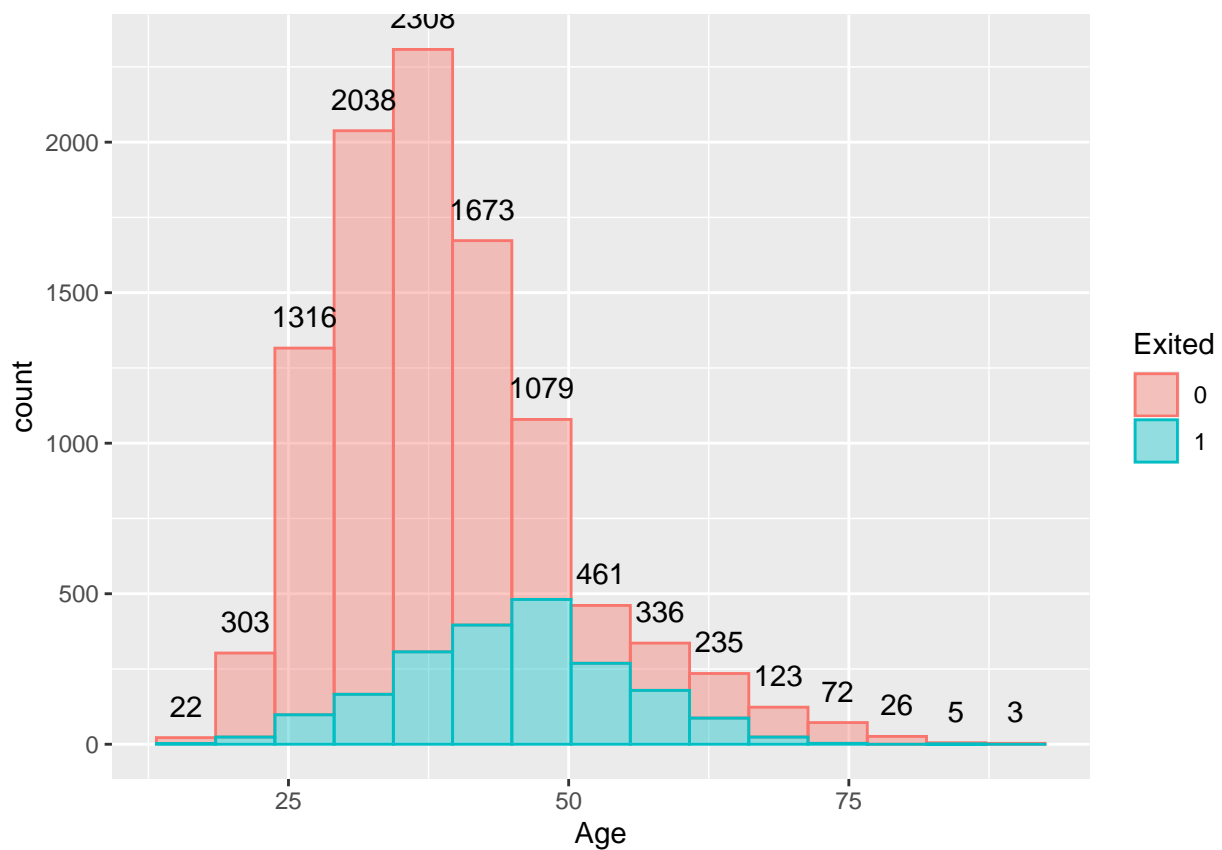
```
##      0      1
## 3404 1139
```

```
#hist(churn$Age,col="blue")
```

```
summary(churn$Age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    18.00   32.00   37.00   38.92   44.00   92.00
```

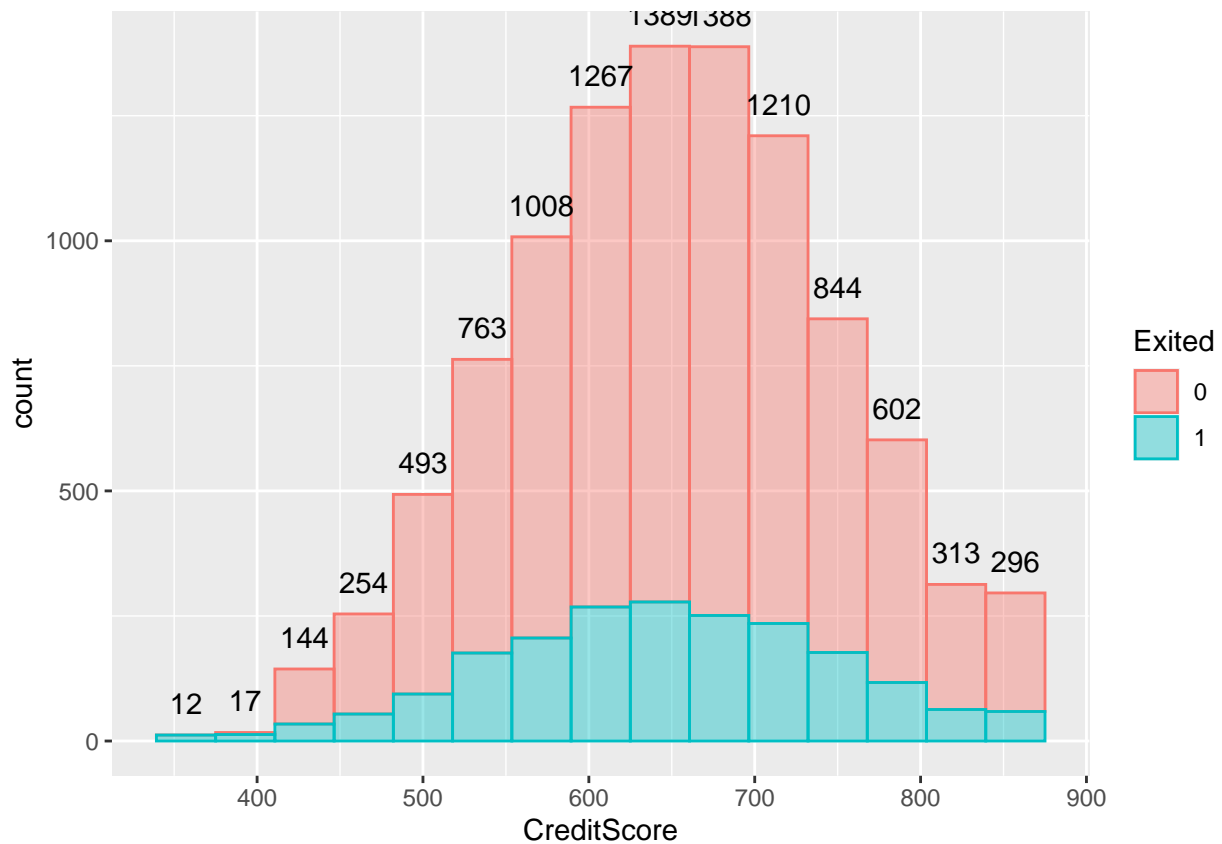
```
ggplot(churn,aes(x=Age))+
  geom_histogram(bins=15,aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  stat_bin(bins=15,aes(y=..count.., label=..count..), geom="text", vjust=-1)
```



```
# Banking information variables
#hist(churn$CreditScore,col="blue")
summary(churn$CreditScore)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    350.0   584.0   652.0   650.5   718.0   850.0
```

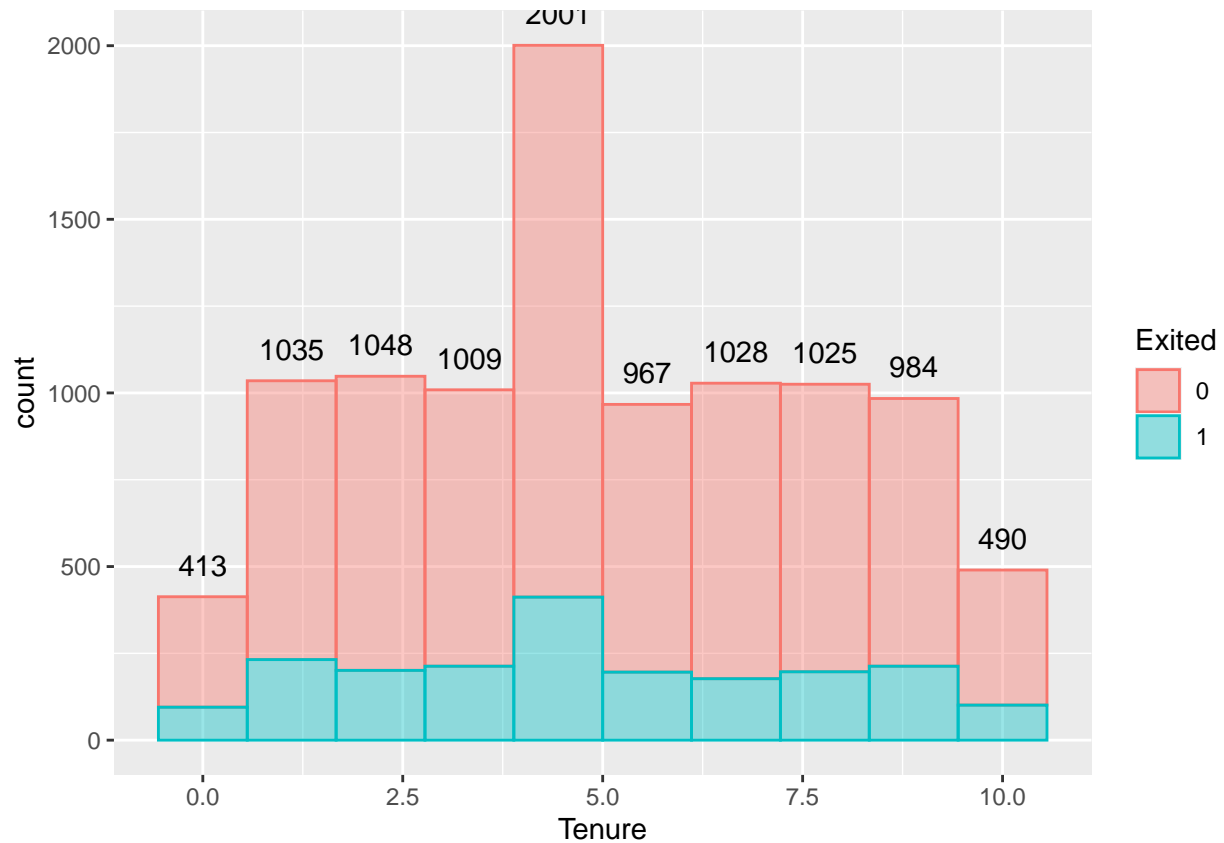
```
ggplot(churn, aes(x = CreditScore)) +
  geom_histogram(bins=15, aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4) +
  stat_bin(bins=15, aes(y=..count.., label=..count..), geom="text", vjust=-1)
```



```
#hist(churn$Tenure,col="blue")
summary(churn$Tenure)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.000   3.000   5.000   5.013   7.000   10.000
```

```
ggplot(churn, aes(x = Tenure)) +
  geom_histogram(bins=10, aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4) +
  stat_bin(bins=10, aes(y=..count.., label=..count..), geom="text", vjust=-1)
```



```
#hist(churn$Balance,col="grey")
summary(churn$Balance)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##         0         0   97199   76486  127644  250898
```

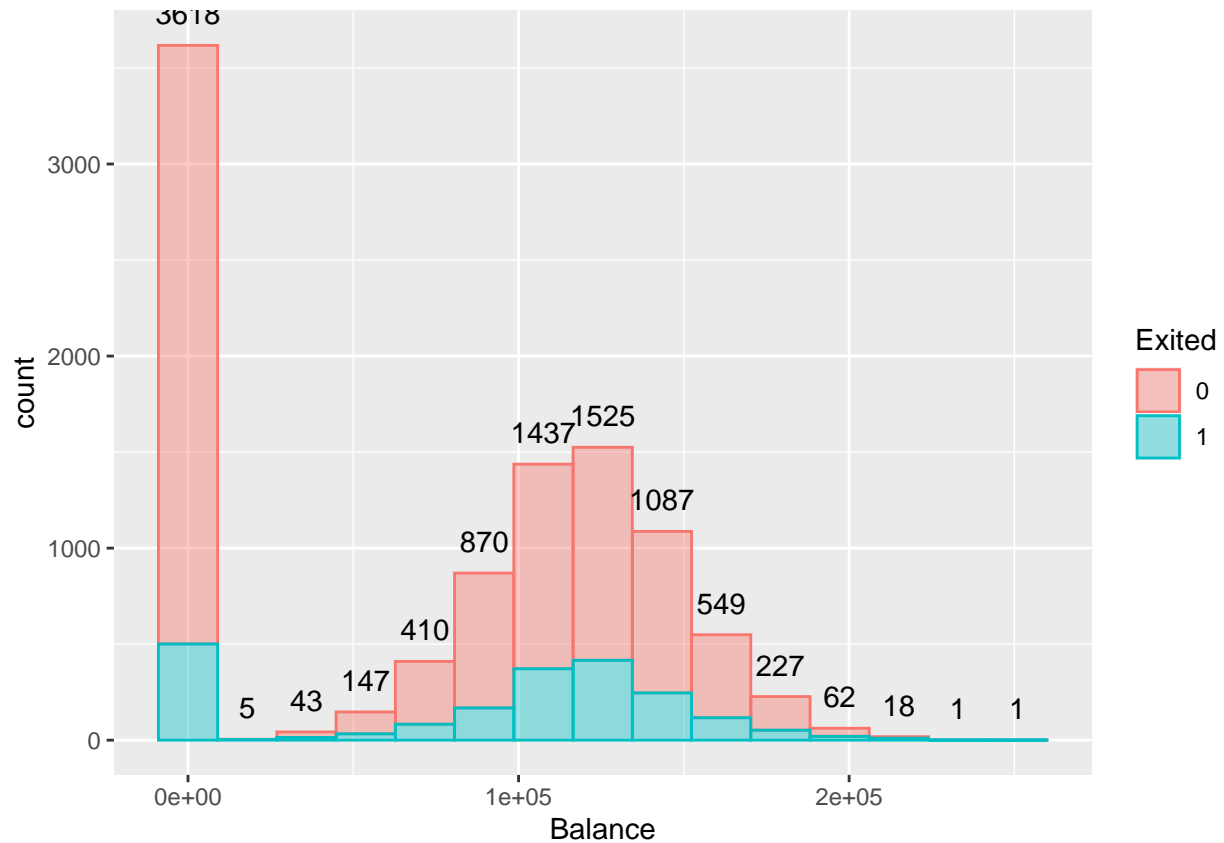
```
summary(churn[churn$Balance==0,]$Exited)
```

```
##      0      1
## 3117   500
```

```
summary(churn[churn$Balance>0,]$Exited)
```

```
##      0      1
## 4846  1537
```

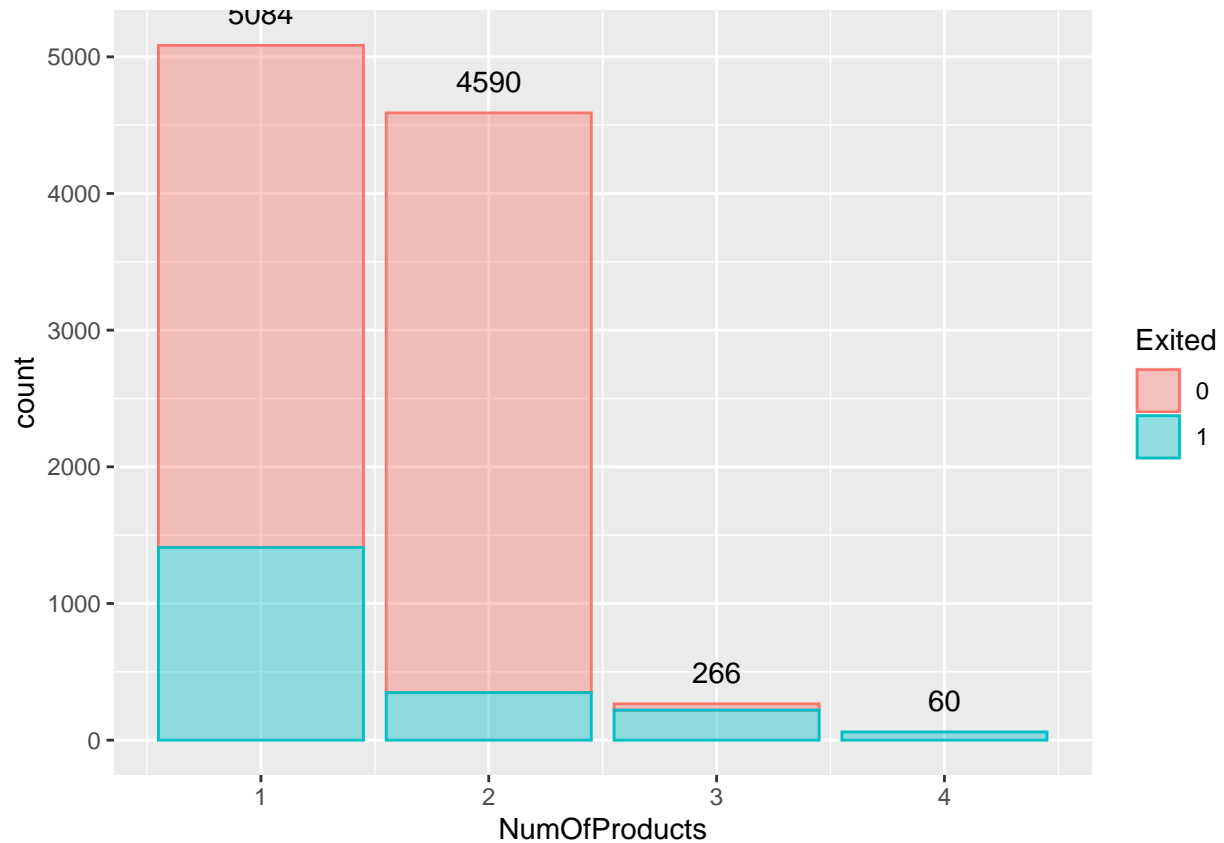
```
ggplot(churn, aes(x = Balance)) +
  geom_histogram(bins=15,aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  stat_bin(bins=15,aes(y=..count.., label=..count..), geom="text", vjust=-1)
```



```
#hist(churn$NumOfProducts,col="blue")
summary(churn$NumOfProducts)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.00   1.00    1.00   1.53   2.00    4.00
```

```
ggplot(churn, aes(x = NumOfProducts)) +
  geom_bar(aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  geom_text(stat='count', aes(label=..count..), vjust=-1)
```



```
summary(churn[churn$NumOfProducts==3,]$Exited)
```

```
##    0    1
##  46  220
```

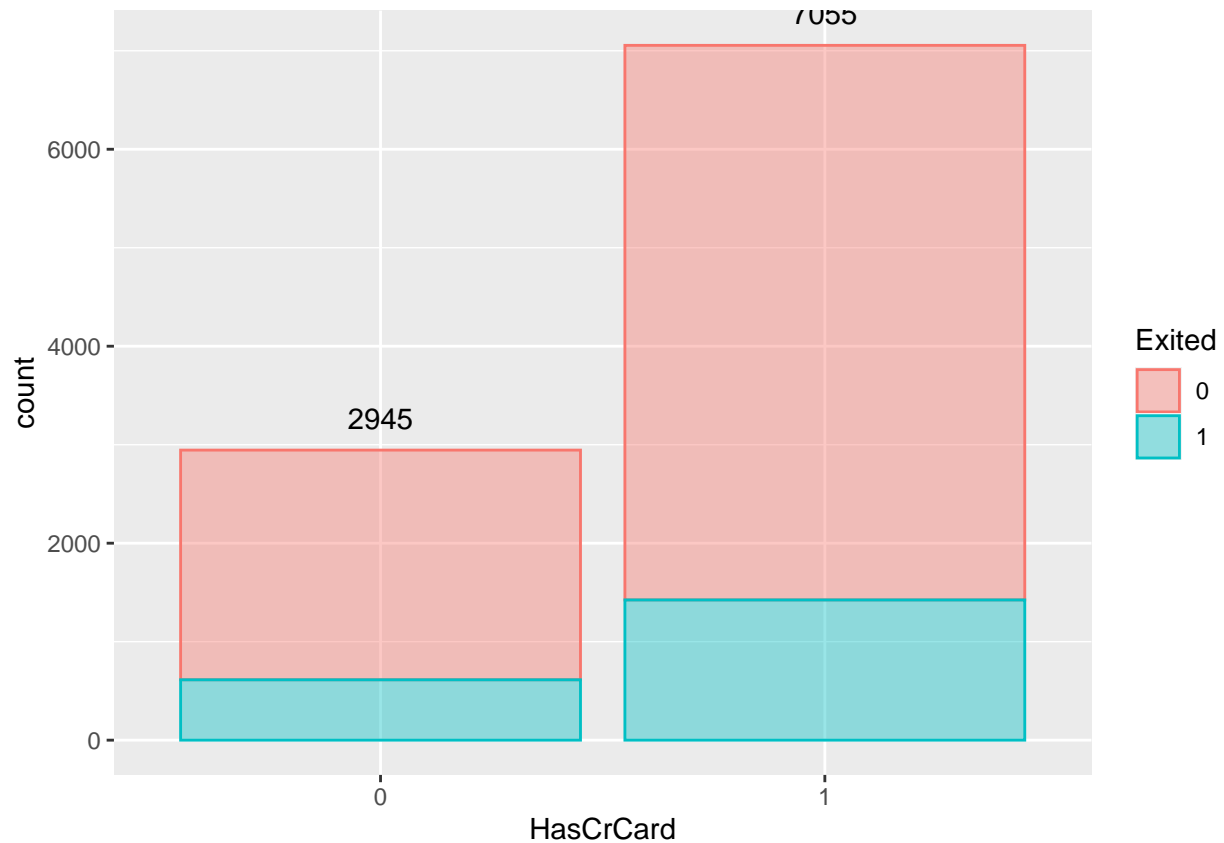
```
summary(churn[churn$NumOfProducts==4,]$Exited)
```

```
##    0    1
##    0    60
```

```
#hist(churn$HasCrCard, col="blue")
summary(churn$HasCrCard)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000  0.0000   1.0000  0.7055  1.0000  1.0000
```

```
churn$HasCrCard <- as.factor(churn$HasCrCard)
ggplot(churn, aes(x = HasCrCard)) +
  geom_bar(aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  geom_text(stat='count', aes(label=..count..), vjust=-1)
```

```
summary(churn[churn$HasCrCard==0,]$Exited)
```

```
##    0    1
## 2332  613
```

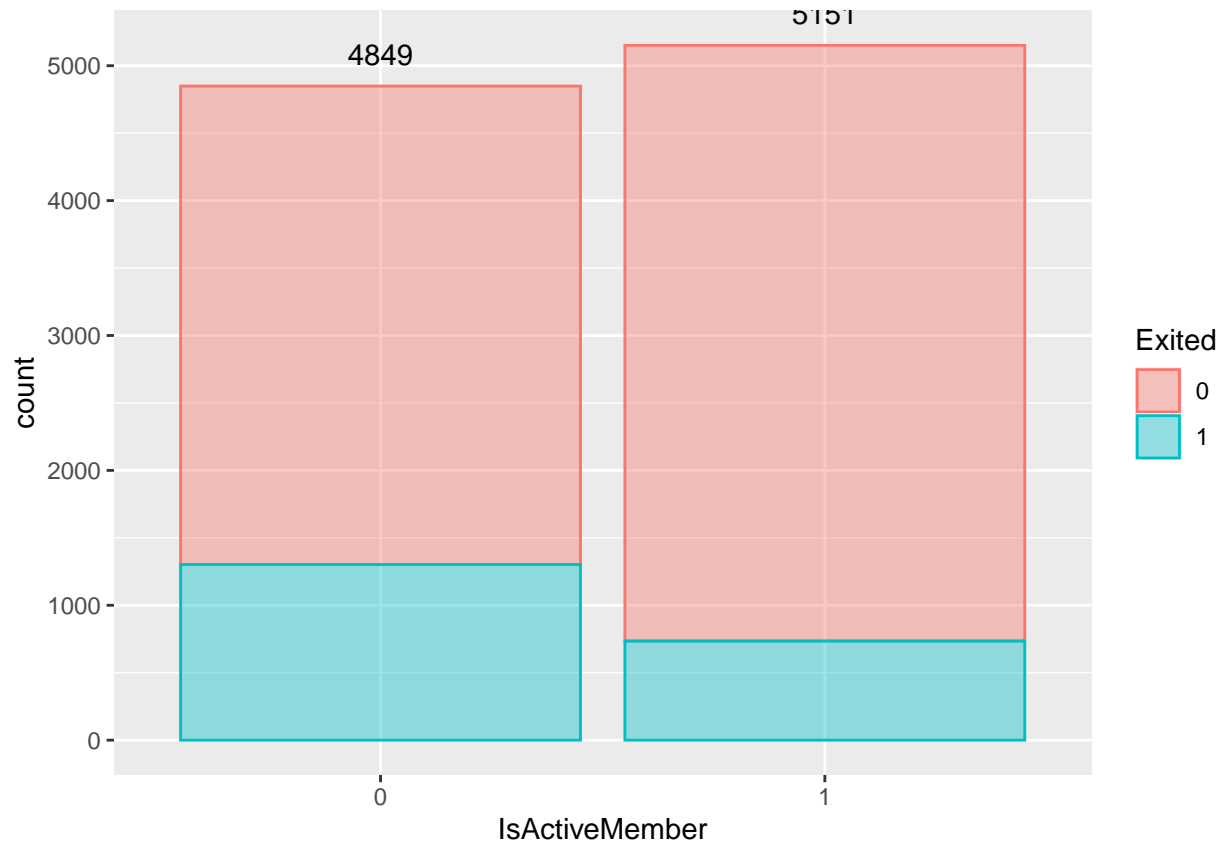
```
summary(churn[churn$HasCrCard==1,]$Exited)
```

```
##    0    1
## 5631 1424
```

```
#hist(churn$IsActiveMember,col="blue")
summary(churn$IsActiveMember)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000 0.0000  1.0000  0.5151  1.0000  1.0000
```

```
churn$IsActiveMember <- as.factor(churn$IsActiveMember)
ggplot(churn, aes(x = IsActiveMember)) +
  geom_bar(aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  geom_text(stat='count', aes(label=..count..), vjust=-1)
```



```
summary(churn[churn$IsActiveMember==0,]$Exited)
```

```
##    0    1
## 3547 1302
```

```
summary(churn[churn$IsActiveMember==1,]$Exited)
```

```
##    0    1
## 4416  735
```

```
#hist(churn$EstimatedSalary,col="blue")
summary(churn$EstimatedSalary)
```

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
##  11.58  51002.11 100193.91 100090.24 149388.25 199992.48
```

```
salary.outlier <- churn[churn$EstimatedSalary<500,]
salary.outlier
```

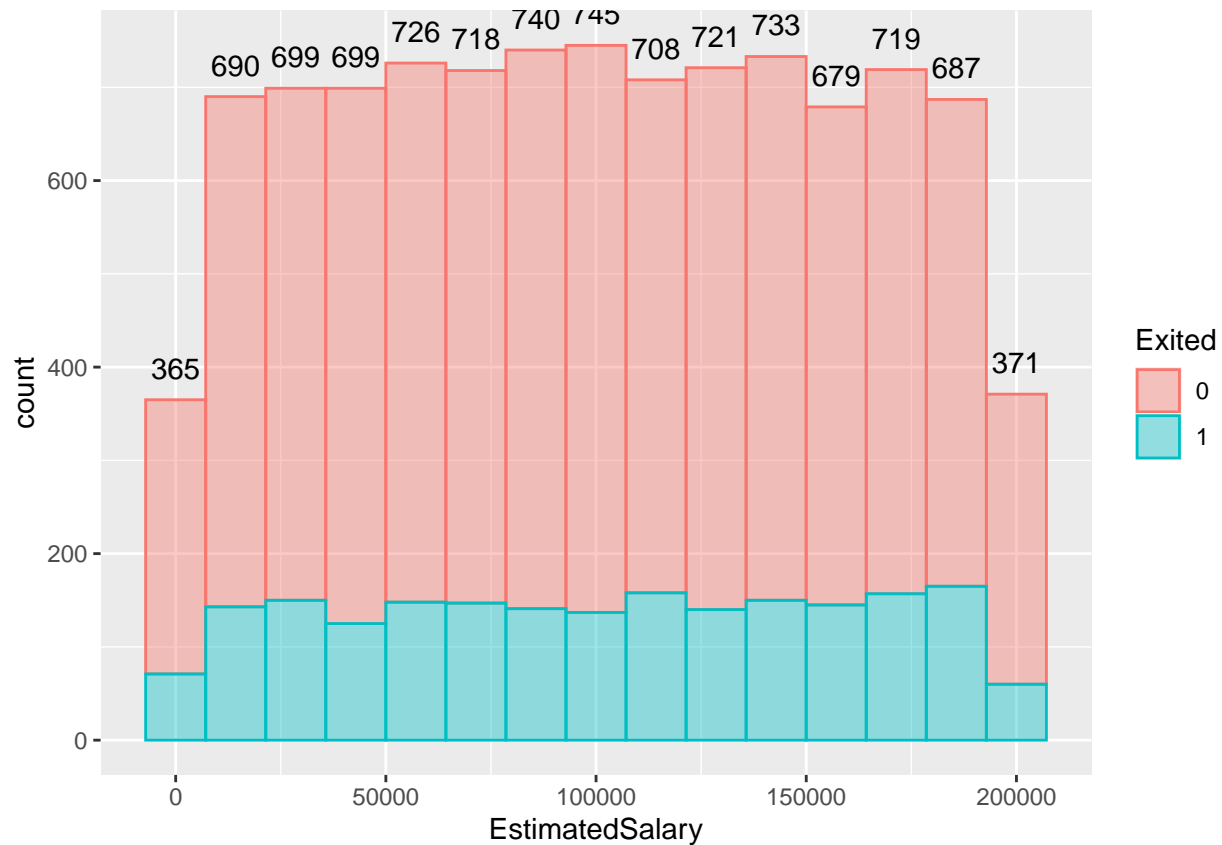
```
##      CreditScore Geography Gender Age Tenure   Balance NumOfProducts HasCrCard
## 394          556     Spain   Male  38      8      0.00             2          0
## 397          557     France   Male  33      3 54503.55             1          1
## 926          714     Germany Female 36      1 101609.01             2          1
```

## 1452	551	France	Female	38	10	0.00	2	1
## 1463	616	Germany	Female	41	1	103560.57	1	1
## 2047	520	Germany	Male	34	8	120018.86	2	1
## 2363	709	Germany	Male	45	4	122917.71	1	1
## 2499	718	France	Male	37	8	0.00	2	1
## 3447	649	France	Male	37	9	87374.88	2	1
## 4775	811	France	Female	35	7	0.00	1	1
## 4990	629	France	Male	40	9	0.00	1	1
## 5121	663	France	Male	26	5	141462.13	1	1
## 5549	710	France	Male	38	2	0.00	2	1
## 5676	766	France	Female	26	3	104258.80	1	1
## 7225	751	Spain	Female	34	8	127095.14	2	0
## 7227	777	Germany	Female	46	0	107362.80	1	1
## 7566	708	France	Male	52	10	105355.81	1	1
## 8503	530	France	Male	37	8	0.00	2	1
## 8672	542	Germany	Female	35	9	127543.11	2	1
## 8723	621	France	Male	35	7	87619.29	1	1
## 9011	645	Germany	Female	59	8	121669.93	2	0
## 9648	625	France	Male	31	5	0.00	2	0
## 9650	715	France	Female	38	0	0.00	2	1
##	IsActiveMember	EstimatedSalary	Exited					
## 394	0	417.41	1					
## 397	1	371.05	0					
## 926	1	447.73	0					
## 1452	1	216.27	0					
## 1463	0	236.45	1					
## 2047	1	343.38	0					
## 2363	1	11.58	1					
## 2499	1	142.81	0					
## 3447	1	247.36	0					
## 4775	1	178.19	0					
## 4990	0	106.67	0					
## 5121	0	440.20	0					
## 5549	0	96.27	0					
## 5676	1	428.23	0					
## 7225	0	479.54	0					
## 7227	0	487.30	0					
## 7566	0	123.07	1					
## 8503	1	287.99	0					
## 8672	0	468.94	1					
## 8723	0	143.34	0					
## 9011	0	91.75	1					
## 9648	1	90.07	0					
## 9650	1	332.81	0					

```
nrow(salary.outlier)
```

```
## [1] 23
```

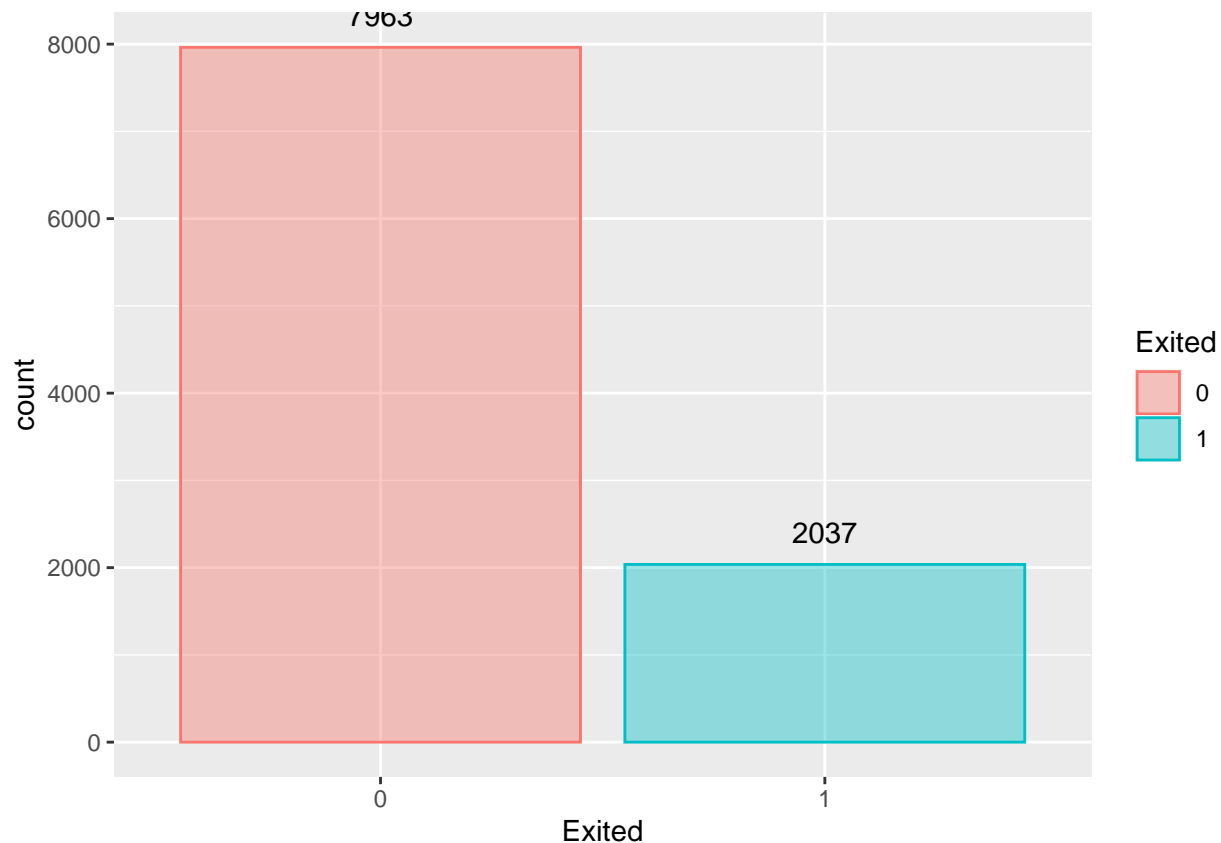
```
ggplot(churn, aes(x = EstimatedSalary)) +
  geom_histogram(bins=15, aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4) +
  stat_bin(bins=15, aes(y=..count.., label=..count..), geom="text", vjust=-1)
```



```
#plot(churn$Exited,col="blue")
summary(churn$Exited)
```

```
##    0    1
## 7963 2037
```

```
ggplot(churn, aes(x = Exited)) +
  geom_bar(aes(color = Exited, fill = Exited), position = "stack", alpha = 0.4)+
  geom_text(stat='count', aes(label=..count..), vjust=-1)
```



Correlations

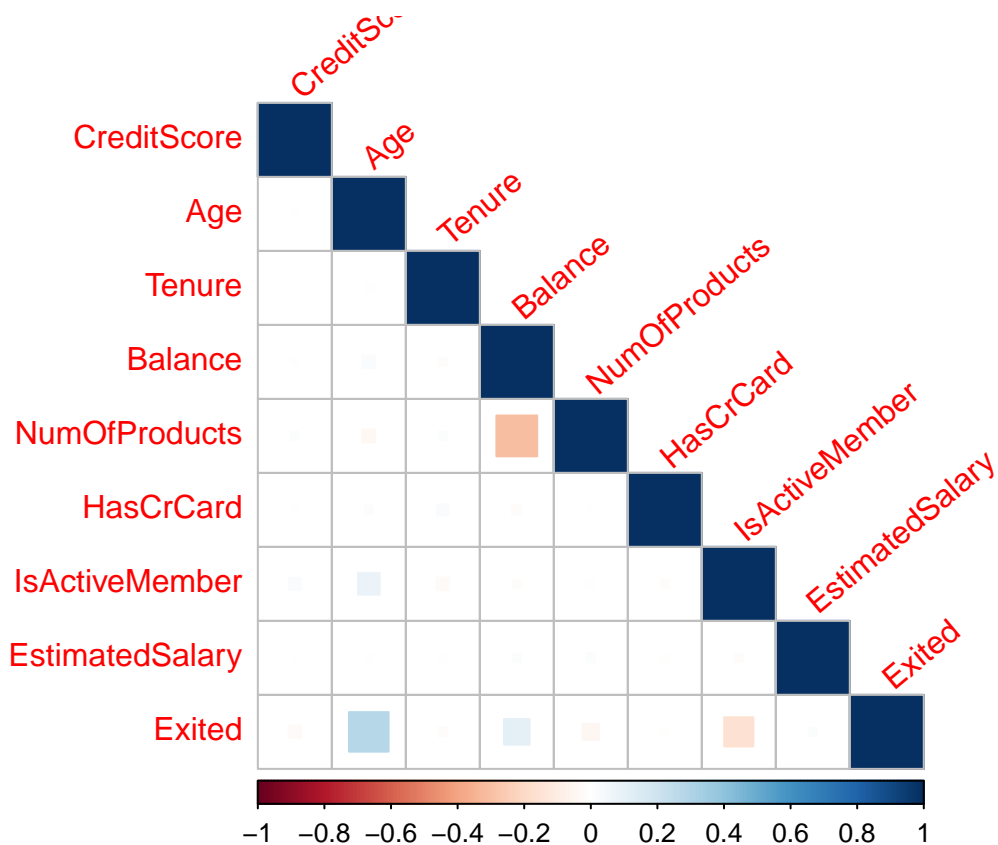
```
setwd("E:/W6400 MV/project")
churn <- read.csv("Churn_Modelling.csv")
churn <- subset(churn, select = -c(LineNumber, CustomerId, Surname))
# churn <- fastDummies::dummy_cols(churn)

# correlation analysis on numeric variables
excl_cor <- names(churn) %in% c('Geography', 'Gender')
cor_df <- churn[!excl_cor]
pearsoncorr<-cor(cor_df, use = "complete.obs") #default->pearson
res <- cor(cor_df)
round(res, 4)
```

	CreditScore	Age	Tenure	Balance	NumOfProducts	HasCrCard
## CreditScore	1.0000	-0.0040	0.0008	0.0063	0.0122	-0.0055
## Age	-0.0040	1.0000	-0.0100	0.0283	-0.0307	-0.0117
## Tenure	0.0008	-0.0100	1.0000	-0.0123	0.0134	0.0226
## Balance	0.0063	0.0283	-0.0123	1.0000	-0.3042	-0.0149
## NumOfProducts	0.0122	-0.0307	0.0134	-0.3042	1.0000	0.0032
## HasCrCard	-0.0055	-0.0117	0.0226	-0.0149	0.0032	1.0000
## IsActiveMember	0.0257	0.0855	-0.0284	-0.0101	0.0096	-0.0119
## EstimatedSalary	-0.0014	-0.0072	0.0078	0.0128	0.0142	-0.0099

```
## Exited          -0.0271  0.2853 -0.0140  0.1185          -0.0478   -0.0071
##               IsActiveMember EstimatedSalary Exited
## CreditScore      0.0257          -0.0014 -0.0271
## Age              0.0855          -0.0072  0.2853
## Tenure           -0.0284          0.0078 -0.0140
## Balance          -0.0101          0.0128  0.1185
## NumOfProducts    0.0096          0.0142 -0.0478
## HasCrCard        -0.0119         -0.0099 -0.0071
## IsActiveMember    1.0000         -0.0114 -0.1561
## EstimatedSalary  -0.0114          1.0000  0.0121
## Exited           -0.1561          0.0121  1.0000
```

```
# visualize correlation matrix
# library(corrplot)
par(mfrow = c(1,1))
corrplot(pearsoncorr, type= "lower", method="square", tl.srt=40)
```



Data transformation

```
churn$age_group <- cut(churn$Age, breaks = c(17, 25, 40, 60, 1000),
  labels=c('college', 'young', 'mid', 'old'))
churn$credit_group <- cut(churn$CreditScore, breaks = c(300, 560, 650, 700, 750, 850),
  labels=c('Very Bad', 'Bad', 'Fair', 'Good', 'Excellent'))
```

```

churn$balance_group <- cut(churn$Balance, breaks = c(-1, 0, 100000, 150000, 300000),
                           labels=c('None', 'Low', 'Mid', 'High'))
churn$product_group <- cut(churn$NumOfProducts, breaks = c(0, 1, 2, 5),
                           labels=c('1', '2', '3_and_over'))
churn$salary_group <- cut(churn$EstimatedSalary, breaks = c(0, 30000, 70000, 100000, 150000, 200000),
                          labels=c('Low', 'Medium_low', 'Medium', 'Medium_high', 'High'))

churn2 <- subset(churn, select = -c(CreditScore, Age, Balance, NumOfProducts, EstimatedSalary))

# churn2 <- subset(churn, select = c(age_group, credit_group, balance_group, Exited))
churn.for.best.logistic <- within(churn2, {
  y <- Exited
  Exited <- NULL
})
names(churn.for.best.logistic)

## [1] "Geography"      "Gender"          "Tenure"          "HasCrCard"
## [5] "IsActiveMember" "age_group"       "credit_group"    "balance_group"
## [9] "product_group"  "salary_group"   "y"

#table(churn.for.best.logistic$salary_group, churn.for.best.logistic$age_group)

```

Stepwise selection

```

# stepwise selection with all interactions
init_glm <- glm(y ~ ., data = churn.for.best.logistic)
stepwise <- step(init_glm, scope = . ~ .^2, direction = 'both')

## Start:  AIC=7021.16
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##       age_group + credit_group + balance_group + product_group +
##       salary_group
##
##              Df Deviance    AIC
## + IsActiveMember:age_group    3   1147.1 6779.2
## + age_group:product_group     6   1154.5 6849.3
## + Geography:balance_group     5   1156.8 6867.2
## + balance_group:product_group  6   1160.8 6904.1
## + IsActiveMember:product_group 2   1169.0 6966.6
## + Geography:product_group     4   1170.3 6981.9
## + Geography:age_group         6   1171.8 6998.0
## + age_group:balance_group     9   1171.4 7001.1
## + Gender:product_group        2   1174.7 7015.4
## + Geography:IsActiveMember    2   1174.8 7016.2
## + Gender:age_group            3   1174.7 7016.7
## + HasCrCard:balance_group     3   1174.8 7017.8
## - salary_group                4   1176.6 7019.3
## + IsActiveMember:balance_group 3   1175.0 7019.5
## - HasCrCard                   1   1175.9 7019.5
## + HasCrCard:IsActiveMember    1   1175.6 7020.9

```

```

## + HasCrCard:salary_group      4  1174.9 7020.9
## + Tenure:age_group            3  1175.2 7020.9
## - credit_group                4  1176.8 7021.1
## <none>                        1175.9 7021.2
## + Tenure:HasCrCard            1  1175.7 7021.4
## + HasCrCard:age_group         3  1175.2 7021.5
## - Tenure                      1  1176.2 7021.5
## + Gender:IsActiveMember       1  1175.8 7022.0
## + Gender:Tenure               1  1175.8 7022.5
## + Geography:Tenure            2  1175.6 7022.6
## + Gender:balance_group        3  1175.4 7022.7
## + Tenure:IsActiveMember       1  1175.9 7022.9
## + Gender:HasCrCard            1  1175.9 7023.2
## + Geography:salary_group      8  1174.2 7023.2
## + Tenure:product_group        2  1175.7 7023.8
## + age_group:credit_group     12  1173.4 7024.3
## + HasCrCard:product_group     2  1175.8 7024.7
## + Geography:Gender            2  1175.9 7025.0
## + Geography:HasCrCard         2  1175.9 7025.1
## + HasCrCard:credit_group      4  1175.5 7026.0
## + Tenure:credit_group         4  1175.5 7026.0
## + IsActiveMember:credit_group 4  1175.6 7026.9
## + Gender:credit_group         4  1175.6 7026.9
## + IsActiveMember:salary_group 4  1175.6 7026.9
## + Tenure:balance_group        3  1175.9 7027.1
## + Tenure:salary_group         4  1175.7 7027.5
## + Gender:salary_group         4  1175.8 7028.1
## + balance_group:salary_group  12  1174.1 7030.2
## + product_group:salary_group   8  1175.1 7030.6
## + credit_group:product_group   8  1175.3 7031.9
## + Geography:credit_group       8  1175.5 7034.0
## + credit_group:balance_group  12  1174.9 7037.1
## + age_group:salary_group      12  1175.1 7038.8
## + credit_group:salary_group   16  1174.3 7040.1
## - balance_group               3  1179.3 7044.5
## - Gender                     1  1185.9 7104.0
## - Geography                   2  1202.2 7238.5
## - IsActiveMember              1  1207.6 7285.0
## - age_group                   3  1294.3 7974.7
## - product_group               2  1347.0 8376.1
##
## Step:  AIC=6779.23
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##      age_group + credit_group + balance_group + product_group +
##      salary_group + IsActiveMember:age_group
##
##
##      Df Deviance    AIC
## + age_group:product_group      6  1125.6 6602.0
## + Geography:balance_group      5  1128.6 6626.8
## + balance_group:product_group  6  1133.2 6669.5
## + IsActiveMember:product_group  2  1141.8 6737.4
## + Geography:product_group      4  1142.0 6742.7
## + Geography:age_group          6  1143.4 6759.0
## + age_group:balance_group      9  1143.5 6765.7

```



```

## + Gender:product_group      2  1146.2 6775.1
## + Gender:age_group          3  1146.0 6775.4
## - salary_group              4  1147.6 6775.9
## + Tenure:age_group           3  1146.0 6776.0
## + HasCrCard:balance_group    3  1146.2 6777.6
## + Geography:IsActiveMember   2  1146.5 6778.0
## - HasCrCard                  1  1147.2 6778.0
## + HasCrCard:IsActiveMember   1  1146.8 6778.3
## - credit_group              4  1147.9 6778.5
## + HasCrCard:age_group        3  1146.3 6778.7
## + HasCrCard:salary_group     4  1146.1 6778.9
## <none>                      1147.1 6779.2
## - Tenure                     1  1147.4 6779.7
## + Geography:Tenure           2  1146.7 6780.0
## + Tenure:HasCrCard           1  1147.0 6780.0
## + IsActiveMember:balance_group 3  1146.5 6780.3
## + Geography:salary_group     8  1145.4 6780.5
## + Gender:Tenure              1  1147.0 6780.7
## + Gender:IsActiveMember      1  1147.0 6780.8
## + Tenure:IsActiveMember      1  1147.1 6780.9
## + Gender:HasCrCard           1  1147.1 6781.2
## + Tenure:product_group       2  1146.9 6781.4
## + Gender:balance_group       3  1146.8 6782.5
## + HasCrCard:product_group    2  1147.1 6783.1
## + Geography:HasCrCard        2  1147.1 6783.1
## + Geography:Gender           2  1147.1 6783.2
## + Tenure:credit_group        4  1146.7 6783.8
## + IsActiveMember:credit_group 4  1146.7 6783.8
## + HasCrCard:credit_group     4  1146.8 6784.4
## + IsActiveMember:salary_group 4  1146.8 6784.7
## + Gender:credit_group        4  1146.8 6784.7
## + Tenure:balance_group       3  1147.1 6785.2
## + Tenure:salary_group        4  1146.9 6785.7
## + Gender:salary_group        4  1146.9 6785.8
## + age_group:credit_group     12  1145.2 6786.6
## + product_group:salary_group  8  1146.5 6789.6
## + balance_group:salary_group 12  1145.6 6789.9
## + credit_group:product_group  8  1146.6 6790.8
## + Geography:credit_group     8  1146.8 6793.1
## + credit_group:balance_group 12  1146.1 6794.9
## + age_group:salary_group     12  1146.5 6797.8
## + credit_group:salary_group  16  1145.7 6798.9
## - balance_group              3  1150.7 6804.2
## - Gender                     1  1156.7 6860.3
## - Geography                   2  1171.9 6989.2
## - IsActiveMember:age_group    3  1175.9 7021.2
## - product_group              2  1310.2 8104.7
##
## Step:  AIC=6601.96
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + credit_group + balance_group + product_group +
##     salary_group + IsActiveMember:age_group + age_group:product_group
##
##                                     Df Deviance    AIC

```

## + Geography:balance_group	5	1107.3	6448.6
## + balance_group:product_group	6	1111.6	6488.7
## + IsActiveMember:product_group	2	1120.4	6560.0
## + Geography:product_group	4	1121.2	6571.1
## + Geography:age_group	6	1122.6	6587.7
## + Gender:age_group	3	1124.5	6598.1
## - salary_group	4	1126.2	6599.4
## + Gender:product_group	2	1124.9	6599.6
## + HasCrCard:balance_group	3	1124.7	6600.0
## + Tenure:age_group	3	1124.7	6600.1
## + Geography:IsActiveMember	2	1125.0	6600.7
## - HasCrCard	1	1125.7	6600.9
## + HasCrCard:age_group	3	1124.8	6601.1
## + HasCrCard:salary_group	4	1124.6	6601.2
## + IsActiveMember:balance_group	3	1124.9	6601.5
## - credit_group	4	1126.4	6601.6
## + HasCrCard:IsActiveMember	1	1125.3	6601.6
## <none>		1125.6	6602.0
## + Geography:Tenure	2	1125.1	6602.0
## - Tenure	1	1125.8	6602.3
## + Geography:salary_group	8	1123.9	6602.6
## + Tenure:HasCrCard	1	1125.5	6603.0
## + Gender:Tenure	1	1125.5	6603.3
## + Tenure:IsActiveMember	1	1125.5	6603.6
## + Gender:IsActiveMember	1	1125.5	6603.7
## + Tenure:product_group	2	1125.4	6603.9
## + Gender:HasCrCard	1	1125.6	6603.9
## + Gender:balance_group	3	1125.2	6604.8
## + HasCrCard:product_group	2	1125.5	6605.6
## + Geography:HasCrCard	2	1125.6	6605.8
## + Geography:Gender	2	1125.6	6606.0
## + Tenure:credit_group	4	1125.2	6606.1
## + IsActiveMember:credit_group	4	1125.2	6606.2
## + HasCrCard:credit_group	4	1125.2	6606.2
## + Gender:credit_group	4	1125.2	6607.0
## + Tenure:salary_group	4	1125.3	6607.3
## + Tenure:balance_group	3	1125.6	6607.8
## + IsActiveMember:salary_group	4	1125.3	6607.8
## + Gender:salary_group	4	1125.5	6608.9
## + age_group:credit_group	12	1123.7	6609.4
## + balance_group:salary_group	12	1124.0	6611.4
## + product_group:salary_group	8	1125.0	6612.4
## + age_group:balance_group	9	1124.8	6612.6
## + credit_group:product_group	8	1125.0	6612.9
## + Geography:credit_group	8	1125.3	6615.5
## + credit_group:balance_group	12	1124.6	6617.0
## + age_group:salary_group	12	1124.8	6619.2
## - balance_group	3	1128.4	6621.0
## + credit_group:salary_group	16	1124.3	6622.6
## - Gender	1	1135.3	6685.8
## - age_group:product_group	6	1147.1	6779.2
## - Geography	2	1148.5	6799.8
## - IsActiveMember:age_group	3	1154.5	6849.3
##			

```

## Step: AIC=6448.6
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + credit_group + balance_group + product_group +
##     salary_group + IsActiveMember:age_group + age_group:product_group +
##     Geography:balance_group
##
##
##           Df Deviance    AIC
## + balance_group:product_group  6   1091.1 6312.7
## + IsActiveMember:product_group  2   1102.3 6406.7
## + Geography:product_group      4   1103.5 6421.6
## + Geography:age_group          6   1104.6 6435.6
## + Gender:age_group             3   1106.2 6444.7
## + Tenure:age_group             3   1106.4 6446.3
## + Gender:product_group         2   1106.7 6446.3
## - salary_group                 4   1108.0 6446.7
## - HasCrCard                    1   1107.5 6447.9
## + Geography:IsActiveMember     2   1106.8 6447.9
## - credit_group                 4   1108.2 6448.1
## + HasCrCard:balance_group      3   1106.6 6448.2
## + HasCrCard:IsActiveMember     1   1107.1 6448.3
## + HasCrCard:salary_group       4   1106.4 6448.4
## + IsActiveMember:balance_group 3   1106.7 6448.6
## <none>                        1107.3 6448.6
## - Tenure                      1   1107.6 6448.8
## + Geography:Tenure             2   1106.9 6448.8
## + HasCrCard:age_group          3   1106.7 6449.0
## + Geography:salary_group       8   1105.7 6449.6
## + Tenure:HasCrCard             1   1107.3 6449.9
## + Gender:Tenure                1   1107.3 6450.0
## + Gender:IsActiveMember        1   1107.3 6450.3
## + Tenure:IsActiveMember        1   1107.3 6450.3
## + Tenure:product_group         2   1107.1 6450.5
## + Gender:HasCrCard             1   1107.3 6450.5
## + Gender:balance_group         3   1107.0 6451.0
## + HasCrCard:product_group      2   1107.3 6451.9
## + Geography:HasCrCard          2   1107.3 6452.2
## + HasCrCard:credit_group       4   1106.9 6452.5
## + Geography:Gender             2   1107.3 6452.6
## + IsActiveMember:credit_group  4   1106.9 6452.6
## + Tenure:credit_group          4   1107.0 6453.7
## + Tenure:salary_group          4   1107.1 6454.1
## + IsActiveMember:salary_group  4   1107.1 6454.1
## + Gender:credit_group          4   1107.1 6454.1
## + Tenure:balance_group         3   1107.3 6454.2
## + age_group:credit_group       12   1105.3 6454.5
## + Gender:salary_group          4   1107.2 6455.2
## + balance_group:salary_group   12   1105.6 6457.2
## + product_group:salary_group   8   1106.7 6458.9
## + credit_group:product_group   8   1106.7 6459.1
## + age_group:balance_group      9   1106.5 6459.4
## + Geography:credit_group       8   1107.0 6461.6
## + credit_group:balance_group   12   1106.4 6464.3
## + age_group:salary_group       12   1106.5 6465.4
## + credit_group:salary_group    16   1106.1 6469.0

```

```

## - Gender                                1    1116.9 6532.5
## - Geography:balance_group               5    1125.6 6602.0
## - age_group:product_group               6    1128.6 6626.8
## - IsActiveMember:age_group              3    1135.6 6694.9
##
## Step: AIC=6312.67
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##      age_group + credit_group + balance_group + product_group +
##      salary_group + IsActiveMember:age_group + age_group:product_group +
##      Geography:balance_group + balance_group:product_group
##
##                                     Df Deviance    AIC
## + Geography:product_group           4    1075.6 6177.7
## + IsActiveMember:product_group       2    1086.0 6269.6
## + Geography:age_group                6    1087.6 6293.0
## + Gender:age_group                   3    1090.1 6309.3
## + Gender:product_group               2    1090.4 6310.7
## - salary_group                       4    1091.8 6311.0
## + IsActiveMember:balance_group        3    1090.3 6311.1
## - credit_group                       4    1091.8 6311.4
## + Geography:IsActiveMember            2    1090.5 6311.5
## + Tenure:age_group                   3    1090.3 6311.6
## - HasCrCard                          1    1091.2 6312.0
## - Tenure                             1    1091.3 6312.4
## <none>                               1091.1 6312.7
## + HasCrCard:IsActiveMember            1    1090.9 6312.7
## + HasCrCard:balance_group             3    1090.4 6312.7
## + Geography:Tenure                    2    1090.7 6313.1
## + HasCrCard:salary_group              4    1090.3 6313.5
## + Gender:Tenure                       1    1091.0 6313.8
## + Geography:salary_group              8    1089.5 6313.8
## + HasCrCard:age_group                 3    1090.6 6314.0
## + Tenure:HasCrCard                    1    1091.0 6314.0
## + Tenure:IsActiveMember               1    1091.0 6314.3
## + Gender:IsActiveMember               1    1091.1 6314.4
## + Gender:HasCrCard                    1    1091.1 6314.6
## + Gender:balance_group                3    1090.7 6314.9
## + Tenure:product_group                2    1090.9 6315.0
## + HasCrCard:product_group             2    1091.0 6316.2
## + Geography:HasCrCard                 2    1091.0 6316.3
## + Geography:Gender                    2    1091.1 6316.6
## + HasCrCard:credit_group              4    1090.7 6316.7
## + IsActiveMember:credit_group          4    1090.7 6316.9
## + Tenure:credit_group                  4    1090.7 6317.4
## + Gender:credit_group                  4    1090.8 6318.0
## + Tenure:balance_group                 3    1091.0 6318.1
## + Tenure:salary_group                  4    1090.8 6318.2
## + IsActiveMember:salary_group          4    1090.8 6318.5
## + age_group:credit_group              12    1089.2 6319.5
## + Gender:salary_group                  4    1091.0 6319.8
## + balance_group:salary_group           12    1089.3 6320.0
## + age_group:balance_group              9    1090.1 6321.5
## + credit_group:product_group           8    1090.5 6323.1
## + product_group:salary_group           8    1090.5 6323.6

```

```

## + Geography:credit_group      8   1090.7 6325.5
## + credit_group:balance_group  12   1090.1 6327.5
## + age_group:salary_group      12   1090.4 6330.0
## + credit_group:salary_group   16   1089.8 6332.8
## - Gender                      1   1100.3 6395.2
## - balance_group:product_group  6   1107.3 6448.6
## - Geography:balance_group      5   1111.6 6488.7
## - age_group:product_group      6   1112.5 6495.3
## - IsActiveMember:age_group     3   1118.2 6552.0
##
## Step:  AIC=6177.68
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##      age_group + credit_group + balance_group + product_group +
##      salary_group + IsActiveMember:age_group + age_group:product_group +
##      Geography:balance_group + balance_group:product_group + Geography:product_group
##
##                                     Df Deviance    AIC
## + IsActiveMember:product_group    2   1070.6 6134.9
## + Geography:age_group              6   1072.6 6162.0
## + Gender:age_group                 3   1074.6 6174.7
## + HasCrCard:balance_group          3   1074.8 6176.3
## - salary_group                    4   1076.3 6176.7
## - HasCrCard                      1   1075.7 6176.7
## + Gender:product_group            2   1075.1 6176.7
## - credit_group                    4   1076.4 6176.8
## + IsActiveMember:balance_group     3   1074.9 6177.5
## - Tenure                         1   1075.8 6177.5
## + Tenure:age_group                3   1075.0 6177.7
## <none>                            1075.6 6177.7
## + Geography:Tenure                2   1075.2 6178.1
## + Geography:IsActiveMember         2   1075.2 6178.2
## + HasCrCard:IsActiveMember         1   1075.4 6178.2
## + HasCrCard:age_group              3   1075.1 6178.7
## + Geography:salary_group           8   1074.0 6178.8
## + Tenure:HasCrCard                 1   1075.5 6179.0
## + Gender:Tenure                    1   1075.5 6179.1
## + Tenure:IsActiveMember            1   1075.5 6179.1
## + HasCrCard:salary_group           4   1074.9 6179.2
## + Gender:IsActiveMember            1   1075.6 6179.6
## + Gender:HasCrCard                1   1075.6 6179.6
## + Tenure:product_group             2   1075.5 6180.5
## + Gender:balance_group             3   1075.3 6180.5
## + Geography:HasCrCard              2   1075.6 6181.4
## + HasCrCard:product_group          2   1075.6 6181.5
## + Geography:Gender                 2   1075.6 6181.6
## + IsActiveMember:credit_group      4   1075.2 6182.0
## + Tenure:credit_group              4   1075.2 6182.2
## + HasCrCard:credit_group           4   1075.3 6182.6
## + Tenure:balance_group             3   1075.5 6183.0
## + Gender:credit_group              4   1075.3 6183.0
## + Tenure:salary_group              4   1075.4 6183.7
## + IsActiveMember:salary_group      4   1075.4 6183.9
## + balance_group:salary_group      12   1073.7 6184.1
## + Gender:salary_group              4   1075.5 6184.9

```

```

## + age_group:credit_group      12  1074.0 6186.3
## + age_group:balance_group     9  1074.7 6186.9
## + credit_group:product_group  8  1075.0 6188.3
## + product_group:salary_group  8  1075.1 6189.3
## + Geography:credit_group      8  1075.2 6190.3
## + credit_group:balance_group  12  1074.5 6191.5
## + age_group:salary_group      12  1074.9 6195.2
## + credit_group:salary_group   16  1074.3 6197.8
## - Gender                      1  1084.1 6254.3
## - Geography:product_group     4  1091.1 6312.7
## - Geography:balance_group     5  1094.1 6338.3
## - age_group:product_group     6  1095.6 6350.1
## - IsActiveMember:age_group    3  1101.6 6410.9
## - balance_group:product_group  6  1103.5 6421.6
##
## Step:  AIC=6134.92
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##      age_group + credit_group + balance_group + product_group +
##      salary_group + IsActiveMember:age_group + age_group:product_group +
##      Geography:balance_group + balance_group:product_group + Geography:product_group +
##      IsActiveMember:product_group
##
##                                     Df Deviance    AIC
## + Geography:age_group              6  1067.6 6119.4
## + Gender:age_group                 3  1069.6 6131.9
## + HasCrCard:balance_group          3  1069.8 6133.5
## - HasCrCard                       1  1070.7 6134.0
## + Gender:product_group             2  1070.1 6134.4
## - credit_group                    4  1071.4 6134.5
## - salary_group                    4  1071.4 6134.6
## + Tenure:age_group                 3  1069.9 6134.6
## <none>                             1  1070.6 6134.9
## - Tenure                          1  1070.8 6134.9
## + Geography:Tenure                 2  1070.2 6135.4
## + HasCrCard:IsActiveMember         1  1070.4 6135.4
## + Geography:IsActiveMember         2  1070.2 6135.5
## + HasCrCard:age_group              3  1070.0 6135.5
## + Geography:salary_group           8  1069.0 6135.8
## + Tenure:IsActiveMember            1  1070.5 6136.2
## + Tenure:HasCrCard                 1  1070.5 6136.2
## + Gender:Tenure                    1  1070.5 6136.5
## + HasCrCard:salary_group           4  1069.9 6136.7
## + Gender:IsActiveMember            1  1070.6 6136.7
## + Gender:HasCrCard                 1  1070.6 6136.8
## + Tenure:product_group             2  1070.5 6137.7
## + Gender:balance_group             3  1070.2 6137.7
## + Geography:HasCrCard              2  1070.5 6138.5
## + HasCrCard:product_group          2  1070.6 6138.8
## + Geography:Gender                 2  1070.6 6138.9
## + Tenure:credit_group              4  1070.2 6139.3
## + IsActiveMember:credit_group      4  1070.2 6139.5
## + HasCrCard:credit_group           4  1070.2 6139.6
## + Tenure:balance_group             3  1070.5 6140.1
## + Gender:credit_group              4  1070.3 6140.4

```

```

## + IsActiveMember:balance_group 3 1070.5 6140.6
## + Tenure:salary_group 4 1070.3 6140.6
## + balance_group:salary_group 12 1068.7 6141.1
## + IsActiveMember:salary_group 4 1070.4 6141.2
## + Gender:salary_group 4 1070.5 6142.1
## + age_group:credit_group 12 1068.9 6143.3
## + age_group:balance_group 9 1069.6 6143.5
## + credit_group:product_group 8 1070.0 6145.8
## + product_group:salary_group 8 1070.1 6146.5
## + Geography:credit_group 8 1070.2 6147.5
## + credit_group:balance_group 12 1069.4 6147.8
## + age_group:salary_group 12 1069.9 6153.0
## + credit_group:salary_group 16 1069.4 6155.6
## - IsActiveMember:product_group 2 1075.6 6177.7
## - Gender 1 1079.2 6212.9
## - Geography:product_group 4 1086.0 6269.6
## - Geography:balance_group 5 1089.0 6295.7
## - age_group:product_group 6 1090.5 6306.9
## - IsActiveMember:age_group 3 1095.0 6354.2
## - balance_group:product_group 6 1098.5 6380.0
##
## Step: AIC=6119.4
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
## age_group + credit_group + balance_group + product_group +
## salary_group + IsActiveMember:age_group + age_group:product_group +
## Geography:balance_group + balance_group:product_group + Geography:product_group +
## IsActiveMember:product_group + Geography:age_group
##
## Df Deviance AIC
## + Gender:age_group 3 1066.8 6117.2
## + HasCrCard:balance_group 3 1066.8 6117.9
## - HasCrCard 1 1067.8 6118.6
## - credit_group 4 1068.4 6118.7
## + Gender:product_group 2 1067.1 6118.7
## + Tenure:age_group 3 1067.0 6119.2
## - salary_group 4 1068.5 6119.3
## + Geography:Tenure 2 1067.2 6119.3
## <none> 1067.6 6119.4
## + HasCrCard:age_group 3 1067.0 6119.6
## - Tenure 1 1067.9 6119.6
## + Geography:IsActiveMember 2 1067.2 6119.8
## + HasCrCard:IsActiveMember 1 1067.5 6119.9
## + Geography:salary_group 8 1066.0 6120.4
## + Tenure:IsActiveMember 1 1067.5 6120.6
## + Tenure:HasCrCard 1 1067.6 6120.7
## + HasCrCard:salary_group 4 1067.0 6120.9
## + Gender:Tenure 1 1067.6 6120.9
## + Gender:IsActiveMember 1 1067.6 6121.2
## + Gender:HasCrCard 1 1067.6 6121.3
## + Tenure:product_group 2 1067.5 6122.3
## + Gender:balance_group 3 1067.3 6122.4
## + Geography:HasCrCard 2 1067.6 6123.1
## + HasCrCard:product_group 2 1067.6 6123.4
## + Geography:Gender 2 1067.6 6123.4

```

```

## + IsActiveMember:credit_group 4 1067.2 6123.7
## + Tenure:credit_group 4 1067.2 6123.7
## + HasCrCard:credit_group 4 1067.3 6123.9
## + age_group:balance_group 9 1066.2 6124.3
## + Tenure:balance_group 3 1067.6 6124.7
## + Gender:credit_group 4 1067.4 6125.0
## + IsActiveMember:balance_group 3 1067.6 6125.0
## + Tenure:salary_group 4 1067.4 6125.2
## + IsActiveMember:salary_group 4 1067.5 6125.8
## + Gender:salary_group 4 1067.5 6126.5
## + balance_group:salary_group 12 1065.9 6126.8
## + age_group:credit_group 12 1066.1 6128.8
## + credit_group:product_group 8 1067.1 6130.0
## + product_group:salary_group 8 1067.2 6131.0
## + Geography:credit_group 8 1067.2 6131.8
## + credit_group:balance_group 12 1066.5 6132.3
## - Geography:age_group 6 1070.6 6134.9
## + age_group:salary_group 12 1067.0 6137.5
## + credit_group:salary_group 16 1066.5 6140.5
## - IsActiveMember:product_group 2 1072.6 6162.0
## - Gender 1 1076.3 6198.0
## - Geography:product_group 4 1082.6 6250.2
## - Geography:balance_group 5 1086.0 6279.8
## - age_group:product_group 6 1086.9 6286.0
## - IsActiveMember:age_group 3 1091.8 6336.9
## - balance_group:product_group 6 1096.1 6370.5
##
## Step: AIC=6117.2
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
## age_group + credit_group + balance_group + product_group +
## salary_group + IsActiveMember:age_group + age_group:product_group +
## Geography:balance_group + balance_group:product_group + Geography:product_group +
## IsActiveMember:product_group + Geography:age_group + Gender:age_group
##
##
## Df Deviance AIC
## + HasCrCard:balance_group 3 1065.9 6115.5
## - HasCrCard 1 1066.9 6116.4
## - credit_group 4 1067.6 6116.7
## + Tenure:age_group 3 1066.1 6117.2
## <none> 1066.8 6117.2
## + Geography:Tenure 2 1066.3 6117.3
## - salary_group 4 1067.6 6117.3
## - Tenure 1 1067.0 6117.4
## + HasCrCard:age_group 3 1066.1 6117.4
## + Gender:product_group 2 1066.4 6117.6
## + Geography:IsActiveMember 2 1066.4 6117.6
## + HasCrCard:IsActiveMember 1 1066.6 6117.7
## + Geography:salary_group 8 1065.2 6118.4
## + Tenure:IsActiveMember 1 1066.7 6118.4
## + Tenure:HasCrCard 1 1066.7 6118.4
## + Gender:Tenure 1 1066.7 6118.7
## + HasCrCard:salary_group 4 1066.1 6118.8
## + Gender:IsActiveMember 1 1066.8 6119.1
## + Gender:HasCrCard 1 1066.8 6119.1

```



```

## - Gender:age_group          3    1067.6 6119.4
## + Tenure:product_group      2    1066.7 6120.1
## + Gender:balance_group      3    1066.5 6120.5
## + Geography:HasCrCard       2    1066.7 6120.9
## + Geography:Gender          2    1066.8 6121.2
## + HasCrCard:product_group   2    1066.8 6121.2
## + HasCrCard:credit_group    4    1066.4 6121.4
## + IsActiveMember:credit_group 4    1066.4 6121.5
## + Tenure:credit_group       4    1066.4 6121.6
## + age_group:balance_group    9    1065.4 6122.3
## + Tenure:balance_group      3    1066.7 6122.5
## + Gender:credit_group       4    1066.5 6122.8
## + IsActiveMember:balance_group 3    1066.7 6122.9
## + Tenure:salary_group       4    1066.5 6123.0
## + IsActiveMember:salary_group 4    1066.6 6123.6
## + Gender:salary_group       4    1066.7 6124.3
## + balance_group:salary_group 12    1065.0 6124.7
## + age_group:credit_group     12    1065.2 6126.2
## + credit_group:product_group 8    1066.2 6128.0
## + product_group:salary_group 8    1066.3 6128.9
## + Geography:credit_group     8    1066.4 6129.6
## + credit_group:balance_group 12    1065.6 6130.2
## - Geography:age_group       6    1069.6 6131.9
## + age_group:salary_group     12    1066.1 6135.1
## + credit_group:salary_group  16    1065.6 6138.3
## - IsActiveMember:product_group 2    1071.8 6159.9
## - Geography:product_group    4    1081.6 6247.6
## - Geography:balance_group    5    1085.1 6277.4
## - age_group:product_group    6    1086.0 6284.0
## - IsActiveMember:age_group   3    1090.8 6334.2
## - balance_group:product_group 6    1095.1 6367.3
##
## Step:  AIC=6115.51
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + credit_group + balance_group + product_group +
##     salary_group + IsActiveMember:age_group + age_group:product_group +
##     Geography:balance_group + balance_group:product_group + Geography:product_group +
##     IsActiveMember:product_group + Geography:age_group + Gender:age_group +
##     HasCrCard:balance_group
##
##              Df Deviance    AIC
## - credit_group      4    1066.8 6115.2
## <none>              1065.9 6115.5
## + Geography:Tenure   2    1065.5 6115.5
## - salary_group      4    1066.8 6115.7
## - Tenure             1    1066.2 6115.7
## + Geography:IsActiveMember 2    1065.5 6115.7
## + HasCrCard:age_group  3    1065.3 6115.7
## + Tenure:age_group    3    1065.3 6115.7
## + Gender:product_group 2    1065.5 6115.8
## + HasCrCard:IsActiveMember 1    1065.8 6116.0
## + Geography:salary_group 8    1064.4 6116.7
## + Tenure:HasCrCard    1    1065.9 6116.7
## + Tenure:IsActiveMember 1    1065.9 6116.7

```

```

## + HasCrCard:salary_group      4  1065.2 6117.0
## + Gender:Tenure                1  1065.9 6117.0
## - HasCrCard:balance_group      3  1066.8 6117.2
## + Gender:IsActiveMember        1  1065.9 6117.3
## + Gender:HasCrCard             1  1065.9 6117.5
## - Gender:age_group             3  1066.8 6117.9
## + Tenure:product_group         2  1065.8 6118.4
## + Gender:balance_group         3  1065.7 6118.9
## + Geography:HasCrCard          2  1065.9 6119.0
## + Geography:Gender             2  1065.9 6119.4
## + HasCrCard:product_group      2  1065.9 6119.5
## + HasCrCard:credit_group       4  1065.5 6119.7
## + IsActiveMember:credit_group  4  1065.6 6119.9
## + Tenure:credit_group          4  1065.6 6119.9
## + age_group:balance_group      9  1064.6 6120.6
## + Tenure:balance_group         3  1065.9 6120.7
## + Gender:credit_group          4  1065.7 6121.0
## + IsActiveMember:balance_group 3  1065.9 6121.1
## + Tenure:salary_group          4  1065.7 6121.4
## + IsActiveMember:salary_group  4  1065.8 6121.9
## + Gender:salary_group          4  1065.8 6122.5
## + balance_group:salary_group   12  1064.2 6122.8
## + age_group:credit_group       12  1064.3 6124.1
## + credit_group:product_group   8  1065.4 6126.1
## + product_group:salary_group   8  1065.5 6127.1
## + Geography:credit_group       8  1065.6 6127.9
## + credit_group:balance_group   12  1064.8 6128.7
## - Geography:age_group          6  1068.8 6130.3
## + age_group:salary_group       12  1065.3 6133.1
## + credit_group:salary_group    16  1064.8 6136.6
## - IsActiveMember:product_group 2  1070.9 6158.3
## - Geography:product_group      4  1081.0 6247.3
## - Geography:balance_group      5  1084.0 6273.7
## - age_group:product_group      6  1085.2 6282.6
## - IsActiveMember:age_group     3  1089.7 6330.0
## - balance_group:product_group  6  1094.3 6366.0
##
## Step:  AIC=6115.2
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + balance_group + product_group + salary_group +
##     IsActiveMember:age_group + age_group:product_group + Geography:balance_group +
##     balance_group:product_group + Geography:product_group + IsActiveMember:product_group +
##     Geography:age_group + Gender:age_group + HasCrCard:balance_group
##
##
##              Df Deviance    AIC
## + Geography:Tenure      2  1066.3 6115.1
## - salary_group          4  1067.6 6115.1
## <none>                  1066.8 6115.2
## + HasCrCard:age_group   3  1066.1 6115.2
## + Geography:IsActiveMember 2  1066.3 6115.3
## - Tenure                1  1067.0 6115.3
## + Gender:product_group  2  1066.4 6115.4
## + Tenure:age_group      3  1066.2 6115.4
## + credit_group          4  1065.9 6115.5

```

```

## + HasCrCard:IsActiveMember      1  1066.6 6115.7
## + Geography:salary_group         8  1065.2 6116.2
## + Tenure:HasCrCard               1  1066.7 6116.4
## + Tenure:IsActiveMember          1  1066.7 6116.4
## - HasCrCard:balance_group        3  1067.6 6116.7
## + Gender:Tenure                  1  1066.7 6116.7
## + HasCrCard:salary_group         4  1066.1 6116.9
## + Gender:IsActiveMember          1  1066.7 6117.0
## + Gender:HasCrCard               1  1066.8 6117.1
## - Gender:age_group               3  1067.6 6117.4
## + Tenure:product_group           2  1066.6 6118.0
## + Gender:balance_group           3  1066.5 6118.7
## + Geography:HasCrCard            2  1066.7 6118.8
## + Geography:Gender               2  1066.8 6119.1
## + HasCrCard:product_group        2  1066.8 6119.2
## + age_group:balance_group        9  1065.3 6119.9
## + Tenure:balance_group           3  1066.7 6120.5
## + IsActiveMember:balance_group   3  1066.7 6120.8
## + Tenure:salary_group            4  1066.5 6121.0
## + IsActiveMember:salary_group     4  1066.6 6121.6
## + Gender:salary_group            4  1066.7 6122.2
## + balance_group:salary_group     12  1065.0 6122.4
## + product_group:salary_group      8  1066.3 6127.0
## - Geography:age_group            6  1069.7 6130.2
## + age_group:salary_group         12  1066.1 6132.7
## - IsActiveMember:product_group   2  1071.7 6157.5
## - Geography:product_group        4  1081.7 6246.6
## - Geography:balance_group        5  1084.9 6273.5
## - age_group:product_group        6  1086.0 6281.9
## - IsActiveMember:age_group       3  1090.6 6330.2
## - balance_group:product_group    6  1095.2 6366.4
##
## Step:  AIC=6115.05
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + balance_group + product_group + salary_group +
##     IsActiveMember:age_group + age_group:product_group + Geography:balance_group +
##     balance_group:product_group + Geography:product_group + IsActiveMember:product_group +
##     Geography:age_group + Gender:age_group + HasCrCard:balance_group +
##     Geography:Tenure
##
##                                     Df Deviance    AIC
## - salary_group                     4   1067.2 6114.9
## <none>                             0   1066.3 6115.1
## + Geography:IsActiveMember         2   1065.9 6115.1
## + HasCrCard:age_group              3   1065.7 6115.2
## - Geography:Tenure                 2   1066.8 6115.2
## + Tenure:age_group                 3   1065.7 6115.2
## + Gender:product_group             2   1065.9 6115.3
## + credit_group                    4   1065.5 6115.5
## + HasCrCard:IsActiveMember         1   1066.2 6115.6
## + Geography:salary_group           8   1064.7 6115.8
## + Tenure:HasCrCard                 1   1066.2 6116.2
## + Tenure:IsActiveMember            1   1066.2 6116.3
## - HasCrCard:balance_group          3   1067.1 6116.6

```

```

## + Gender:Tenure          1  1066.3 6116.6
## + HasCrCard:salary_group  4  1065.7 6116.8
## + Gender:IsActiveMember   1  1066.3 6116.9
## + Gender:HasCrCard        1  1066.3 6117.0
## - Gender:age_group        3  1067.2 6117.2
## + Tenure:product_group    2  1066.2 6117.9
## + Gender:balance_group    3  1066.1 6118.6
## + Geography:HasCrCard     2  1066.3 6118.6
## + Geography:Gender        2  1066.3 6119.0
## + HasCrCard:product_group  2  1066.3 6119.0
## + age_group:balance_group  9  1064.9 6119.7
## + Tenure:balance_group    3  1066.2 6120.2
## + IsActiveMember:balance_group 3  1066.3 6120.6
## + Tenure:salary_group     4  1066.1 6120.9
## + IsActiveMember:salary_group 4  1066.2 6121.4
## + Gender:salary_group     4  1066.2 6122.1
## + balance_group:salary_group 12 1064.5 6122.4
## + product_group:salary_group 8  1065.9 6126.9
## - Geography:age_group     6  1069.3 6130.6
## + age_group:salary_group   12 1065.6 6132.5
## - IsActiveMember:product_group 2  1071.3 6157.4
## - Geography:product_group  4  1081.3 6246.4
## - Geography:balance_group  5  1084.4 6273.2
## - age_group:product_group  6  1085.6 6282.5
## - IsActiveMember:age_group  3  1090.2 6330.8
## - balance_group:product_group 6  1094.7 6365.9
##
## Step:  AIC=6114.91
## y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + balance_group + product_group + IsActiveMember:age_group +
##     age_group:product_group + Geography:balance_group + balance_group:product_group +
##     Geography:product_group + IsActiveMember:product_group +
##     Geography:age_group + Gender:age_group + HasCrCard:balance_group +
##     Geography:Tenure
##
##              Df Deviance    AIC
## <none>              1067.2 6114.9
## + salary_group      4  1066.3 6115.1
## - Geography:Tenure   2  1067.6 6115.1
## + Gender:product_group 2  1066.8 6115.1
## + Geography:IsActiveMember 2  1066.8 6115.1
## + HasCrCard:age_group  3  1066.5 6115.1
## + Tenure:age_group     3  1066.5 6115.2
## + HasCrCard:IsActiveMember 1  1067.0 6115.3
## + credit_group       4  1066.4 6115.6
## + Tenure:HasCrCard    1  1067.1 6116.0
## + Tenure:IsActiveMember 1  1067.1 6116.1
## - HasCrCard:balance_group 3  1068.0 6116.4
## + Gender:Tenure      1  1067.1 6116.5
## + Gender:IsActiveMember 1  1067.1 6116.8
## + Gender:HasCrCard    1  1067.2 6116.8
## - Gender:age_group    3  1068.0 6116.9
## + Tenure:product_group 2  1067.0 6117.7
## + Gender:balance_group 3  1066.9 6118.5

```

```
## + Geography:HasCrCard          2   1067.1 6118.5
## + Geography:Gender              2   1067.2 6118.8
## + HasCrCard:product_group       2   1067.2 6118.9
## + age_group:balance_group       9   1065.7 6119.6
## + Tenure:balance_group          3   1067.1 6120.0
## + IsActiveMember:balance_group  3   1067.1 6120.5
## - Geography:age_group           6   1070.1 6130.2
## - IsActiveMember:product_group  2   1072.0 6156.6
## - Geography:product_group       4   1082.1 6245.6
## - Geography:balance_group       5   1085.2 6272.5
## - age_group:product_group       6   1086.4 6281.5
## - IsActiveMember:age_group      3   1091.3 6332.4
## - balance_group:product_group   6   1095.5 6364.9
```

```
#par(mfrow =c(2,2))
#plot(stepwise)
summary(stepwise)
```

```
##
## Call:
## glm(formula = y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##      age_group + balance_group + product_group + IsActiveMember:age_group +
##      age_group:product_group + Geography:balance_group + balance_group:product_group +
##      Geography:product_group + IsActiveMember:product_group +
##      Geography:age_group + Gender:age_group + HasCrCard:balance_group +
##      Geography:Tenure, data = churn.for.best.logistic)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03245  -0.15591  -0.04829   0.01858   1.02851
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value
## (Intercept)      0.2795075  0.0359634   7.772
## GeographyGermany -0.0356985  0.0435002  -0.821
## GeographySpain    0.0282177  0.0391330   0.721
## GenderMale        -0.0256926  0.0267857  -0.959
## Tenure            -0.0005017  0.0016050  -0.313
## HasCrCard         -0.0064802  0.0121012  -0.535
## IsActiveMember    -0.1127684  0.0275656  -4.091
## age_groupyoung     0.0575530  0.0332153   1.733
## age_groupmid       0.4308832  0.0344330  12.514
## age_groupold       0.6359256  0.0541604  11.742
## balance_groupLow   -0.1657709  0.0236138  -7.020
## balance_groupMid   -0.1901220  0.0193133  -9.844
## balance_groupHigh  -0.0882108  0.0278807  -3.164
## product_group2     -0.2355935  0.0305428  -7.714
## product_group3_and_over 0.2494432  0.1131835   2.204
## IsActiveMember:age_groupyoung 0.0193102  0.0279710   0.690
## IsActiveMember:age_groupmid -0.0991745  0.0292019  -3.396
## IsActiveMember:age_groupold -0.5364142  0.0481388 -11.143
## age_groupyoung:product_group2 -0.0574042  0.0282376  -2.033
## age_groupmid:product_group2 -0.2415484  0.0295985  -8.161
## age_groupold:product_group2 -0.0650227  0.0417557  -1.557
```

## age_groupyoung:product_group3_and_over	0.0448016	0.1112135	0.403
## age_groupmid:product_group3_and_over	-0.0933114	0.1102705	-0.846
## age_groupold:product_group3_and_over	-0.2568321	0.1334731	-1.924
## GeographyGermany:balance_groupLow	0.0740034	0.0314573	2.353
## GeographySpain:balance_groupLow	0.0810424	0.0259618	3.122
## GeographyGermany:balance_groupMid	0.2481801	0.0276188	8.986
## GeographySpain:balance_groupMid	0.0093290	0.0210091	0.444
## GeographyGermany:balance_groupHigh	NA	NA	NA
## GeographySpain:balance_groupHigh	0.0424415	0.0306579	1.384
## balance_groupLow:product_group2	0.2904972	0.0234404	12.393
## balance_groupMid:product_group2	0.2330857	0.0195779	11.906
## balance_groupHigh:product_group2	0.2843683	0.0272344	10.442
## balance_groupLow:product_group3_and_over	0.4430065	0.0630665	7.024
## balance_groupMid:product_group3_and_over	0.4062677	0.0530081	7.664
## balance_groupHigh:product_group3_and_over	0.5165661	0.0731387	7.063
## GeographyGermany:product_group2	-0.1881491	0.0195572	-9.620
## GeographySpain:product_group2	0.0305775	0.0187202	1.633
## GeographyGermany:product_group3_and_over	-0.2822204	0.0502185	-5.620
## GeographySpain:product_group3_and_over	0.0090144	0.0501226	0.180
## IsActiveMember:product_group2	0.0852167	0.0136106	6.261
## IsActiveMember:product_group3_and_over	0.1324298	0.0382624	3.461
## GeographyGermany:age_groupyoung	0.0307387	0.0350596	0.877
## GeographySpain:age_groupyoung	-0.0283123	0.0336447	-0.842
## GeographyGermany:age_groupmid	0.0954969	0.0362541	2.634
## GeographySpain:age_groupmid	-0.0518019	0.0352740	-1.469
## GeographyGermany:age_groupold	0.0470734	0.0505553	0.931
## GeographySpain:age_groupold	-0.0035854	0.0489638	-0.073
## GenderMale:age_groupyoung	-0.0236063	0.0281525	-0.839
## GenderMale:age_groupmid	-0.0593680	0.0292719	-2.028
## GenderMale:age_groupold	-0.0179202	0.0407719	-0.440
## HasCrCard:balance_groupLow	-0.0171057	0.0215086	-0.795
## HasCrCard:balance_groupMid	0.0164539	0.0167935	0.980
## HasCrCard:balance_groupHigh	-0.0491173	0.0259949	-1.889
## GeographyGermany:Tenure	0.0003742	0.0027490	0.136
## GeographySpain:Tenure	-0.0052646	0.0028157	-1.870
##	Pr(> t)		
## (Intercept)	8.49e-15	***	
## GeographyGermany	0.411865		
## GeographySpain	0.470883		
## GenderMale	0.337485		
## Tenure	0.754612		
## HasCrCard	0.592319		
## IsActiveMember	4.33e-05	***	
## age_groupyoung	0.083175	.	
## age_groupmid	< 2e-16	***	
## age_groupold	< 2e-16	***	
## balance_groupLow	2.36e-12	***	
## balance_groupMid	< 2e-16	***	
## balance_groupHigh	0.001562	**	
## product_group2	1.34e-14	***	
## product_group3_and_over	0.027555	*	
## IsActiveMember:age_groupyoung	0.489980		
## IsActiveMember:age_groupmid	0.000686	***	
## IsActiveMember:age_groupold	< 2e-16	***	

```

## age_groupyoung:product_group2          0.042090 *
## age_groupmid:product_group2             3.73e-16 ***
## age_groupold:product_group2             0.119451
## age_groupyoung:product_group3_and_over  0.687072
## age_groupmid:product_group3_and_over    0.397459
## age_groupold:product_group3_and_over    0.054355 .
## GeographyGermany:balance_groupLow       0.018667 *
## GeographySpain:balance_groupLow         0.001804 **
## GeographyGermany:balance_groupMid       < 2e-16 ***
## GeographySpain:balance_groupMid         0.657019
## GeographyGermany:balance_groupHigh      NA
## GeographySpain:balance_groupHigh        0.166281
## balance_groupLow:product_group2        < 2e-16 ***
## balance_groupMid:product_group2        < 2e-16 ***
## balance_groupHigh:product_group2       < 2e-16 ***
## balance_groupLow:product_group3_and_over 2.29e-12 ***
## balance_groupMid:product_group3_and_over 1.97e-14 ***
## balance_groupHigh:product_group3_and_over 1.74e-12 ***
## GeographyGermany:product_group2        < 2e-16 ***
## GeographySpain:product_group2          0.102417
## GeographyGermany:product_group3_and_over 1.96e-08 ***
## GeographySpain:product_group3_and_over  0.857277
## IsActiveMember:product_group2         3.98e-10 ***
## IsActiveMember:product_group3_and_over  0.000540 ***
## GeographyGermany:age_groupyoung        0.380640
## GeographySpain:age_groupyoung          0.400084
## GeographyGermany:age_groupmid          0.008449 **
## GeographySpain:age_groupmid            0.141984
## GeographyGermany:age_groupold          0.351811
## GeographySpain:age_groupold            0.941629
## GenderMale:age_groupyoung             0.401762
## GenderMale:age_groupmid               0.042571 *
## GenderMale:age_groupold               0.660291
## HasCrCard:balance_groupLow            0.426459
## HasCrCard:balance_groupMid            0.327219
## HasCrCard:balance_groupHigh           0.058855 .
## GeographyGermany:Tenure               0.891733
## GeographySpain:Tenure                 0.061553 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.107306)
##
##    Null deviance: 1622.1  on 9999  degrees of freedom
## Residual deviance: 1067.2  on 9945  degrees of freedom
## AIC: 6114.9
##
## Number of Fisher Scoring iterations: 2

```

Train test split

```
set.seed(123)
inds = sample(1:nrow(churn.for.best.logistic), 0.80 * nrow(churn.for.best.logistic))
train_churn <- churn.for.best.logistic[inds,]
test_churn <- churn.for.best.logistic[-inds,]
```

Logistic without interaction

```
glm1 <- glm(y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
  age_group + balance_group + product_group, data = train_churn, family = binomial)
summary(glm1)
```

```
##
## Call:
## glm(formula = y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##     age_group + balance_group + product_group, family = binomial,
##     data = train_churn)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7935  -0.5790  -0.3599  -0.1761   3.0995
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.45686    0.22121  -6.586 4.52e-11 ***
## GeographyGermany    0.89877    0.08435  10.655 < 2e-16 ***
## GeographySpain     0.03255    0.08517   0.382 0.70231
## GenderMale        -0.53460    0.06627  -8.068 7.17e-16 ***
## Tenure            -0.01889    0.01136  -1.663 0.09635 .
## HasCrCard         -0.06716    0.07194  -0.934 0.35053
## IsActiveMember    -0.95218    0.06861 -13.878 < 2e-16 ***
## age_groupyoung     0.45541    0.19695   2.312 0.02076 *
## age_groupmid       2.10763    0.19639  10.732 < 2e-16 ***
## age_groupold       1.66327    0.23681   7.024 2.16e-12 ***
## balance_groupLow   -0.33749    0.11159  -3.024 0.00249 **
## balance_groupMid   -0.06540    0.09422  -0.694 0.48761
## balance_groupHigh  -0.18944    0.12506  -1.515 0.12985
## product_group2     -1.53303    0.08114 -18.894 < 2e-16 ***
## product_group3_and_over 2.75549    0.20085  13.719 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 8099.8  on 7999  degrees of freedom
## Residual deviance: 5902.8  on 7985  degrees of freedom
## AIC: 5932.8
##
## Number of Fisher Scoring iterations: 5
```


Without interaction predictions

```
preds1 <- predict(glm1, newdata = test_churn, type = 'response')
preds1_encode <- ifelse(preds1 >= 0.5, 1, 0)
sum(preds1_encode == test_churn$y)
```

```
## [1] 1697
```

```
sum(preds1_encode == test_churn$y) / nrow(test_churn)
```

```
## [1] 0.8485
```

```
table(Predicted=preds1_encode, Actual=test_churn$y) #confusion matrix
```

```
##           Actual
## Predicted    0    1
##           0 1532  238
##           1   65  165
```

Logistic with interaction

```
glm2 <- glm(y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
  age_group + balance_group + product_group + IsActiveMember:age_group +
  age_group:product_group + Geography:balance_group + balance_group:product_group +
  Geography:product_group + IsActiveMember:product_group +
  Geography:age_group + Gender:age_group + HasCrCard:balance_group +
  Geography:Tenure, data = train_churn, family = binomial)
summary(glm2)
```

```
##
## Call:
## glm(formula = y ~ Geography + Gender + Tenure + HasCrCard + IsActiveMember +
##   age_group + balance_group + product_group + IsActiveMember:age_group +
##   age_group:product_group + Geography:balance_group + balance_group:product_group +
##   Geography:product_group + IsActiveMember:product_group +
##   Geography:age_group + Gender:age_group + HasCrCard:balance_group +
##   Geography:Tenure, family = binomial, data = train_churn)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8096  -0.5331  -0.3148  -0.1530   3.1470
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.00237    0.44478  -2.254  0.024220
## GeographyGermany -0.73495    0.54684  -1.344  0.178957
## GeographySpain   0.45593    0.51354   0.888  0.374634
## GenderMale     -0.42494    0.38409  -1.106  0.268573
```

## Tenure	-0.01501	0.01785	-0.841	0.400399
## HasCrCard	-0.12249	0.15138	-0.809	0.418442
## IsActiveMember	-1.11630	0.41123	-2.715	0.006636
## age_groupyoung	0.50057	0.42612	1.175	0.240106
## age_groupmid	2.36792	0.42674	5.549	2.87e-08
## age_groupold	3.73337	0.62567	5.967	2.42e-09
## balance_groupLow	-0.89541	0.23345	-3.836	0.000125
## balance_groupMid	-1.31029	0.19451	-6.736	1.62e-11
## balance_groupHigh	-0.40656	0.25212	-1.613	0.106843
## product_group2	-2.38947	0.45198	-5.287	1.25e-07
## product_group3_and_over	1.61793	0.98375	1.645	0.100040
## IsActiveMember:age_groupyoung	0.40182	0.42135	0.954	0.340263
## IsActiveMember:age_groupmid	0.08813	0.42013	0.210	0.833855
## IsActiveMember:age_groupold	-2.78613	0.58786	-4.739	2.14e-06
## age_groupyoung:product_group2	-0.43520	0.44132	-0.986	0.324058
## age_groupmid:product_group2	-0.62641	0.43942	-1.426	0.154000
## age_groupold:product_group2	-0.03415	0.58225	-0.059	0.953230
## age_groupyoung:product_group3_and_over	-0.44656	0.95228	-0.469	0.639111
## age_groupmid:product_group3_and_over	-0.03354	1.00661	-0.033	0.973416
## age_groupold:product_group3_and_over	-0.86230	1.24766	-0.691	0.489483
## GeographyGermany:balance_groupLow	0.81035	0.32861	2.466	0.013663
## GeographySpain:balance_groupLow	0.61513	0.24883	2.472	0.013432
## GeographyGermany:balance_groupMid	2.21496	0.28616	7.740	9.92e-15
## GeographySpain:balance_groupMid	-0.14391	0.22796	-0.631	0.527846
## GeographyGermany:balance_groupHigh	NA	NA	NA	NA
## GeographySpain:balance_groupHigh	0.25468	0.28089	0.907	0.364564
## balance_groupLow:product_group2	2.43802	0.25662	9.500	< 2e-16
## balance_groupMid:product_group2	2.54648	0.22753	11.192	< 2e-16
## balance_groupHigh:product_group2	2.50210	0.28390	8.813	< 2e-16
## balance_groupLow:product_group3_and_over	16.85306	248.66420	0.068	0.945965
## balance_groupMid:product_group3_and_over	16.28935	248.66283	0.066	0.947770
## balance_groupHigh:product_group3_and_over	17.63116	248.66459	0.071	0.943475
## GeographyGermany:product_group2	-1.42658	0.21371	-6.675	2.47e-11
## GeographySpain:product_group2	0.11158	0.20430	0.546	0.584965
## GeographyGermany:product_group3_and_over	-15.30765	248.66304	-0.062	0.950913
## GeographySpain:product_group3_and_over	0.21723	0.58225	0.373	0.709084
## IsActiveMember:product_group2	0.21594	0.16213	1.332	0.182903
## IsActiveMember:product_group3_and_over	0.23310	0.46501	0.501	0.616169
## GeographyGermany:age_groupyoung	0.20594	0.48023	0.429	0.668037
## GeographySpain:age_groupyoung	-0.34381	0.48206	-0.713	0.475718
## GeographyGermany:age_groupmid	0.23027	0.48074	0.479	0.631942
## GeographySpain:age_groupmid	-0.60420	0.48025	-1.258	0.208359
## GeographyGermany:age_groupold	0.20802	0.63851	0.326	0.744577
## GeographySpain:age_groupold	0.14029	0.62380	0.225	0.822058
## GenderMale:age_groupyoung	-0.13261	0.39785	-0.333	0.738889
## GenderMale:age_groupmid	-0.07749	0.39700	-0.195	0.845254
## GenderMale:age_groupold	-0.03281	0.52219	-0.063	0.949895
## HasCrCard:balance_groupLow	-0.17307	0.23077	-0.750	0.453266
## HasCrCard:balance_groupMid	0.22490	0.18948	1.187	0.235263
## HasCrCard:balance_groupHigh	-0.28918	0.25920	-1.116	0.264556
## GeographyGermany:Tenure	0.01439	0.02721	0.529	0.596820
## GeographySpain:Tenure	-0.03250	0.03090	-1.052	0.292879
##				
## (Intercept)	*			

```

## GeographyGermany
## GeographySpain
## GenderMale
## Tenure
## HasCrCard
## IsActiveMember          **
## age_groupyoung
## age_groupmid            ***
## age_groupold            ***
## balance_groupLow        ***
## balance_groupMid        ***
## balance_groupHigh
## product_group2          ***
## product_group3_and_over
## IsActiveMember:age_groupyoung
## IsActiveMember:age_groupmid
## IsActiveMember:age_groupold      ***
## age_groupyoung:product_group2
## age_groupmid:product_group2
## age_groupold:product_group2
## age_groupyoung:product_group3_and_over
## age_groupmid:product_group3_and_over
## age_groupold:product_group3_and_over
## GeographyGermany:balance_groupLow      *
## GeographySpain:balance_groupLow        *
## GeographyGermany:balance_groupMid      ***
## GeographySpain:balance_groupMid
## GeographyGermany:balance_groupHigh
## GeographySpain:balance_groupHigh
## balance_groupLow:product_group2        ***
## balance_groupMid:product_group2        ***
## balance_groupHigh:product_group2       ***
## balance_groupLow:product_group3_and_over
## balance_groupMid:product_group3_and_over
## balance_groupHigh:product_group3_and_over
## GeographyGermany:product_group2      ***
## GeographySpain:product_group2
## GeographyGermany:product_group3_and_over
## GeographySpain:product_group3_and_over
## IsActiveMember:product_group2
## IsActiveMember:product_group3_and_over
## GeographyGermany:age_groupyoung
## GeographySpain:age_groupyoung
## GeographyGermany:age_groupmid
## GeographySpain:age_groupmid
## GeographyGermany:age_groupold
## GeographySpain:age_groupold
## GenderMale:age_groupyoung
## GenderMale:age_groupmid
## GenderMale:age_groupold
## HasCrCard:balance_groupLow
## HasCrCard:balance_groupMid
## HasCrCard:balance_groupHigh
## GeographyGermany:Tenure

```

```
## GeographySpain:Tenure
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 8099.8  on 7999  degrees of freedom
## Residual deviance: 5462.4  on 7945  degrees of freedom
## AIC: 5572.4
##
## Number of Fisher Scoring iterations: 15
```

With interaction predictions

```
preds2 <- predict(glm2, newdata = test_churn, type = 'response')
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type == :
## prediction from a rank-deficient fit may be misleading
```

```
preds2_encode <- ifelse(preds2 >= 0.5, 1, 0)
sum(preds2_encode == test_churn$y)
```

```
## [1] 1728
```

```
sum(preds2_encode == test_churn$y) / nrow(test_churn)
```

```
## [1] 0.864
```

```
table(Predicted=preds2_encode, Actual=test_churn$y) # confusion matrix
```

```
##           Actual
## Predicted    0    1
##           0 1533 208
##           1   64 195
```

Interaction plot example

```
inter1 <- glm(y ~ balance_group*product_group, data = train_churn)
summary(inter1)
```

```
##
## Call:
## glm(formula = y ~ balance_group * product_group, data = train_churn)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.95370  -0.23819  -0.13386  -0.03465   0.96535
```

```
##
## Coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.36981    0.01373   26.932 < 2e-16
## balance_groupLow -0.15000    0.01879   -7.984 1.61e-15
## balance_groupMid -0.09487    0.01596   -5.943 2.91e-09
## balance_groupHigh -0.13162    0.02137   -6.160 7.63e-10
## product_group2    -0.33516    0.01594  -21.027 < 2e-16
## product_group3_and_over 0.33134    0.04187    7.913 2.85e-15
## balance_groupLow:product_group2 0.22508    0.02753    8.175 3.40e-16
## balance_groupMid:product_group2 0.20371    0.02172    9.379 < 2e-16
## balance_groupHigh:product_group2 0.23083    0.03253    7.096 1.39e-12
## balance_groupLow:product_group3_and_over 0.39329    0.07549    5.210 1.94e-07
## balance_groupMid:product_group3_and_over 0.34742    0.05550    6.260 4.05e-10
## balance_groupHigh:product_group3_and_over 0.34713    0.08771    3.958 7.64e-05
##
## (Intercept)      ***
## balance_groupLow  ***
## balance_groupMid  ***
## balance_groupHigh ***
## product_group2    ***
## product_group3_and_over ***
## balance_groupLow:product_group2 ***
## balance_groupMid:product_group2 ***
## balance_groupHigh:product_group2 ***
## balance_groupLow:product_group3_and_over ***
## balance_groupMid:product_group3_and_over ***
## balance_groupHigh:product_group3_and_over ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.1361332)
##
##    Null deviance: 1300.3  on 7999  degrees of freedom
## Residual deviance: 1087.4  on 7988  degrees of freedom
## AIC: 6764
##
## Number of Fisher Scoring iterations: 2
```

```
plot_model(inter1, type = "int")
```

