AED

Wenjia Ye

2025-10-08

```r
load("C:/Users/34688/OneDrive - Universitat de Barcelona/Escritorio/Mineria/DATA/data.RData")
library(ggplot2)
library(psych)
```

Anàlisis univariante

Univariante Numèrica

Univariante Categòrica

Anàlisis bivariante

Es realizarà un estudi a dos dimensions per trobar possibles relacions interessants.

Numèrica i numèrica

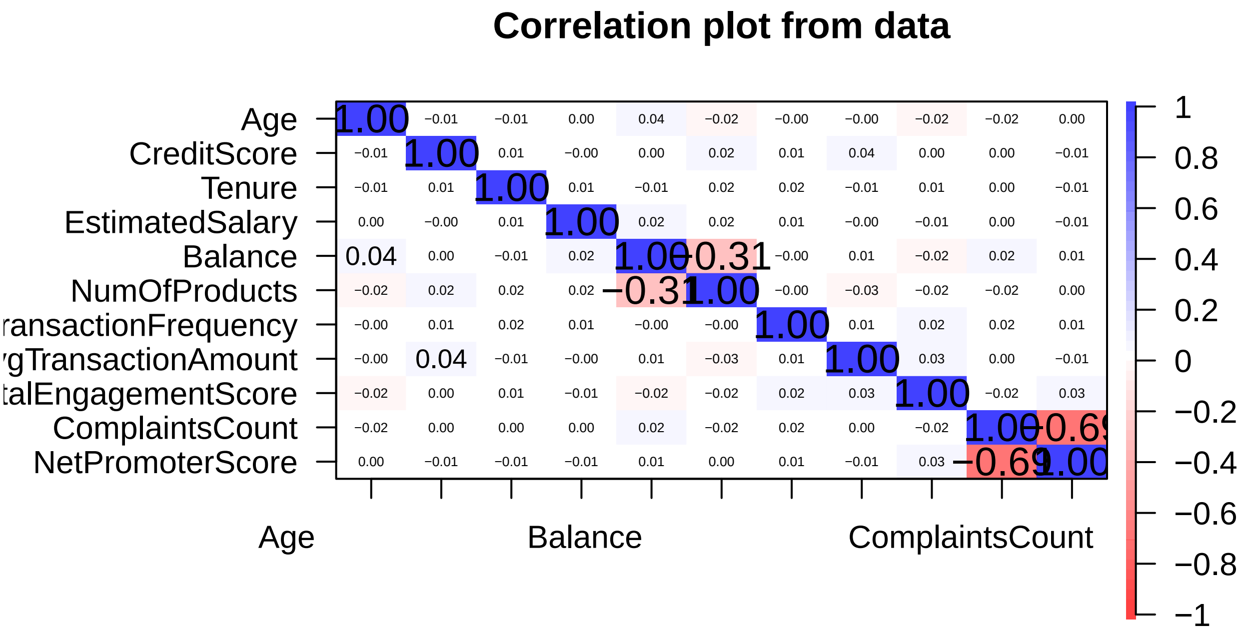
Estudiarem la correlació entre les varibales numèriques.

```r
cor(data[, varNum],use = "pairwise.complete.obs")
```

```
##                                 Age   CreditScore       Tenure EstimatedSalary
## Age                    1.0000000000 -1.288793e-02 -0.011927920    0.0043640775
## CreditScore           -0.0128879311  1.000000e+00  0.012452797   -0.0004280996
## Tenure                -0.0119279205  1.245280e-02  1.000000000    0.0123469012
## EstimatedSalary        0.0043640775 -4.280996e-04  0.012346901    1.0000000000
## Balance                0.0370722102  3.365710e-03 -0.007427842    0.0233364866
## NumOfProducts         -0.0236431172  2.147604e-02  0.017258959    0.0185216576
## TransactionFrequency  -0.0015204986  7.914231e-03  0.018024611    0.0086363209
## AvgTransactionAmount  -0.0011060698  3.931344e-02 -0.006490741   -0.0002324889
## DigitalEngagementScore -0.0208434404  9.183578e-05  0.008434622   -0.0085572580
## ComplaintsCount       -0.0150222405  3.475790e-04  0.002818873    0.0007486169
## NetPromoterScore       0.0009555647 -1.192457e-02 -0.006851175   -0.0099860853
##                             Balance NumOfProducts TransactionFrequency
## Age                     0.037072210  -0.023643117          -0.001520499
## CreditScore             0.003365710   0.021476040           0.007914231
## Tenure                 -0.007427842   0.017258959           0.018024611
## EstimatedSalary         0.023336487   0.018521658           0.008636321
## Balance                 1.000000000  -0.314486243          -0.003322540
## NumOfProducts          -0.314486243   1.000000000          -0.002818329
## TransactionFrequency   -0.003322540  -0.002818329           1.000000000
## AvgTransactionAmount    0.005062801  -0.025388200           0.006821174
## DigitalEngagementScore -0.019772640  -0.015163911           0.020503983
## ComplaintsCount         0.021986472  -0.018400137           0.016389856
```

```
## NetPromoterScore        0.012125535  0.002440939          0.007265435
##                        AvgTransactionAmount DigitalEngagementScore
## Age                            -0.0011060698          -2.084344e-02
## CreditScore                     0.0393134350           9.183578e-05
## Tenure                         -0.0064907411           8.434622e-03
## EstimatedSalary                -0.0002324889          -8.557258e-03
## Balance                         0.0050628014          -1.977264e-02
## NumOfProducts                  -0.0253881998          -1.516391e-02
## TransactionFrequency            0.0068211743           2.050398e-02
## AvgTransactionAmount            1.0000000000           2.727683e-02
## DigitalEngagementScore          0.0272768254           1.000000e+00
## ComplaintsCount                 0.0024675876          -1.649286e-02
## NetPromoterScore               -0.0125723111           2.514659e-02
##                        ComplaintsCount NetPromoterScore
## Age                      -0.0150222405     0.0009555647
## CreditScore               0.0003475790    -0.0119245738
## Tenure                    0.0028188733    -0.0068511752
## EstimatedSalary           0.0007486169    -0.0099860853
## Balance                   0.0219864718     0.0121255346
## NumOfProducts            -0.0184001371     0.0024409386
## TransactionFrequency      0.0163898562     0.0072654348
## AvgTransactionAmount      0.0024675876    -0.0125723111
## DigitalEngagementScore   -0.0164928593     0.0251465852
## ComplaintsCount           1.0000000000    -0.6914615774
## NetPromoterScore         -0.6914615774     1.0000000000
```

```r
corPlot(data[, varNum])
```



S'observa que majoritariament les variables no estan correlacionades, totes amb valors propers a 0. Excepte dos casos, una correlació negativa entre ComplaintsCount i NetPromoterScore de -0.69 i Balance amb NumOfProducts de -0.31. Com s'ha comentat previament, la varible NumOfProducts i ComplaintsCount possiblement es transformarà com a variable categòrica, llavors, no es realitzà cap canvi aqui.

Categòrica i categòrica

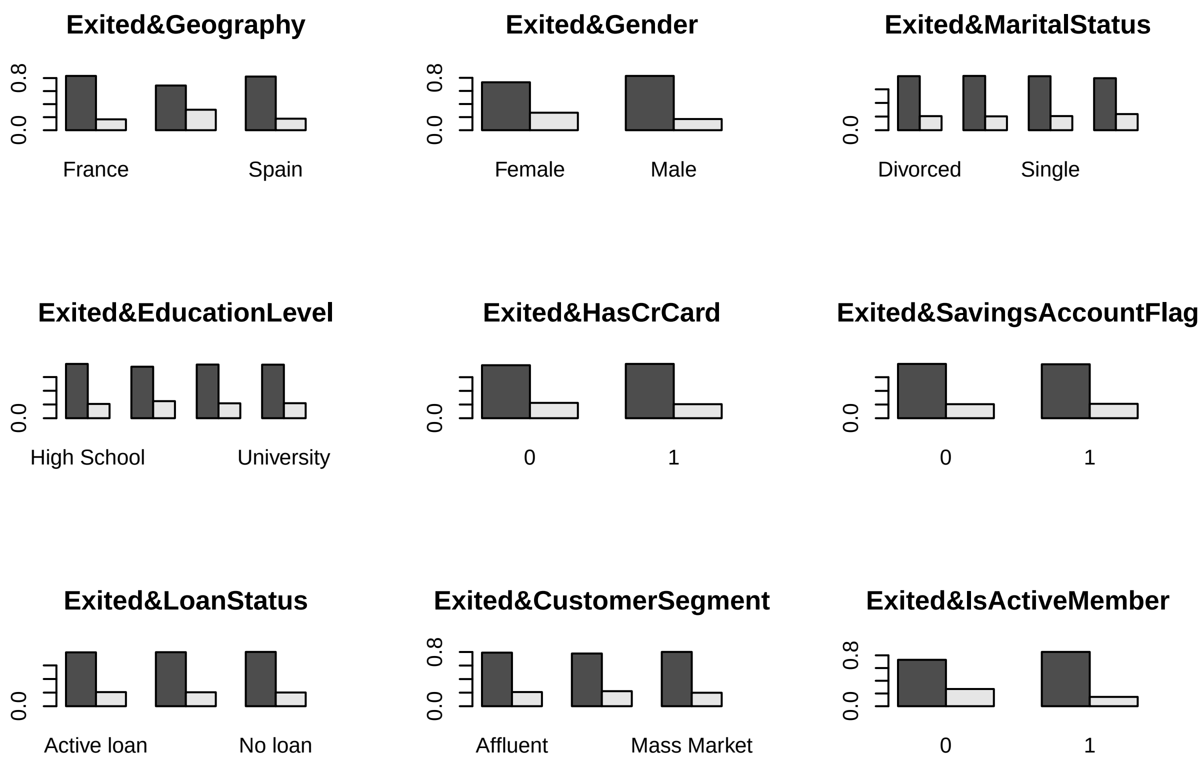
S'ha estudiat la relació entre la varible resposta Exited i la resta de variables categòriques.

```
cat<-varCat[-10]
v<-"Exited"
par(mfrow = c(3, 3))
for (varc1 in cat) {
    if (varc1 != v) {
      prop_table <- prop.table(table(data[, v], data[, varc1]),margin = 2)
      print(prop_table)
      barplot(prop_table, beside = TRUE,main = paste0(v,"&",varc1))
    }
}
```

```
##
##        France   Germany     Spain
##   0 0.8331288 0.6858766 0.8233851
##   1 0.1668712 0.3141234 0.1766149
```

```
##
##        Female      Male
##   0 0.7326467 0.8293963
```

```
##   1 0.2673533 0.1706037
##
##      Divorced   Married    Single   Widowed
##   0 0.7932489 0.7966903 0.7925718 0.7633929
##   1 0.2067511 0.2033097 0.2074282 0.2366071
##
##    High School     Other Postgraduate University
##   0   0.7919180 0.7511521    0.7826087  0.7811271
##   1   0.2080820 0.2488479    0.2173913  0.2188729
##
##             0         1
##   0 0.7756498 0.7946481
##   1 0.2243502 0.2053519
##
##             0         1
##   0 0.7945455 0.7895385
##   1 0.2054545 0.2104615
##
##    Active loan Default risk   No loan
##   0   0.7923729    0.7947154 0.7984626
##   1   0.2076271    0.2052846 0.2015374
##
##      Affluent High Net Worth Mass Market
##   0 0.7908127      0.7778970   0.8014101
##   1 0.2091873      0.2221030   0.1985899
##
##             0         1
##   0 0.7297412 0.8537161
##   1 0.2702588 0.1462839
```



S'observa que la distribució de la variable respota és molt similar dins de cada grup excepte alguns. A través del gràfic de barres es veu la variable Geography, IsActiveMember i Gender tenen una diferencia considerable i obtenim les conclusions següents.

- Els d'origen alemany tenen aproximandament 1/3 de probalitat de marxar.
- Les dones marxen amb més probabilitat.
- Els membres no actius tenen una prob més elevada de marxar.

S'ha realitzat el test de chi-quadrat per tal de veure si la diferencia observada és significativa.

```r
for (varc1 in cat) {
  if (varc1 != v) {
    tab <- table(data[[v]], data[[varc1]])
    test <- chisq.test(tab, correct = FALSE)

    cat("\nVariable:", varc1, "\n")
    cat("Chi-squared =", round(test$statistic, 3),
        "df =", test$parameter,
        "p-value =", signif(test$p.value, 5), "\n")

    if (test$p.value < 0.05) cat("-> Diferencias significativas entre columnas\n")
  }
}
```

```
##
## Variable: Geography
## Chi-squared = 117.259 df = 2 p-value = 3.4475e-26
## -> Diferencias significativas entre columnas
```

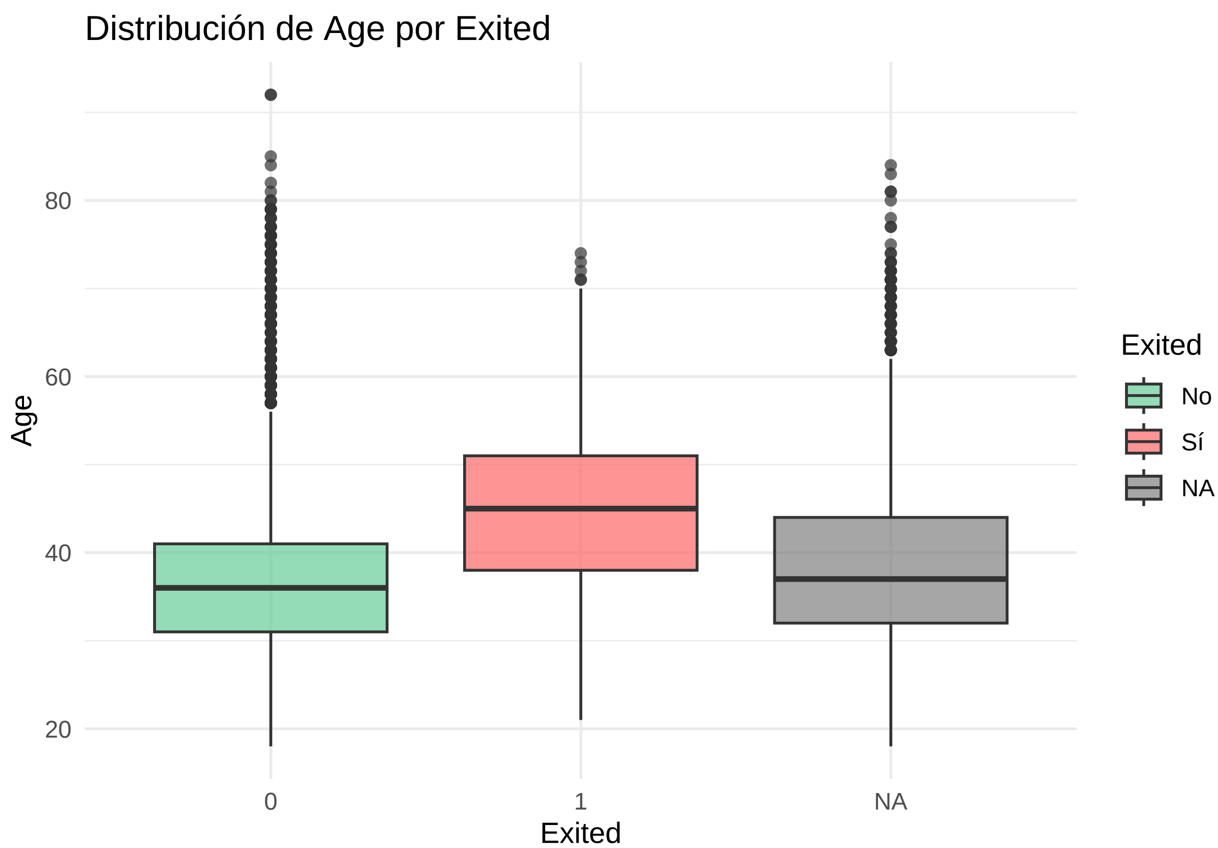
```
##
## Variable: Gender
## Chi-squared = 67.477 df = 1 p-value = 2.1316e-16
## -> Diferencias significativas entre columnas
##
## Variable: MaritalStatus
## Chi-squared = 1.404 df = 3 p-value = 0.70451
##
## Variable: EducationLevel
## Chi-squared = 2.111 df = 3 p-value = 0.54968
##
## Variable: HasCrCard
## Chi-squared = 2.224 df = 1 p-value = 0.1359
##
## Variable: SavingsAccountFlag
## Chi-squared = 0.166 df = 1 p-value = 0.68361
##
## Variable: LoanStatus
## Chi-squared = 0.229 df = 2 p-value = 0.89203
##
## Variable: CustomerSegment
## Chi-squared = 2.421 df = 2 p-value = 0.29805
##
## Variable: IsActiveMember
## Chi-squared = 114.979 df = 1 p-value = 7.9551e-27
## -> Diferencias significativas entre columnas
```

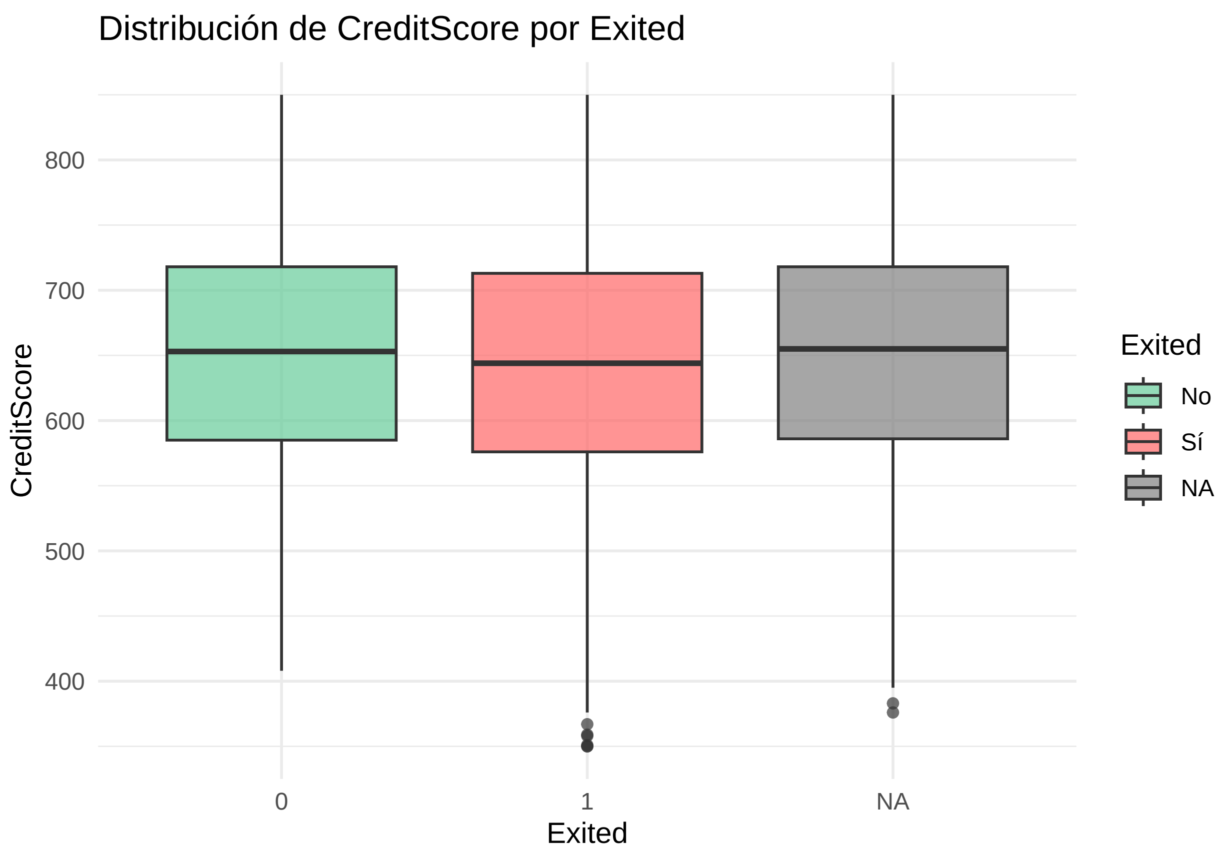
Un cop fet el test, s'afirma lo que s'ha concluït anteriorment.

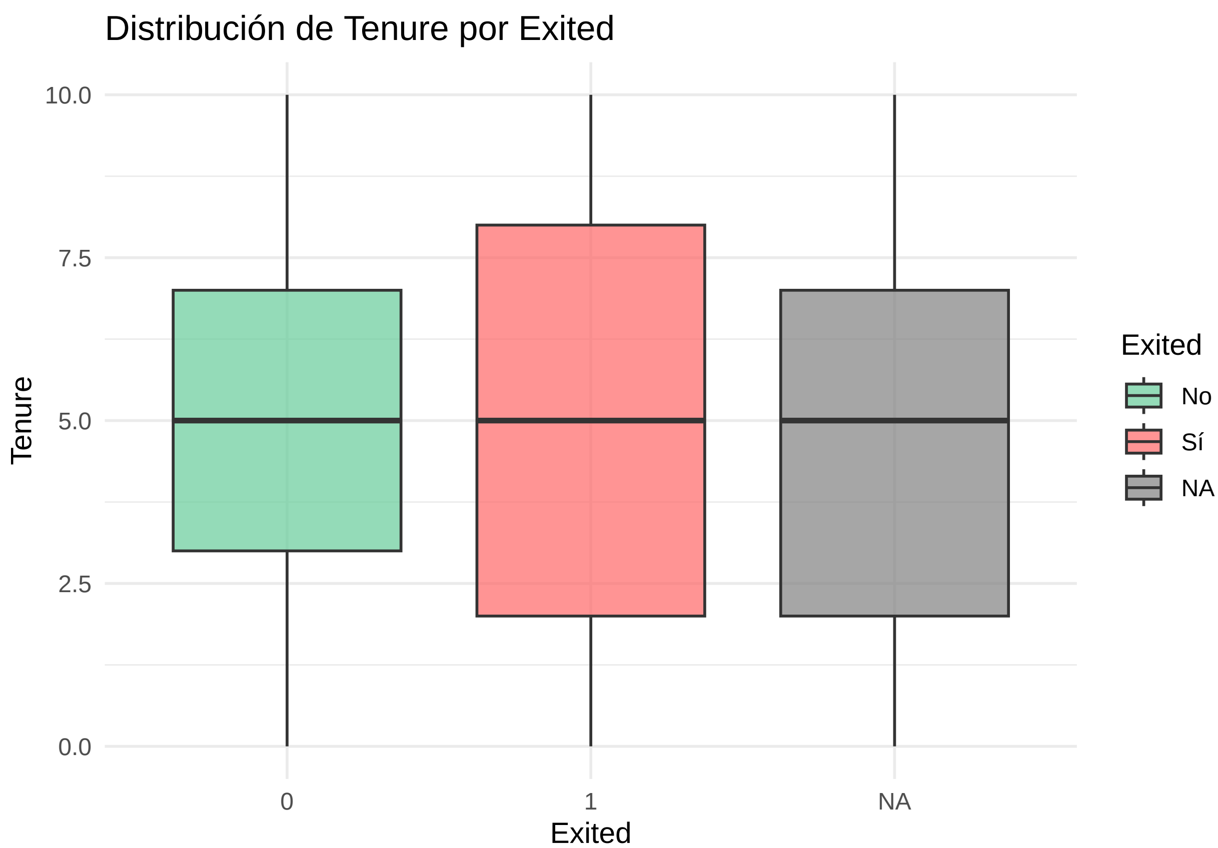
Numèrica i categòrica

Estudiem la variable resposta amb la resta de variables numèriques.

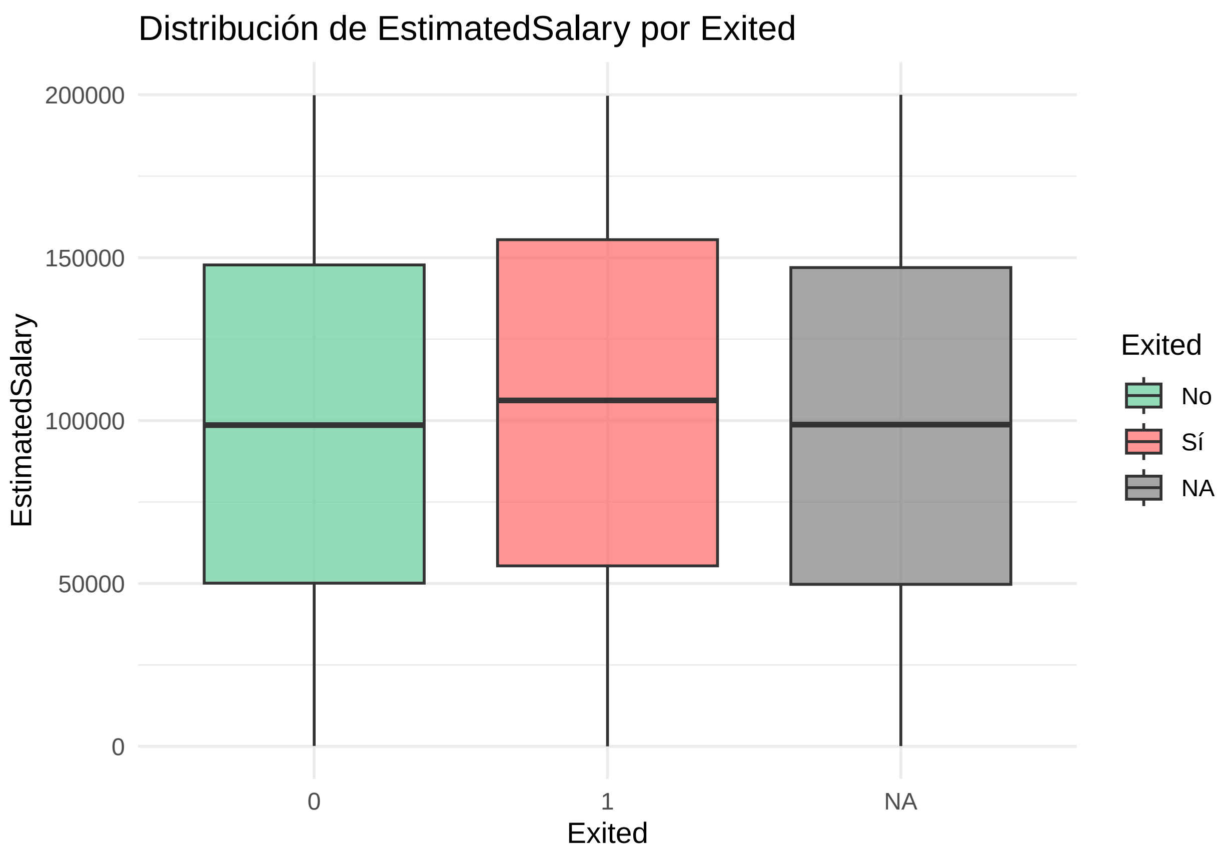
```
for (var in varNum) {
  p <- ggplot(data, aes(x = factor(Exited), y = .data[[var]], fill = factor(Exited))) +
    geom_boxplot(alpha = 0.7, na.rm = TRUE) +
    labs(x = "Exited",
         y = var,
         title = paste("Distribución de", var, "por Exited")) +
    scale_fill_manual(values = c("#66CC99", "#FF6666"),
                      name = "Exited",
                      labels = c("No", "Sí")) +
    theme_minimal()

  print(p)
}
```

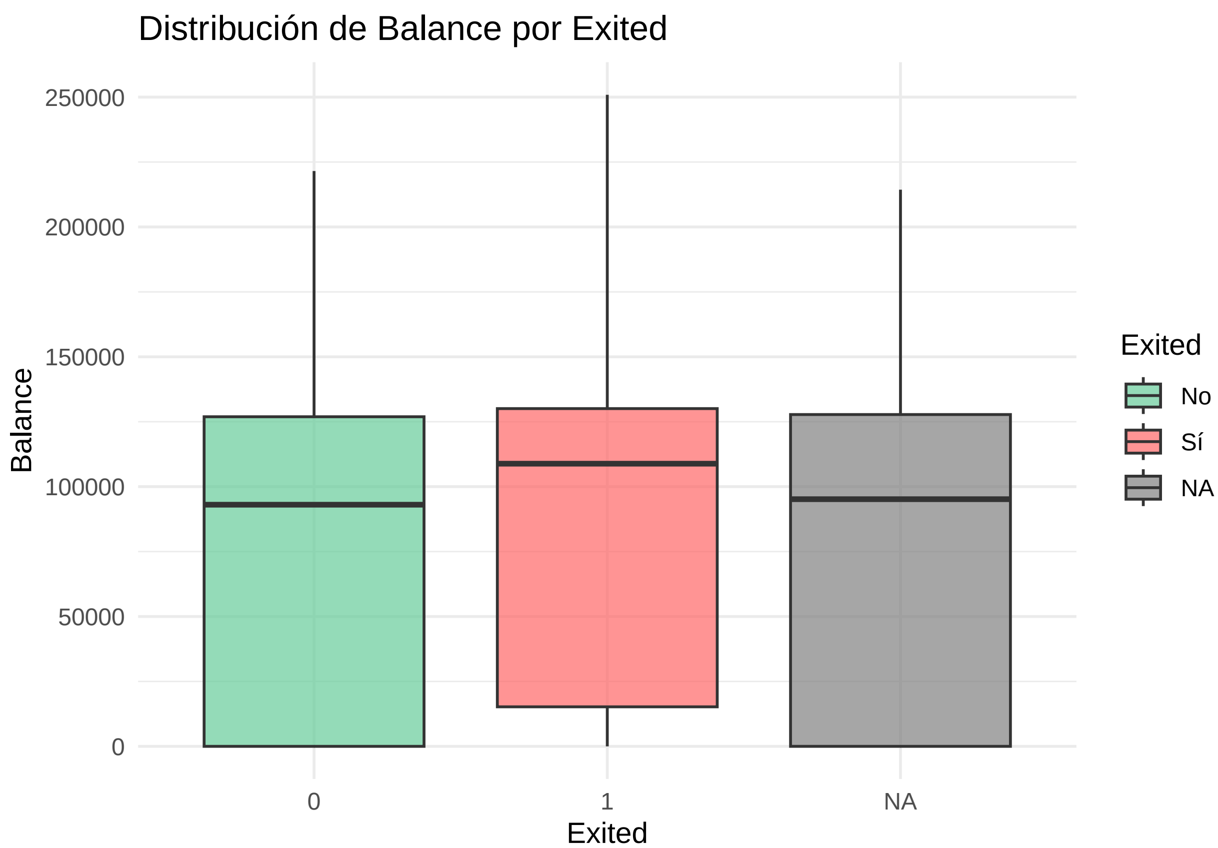

Distribución de Age por Exited
Age
80
60
40
20
0
1
NA
Exited
Exited
No
Sí
NA

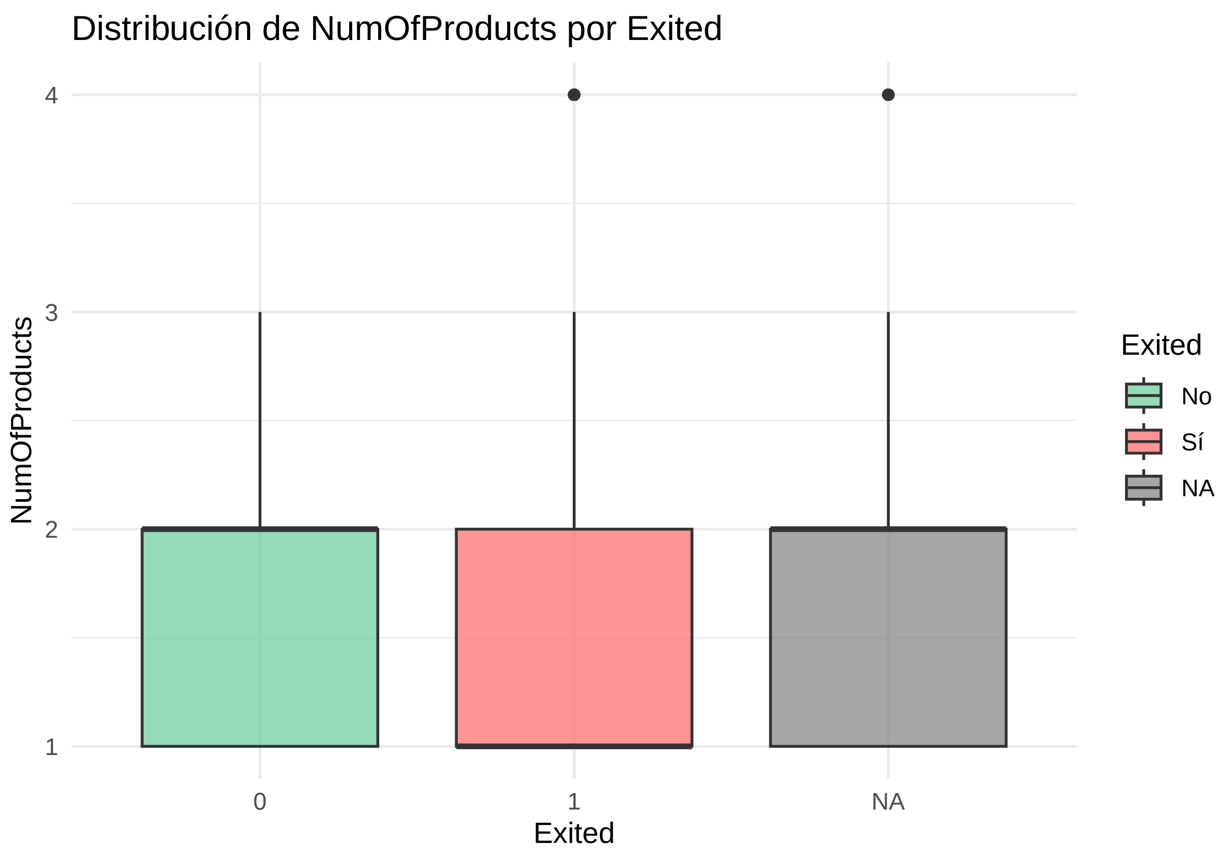

Distribución de CreditScore por Exited
800
700
600
500
400
CreditScore
0
1
NA
Exited
Exited
No
Sí
NA

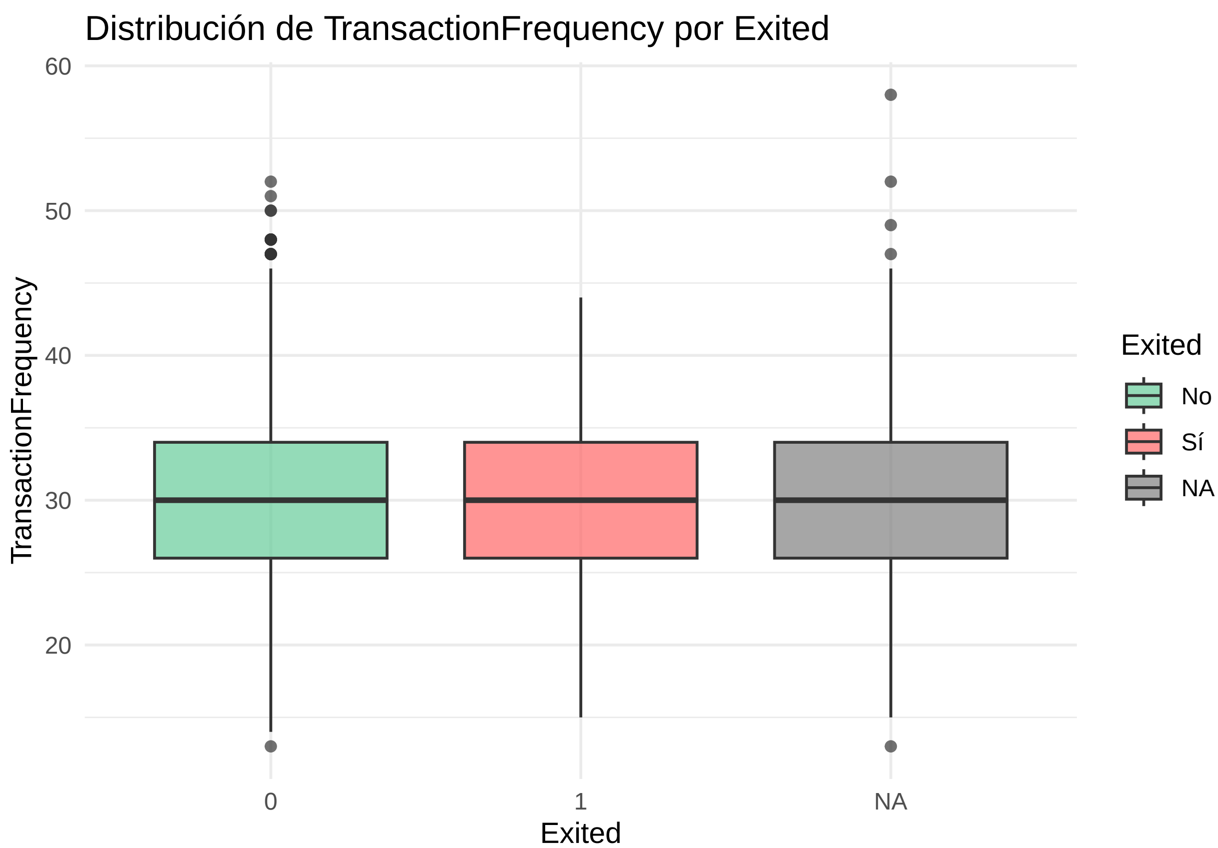


Distribución de Tenure por Exited
Tenure
10.0
7.5
5.0
2.5
0.0
0
1
NA
Exited
Exited
No
Sí
NA

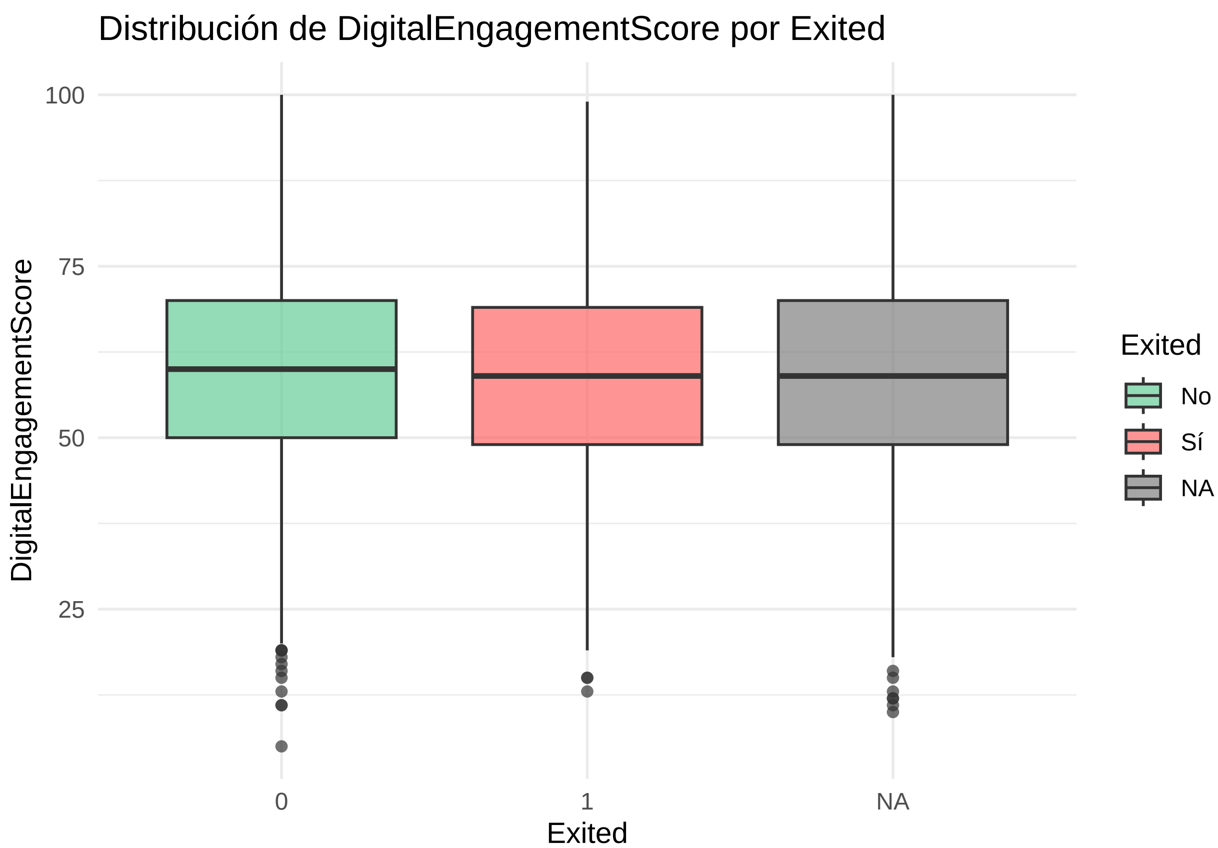

Distribución de EstimatedSalary por Exited
EstimatedSalary
200000
150000
100000
50000
0
0
1
NA
Exited
Exited
No
Sí
NA

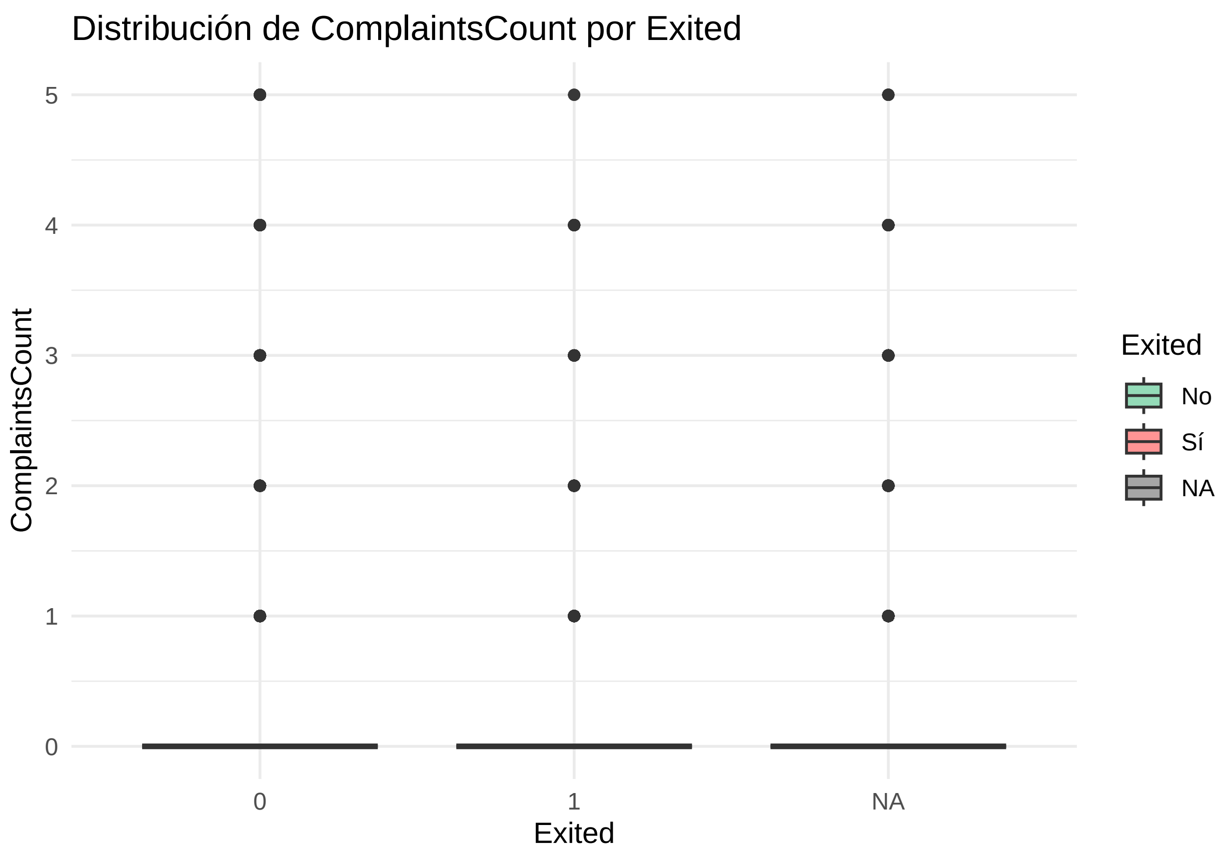

Distribución de Balance por Exited
250000
200000
150000
100000
50000
0
Balance
0
1
NA
Exited
Exited
No
Sí
NA

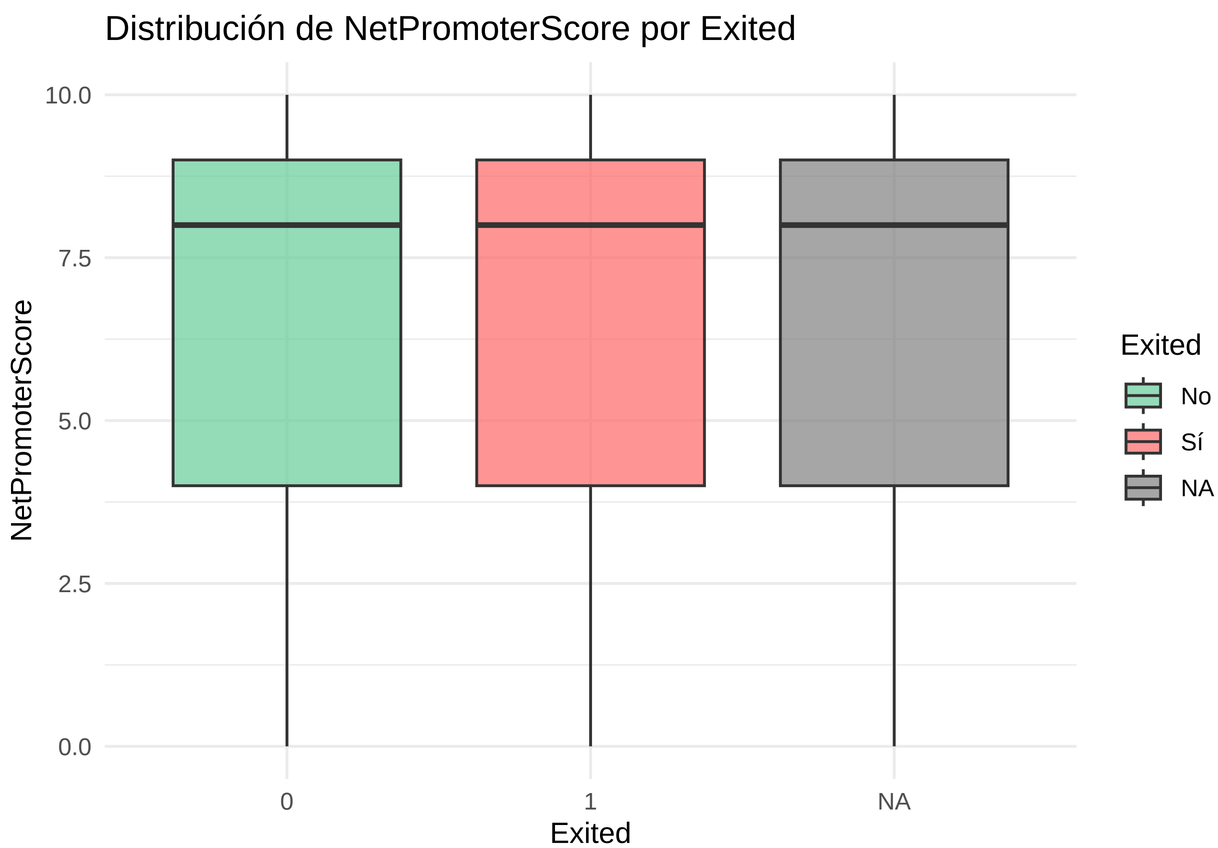

Distribución de NumOfProducts por Exited
NumOfProducts
4
3
2
1
0
1
NA
Exited
Exited
No
Sí
NA


Distribución de TransactionFrequency por Exited
TransactionFrequency
60
50
40
30
20
0
1
NA
Exited
Exited
No
Sí
NA


Distribución de AvgTransactionAmount por Exited
AvgTransactionAmount
600
400
200
0
0
1
NA
Exited
Exited
No
Sí
NA


Distribución de DigitalEngagementScore por Exited
DigitalEngagementScore
100
75
50
25
0
1
NA
Exited
Exited
No
Sí
NA


Distribución de ComplaintsCount por Exited
ComplaintsCount
5
4
3
2
1
0
0
1
NA
Exited
Exited
No
Sí
NA



Observem la mateixa distribució de les varibles numèriques segons si marxa o no del banc excepte les variables age,creditscore,numofproducts, Balance i Estimated salary.

Realitzem el test de medianas per veure si la diferencia és significativa.

```
resultados_mediana <- data.frame(
  Variable = character(),
  Mediana_Exited0 = numeric(),
  Mediana_Exited1 = numeric(),
  p_value = numeric(),
  stringsAsFactors = FALSE
)

for (var in varNum) {
  medianas <- tapply(data[[var]], data$Exited, median, na.rm = TRUE)
  p_val <- wilcox.test(data[[var]] ~ data$Exited)$p.value

  resultados_mediana <- rbind(resultados_mediana, data.frame(
    Variable = var,
    Mediana_Exited0 = medianas["0"],
    Mediana_Exited1 = medianas["1"],
    p_value = p_val
  ))
}
resultados_mediana
```

```
##                     Variable Mediana_Exited0 Mediana_Exited1       p_value
```

```
## 0                       Age        36.00000        45.00000 1.973454e-114
## 01              CreditScore       653.00000       644.00000  1.860831e-02
## 02                   Tenure         5.00000         5.00000  3.200349e-01
## 03          EstimatedSalary     98602.46500    106188.38500  4.059852e-02
## 04                  Balance     93052.50000    108844.05000  5.804964e-12
## 05            NumOfProducts         2.00000         1.00000  9.255184e-22
## 06     TransactionFrequency        30.00000        30.00000  5.593502e-01
## 07     AvgTransactionAmount        99.16118        96.60327  3.556860e-01
## 08  DigitalEngagementScore        60.00000        59.00000  2.513216e-01
## 09          ComplaintsCount         0.00000         0.00000  7.733715e-01
## 010        NetPromoterScore         8.00000         8.00000  9.145862e-01
```

Com a resultat s'observa que ens surt significatius aquells vaariables mencionades anteriorment. LLavors podem concluir lo següent:

- Els que tenen una edat més elevada tendeixen a marxar del banc amb més probabilitat.
- Els que tenen un creditScore baix tendeixen a marxar del banc amb més probabilitat.
- Els que tenen un salary estimat més elevat tendeixen a marxar del banc amb més probabilitat.
- Els que tenen només un producte tendeixen a marxar amb més probabilitat.
- Els que tenen un saldo més elevat tendeixen a marxar amb més probabilitat.

Missings

Annex

```
## ============= Geography  vs.  Gender =========================
##
##              Female      Male
##   France  0.4879249 0.5097015
##   Germany 0.2629696 0.2402985
##   Spain   0.2491055 0.2500000
## ============= Geography  vs.  MaritalStatus =========================
##
##            Divorced   Married    Single   Widowed
##   France  0.4829468 0.5075356 0.5131397 0.5387597
##   Germany 0.2701228 0.2533605 0.2392808 0.2325581
##   Spain   0.2469304 0.2391039 0.2475795 0.2286822
## ============= Geography  vs.  EducationLevel =========================
##
##           High School     Other Postgraduate University
##   France    0.5100473 0.5260870    0.4808959  0.5092135
##   Germany   0.2446809 0.2304348    0.2463768  0.2471910
##   Spain     0.2452719 0.2434783    0.2727273  0.2435955
## ============= Geography  vs.  HasCrCard =========================
##
##                   0         1
##   France  0.4996521 0.5030303
##   Germany 0.2498260 0.2502165
##   Spain   0.2505219 0.2467532
## ============= Geography  vs.  SavingsAccountFlag =========================
##
##                   0         1
```

```
##   France  0.4920260 0.5068836
##   Germany 0.2510337 0.2537547
##   Spain   0.2569403 0.2393617
## ============= Geography  vs.  LoanStatus =========================
##
##           Active loan Default risk   No loan
##   France    0.4921179    0.4923372 0.5020590
##   Germany   0.2529130    0.2681992 0.2512011
##   Spain     0.2549692    0.2394636 0.2467399
## ============= Geography  vs.  CustomerSegment =========================
##
##            Affluent High Net Worth Mass Market
##   France  0.5058500      0.4902163   0.4925613
##   Germany 0.2573985      0.2533471   0.2597507
##   Spain   0.2367515      0.2564367   0.2476880
## ============= Geography  vs.  Exited =========================
##
##                   0         1
##   France  0.5237850 0.4035608
##   Germany 0.2172795 0.3827893
##   Spain   0.2589355 0.2136499
## ============= Geography  vs.  group =========================
##
##                test     train
##   France  0.5128571 0.4989796
##   Germany 0.2490476 0.2514286
##   Spain   0.2380952 0.2495918
## ============= Geography  vs.  IsActiveMember =========================
##
##                   0         1
##   France  0.5042194 0.5128205
##   Germany 0.2552743 0.2370809
##   Spain   0.2405063 0.2500986
## ============= Gender  vs.  Geography =========================
##
##             France   Germany     Spain
##   Female 0.4440374 0.4772727 0.4539527
##   Male   0.5559626 0.5227273 0.5460473
## ============= Gender  vs.  MaritalStatus =========================
##
##           Divorced   Married    Single   Widowed
##   Female 0.4534247 0.4524577 0.4758621 0.4440154
##   Male   0.5465753 0.5475423 0.5241379 0.5559846
## ============= Gender  vs.  EducationLevel =========================
##
##          High School     Other Postgraduate University
##   Female   0.4752829 0.4110169    0.4403183  0.4562028
##   Male     0.5247171 0.5889831    0.5596817  0.5437972
## ============= Gender  vs.  HasCrCard =========================
##
##                  0         1
##   Female 0.4645341 0.4563700
##   Male   0.5354659 0.5436300
## ============= Gender  vs.  SavingsAccountFlag =========================
```

```
##
##                 0         1
##   Female 0.4636038 0.4539228
##   Male   0.5363962 0.5460772
## ============= Gender  vs.  LoanStatus =========================
##
##          Active loan Default risk   No loan
##   Female   0.4863350    0.4504854 0.4399864
##   Male     0.5136650    0.5495146 0.5600136
## ============= Gender  vs.  CustomerSegment =========================
##
##           Affluent High Net Worth Mass Market
##   Female 0.4658856      0.4385417   0.4541101
##   Male   0.5341144      0.5614583   0.5458899
## ============= Gender  vs.  Exited =========================
##
##                 0         1
##   Female 0.4251559 0.5674905
##   Male   0.5748441 0.4325095
## ============= Gender  vs.  group =========================
##
##               test     train
##   Female 0.4542857 0.4557143
##   Male   0.5457143 0.5442857
## ============= Gender  vs.  IsActiveMember =========================
##
##                 0         1
##   Female 0.4625850 0.4349009
##   Male   0.5374150 0.5650991
## ============= MaritalStatus  vs.  Geography =========================
##
##                France    Germany      Spain
##   Divorced 0.14268440 0.16150082 0.15274262
##   Married  0.50221685 0.50734095 0.49535865
##   Single   0.29907295 0.28221860 0.30210970
##   Widowed  0.05602580 0.04893964 0.04978903
## ============= MaritalStatus  vs.  Gender =========================
##
##                Female       Male
##   Divorced 0.14652501 0.14988730
##   Married  0.49712262 0.51051841
##   Single   0.30544489 0.28549962
##   Widowed  0.05090748 0.05409467
## ============= MaritalStatus  vs.  EducationLevel =========================
##
##            High School     Other Postgraduate University
##   Divorced   0.0000000 0.0000000    1.0000000  0.0000000
##   Married    0.1352335 0.0000000    0.0000000  1.0000000
##   Single     0.8647665 0.0000000    0.0000000  0.0000000
##   Widowed    0.0000000 1.0000000    0.0000000  0.0000000
## ============= MaritalStatus  vs.  HasCrCard =========================
##
##                    0          1
##   Divorced 0.13969107 0.14577259
```

```
##   Married  0.50839490 0.50991254
##   Single   0.30020148 0.29533528
##   Widowed  0.05171256 0.04897959
## ============= MaritalStatus  vs.  SavingsAccountFlag =========================
##
##                    0         1
##   Divorced 0.4480480 0.0000000
##   Married  0.0000000 0.7722256
##   Single   0.4102102 0.2277744
##   Widowed  0.1417417 0.0000000
## ============= MaritalStatus  vs.  LoanStatus =========================
##
##            Active loan Default risk   No loan
##   Divorced   0.3251404    0.5183752 0.0000000
##   Married    0.0000000    0.0000000 0.8298376
##   Single     0.6748596    0.0000000 0.1701624
##   Widowed    0.0000000    0.4816248 0.0000000
## ============= MaritalStatus  vs.  CustomerSegment =========================
##
##             Affluent High Net Worth Mass Market
##   Divorced 0.0000000      0.7645195   0.0000000
##   Married  0.0000000      0.0000000   1.0000000
##   Single   1.0000000      0.0000000   0.0000000
##   Widowed  0.0000000      0.2354805   0.0000000
## ============= MaritalStatus  vs.  Exited =========================
##
##                     0          1
##   Divorced 0.14506173 0.14525692
##   Married  0.52006173 0.50988142
##   Single   0.29089506 0.29249012
##   Widowed  0.04398148 0.05237154
## ============= MaritalStatus  vs.  group =========================
##
##                  test      train
##   Divorced 0.15142857 0.14510204
##   Married  0.48142857 0.51795918
##   Single   0.30809524 0.29122449
##   Widowed  0.05904762 0.04571429
## ============= MaritalStatus  vs.  IsActiveMember =========================
##
##                     0          1
##   Divorced 0.14123581 0.15461049
##   Married  0.50861707 0.50834658
##   Single   0.29340059 0.29451510
##   Widowed  0.05674653 0.04252782
## ============= EducationLevel  vs.  Geography =========================
##
##                    France    Germany      Spain
##   High School  0.34770346 0.34385382 0.34016393
##   Other        0.04875101 0.04401993 0.04590164
##   Postgraduate 0.14705882 0.15531561 0.16967213
##   University   0.45648670 0.45681063 0.44426230
## ============= EducationLevel  vs.  Gender =========================
##
```

```
##                    Female       Male
##   High School  0.35419441 0.33058161
##   Other        0.04305371 0.05215760
##   Postgraduate 0.14735908 0.15834897
##   University   0.45539281 0.45891182
## ============= EducationLevel  vs.  MaritalStatus =========================
##
##                  Divorced    Married     Single    Widowed
##   High School  0.00000000 0.09046653 1.00000000 0.00000000
##   Other        0.00000000 0.00000000 0.00000000 1.00000000
##   Postgraduate 1.00000000 0.00000000 0.00000000 0.00000000
##   University   0.00000000 0.90953347 0.00000000 0.00000000
## ============= EducationLevel  vs.  HasCrCard =========================
##
##                         0          1
##   High School  0.34556787 0.33936782
##   Other        0.05263158 0.04655172
##   Postgraduate 0.15235457 0.15057471
##   University   0.44944598 0.46350575
## ============= EducationLevel  vs.  SavingsAccountFlag =========================
##
##                        0         1
##   High School  0.4325280 0.2929956
##   Other        0.1249263 0.0000000
##   Postgraduate 0.4425457 0.0000000
##   University   0.0000000 0.7070044
## ============= EducationLevel  vs.  LoanStatus =========================
##
##                Active loan Default risk   No loan
##   High School    0.6677943    0.0000000 0.2421875
##   Other          0.0000000    0.4688129 0.0000000
##   Postgraduate   0.3322057    0.5311871 0.0000000
##   University     0.0000000    0.0000000 0.7578125
## ============= EducationLevel  vs.  CustomerSegment =========================
##
##                  Affluent High Net Worth Mass Market
##   High School  1.00000000     0.00000000  0.09716599
##   Other        0.00000000     0.22966014  0.00000000
##   Postgraduate 0.00000000     0.77033986  0.00000000
##   University   0.00000000     0.00000000  0.90283401
## ============= EducationLevel  vs.  Exited =========================
##
##                         0          1
##   High School  0.34192708 0.32547170
##   Other        0.04244792 0.05094340
##   Postgraduate 0.15000000 0.15094340
##   University   0.46562500 0.47264151
## ============= EducationLevel  vs.  group =========================
##
##                      test      train
##   High School  0.35333333 0.33836735
##   Other        0.05380952 0.04428571
##   Postgraduate 0.15619048 0.15020408
##   University   0.43666667 0.46714286
```

```
## ============= EducationLevel  vs.  IsActiveMember ========================
##
##                          0          1
##   High School  0.33192210 0.34816857
##   Other        0.05376799 0.04135486
##   Postgraduate 0.15283658 0.16069319
##   University   0.46147333 0.44978338
## ============= HasCrCard  vs.  Geography ========================
##
##        France   Germany     Spain
##   0 0.2917513 0.2928222 0.2962963
##   1 0.7082487 0.7071778 0.7037037
## ============= HasCrCard  vs.  Gender ========================
##
##        Female      Male
##   0 0.2986142 0.2917772
##   1 0.7013858 0.7082228
## ============= HasCrCard  vs.  MaritalStatus ========================
##
##      Divorced   Married    Single   Widowed
##   0 0.2937853 0.3020750 0.3061644 0.3142857
##   1 0.7062147 0.6979250 0.6938356 0.6857143
## ============= HasCrCard  vs.  EducationLevel ========================
##
##     High School     Other Postgraduate University
##   0   0.2970238 0.3193277    0.2956989  0.2869142
##   1   0.7029762 0.6806723    0.7043011  0.7130858
## ============= HasCrCard  vs.  SavingsAccountFlag ========================
##
##             0         1
##   0 0.2991607 0.2965261
##   1 0.7008393 0.7034739
## ============= HasCrCard  vs.  LoanStatus ========================
##
##     Active loan Default risk   No loan
##   0   0.2968641    0.3070866 0.2941975
##   1   0.7031359    0.6929134 0.7058025
## ============= HasCrCard  vs.  CustomerSegment ========================
##
##      Affluent High Net Worth Mass Market
##   0 0.3102740      0.3072862   0.2893258
##   1 0.6897260      0.6927138   0.7106742
## ============= HasCrCard  vs.  Exited ========================
##
##             0         1
##   0 0.2933264 0.3172147
##   1 0.7066736 0.6827853
## ============= HasCrCard  vs.  group ========================
##
##          test     train
##   0 0.2952381 0.2983673
##   1 0.7047619 0.7016327
## ============= HasCrCard  vs.  IsActiveMember ========================
##
```

```
##              0         1
##   0 0.3015013 0.3013752
##   1 0.6984987 0.6986248
## ============= SavingsAccountFlag  vs.  Geography =========================
##
##        France   Germany     Spain
##   0 0.3395842 0.3438511 0.3625000
##   1 0.6604158 0.6561489 0.6375000
## ============= SavingsAccountFlag  vs.  Gender =========================
##
##        Female      Male
##   0 0.3476510 0.3388617
##   1 0.6523490 0.6611383
## ============= SavingsAccountFlag  vs.  MaritalStatus =========================
##
##      Divorced   Married    Single   Widowed
##   0 1.0000000 0.0000000 0.4789621 1.0000000
##   1 0.0000000 1.0000000 0.5210379 0.0000000
## ============= SavingsAccountFlag  vs.  EducationLevel =========================
##
##     High School     Other Postgraduate University
##   0   0.4392579 1.0000000    1.0000000  0.0000000
##   1   0.5607421 0.0000000    0.0000000  1.0000000
## ============= SavingsAccountFlag  vs.  HasCrCard =========================
##
##              0         1
##   0 0.3429553 0.3401222
##   1 0.6570447 0.6598778
## ============= SavingsAccountFlag  vs.  LoanStatus =========================
##
##     Active loan Default risk   No loan
##   0   0.8218157    1.0000000 0.0000000
##   1   0.1781843    0.0000000 1.0000000
## ============= SavingsAccountFlag  vs.  CustomerSegment =========================
##
##     Affluent High Net Worth Mass Market
##   0 0.472067       1.000000    0.000000
##   1 0.527933       0.000000    1.000000
## ============= SavingsAccountFlag  vs.  Exited =========================
##
##              0         1
##   0 0.3381481 0.3313783
##   1 0.6618519 0.6686217
## ============= SavingsAccountFlag  vs.  group =========================
##
##          test     train
##   0 0.3566667 0.3367347
##   1 0.6433333 0.6632653
## ============= SavingsAccountFlag  vs.  IsActiveMember =========================
##
##              0         1
##   0 0.3438023 0.3405342
##   1 0.6561977 0.6594658
## ============= LoanStatus  vs.  Geography =========================
```

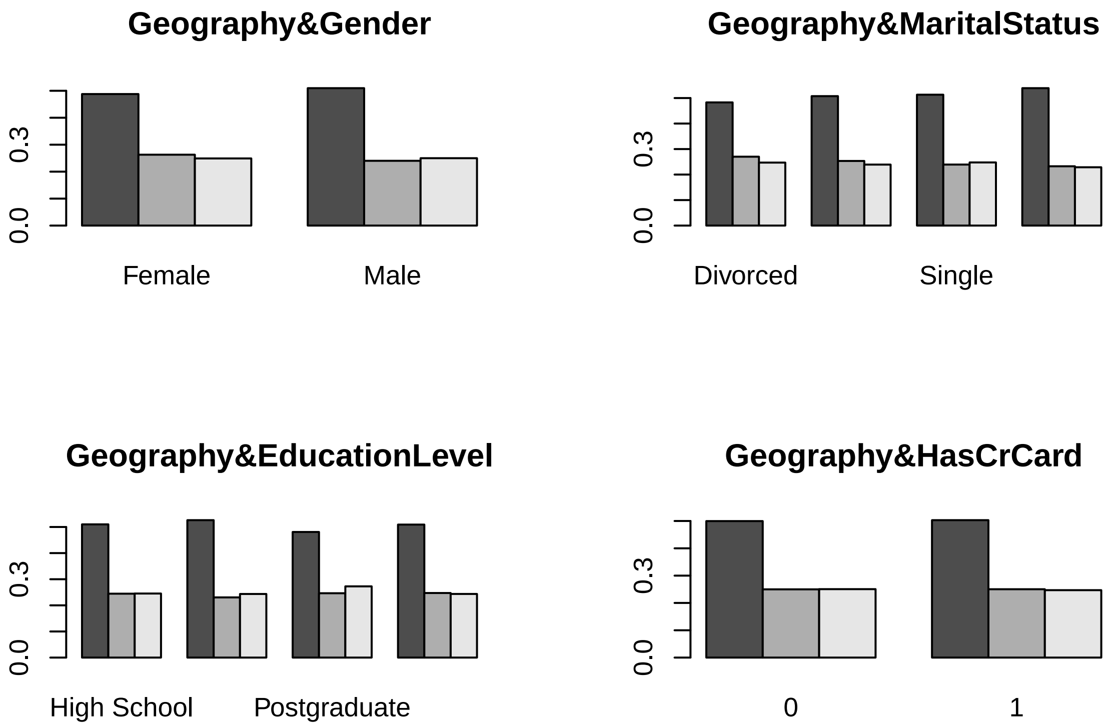
```
##
##                  France   Germany     Spain
##   Active loan  0.2945037 0.2973409 0.3059211
##   Default risk 0.1054143 0.1128122 0.1027961
##   No loan      0.6000820 0.5898469 0.5912829
## ============= LoanStatus  vs.  Gender =========================
##
##                  Female      Male
##   Active loan  0.3126126 0.2752535
##   Default risk 0.1045045 0.1062711
##   No loan      0.5828829 0.6184754
## ============= LoanStatus  vs.  MaritalStatus =========================
##
##                Divorced   Married    Single   Widowed
##   Active loan  0.6333789 0.0000000 0.6564208 0.0000000
##   Default risk 0.3666211 0.0000000 0.0000000 1.0000000
##   No loan      0.0000000 1.0000000 0.3435792 0.0000000
## ============= LoanStatus  vs.  EducationLevel =========================
##
##               High School     Other Postgraduate University
##   Active loan    0.5805882 0.0000000    0.6503311  0.0000000
##   Default risk   0.0000000 1.0000000    0.3496689  0.0000000
##   No loan        0.4194118 0.0000000    0.0000000  1.0000000
## ============= LoanStatus  vs.  HasCrCard =========================
##
##                       0         1
##   Active loan  0.2939959 0.2932287
##   Default risk 0.1076605 0.1022958
##   No loan      0.5983437 0.6044754
## ============= LoanStatus  vs.  SavingsAccountFlag =========================
##
##                       0         1
##   Active loan  0.7081144 0.0821875
##   Default risk 0.2918856 0.0000000
##   No loan      0.0000000 0.9178125
## ============= LoanStatus  vs.  CustomerSegment =========================
##
##                Affluent High Net Worth Mass Market
##   Active loan  0.6473001      0.4893617   0.0000000
##   Default risk 0.0000000      0.5106383   0.0000000
##   No loan      0.3526999      0.0000000   1.0000000
## ============= LoanStatus  vs.  Exited =========================
##
##                       0         1
##   Active loan  0.2875448 0.2945892
##   Default risk 0.1002050 0.1012024
##   No loan      0.6122501 0.6042084
## ============= LoanStatus  vs.  group =========================
##
##                    test     train
##   Active loan  0.3090476 0.2889796
##   Default risk 0.1080952 0.1004082
##   No loan      0.5828571 0.6106122
## ============= LoanStatus  vs.  IsActiveMember =========================
```

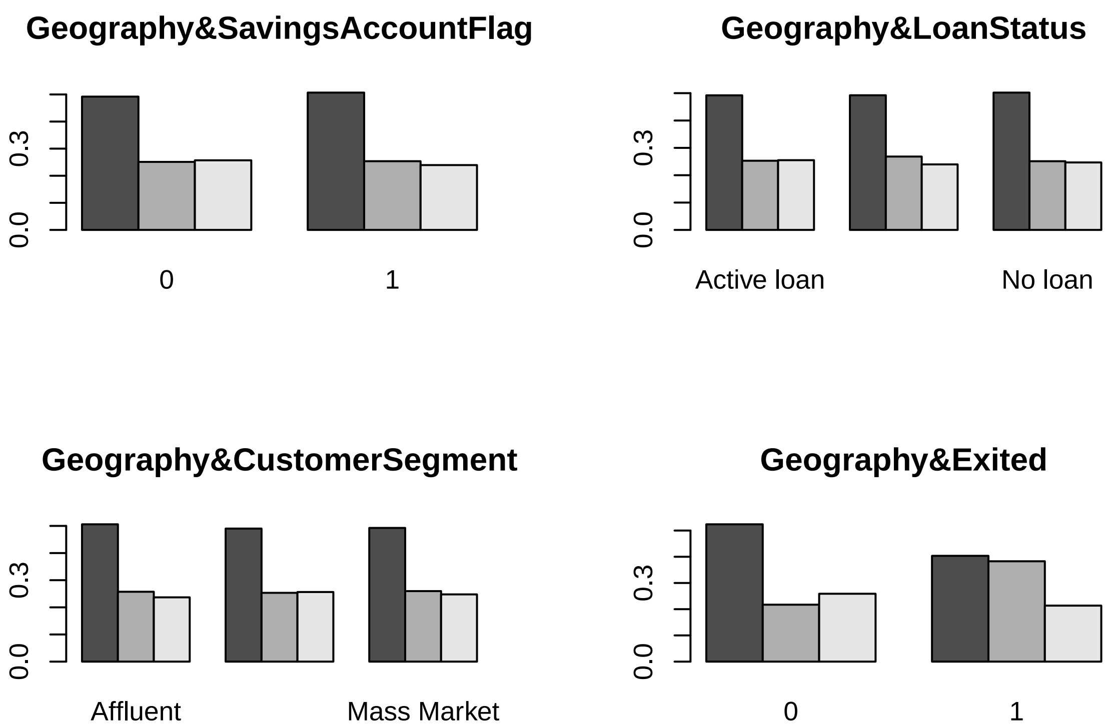
```
##
##                     0         1
##   Active loan  0.2984823 0.2997199
##   Default risk 0.0994941 0.1060424
##   No loan      0.6020236 0.5942377
## ============= CustomerSegment  vs.  Geography =========================
##
##                   France   Germany     Spain
##   Affluent       0.3017241 0.2954186 0.2845327
##   High Net Worth 0.1954023 0.1943128 0.2059553
##   Mass Market    0.5028736 0.5102686 0.5095120
## ============= CustomerSegment  vs.  Gender =========================
##
##                   Female      Male
##   Affluent       0.3024609 0.2889635
##   High Net Worth 0.1883669 0.2009694
##   Mass Market    0.5091723 0.5100671
## ============= CustomerSegment  vs.  MaritalStatus =========================
##
##                 Divorced Married Single Widowed
##   Affluent              0       0      1       0
##   High Net Worth        1       0      0       1
##   Mass Market           0       1      0       0
## ============= CustomerSegment  vs.  EducationLevel =========================
##
##                 High School     Other Postgraduate University
##   Affluent         0.8570578 0.0000000    0.0000000  0.0000000
##   High Net Worth   0.0000000 1.0000000    1.0000000  0.0000000
##   Mass Market      0.1429422 0.0000000    0.0000000  1.0000000
## ============= CustomerSegment  vs.  HasCrCard =========================
##
##                         0         1
##   Affluent       0.3092150 0.2932440
##   High Net Worth 0.1986348 0.1910309
##   Mass Market    0.4921502 0.5157251
## ============= CustomerSegment  vs.  SavingsAccountFlag =========================
##
##                         0         1
##   Affluent       0.4119439 0.2308397
##   High Net Worth 0.5880561 0.0000000
##   Mass Market    0.0000000 0.7691603
## ============= CustomerSegment  vs.  LoanStatus =========================
##
##                 Active loan Default risk   No loan
##   Affluent         0.6622378    0.0000000 0.1720000
##   High Net Worth   0.3377622    1.0000000 0.0000000
##   Mass Market      0.0000000    0.0000000 0.8280000
## ============= CustomerSegment  vs.  Exited =========================
##
##                         0         1
##   Affluent       0.2876607 0.2930693
##   High Net Worth 0.1863753 0.2049505
##   Mass Market    0.5259640 0.5019802
## ============= CustomerSegment  vs.  group =========================
```

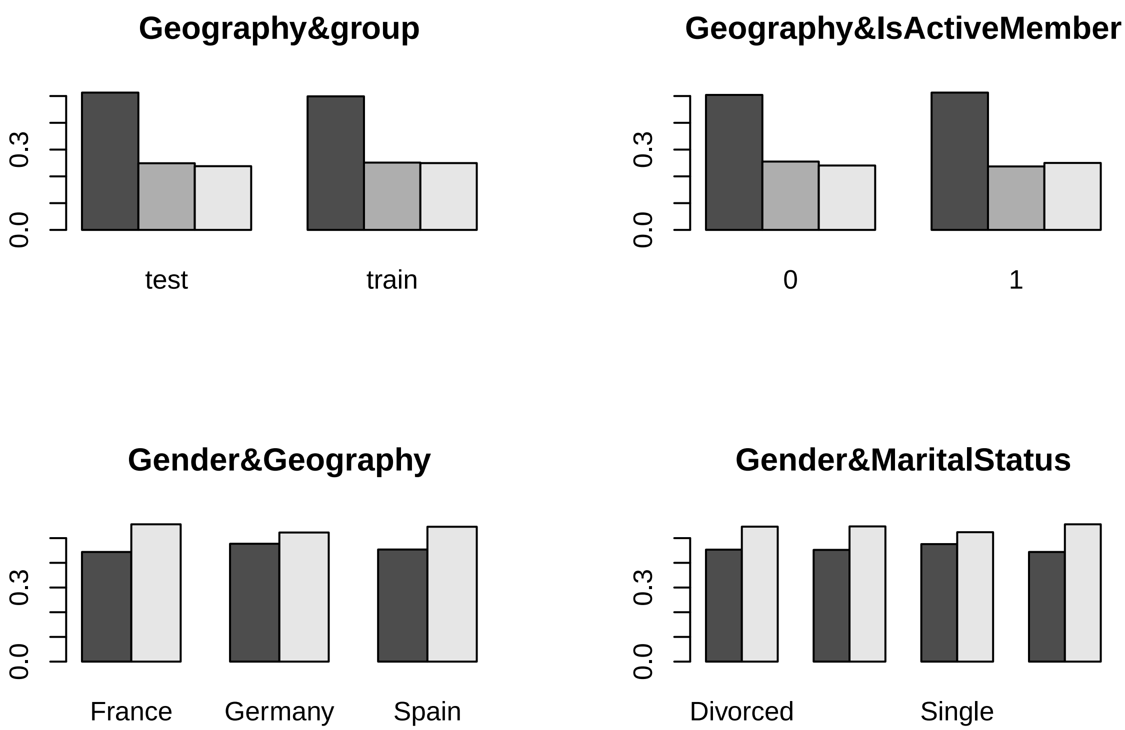
```
##
##                        test     train
##   Affluent       0.3114286 0.2887755
##   High Net Worth 0.2061905 0.1902041
##   Mass Market    0.4823810 0.5210204
## ============= CustomerSegment  vs.  IsActiveMember =========================
##
##                          0         1
##   Affluent       0.2910416 0.3001949
##   High Net Worth 0.2018860 0.1925926
##   Mass Market    0.5070724 0.5072125
## ============= Exited  vs.  Geography =========================
##
##        France   Germany     Spain
##   0 0.8331288 0.6858766 0.8233851
##   1 0.1668712 0.3141234 0.1766149
## ============= Exited  vs.  Gender =========================
##
##        Female      Male
##   0 0.7326467 0.8293963
##   1 0.2673533 0.1706037
## ============= Exited  vs.  MaritalStatus =========================
##
##      Divorced   Married    Single   Widowed
##   0 0.7932489 0.7966903 0.7925718 0.7633929
##   1 0.2067511 0.2033097 0.2074282 0.2366071
## ============= Exited  vs.  EducationLevel =========================
##
##     High School     Other Postgraduate University
##   0   0.7919180 0.7511521    0.7826087  0.7811271
##   1   0.2080820 0.2488479    0.2173913  0.2188729
## ============= Exited  vs.  HasCrCard =========================
##
##             0         1
##   0 0.7756498 0.7946481
##   1 0.2243502 0.2053519
## ============= Exited  vs.  SavingsAccountFlag =========================
##
##             0         1
##   0 0.7945455 0.7895385
##   1 0.2054545 0.2104615
## ============= Exited  vs.  LoanStatus =========================
##
##     Active loan Default risk   No loan
##   0   0.7923729    0.7947154 0.7984626
##   1   0.2076271    0.2052846 0.2015374
## ============= Exited  vs.  CustomerSegment =========================
##
##      Affluent High Net Worth Mass Market
##   0 0.7908127      0.7778970   0.8014101
##   1 0.2091873      0.2221030   0.1985899
## ============= Exited  vs.  group =========================
##
##     test     train
```

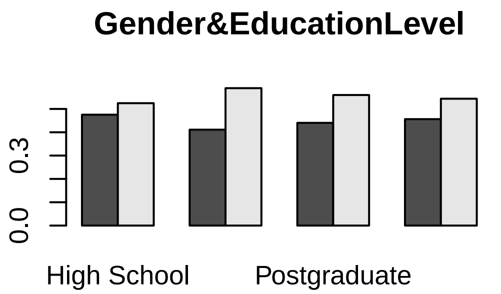
```
##   0      0.7928571
##   1      0.2071429
## ============= Exited  vs.  IsActiveMember =========================
##
##             0         1
##   0 0.7297412 0.8537161
##   1 0.2702588 0.1462839
## ============= group  vs.  Geography =========================
##
##            France   Germany     Spain
##   test  0.3057922 0.2980057 0.2901915
##   train 0.6942078 0.7019943 0.7098085
## ============= group  vs.  Gender =========================
##
##            Female      Male
##   test  0.2993411 0.3005507
##   train 0.7006589 0.6994493
## ============= group  vs.  MaritalStatus =========================
##
##          Divorced   Married    Single   Widowed
##   test  0.3090379 0.2848690 0.3119576 0.3563218
##   train 0.6909621 0.7151310 0.6880424 0.6436782
## ============= group  vs.  EducationLevel =========================
##
##         High School     Other Postgraduate University
##   test    0.3091667 0.3424242    0.3082707  0.2860262
##   train   0.6908333 0.6575758    0.6917293  0.7139738
## ============= group  vs.  HasCrCard =========================
##
##                 0         1
##   test  0.2977906 0.3009353
##   train 0.7022094 0.6990647
## ============= group  vs.  SavingsAccountFlag =========================
##
##                 0         1
##   test  0.3122134 0.2936318
##   train 0.6877866 0.7063682
## ============= group  vs.  LoanStatus =========================
##
##         Active loan Default risk   No loan
##   test    0.3142857    0.3157163 0.2903226
##   train   0.6857143    0.6842837 0.7096774
## ============= group  vs.  CustomerSegment =========================
##
##          Affluent High Net Worth Mass Market
##   test  0.3160947      0.3172161   0.2840718
##   train 0.6839053      0.6827839   0.7159282
## ============= group  vs.  Exited =========================
##
##         0 1
##   test  0 0
##   train 1 1
## ============= group  vs.  IsActiveMember =========================
##
```

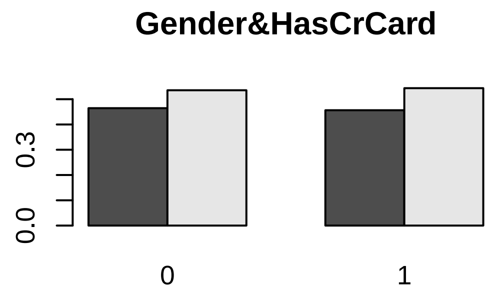
```
##                 0         1
##   test  0.3032811 0.2969312
##   train 0.6967189 0.7030688
## ============= IsActiveMember  vs.  Geography =========================
##
##        France   Germany     Spain
##   0 0.4789579 0.5016584 0.4734219
##   1 0.5210421 0.4983416 0.5265781
## ============= IsActiveMember  vs.  Gender =========================
##
##        Female      Male
##   0 0.4929769 0.4650478
##   1 0.5070231 0.5349522
## ============= IsActiveMember  vs.  MaritalStatus =========================
##
##      Divorced   Married    Single   Widowed
##   0 0.4634483 0.4861390 0.4850591 0.5578512
##   1 0.5365517 0.5138610 0.5149409 0.4421488
## ============= IsActiveMember  vs.  EducationLevel =========================
##
##     High School     Other Postgraduate University
##   0   0.4700240 0.5474138    0.4694408  0.4883513
##   1   0.5299760 0.4525862    0.5305592  0.5116487
## ============= IsActiveMember  vs.  HasCrCard =========================
##
##             0         1
##   0 0.4852349 0.4850854
##   1 0.5147651 0.5149146
## ============= IsActiveMember  vs.  SavingsAccountFlag =========================
##
##             0         1
##   0 0.4863744 0.4827480
##   1 0.5136256 0.5172520
## ============= IsActiveMember  vs.  LoanStatus =========================
##
##     Active loan Default risk   No loan
##   0   0.4859300    0.4710579 0.4902163
##   1   0.5140700    0.5289421 0.5097837
## ============= IsActiveMember  vs.  CustomerSegment =========================
##
##      Affluent High Net Worth Mass Market
##   0 0.4685990      0.4880829   0.4762480
##   1 0.5314010      0.5119171   0.5237520
## ============= IsActiveMember  vs.  Exited =========================
##
##             0         1
##   0 0.4420457 0.6313181
##   1 0.5579543 0.3686819
## ============= IsActiveMember  vs.  group =========================
##
##          test     train
##   0 0.4885714 0.4810204
##   1 0.5114286 0.5189796
```


Geography&Gender
0.0
0.3
Female
Male
Geography&MaritalStatus
0.0
0.3
Divorced
Single
Geography&EducationLevel
0.0
0.3
High School
Postgraduate
Geography&HasCrCard
0.0
0.3
0
1

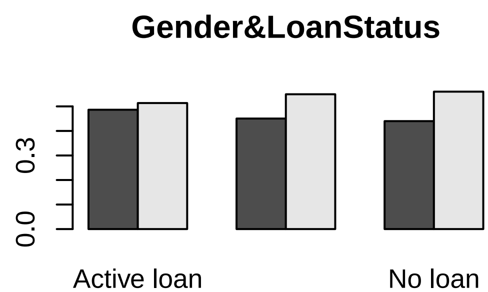

Geography&SavingsAccountFlag
0.0
0.3
0
1
Geography&LoanStatus
0.0
0.3
Active loan
No loan
Geography&CustomerSegment
0.0
0.3
Affluent
Mass Market
Geography&Exited
0.0
0.3
0
1


Geography&group
0.0
0.3
test
train
Geography&IsActiveMember
0.0
0.3
0
1
Gender&Geography
0.0
0.3
France
Germany
Spain
Gender&MaritalStatus
0.0
0.3
Divorced
Single

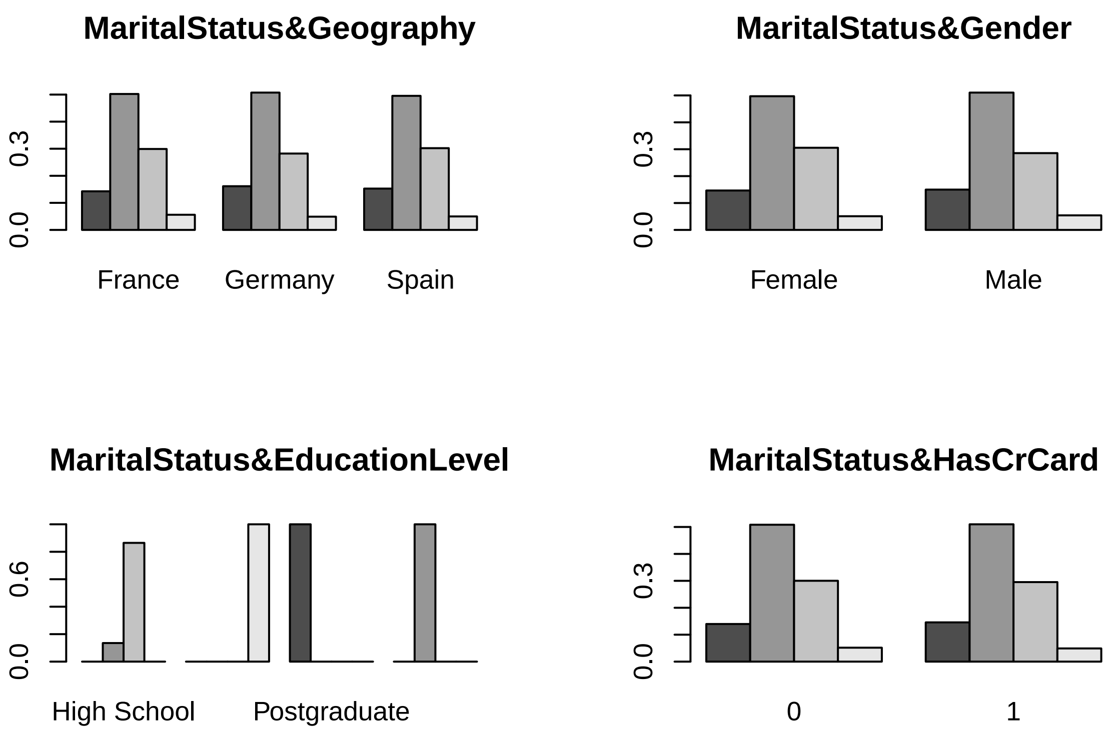

Gender&EducationLevel
0.0
0.3
High School
Postgraduate


Gender&HasCrCard
0.0
0.3
0
1

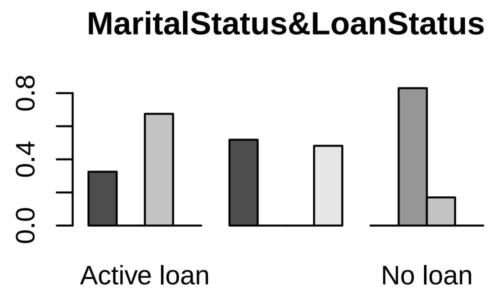

Gender&SavingsAccountFlag
0.0
0.3
0
1


Gender&LoanStatus
0.0
0.3
Active loan
No loan

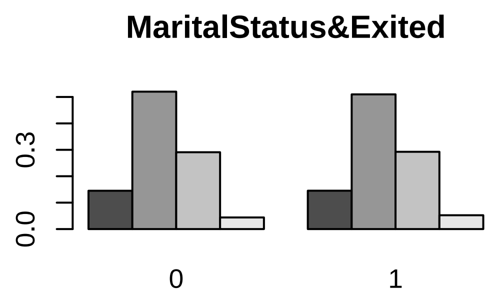
Gender&CustomerSegment
0.0
0.3
Affluent
Mass Market
Gender&Exited
0.0
0.3
0
1
Gender&group
0.0
0.3
test
train
Gender&IsActiveMember
0.0
0.3
0
1

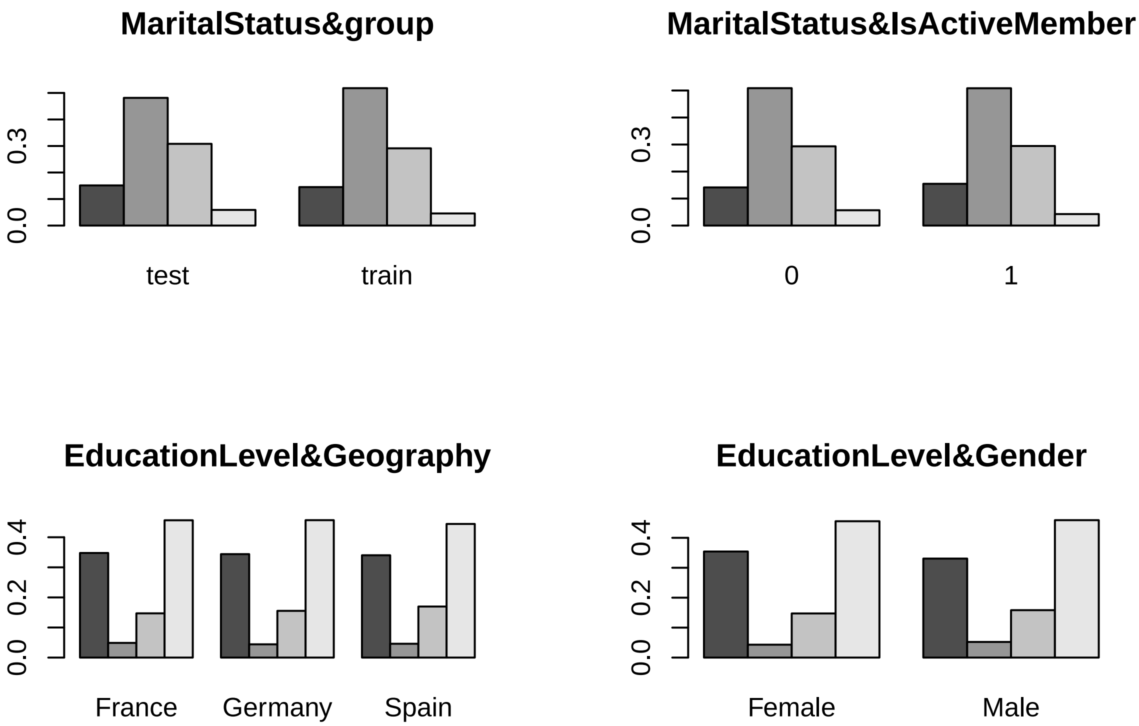

MaritalStatus&Geography
0.0
0.3
France
Germany
Spain
MaritalStatus&Gender
0.0
0.3
Female
Male
MaritalStatus&EducationLevel
0.0
0.6
High School
Postgraduate
MaritalStatus&HasCrCard
0.0
0.3
0
1

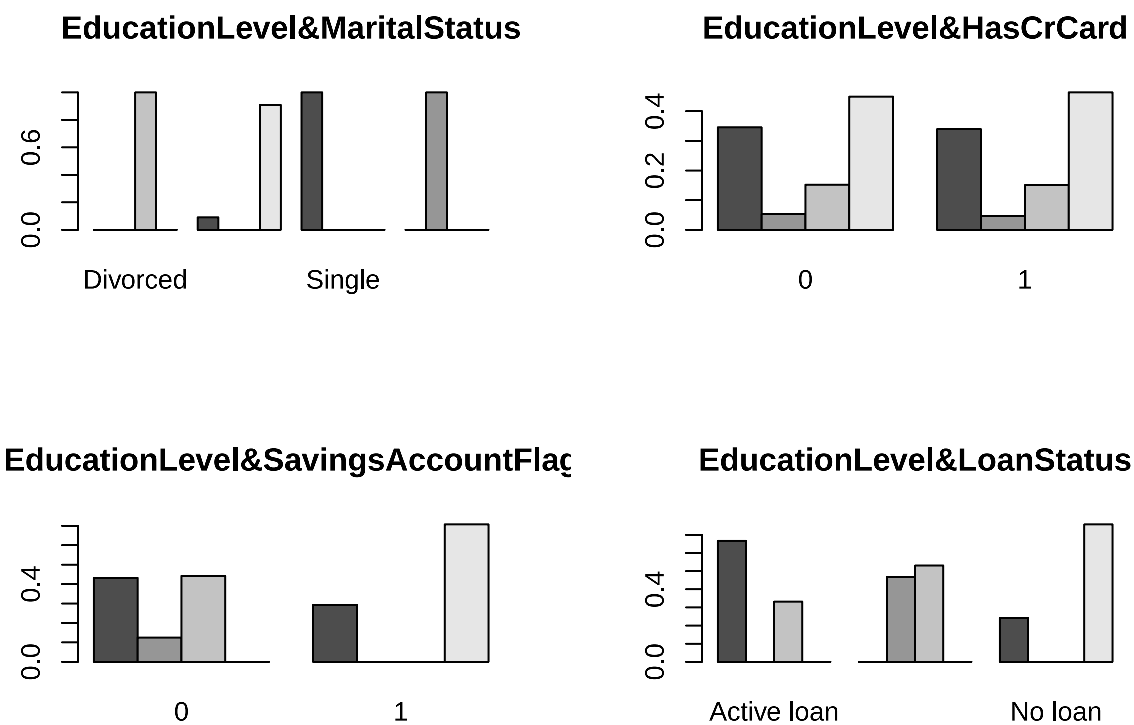

MaritalStatus&SavingsAccountFlag
0.0
0.4
0
1

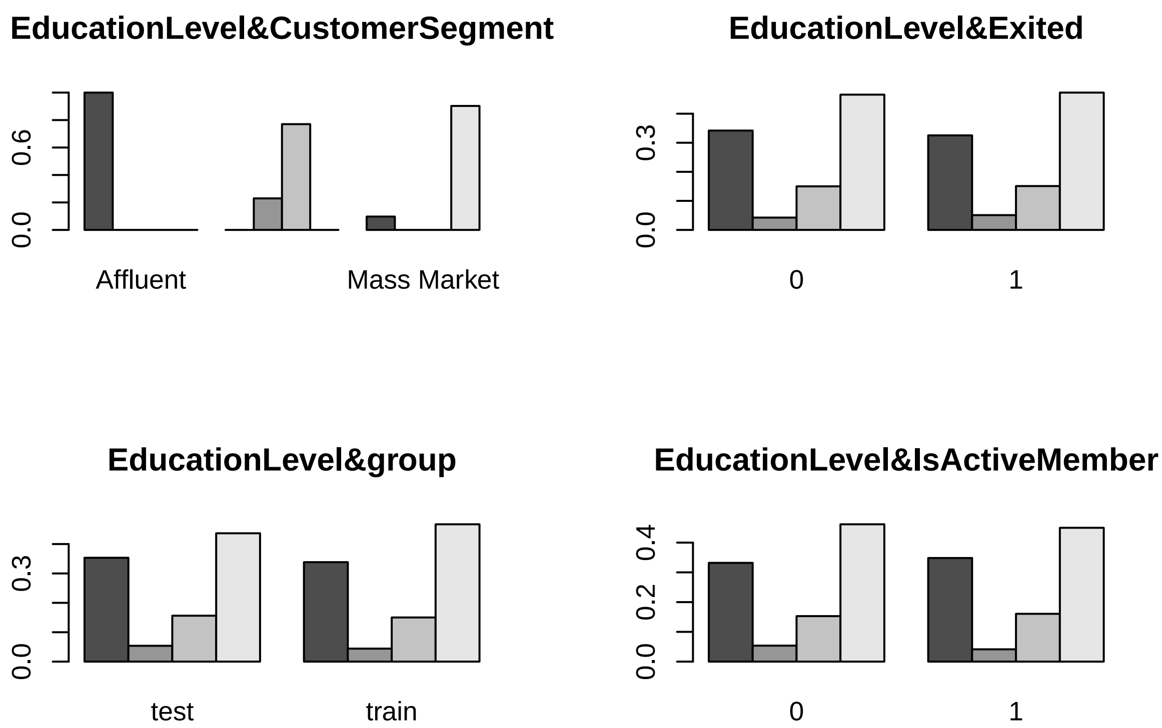

MaritalStatus&LoanStatus
0.0
0.4
0.8
Active loan
No loan

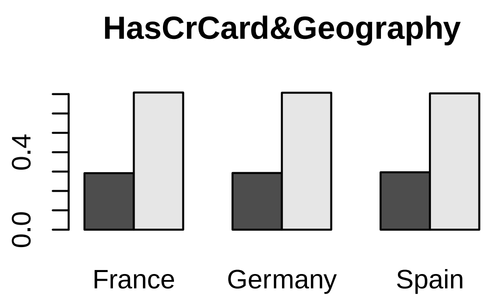

MaritalStatus&CustomerSegment
0.0
0.6
Affluent
Mass Market

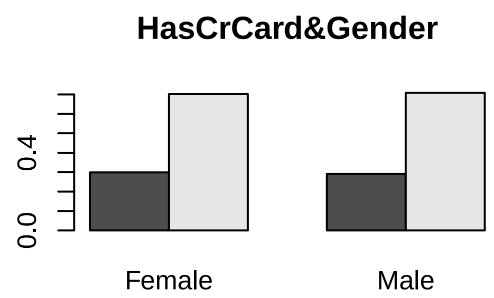

MaritalStatus&Exited
0.0
0.3
0
1

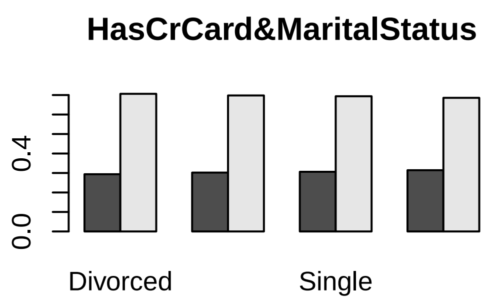

MaritalStatus&group
0.0
0.3
test
train
MaritalStatus&IsActiveMember
0.0
0.3
0
1
EducationLevel&Geography
0.0
0.2
0.4
France
Germany
Spain
EducationLevel&Gender
0.0
0.2
0.4
Female
Male

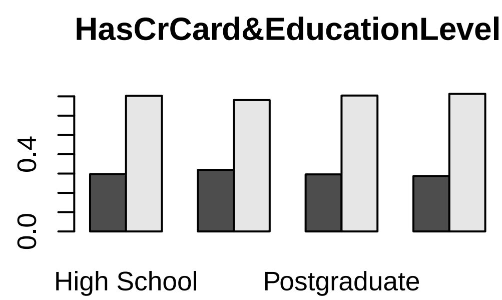

EducationLevel&MaritalStatus
0.0
0.6
Divorced
Single
EducationLevel&HasCrCard
0.0
0.2
0.4
0
1
EducationLevel&SavingsAccountFlag
0.0
0.4
0
1
EducationLevel&LoanStatus
0.0
0.4
Active loan
No loan

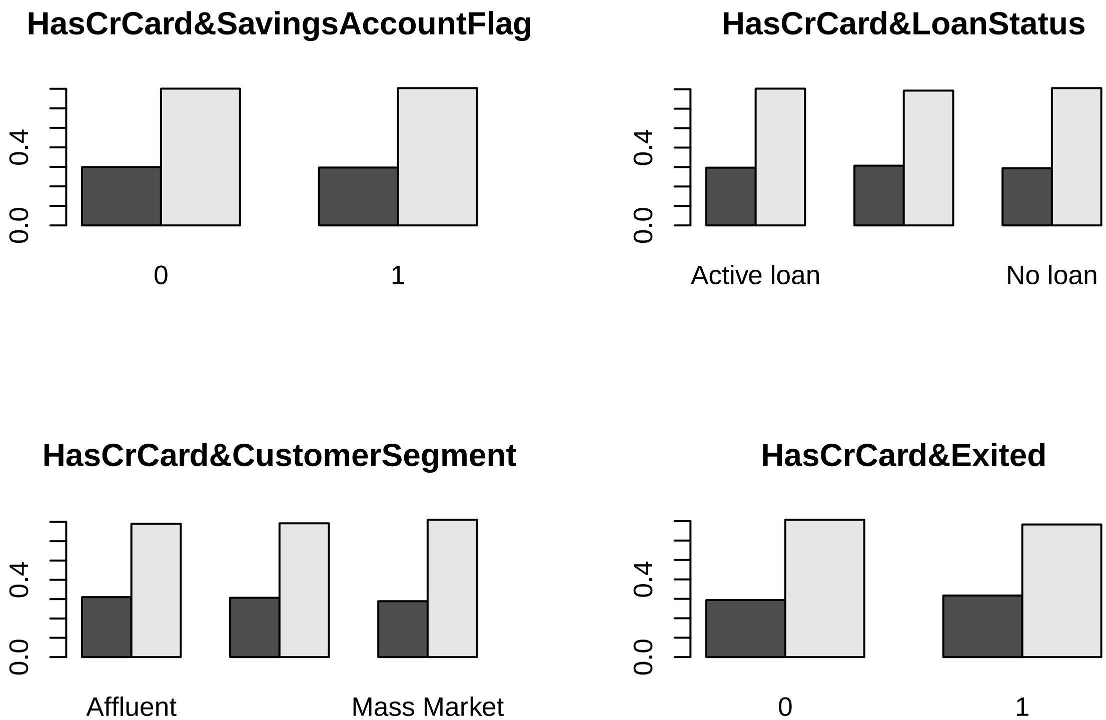

EducationLevel&CustomerSegment
0.0
0.6
Affluent
Mass Market
EducationLevel&Exited
0.0
0.3
0
1
EducationLevel&group
0.0
0.3
test
train
EducationLevel&IsActiveMember
0.0
0.2
0.4
0
1

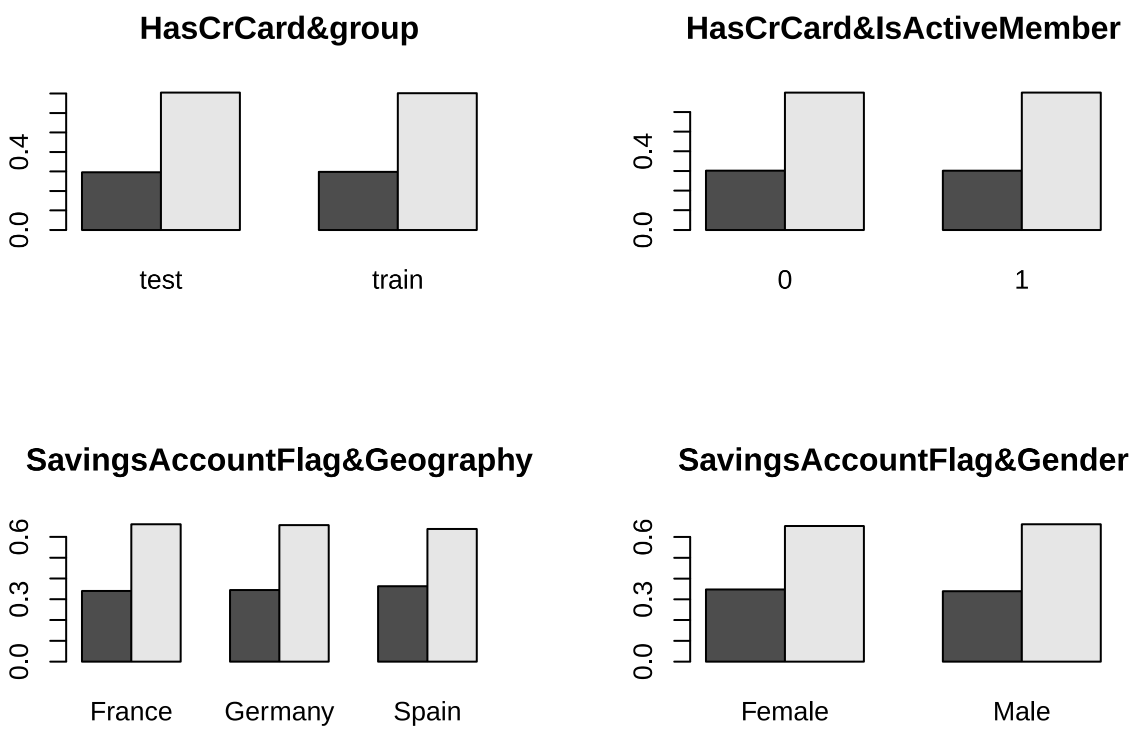

HasCrCard&Geography
0.0
0.4
France
Germany
Spain

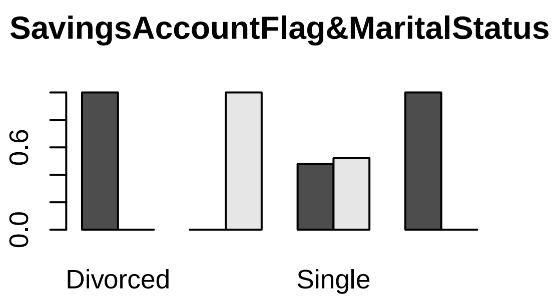

HasCrCard&Gender
0.0
0.4
Female
Male

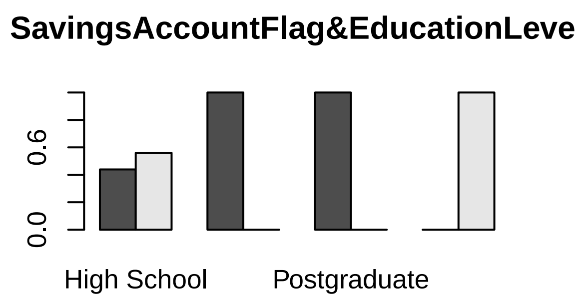

HasCrCard&MaritalStatus
0.0
0.4
Divorced
Single

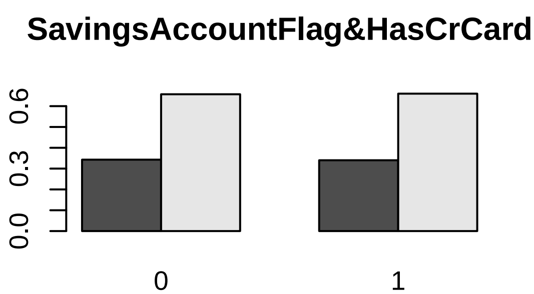

HasCrCard&EducationLevel
0.0
0.4
High School
Postgraduate

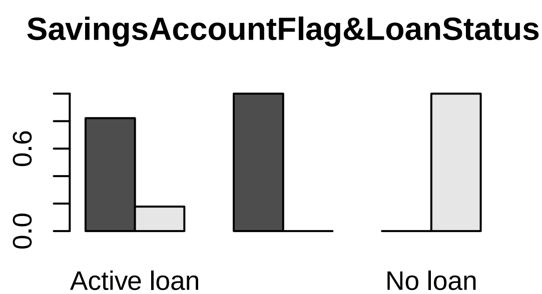

HasCrCard&SavingsAccountFlag
0.0
0.4
0
1
HasCrCard&LoanStatus
0.0
0.4
Active loan
No loan
HasCrCard&CustomerSegment
0.0
0.4
Affluent
Mass Market
HasCrCard&Exited
0.0
0.4
0
1


HasCrCard&group
0.0
0.4
test
train
HasCrCard&IsActiveMember
0.0
0.4
0
1
SavingsAccountFlag&Geography
0.0
0.3
0.6
France
Germany
Spain
SavingsAccountFlag&Gender
0.0
0.3
0.6
Female
Male

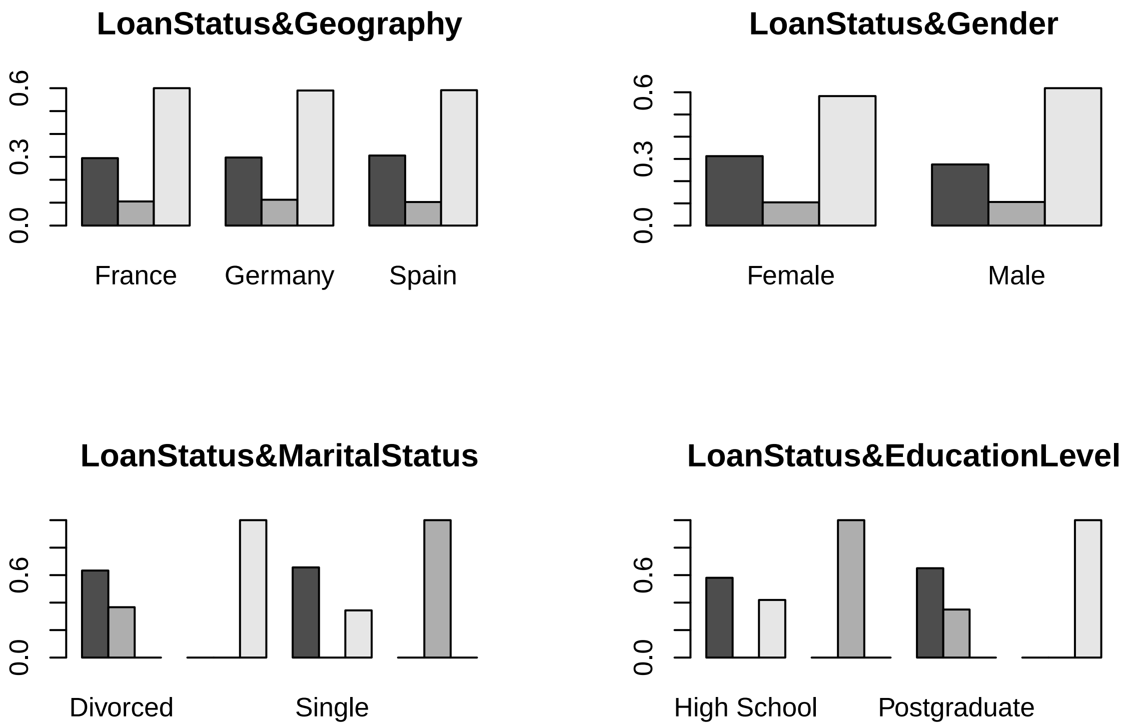

SavingsAccountFlag&MaritalStatus
0.0
0.6
Divorced
Single

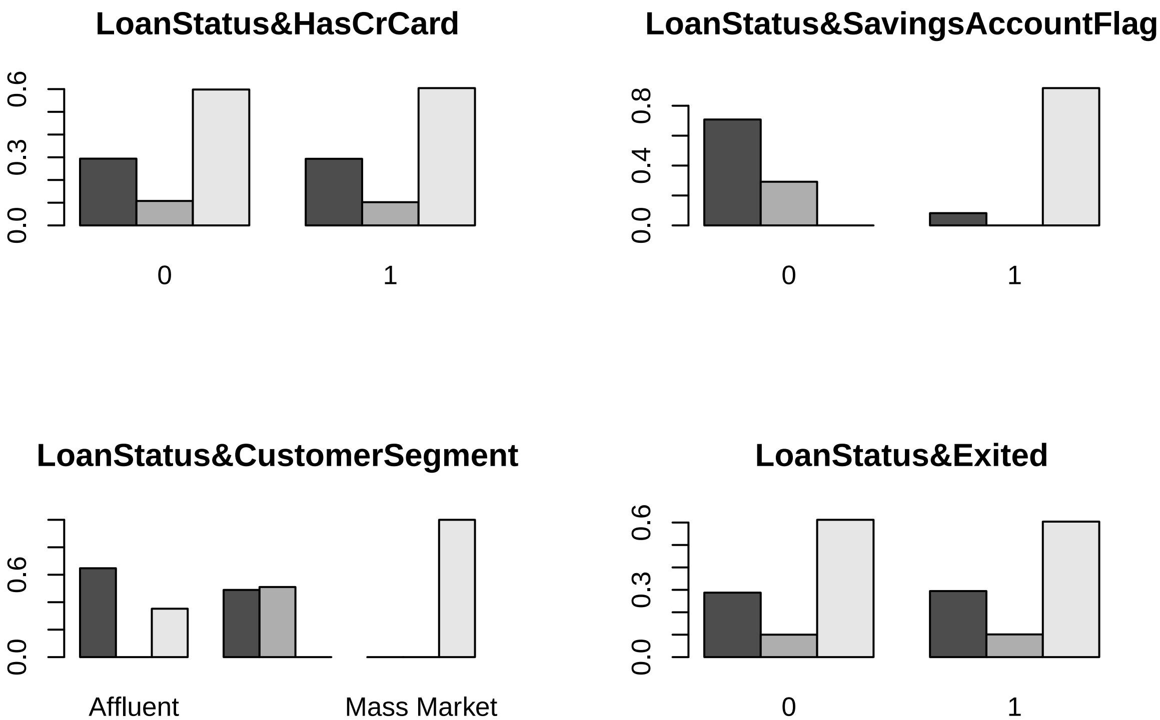

SavingsAccountFlag&EducationLeve
0.0
0.6
High School
Postgraduate

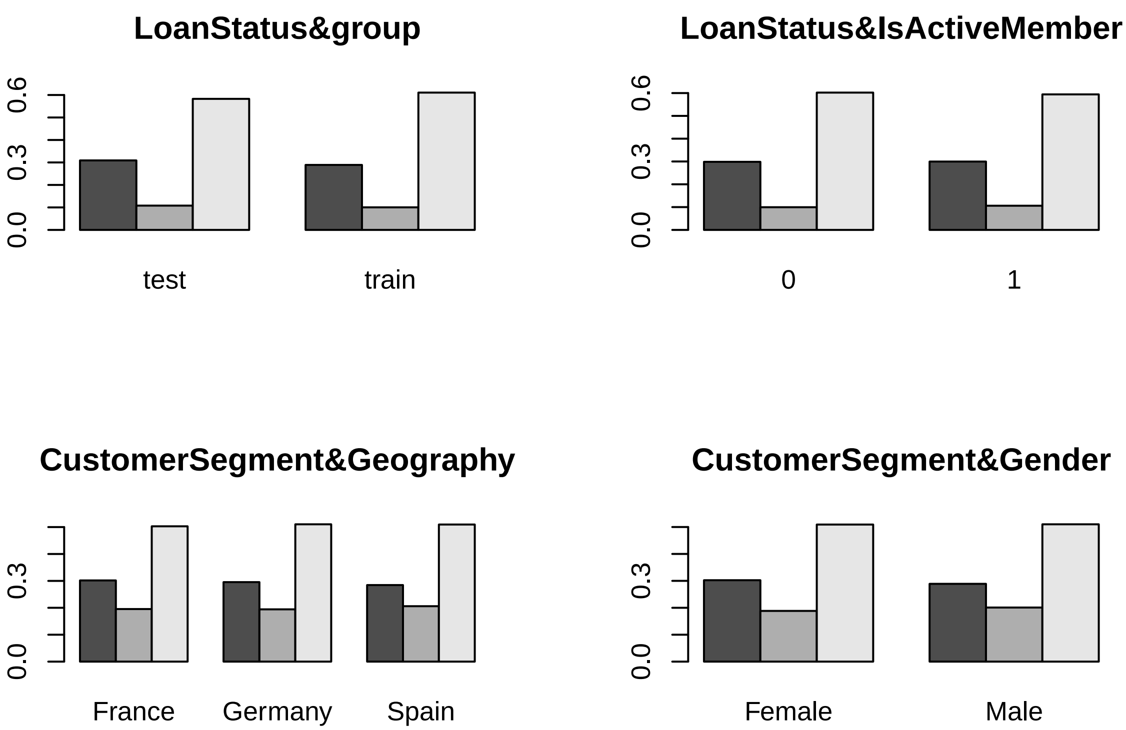

SavingsAccountFlag&HasCrCard
0.0
0.3
0.6
0
1

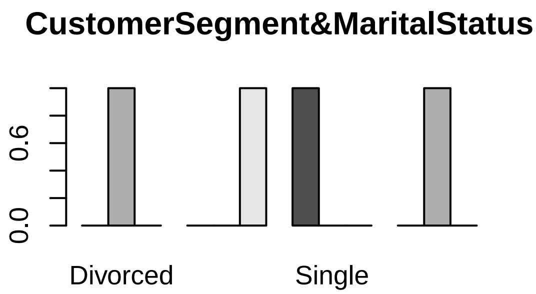

SavingsAccountFlag&LoanStatus
0.0
0.6
Active loan
No loan

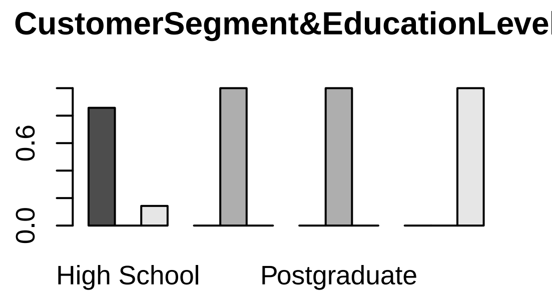

SavingsAccountFlag&CustomerSegme
0.0
0.6
Affluent
Mass Market
SavingsAccountFlag&Exited
0.0
0.3
0.6
0
1
SavingsAccountFlag&group
0.0
0.3
0.6
test
train
SavingsAccountFlag&IsActiveMembe
0.0
0.3
0.6
0
1

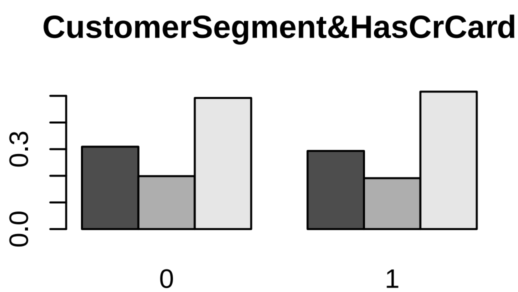

LoanStatus&Geography
0.0
0.3
0.6
France
Germany
Spain
LoanStatus&Gender
0.0
0.3
0.6
Female
Male
LoanStatus&MaritalStatus
0.0
0.6
Divorced
Single
LoanStatus&EducationLevel
0.0
0.6
High School
Postgraduate

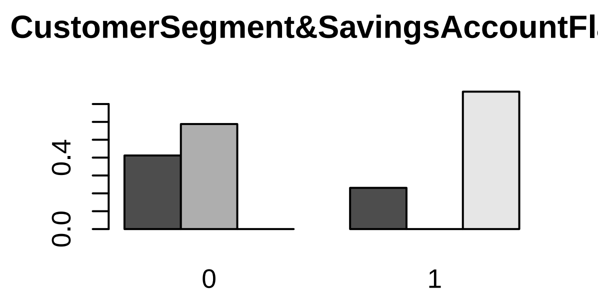

LoanStatus&HasCrCard
0.0
0.3
0.6
0
1
LoanStatus&SavingsAccountFlag
0.0
0.4
0.8
0
1
LoanStatus&CustomerSegment
0.0
0.6
Affluent
Mass Market
LoanStatus&Exited
0.0
0.3
0.6
0
1

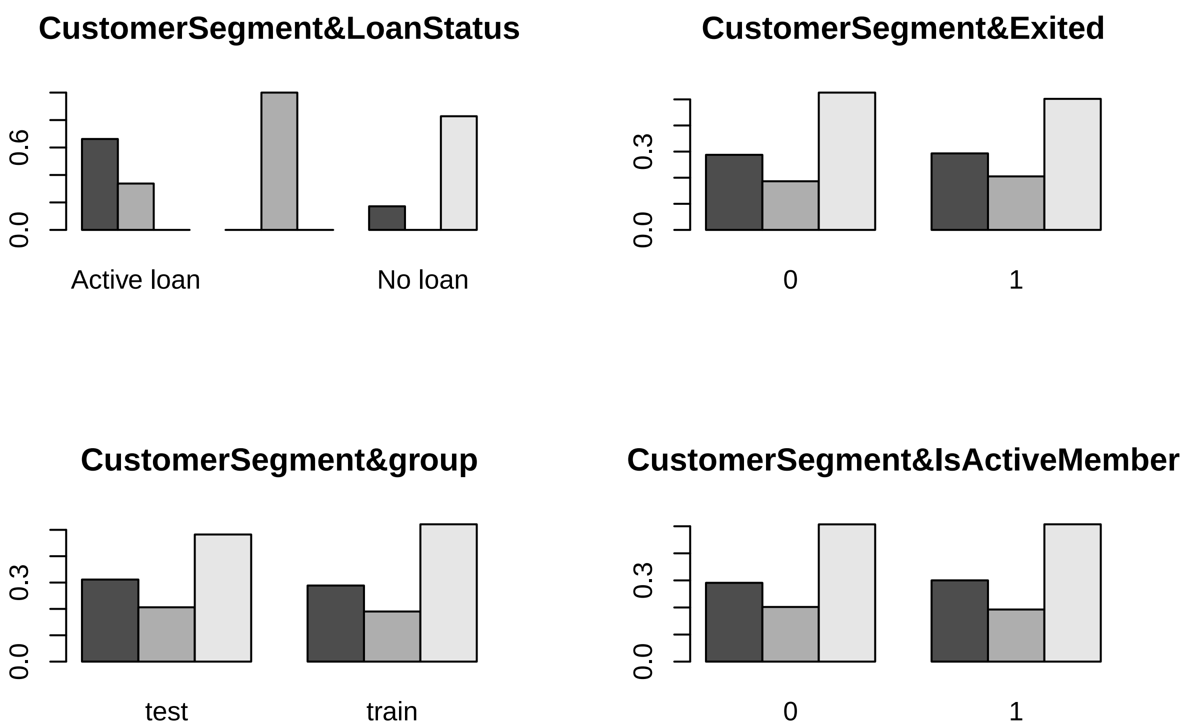

LoanStatus&group
0.0
0.3
0.6
test
train
LoanStatus&IsActiveMember
0.0
0.3
0.6
0
1
CustomerSegment&Geography
0.0
0.3
France
Germany
Spain
CustomerSegment&Gender
0.0
0.3
Female
Male

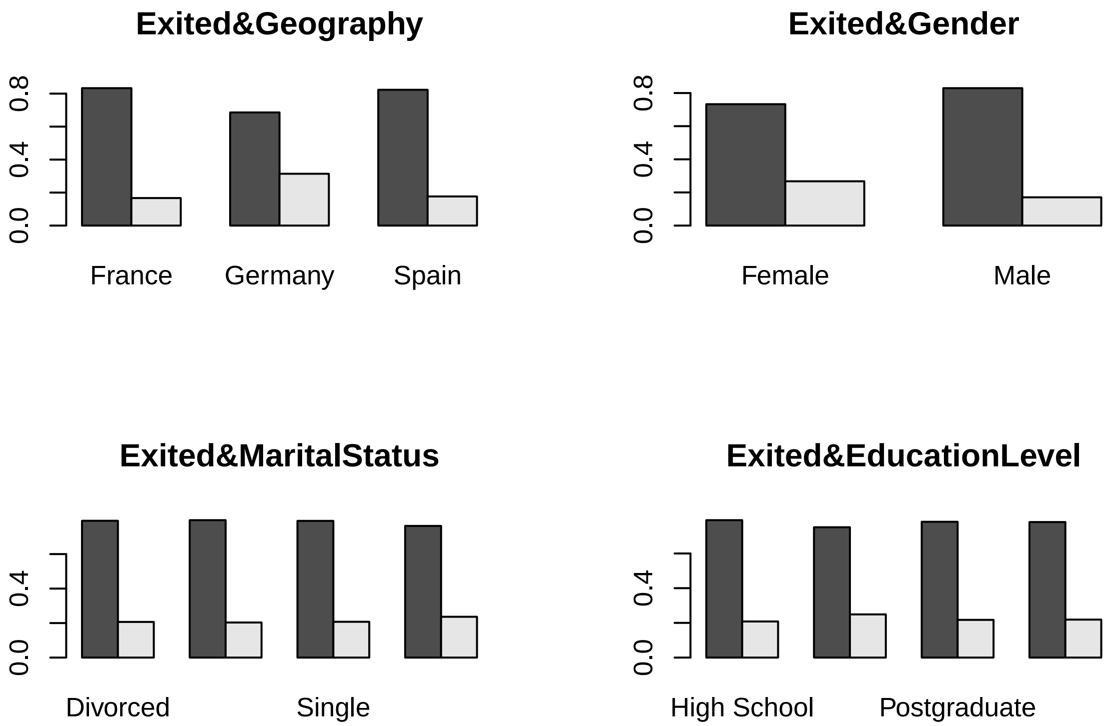

CustomerSegment&MaritalStatus
0.0
0.6
Divorced
Single

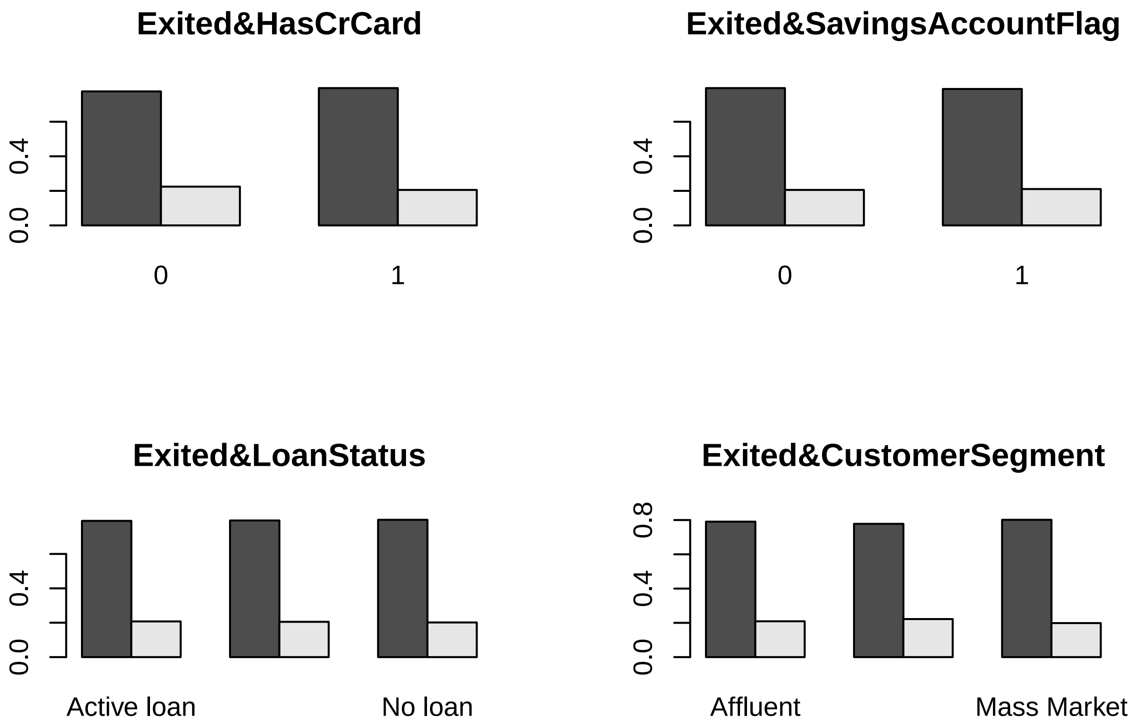

CustomerSegment&EducationLevel
0.0
0.6
High School
Postgraduate

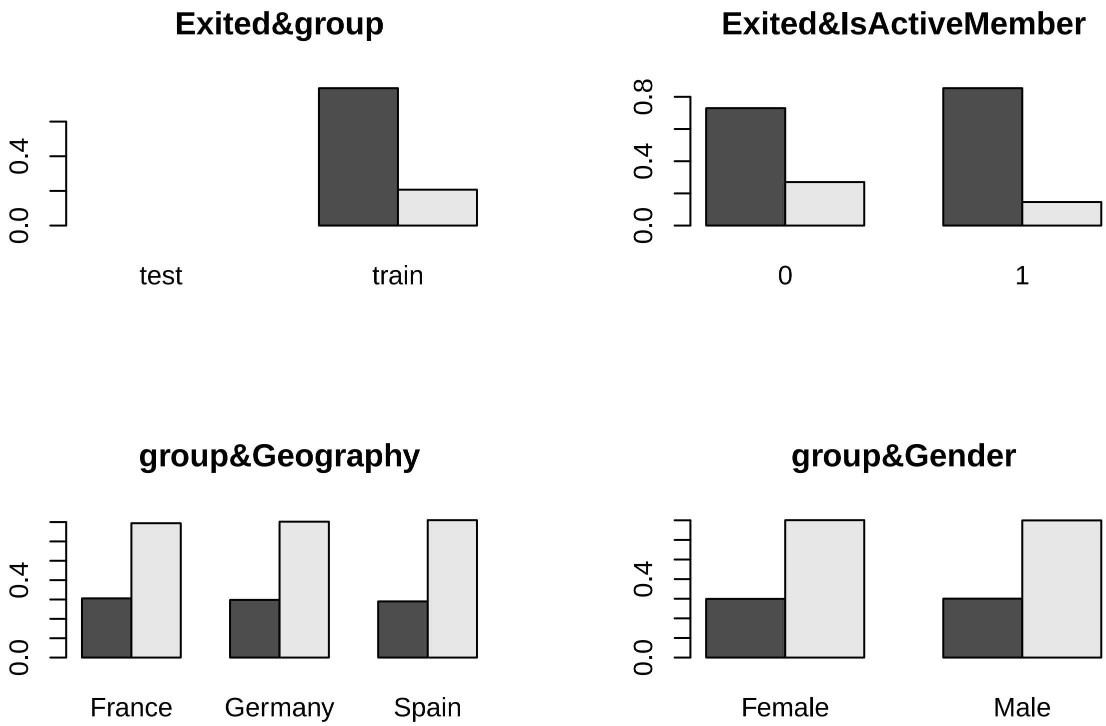

CustomerSegment&HasCrCard
0.0
0.3
0
1

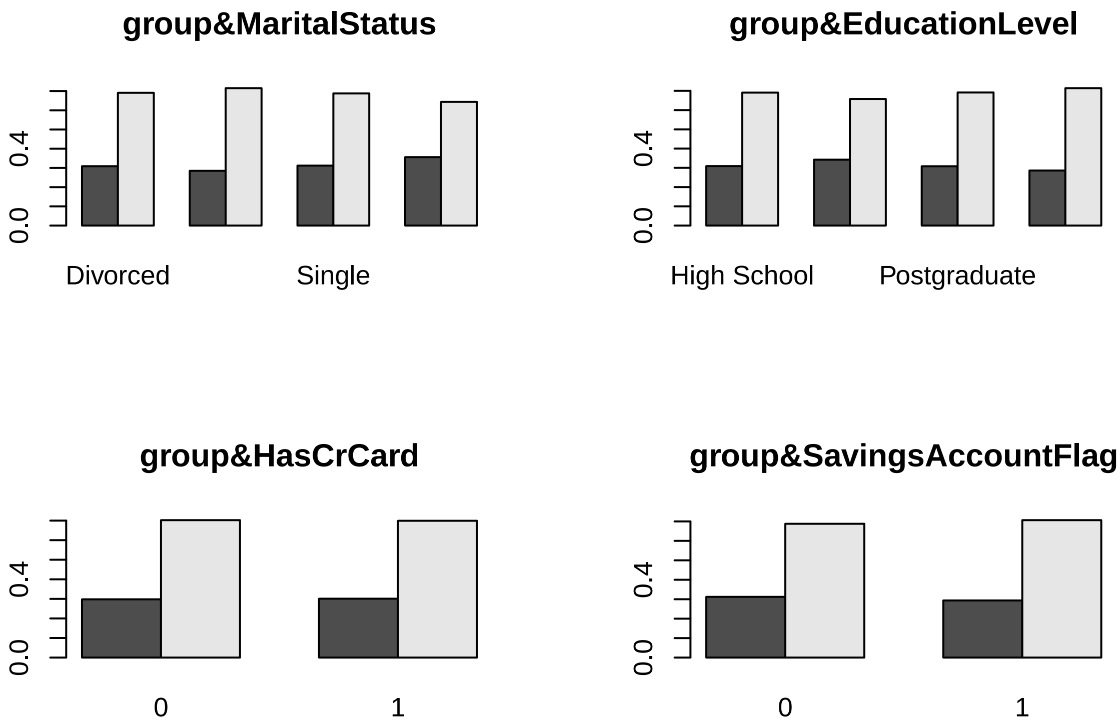

CustomerSegment&SavingsAccountFla
0.0
0.4
0
1

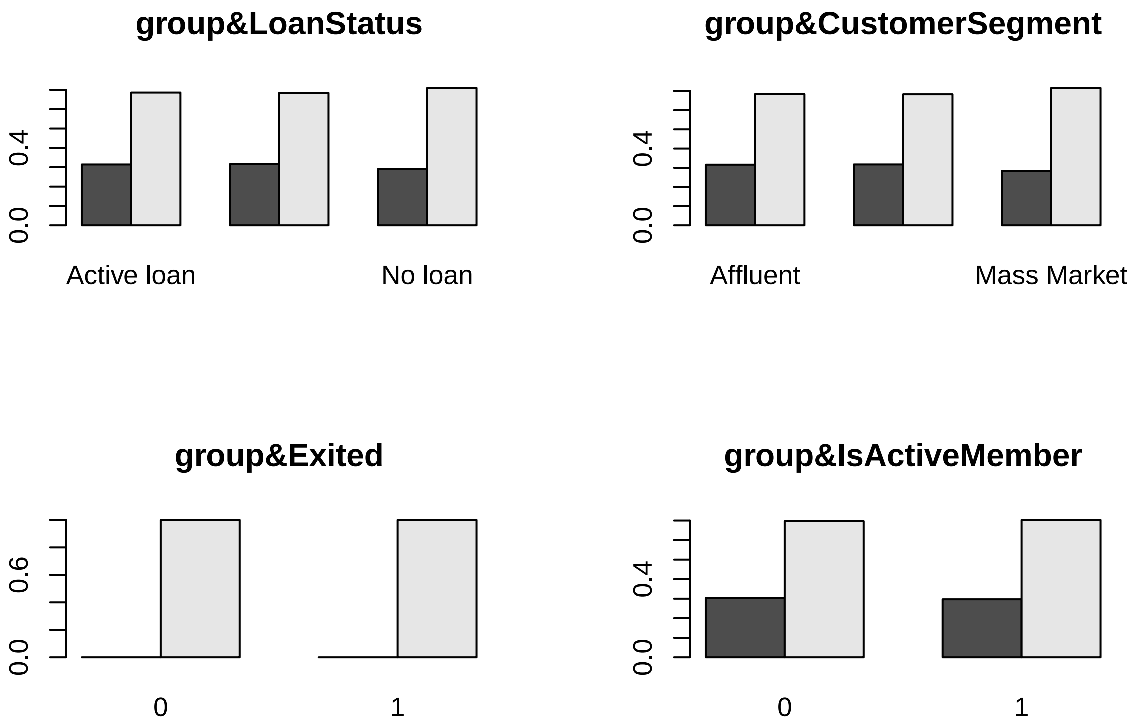

CustomerSegment&LoanStatus
0.6
0.0
Active loan
No loan
CustomerSegment&Exited
0.3
0.0
0
1
CustomerSegment&group
0.3
0.0
test
train
CustomerSegment&IsActiveMember
0.3
0.0
0
1

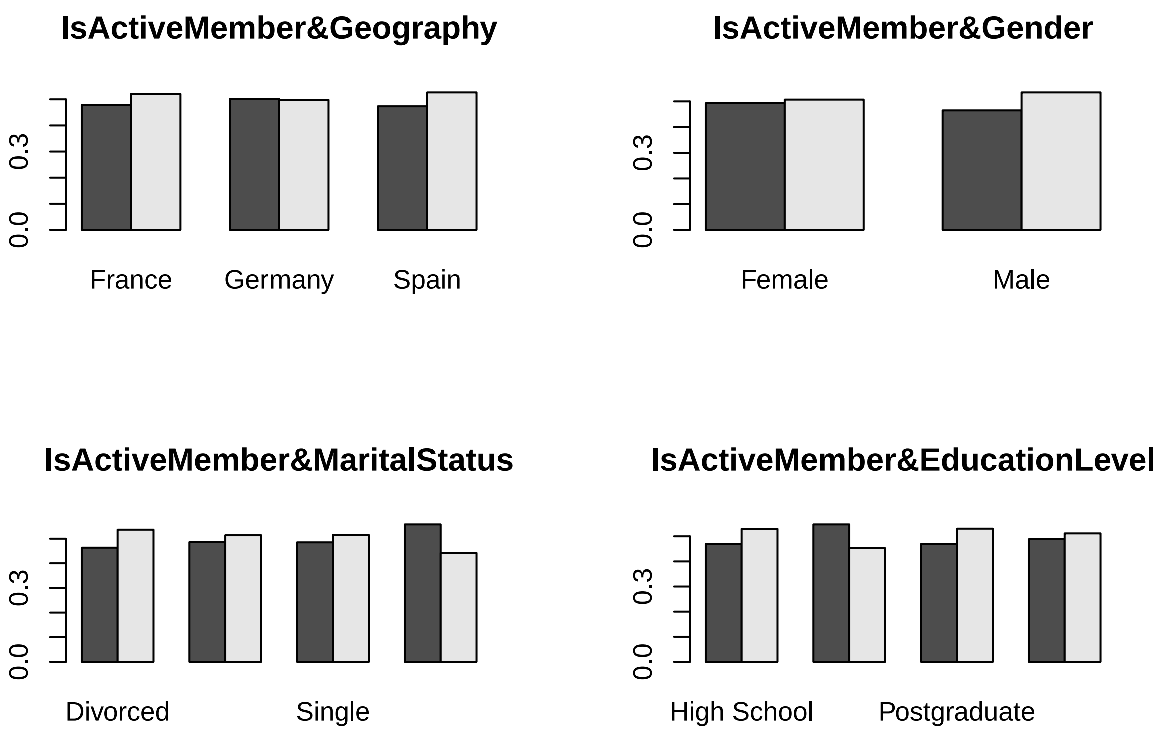

Exited&Geography
0.0
0.4
0.8
France
Germany
Spain
Exited&Gender
0.0
0.4
0.8
Female
Male
Exited&MaritalStatus
0.0
0.4
Divorced
Single
Exited&EducationLevel
0.0
0.4
High School
Postgraduate


Exited&HasCrCard
0.0
0.4
0
1
Exited&SavingsAccountFlag
0.0
0.4
0
1
Exited&LoanStatus
0.0
0.4
Active loan
No loan
Exited&CustomerSegment
0.0
0.4
0.8
Affluent
Mass Market

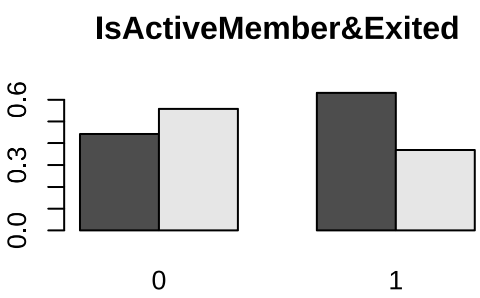

Exited&group
0.0
0.4
test
train
Exited&IsActiveMember
0.0
0.4
0.8
0
1
group&Geography
0.0
0.4
France
Germany
Spain
group&Gender
0.0
0.4
Female
Male

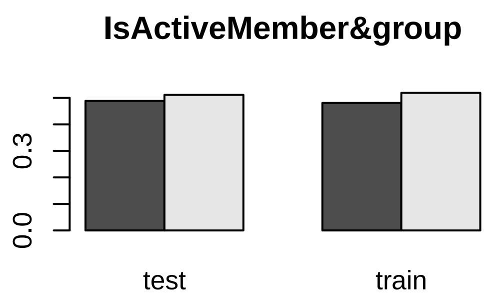

group&MaritalStatus
0.0
0.4
Divorced
Single
group&EducationLevel
0.0
0.4
High School
Postgraduate
group&HasCrCard
0.0
0.4
0
1
group&SavingsAccountFlag
0.0
0.4
0
1


group&LoanStatus
0.0
0.4
Active loan
No loan
group&CustomerSegment
0.0
0.4
Affluent
Mass Market
group&Exited
0.0
0.6
0
1
group&IsActiveMember
0.0
0.4
0
1


IsActiveMember&Geography
0.0
0.3
France
Germany
Spain
IsActiveMember&Gender
0.0
0.3
Female
Male
IsActiveMember&MaritalStatus
0.0
0.3
Divorced
Single
IsActiveMember&EducationLevel
0.0
0.3
High School
Postgraduate


IsActiveMember&HasCrCard
0.0
0.3
0
1
IsActiveMember&SavingsAccountFlag
0.0
0.3
0
1
IsActiveMember&LoanStatus
0.0
0.3
Active loan
No loan
IsActiveMember&CustomerSegment
0.0
0.3
Affluent
Mass Market


IsActiveMember&Exited
0.0
0.3
0.6
0
1


IsActiveMember&group
0.0
0.3
test
train