BCG - Task 4

Lingyu Tan

First, load and process the dataset (We could have saved the R workspace so we don't need to process the data every time we save it as a csv file).

```r
data_combo <- read.csv("data_combo.csv")
data_combo_new <- data_combo[, c(2:6, 11:36)]

# Too many categories in `activity_new`.
# Replace all NA and those whose occurrences are less than 30 with "Others"

data_combo_new$activity_new[is.na(data_combo_new$activity_new)] <- "Others"
data_combo_new$activity_new[data_combo_new$activity_new %in% names(which(table(data_combo_new$activity_n

data_combo_new$activity_new <- as.factor(data_combo_new$activity_new)
data_combo_new$channel_sales <- as.factor(data_combo_new$channel_sales)
data_combo_new$origin_up <- as.factor(data_combo_new$origin_up)
```

Then we use a full dataset with all possible explanatory variables to fit a generalized linear model using logit link function as follows:

```r
churn_logit_full <- glm(churn ~ ., data = data_combo_new, family = binomial(link = logit))
summary(churn_logit_full)
```

```
##
## Call:
## glm(formula = churn ~ ., family = binomial(link = logit), data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1729  -0.5010  -0.4103  -0.3307   3.0020
##
## Coefficients:
##                                               Estimate Std. Error z value
## (Intercept)                                  -3.864e+00  8.167e-01  -4.732
## activity_newapdekpcbwosbxepsfxclislboipuxpop -1.065e-02  1.497e-01  -0.071
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc -1.513e-01  6.108e-01  -0.248
## activity_newbalskueexlmuccwdffilikwxasupasxf  2.993e-01  4.476e-01   0.669
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo  2.624e-01  4.444e-01   0.590
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf  4.199e-02  4.463e-01   0.094
## activity_newcfdsselwimsklimddecfifseabdkxfcs -1.778e+00  1.013e+00  -1.755
## activity_newckadsdebplpkplelfspfoiucmxkeppus  3.099e-02  6.116e-01   0.051
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa -4.425e-01  2.937e-01  -1.507
## activity_newcluecxlameloamldmasudocsbmaoamdw -1.846e-01  3.238e-01  -0.570
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu -3.662e-01  3.467e-01  -1.056
```

1

```
## activity_newdaobdssbkieoukwxbopxiiospudkopwl  -4.777e-01  5.241e-01  -0.911
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl  -1.050e+00  7.333e-01  -1.432
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk  -8.946e-02  6.159e-01  -0.145
## activity_newdupxuibdflmskeieweeofcaluuuiioix  -6.374e-01  7.324e-01  -0.870
## activity_newfkmblacmaapkaoauabpwpuweokkeiali   5.490e-02  5.494e-01   0.100
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb   1.785e-01  2.055e-01   0.869
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm   7.412e-02  4.421e-01   0.168
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi  -3.105e-01  5.374e-01  -0.578
## activity_newipdldckuswupeifllfbwccfpeafludfi   2.101e-01  3.867e-01   0.543
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce  -3.139e-02  1.773e-01  -0.177
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca  -1.364e+00  1.017e+00  -1.340
## activity_newkwuslieomapmswolewpobpplkaooaaew   4.449e-02  2.135e-01   0.208
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx  -1.216e-01  5.338e-01  -0.228
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil   9.729e-02  4.501e-01   0.216
## activity_newOthers                            7.984e-02  7.942e-02   1.005
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae   4.421e-01  4.254e-01   1.039
## activity_newsaxlifeumaobawxpemwuopbwwldlucff   4.839e-01  3.689e-01   1.312
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc  -2.402e-01  4.314e-01  -0.557
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu  -4.140e-01  4.664e-01  -0.888
## activity_newsumdxiaiudmaioicexmiwuudlblkissm  -2.693e-01  6.075e-01  -0.443
## activity_newsxublbwoeuckkocekklxkllcdxxaisop   4.745e-01  3.474e-01   1.366
## activity_newwxemiwkumpibllwklfbcooafckufkdlm   3.342e-01  2.692e-01   1.242
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.177e+01  2.663e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.184e-01  1.450e-01  -1.506
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  2.485e-02  9.259e-02   0.268
## channel_saleslmkebamcaaclubfxadlmueccxoimlema -3.785e-01  1.239e-01  -3.054
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds -1.144e+01  1.691e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf -9.284e-02  1.237e-01  -0.750
## cons_12m                                      -1.207e-07  2.564e-07  -0.471
## cons_gas_12m                                  -1.209e-06  4.662e-07  -2.594
## cons_last_month                              -1.674e-06  2.480e-06  -0.675
## forecast_cons_12m                            -9.010e-06  1.697e-05  -0.531
## forecast_cons_year                            1.204e-05  2.961e-05   0.407
## forecast_discount_energy                      6.656e-03  6.922e-03   0.962
## forecast_meter_rent_12m                       7.890e-04  5.648e-04   1.397
## forecast_price_energy_p1                     -8.019e+00  4.808e+00  -1.668
## forecast_price_energy_p2                     -1.588e+00  2.792e+00  -0.569
## forecast_price_pow_p1                         3.599e-02  2.245e-02   1.603
## has_gas                                      -8.927e-02  1.145e-01  -0.779
## imp_cons                                     -5.761e-05  3.126e-04  -0.184
## margin_gross_pow_ele                          9.700e-03  2.223e-03   4.363
## margin_net_pow_ele                            1.446e-03  1.983e-03   0.729
## nb_prod_act                                  -3.008e-02  5.749e-02  -0.523
## net_margin                                    3.748e-04  1.595e-04   2.351
## num_years_antig                              -1.337e-01  6.116e-02  -2.187
## origin_upewxeelcelemmiwuafmddpobolfuxioce    -1.033e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws     3.235e-01  4.748e-01   0.681
## origin_upldkssxwpmemidmecebumciepifcamkci     6.791e-01  4.741e-01   1.432
## origin_uplxidpiddsbxsbosboudacockeimpuepw     8.982e-01  4.713e-01   1.906
## origin_upusapbepcfoloekilkwsdiboslwaxobdp    -1.057e+01  3.785e+02  -0.028
## pow_max                                      -8.436e-03  2.364e-03  -3.569
## p1_var                                       -1.674e+00  6.256e+00  -0.268
## p2_var                                        5.517e+00  3.371e+00   1.636
## p3_var                                       -3.856e+01  1.314e+01  -2.934
```

```
## p1_fix                                               2.650e-02  2.783e-02   0.952
## p2_fix                                              -2.287e-03  2.972e-02  -0.077
## p3_fix                                               1.640e-01  4.155e-02   3.948
## day                                                  1.092e-04  1.645e-04   0.663
## day_modif                                           -4.755e-05  3.773e-05  -1.260
## day_renewal                                         -3.877e-04  5.502e-04  -0.705
##                                                     Pr(>|z|)
## (Intercept)                                         2.22e-06 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop        0.943294
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc        0.804361
## activity_newbalskueexlmuccwdffilikwxasupasxf        0.503773
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo        0.554900
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf        0.925053
## activity_newcfdsselwimsklimddecfifseabdkxfcs        0.079271 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus        0.959587
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa        0.131874
## activity_newcluecxlameloamldmasudocsbmaoamdw        0.568513
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu        0.290839
## activity_newdaobdssbkieoukwxbopxiiospudkopwl        0.362053
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl        0.152184
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk        0.884519
## activity_newdupxuibdflmskeieweeofcaluuuiioix        0.384188
## activity_newfkmblacmaapkaoauabpwpuweokkeiali        0.920393
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb        0.385074
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm        0.866861
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi        0.563443
## activity_newipdldckuswupeifllfbwccfpeafludfi        0.587020
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce        0.859486
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca        0.180171
## activity_newkwuslieomapmswolewpobpplkaooaaew        0.834970
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx        0.819757
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil        0.828882
## activity_newOthers                                  0.314738
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae        0.298740
## activity_newsaxlifeumaobawxpemwuopbwwldlucff        0.189652
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc        0.577750
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu        0.374658
## activity_newsumdxiaiudmaioicexmiwuudlblkissm        0.657525
## activity_newsxublbwoeuckkocekklxkllcdxxaisop        0.172002
## activity_newwxemiwkumpibllwklfbcooafckufkdlm        0.214361
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu       0.964740
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci       0.131992
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua       0.788411
## channel_saleslmkebamcaaclubfxadlmueccxoimlema       0.002258 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds       0.946054
## channel_salesusilxuppasemubllopkaafesmlibmsdf       0.453033
## cons_12m                                            0.637908
## cons_gas_12m                                        0.009495 **
## cons_last_month                                     0.499638
## forecast_cons_12m                                   0.595474
## forecast_cons_year                                  0.684352
## forecast_discount_energy                            0.336245
## forecast_meter_rent_12m                             0.162422
## forecast_price_energy_p1                            0.095370 .
```

```
## forecast_price_energy_p2                          0.569416
## forecast_price_pow_p1                             0.108913
## has_gas                                           0.435775
## imp_cons                                          0.853763
## margin_gross_pow_ele                              1.28e-05 ***
## margin_net_pow_ele                                0.465803
## nb_prod_act                                       0.600817
## net_margin                                        0.018737 *
## num_years_antig                                   0.028765 *
## origin_upewxeelcelemmiwuafmddpobolfuxioce         0.984609
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws         0.495597
## origin_upldkssxwpmemidmecebumciepifcamkci         0.152062
## origin_uplxidpiddsbxsbosboudacockeimpuepw         0.056687 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp         0.977722
## pow_max                                           0.000358 ***
## p1_var                                            0.789045
## p2_var                                            0.101767
## p3_var                                            0.003349 **
## p1_fix                                            0.341037
## p2_fix                                            0.938672
## p3_fix                                            7.89e-05 ***
## day                                               0.507082
## day_modif                                         0.207569
## day_renewal                                       0.481004
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10021  on 15760  degrees of freedom
## Residual deviance:  9579  on 15690  degrees of freedom
## AIC: 9721
##
## Number of Fisher Scoring iterations: 12
```

By observing the p-values we can see there are quite a lot insignificant variables to deal with and perhaps multicollinearity issues. We need to handle multicollinearity first as it will greatly influence the accuracy of p-values. Here we can find corresponding VIF:

```
vif(churn_logit_full)
```

```
##                            GVIF Df GVIF^(1/(2*Df))
## activity_new           2.364195 32        1.013535
## channel_sales          2.432352  6        1.076884
## cons_12m               9.704898  1        3.115268
## cons_gas_12m           1.248247  1        1.117250
## cons_last_month        9.494067  1        3.081244
## forecast_cons_12m      6.307682  1        2.511510
## forecast_cons_year    39.133510  1        6.255678
## forecast_discount_energy 1.937995 1       1.392119
## forecast_meter_rent_12m  2.981016 1       1.726562
## forecast_price_energy_p1 21.566800 1      4.644007
## forecast_price_energy_p2 24.991130 1      4.999113
```

4

```
## forecast_price_pow_p1        13.220558  1          3.636009
## has_gas                       2.215091  1          1.488318
## imp_cons                     37.033959  1          6.085553
## margin_gross_pow_ele          5.748865  1          2.397679
## margin_net_pow_ele            5.557569  1          2.357450
## nb_prod_act                   2.030850  1          1.425079
## net_margin                    3.939679  1          1.984862
## num_years_antig              12.040452  1          3.469935
## origin_up                     1.716355  5          1.055506
## pow_max                       2.619233  1          1.618404
## p1_var                       25.305967  1          5.030504
## p2_var                       32.337088  1          5.686571
## p3_var                      264.299909  1         16.257303
## p1_fix                       17.316102  1          4.161262
## p2_fix                      165.028559  1         12.846344
## p3_fix                      121.965850  1         11.043815
## day                          12.270137  1          3.502876
## day_modif                     1.383430  1          1.176193
## day_renewal                   1.146651  1          1.070818
```

By observing GVIF^(1/(2*Df)), the prices of energy/power seem highly correlated, which can also be confirmed by the following correlation matrices:

```
cor(data_combo_new[, 23:28])
```

```
##                p1_var     p2_var     p3_var         p1_fix         p2_fix     p3_fix
## p1_var  1.0000000 -0.3475132 -0.6157592  0.4159744126 -0.6501226848 -0.5900898
## p2_var -0.3475132  1.0000000  0.8276109 -0.1025347316  0.8025554021  0.8141300
## p3_var -0.6157592  0.8276109  1.0000000 -0.1406708577  0.9745064009  0.9803801
## p1_fix  0.4159744 -0.1025347 -0.1406709  1.0000000000 -0.0005740018 -0.2573205
## p2_fix -0.6501227  0.8025554  0.9745064 -0.0005740018  1.0000000000  0.9265031
## p3_fix -0.5900898  0.8141300  0.9803801 -0.2573205096  0.9265031067  1.0000000
```

```
# Ignore the rows if `p3_var` is 0, which is similar to the above one
# cor(data_combo_new[which(data_combo_new[, 25] != 0), 23:28])
```

Cor(p3_var, p3_fix) = 0.98 indicates a strong positive correlation between `p3_var` and `p3_fix`. Considering the GVIF, we can delete the variable with largest GVIF^(1/(2*Df)) value one by one until the model keeps all marginal significant explanatory variables.

```
# m1: remove p3_var
m1 <- update(churn_logit_full, . ~ . - p3_var)
vif(m1)
```

```
##                              GVIF Df GVIF^(1/(2*Df))
## activity_new             2.344427 32          1.013402
## channel_sales            2.426823  6          1.076680
## cons_12m                 9.721224  1          3.117888
## cons_gas_12m             1.245759  1          1.116136
## cons_last_month          9.507032  1          3.083347
## forecast_cons_12m        6.390602  1          2.527964
## forecast_cons_year      39.824770  1          6.310687
```

5

```
## forecast_discount_energy   1.702267  1      1.304710
## forecast_meter_rent_12m     2.994089  1      1.730344
## forecast_price_energy_p1   20.383061  1      4.514760
## forecast_price_energy_p2   25.113746  1      5.011362
## forecast_price_pow_p1      13.097300  1      3.619019
## has_gas                     2.190886  1      1.480164
## imp_cons                   37.613731  1      6.133003
## margin_gross_pow_ele        5.585305  1      2.363325
## margin_net_pow_ele          5.573627  1      2.360853
## nb_prod_act                 2.008789  1      1.417318
## net_margin                  3.975399  1      1.993840
## num_years_antig            11.989626  1      3.462604
## origin_up                   1.715249  5      1.055438
## pow_max                     2.556649  1      1.598952
## p1_var                     24.641322  1      4.964003
## p2_var                     32.433451  1      5.695037
## p1_fix                     17.131125  1      4.138976
## p2_fix                    100.637768  1     10.031838
## p3_fix                     66.127547  1      8.131885
## day                        12.246821  1      3.499546
## day_modif                   1.378016  1      1.173889
## day_renewal                 1.145907  1      1.070470
```

```r
# m2: remove p2_fix
m2 <- update(m1, . ~ . - p2_fix)
vif(m2)
```

```
##                                GVIF Df GVIF^(1/(2*Df))
## activity_new                2.337687 32       1.013357
## channel_sales               2.448344  6       1.077472
## cons_12m                    9.688780  1       3.112680
## cons_gas_12m                1.245591  1       1.116061
## cons_last_month             9.463888  1       3.076343
## forecast_cons_12m           6.207207  1       2.491427
## forecast_cons_year         35.328824  1       5.943806
## forecast_discount_energy    1.442279  1       1.200949
## forecast_meter_rent_12m     3.017301  1       1.737038
## forecast_price_energy_p1   20.889662  1       4.570521
## forecast_price_energy_p2   24.573872  1       4.957204
## forecast_price_pow_p1      12.739231  1       3.569206
## has_gas                     2.202977  1       1.484243
## imp_cons                   33.088899  1       5.752295
## margin_gross_pow_ele        5.597155  1       2.365831
## margin_net_pow_ele          5.544181  1       2.354608
## nb_prod_act                 2.021369  1       1.421749
## net_margin                  3.724264  1       1.929835
## num_years_antig            12.043692  1       3.470402
## origin_up                   1.706786  5       1.054916
## pow_max                     2.514386  1       1.585682
## p1_var                     18.282490  1       4.275803
## p2_var                     29.244581  1       5.407826
## p1_fix                     12.525951  1       3.539202
## p3_fix                      7.322690  1       2.706047
## day                        12.291290  1       3.505894
```

```
## day_modif                    1.371858  1        1.171263
## day_renewal                  1.136152  1        1.065904
```

```r
# m3: remove forecast_cons_year
m3 <- update(m2, . ~ . - forecast_cons_year)
vif(m3)
```

```
##                                 GVIF Df GVIF^(1/(2*Df))
## activity_new             2.315369 32        1.013205
## channel_sales            2.448636  6        1.077483
## cons_12m                 9.672882  1        3.110126
## cons_gas_12m             1.245919  1        1.116207
## cons_last_month          9.423075  1        3.069703
## forecast_cons_12m        5.847704  1        2.418203
## forecast_discount_energy 1.439378  1        1.199741
## forecast_meter_rent_12m  3.013502  1        1.735944
## forecast_price_energy_p1 20.848237  1       4.565987
## forecast_price_energy_p2 24.555357  1       4.955336
## forecast_price_pow_p1    12.639603  1        3.555222
## has_gas                  2.203412  1        1.484390
## imp_cons                 2.354553  1        1.534455
## margin_gross_pow_ele     5.592941  1        2.364940
## margin_net_pow_ele       5.537262  1        2.353139
## nb_prod_act              2.022676  1        1.422208
## net_margin               3.607791  1        1.899419
## num_years_antig          12.043067  1       3.470312
## origin_up                1.706299  5        1.054886
## pow_max                  2.495573  1        1.579738
## p1_var                   18.249057  1        4.271891
## p2_var                   29.227273  1        5.406225
## p1_fix                   12.465745  1        3.530686
## p3_fix                   7.341712  1        2.709559
## day                      12.290780  1        3.505821
## day_modif                1.371503  1        1.171112
## day_renewal              1.136145  1        1.065901
```

```r
# m4: remove p2_var
m4 <- update(m3, . ~ . - p2_var)
vif(m4)
```

```
##                                 GVIF Df GVIF^(1/(2*Df))
## activity_new             2.300671 32        1.013104
## channel_sales            2.443609  6        1.077298
## cons_12m                 9.653728  1        3.107045
## cons_gas_12m             1.246255  1        1.116358
## cons_last_month          9.404132  1        3.066616
## forecast_cons_12m        5.816443  1        2.411730
## forecast_discount_energy 1.428232  1        1.195087
## forecast_meter_rent_12m  3.018473  1        1.737375
## forecast_price_energy_p1 20.135509  1       4.487261
## forecast_price_energy_p2 3.630494  1        1.905386
## forecast_price_pow_p1    12.508140  1        3.536685
## has_gas                  2.204199  1        1.484654
```

```
## imp_cons                     2.356496   1         1.535088
## margin_gross_pow_ele         5.613943   1         2.369376
## margin_net_pow_ele          5.553869   1         2.356665
## nb_prod_act                  2.024013   1         1.422678
## net_margin                   3.568729   1         1.889108
## num_years_antig             12.047049   1         3.470886
## origin_up                    1.704683   5         1.054786
## pow_max                      2.494301   1         1.579336
## p1_var                      17.871581   1         4.227479
## p1_fix                      12.267888   1         3.502555
## p3_fix                       6.058301   1         2.461362
## day                         12.295868   1         3.506546
## day_modif                    1.370620   1         1.170735
## day_renewal                  1.137559   1         1.066564
```

```r
# m5: remove forecast_price_energy_p1
m5 <- update(m4, . ~ . - forecast_price_energy_p1)
vif(m5)
```

```
##                              GVIF Df GVIF^(1/(2*Df))
## activity_new             2.295468 32        1.013068
## channel_sales            2.442377  6        1.077253
## cons_12m                 9.679683  1        3.111219
## cons_gas_12m             1.246576  1        1.116502
## cons_last_month          9.431255  1        3.071035
## forecast_cons_12m        5.803180  1        2.408979
## forecast_discount_energy 1.328457  1        1.152587
## forecast_meter_rent_12m  2.960924  1        1.720734
## forecast_price_energy_p2 3.601503  1        1.897763
## forecast_price_pow_p1    7.717645  1        2.778065
## has_gas                  2.237959  1        1.495981
## imp_cons                 2.356416  1        1.535062
## margin_gross_pow_ele     5.644690  1        2.375856
## margin_net_pow_ele       5.562506  1        2.358496
## nb_prod_act              2.057153  1        1.434278
## net_margin               3.556265  1        1.885806
## num_years_antig         11.996967  1        3.463664
## origin_up                1.703238  5        1.054697
## pow_max                  2.478244  1        1.574244
## p1_var                   2.880435  1        1.697184
## p1_fix                   7.772736  1        2.787963
## p3_fix                   5.970953  1        2.443553
## day                     12.228293  1        3.496898
## day_modif                1.364399  1        1.168075
## day_renewal              1.137136  1        1.066366
```

```r
# m6: remove day
m6 <- update(m5, . ~ . - day)
vif(m6)
```

```
##                         GVIF Df GVIF^(1/(2*Df))
## activity_new        2.280997 32        1.012968
## channel_sales       2.370714  6        1.074583
```

```
## cons_12m               9.692013  1        3.113200
## cons_gas_12m           1.246371  1        1.116410
## cons_last_month        9.442466  1        3.072860
## forecast_cons_12m      5.795653  1        2.407416
## forecast_discount_energy 1.326202  1      1.151608
## forecast_meter_rent_12m 2.956216  1       1.719365
## forecast_price_energy_p2 3.599875  1      1.897334
## forecast_price_pow_p1  7.558209  1        2.749220
## has_gas               2.239765  1        1.496584
## imp_cons              2.355434  1        1.534742
## margin_gross_pow_ele   5.644640  1        2.375845
## margin_net_pow_ele     5.569032  1        2.359880
## nb_prod_act           2.058512  1        1.434752
## net_margin            3.546637  1        1.883252
## num_years_antig       1.901350  1        1.378895
## origin_up             1.683844  5        1.053489
## pow_max               2.472069  1        1.572281
## p1_var                2.869888  1        1.694075
## p1_fix                7.613840  1        2.759319
## p3_fix                5.974667  1        2.444313
## day_modif             1.365498  1        1.168545
## day_renewal           1.134403  1        1.065084
```

```
# m7: remove cons_12m
m7 <- update(m6, . ~ . - cons_12m)
vif(m7)
```

```
##                         GVIF Df GVIF^(1/(2*Df))
## activity_new         2.199580 32        1.012393
## channel_sales        2.365189  6        1.074374
## cons_gas_12m         1.243206  1        1.114992
## cons_last_month      1.665863  1        1.290683
## forecast_cons_12m    5.599148  1        2.366252
## forecast_discount_energy 1.326291  1      1.151647
## forecast_meter_rent_12m 2.957105  1       1.719624
## forecast_price_energy_p2 3.599640  1      1.897272
## forecast_price_pow_p1 7.542153  1        2.746298
## has_gas              2.233974  1        1.494648
## imp_cons             2.138574  1        1.462386
## margin_gross_pow_ele  5.644331  1        2.375780
## margin_net_pow_ele    5.568456  1        2.359758
## nb_prod_act          2.054077  1        1.433205
## net_margin           3.548330  1        1.883701
## num_years_antig      1.900813  1        1.378700
## origin_up            1.683192  5        1.053449
## pow_max              2.473163  1        1.572629
## p1_var               2.869167  1        1.693862
## p1_fix               7.599146  1        2.756655
## p3_fix               5.976052  1        2.444597
## day_modif            1.363757  1        1.167800
## day_renewal          1.134327  1        1.065048
```

```
# m8: remove p1_fix
m8 <- update(m7, . ~ . - p1_fix)
vif(m8)
```

```
##                             GVIF Df GVIF^(1/(2*Df))
## activity_new            2.195898 32        1.012366
## channel_sales           2.361504  6        1.074234
## cons_gas_12m            1.243165  1        1.114973
## cons_last_month         1.666208  1        1.290817
## forecast_cons_12m       5.595505  1        2.365482
## forecast_discount_energy 1.320340 1        1.149061
## forecast_meter_rent_12m 2.919548  1        1.708668
## forecast_price_energy_p2 3.587181 1        1.893985
## forecast_price_pow_p1   1.474610  1        1.214335
## has_gas                 2.238421  1        1.496135
## imp_cons                2.138750  1        1.462447
## margin_gross_pow_ele    5.653314  1        2.377670
## margin_net_pow_ele      5.578130  1        2.361806
## nb_prod_act             2.058444  1        1.434728
## net_margin              3.540745  1        1.881687
## num_years_antig         1.901170  1        1.378829
## origin_up               1.675595  5        1.052972
## pow_max                 2.467358  1        1.570782
## p1_var                  2.684149  1        1.638337
## p3_fix                  5.928153  1        2.434780
## day_modif               1.363984  1        1.167897
## day_renewal             1.134297  1        1.065034
```

```
# m9: remove p3_fix
m9 <- update(m8, . ~ . - p3_fix)
vif(m9)
```

```
##                             GVIF Df GVIF^(1/(2*Df))
## activity_new            2.168166 32        1.012165
## channel_sales           2.348291  6        1.073732
## cons_gas_12m            1.242821  1        1.114819
## cons_last_month         1.667070  1        1.291151
## forecast_cons_12m       5.677012  1        2.382648
## forecast_discount_energy 1.319153 1        1.148544
## forecast_meter_rent_12m 2.585531  1        1.607959
## forecast_price_energy_p2 1.865385 1        1.365791
## forecast_price_pow_p1   1.378190  1        1.173963
## has_gas                 2.247382  1        1.499127
## imp_cons                2.135620  1        1.461376
## margin_gross_pow_ele    5.558794  1        2.357709
## margin_net_pow_ele      5.529002  1        2.351383
## nb_prod_act             2.069544  1        1.438591
## net_margin              3.631920  1        1.905760
## num_years_antig         1.895844  1        1.376897
## origin_up               1.673901  5        1.052866
## pow_max                 2.462268  1        1.569162
## p1_var                  2.168094  1        1.472445
## day_modif               1.360398  1        1.166361
```

```
## day_renewal               1.132054  1          1.063980
```

```
# m10: remove forecast_cons_12m
m10 <- update(m9, . ~ . - forecast_cons_12m)
vif(m10)
```

```
##                            GVIF Df GVIF^(1/(2*Df))
## activity_new            2.150260 32        1.012034
## channel_sales           2.345289  6        1.073618
## cons_gas_12m            1.242549  1        1.114697
## cons_last_month         1.665850  1        1.290678
## forecast_discount_energy 1.316163 1        1.147241
## forecast_meter_rent_12m 2.586141  1        1.608148
## forecast_price_energy_p2 1.852925 1        1.361222
## forecast_price_pow_p1   1.372893  1        1.171705
## has_gas                 2.248653  1        1.499551
## imp_cons                1.659058  1        1.288044
## margin_gross_pow_ele    5.527130  1        2.350985
## margin_net_pow_ele      5.508133  1        2.346941
## nb_prod_act             2.071155  1        1.439151
## net_margin              1.842217  1        1.357283
## num_years_antig         1.895451  1        1.376754
## origin_up               1.673458  5        1.052838
## pow_max                 2.120543  1        1.456208
## p1_var                  2.165372  1        1.471520
## day_modif               1.360325  1        1.166330
## day_renewal             1.129667  1        1.062858
```

```
# m11: remove margin_gross_pow_ele
m11 <- update(m10, . ~ . - margin_gross_pow_ele)
vif(m11)
```

```
##                            GVIF Df GVIF^(1/(2*Df))
## activity_new            2.139510 32        1.011955
## channel_sales           2.346488  6        1.073664
## cons_gas_12m            1.238899  1        1.113058
## cons_last_month         1.658040  1        1.287649
## forecast_discount_energy 1.300253 1        1.140286
## forecast_meter_rent_12m 2.606400  1        1.614435
## forecast_price_energy_p2 1.830849 1        1.353089
## forecast_price_pow_p1   1.343485  1        1.159088
## has_gas                 2.226858  1        1.492266
## imp_cons                1.652344  1        1.285435
## margin_net_pow_ele      1.301654  1        1.140900
## nb_prod_act             2.050562  1        1.431978
## net_margin              1.784016  1        1.335671
## num_years_antig         1.903460  1        1.379659
## origin_up               1.673618  5        1.052848
## pow_max                 2.037040  1        1.427249
## p1_var                  2.143563  1        1.464091
## day_modif               1.364961  1        1.168315
## day_renewal             1.127979  1        1.062064
```

Now every GVIF is less than 4 indicating removal of multicollinearity. Then we should deal with the insignificant explanatory variables by looking at their p-values.

```
summary(m11)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_discount_energy + forecast_meter_rent_12m +
##     forecast_price_energy_p2 + forecast_price_pow_p1 + has_gas +
##     imp_cons + margin_net_pow_ele + nb_prod_act + net_margin +
##     num_years_antig + origin_up + pow_max + p1_var + day_modif +
##     day_renewal, family = binomial(link = logit), data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0319  -0.5028  -0.4122  -0.3365   2.9585
##
## Coefficients:
##                                            Estimate Std. Error z value
## (Intercept)                               -2.723e+00  6.148e-01  -4.428
## activity_newapdekpcbwosbxepsfxclislboipuxpop -2.103e-02  1.462e-01  -0.144
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc -1.366e-01  6.106e-01  -0.224
## activity_newbalskueexlmuccwdffilikwxasupasxf  3.407e-01  4.462e-01   0.763
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo  1.989e-01  4.487e-01   0.443
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf  5.805e-02  4.448e-01   0.131
## activity_newcfdsselwimsklimddecfifseabdkxfcs -1.768e+00  1.013e+00  -1.746
## activity_newckadsdebplpkplelfspfoiucmxkeppus  2.720e-02  6.112e-01   0.044
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa -4.170e-01  2.923e-01  -1.427
## activity_newcluecxlameloamldmasudocsbmaoamdw -8.289e-02  3.157e-01  -0.263
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu -3.058e-01  3.393e-01  -0.901
## activity_newdaobdssbkieoukwxbopxiiospudkopwl -4.652e-01  5.229e-01  -0.890
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl -9.598e-01  7.312e-01  -1.313
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk -6.738e-02  6.149e-01  -0.110
## activity_newdupxuibdflmskeieweeofcaluuuiioix -6.104e-01  7.317e-01  -0.834
## activity_newfkmblacmaapkaoauabpwpuweokkeiali  7.971e-02  5.507e-01   0.145
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb  2.029e-01  2.042e-01   0.993
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm  9.079e-02  4.403e-01   0.206
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi -2.823e-01  5.337e-01  -0.529
## activity_newipdldckuswupeifllfbwccfpeafludfi  2.473e-01  3.860e-01   0.641
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce -1.506e-02  1.769e-01  -0.085
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca -1.324e+00  1.017e+00  -1.302
## activity_newkwuslieomapmswolewpobpplkaooaaew  6.786e-02  2.124e-01   0.320
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx -9.620e-02  5.333e-01  -0.180
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil  1.652e-01  4.467e-01   0.370
## activity_newOthers                            9.872e-02  7.874e-02   1.254
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae  4.446e-01  4.252e-01   1.045
## activity_newsaxlifeumaobawxpemwuopbwwldlucff  4.400e-01  3.702e-01   1.189
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc -2.454e-01  4.309e-01  -0.570
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu -3.884e-01  4.658e-01  -0.834
## activity_newsumdxiaiudmaioicexmiwuudlblkissm -3.183e-01  6.066e-01  -0.525
## activity_newsxublbwoeuckkocekklxkllcdxxaisop  4.932e-01  3.463e-01   1.424
## activity_newwxemiwkumpibllwklfbcooafckufkdlm  3.465e-01  2.682e-01   1.292
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.174e+01  2.669e+02  -0.044
```

12

```
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.028e-01  1.430e-01  -1.418
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  3.295e-02  9.147e-02   0.360
## channel_saleslmkebamcaaclubfxadlmueccxoimlema -3.666e-01  1.228e-01  -2.986
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds -1.145e+01  1.691e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf -9.482e-02  1.233e-01  -0.769
## cons_gas_12m                                   -1.175e-06  4.577e-07  -2.568
## cons_last_month                                -2.657e-06  1.030e-06  -2.578
## forecast_discount_energy                       -7.164e-03  5.654e-03  -1.267
## forecast_meter_rent_12m                         8.095e-04  5.179e-04   1.563
## forecast_price_energy_p2                        1.848e+00  7.524e-01   2.457
## forecast_price_pow_p1                           1.227e-02  6.500e-03   1.887
## has_gas                                        -8.632e-02  1.147e-01  -0.752
## imp_cons                                        4.585e-05  6.624e-05   0.692
## margin_net_pow_ele                              6.447e-03  8.818e-04   7.311
## nb_prod_act                                    -3.217e-02  5.794e-02  -0.555
## net_margin                                      2.390e-04  9.971e-05   2.397
## num_years_antig                                -9.147e-02  2.421e-02  -3.779
## origin_upewxeelcelemmiwuafmddpobolfuxioce      -1.038e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       2.911e-01  4.689e-01   0.621
## origin_upldkssxwpmemidmecebumciepifcamkci       6.147e-01  4.679e-01   1.314
## origin_uplxidpiddsbxsbosboudacockeimpuepw       8.457e-01  4.652e-01   1.818
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      -1.062e+01  3.782e+02  -0.028
## pow_max                                        -5.488e-03  1.949e-03  -2.816
## p1_var                                         -1.273e+00  1.725e+00  -0.738
## day_modif                                      -5.698e-05  3.737e-05  -1.525
## day_renewal                                    -3.813e-04  5.444e-04  -0.700
##                                                 Pr(>|z|)
## (Intercept)                                     9.49e-06 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.885659
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.822936
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.445167
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.657464
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.896146
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.080737 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.964506
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.153671
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.792885
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.367539
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.373603
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.189288
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.912742
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.404102
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.884923
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.320483
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.836649
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.596812
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.521735
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.932174
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.193055
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.749327
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.856848
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.711556
## activity_newOthers                              0.209940
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.295792
```

```
## activity_newsaxlifeumaobawxpemwuopbwwldlucff  0.234585
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc  0.568960
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu  0.404355
## activity_newsumdxiaiudmaioicexmiwuudlblkissm  0.599802
## activity_newsxublbwoeuckkocekklxkllcdxxaisop  0.154395
## activity_newwxemiwkumpibllwklfbcooafckufkdlm  0.196409
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu 0.964901
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci 0.156168
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua 0.718699
## channel_saleslmkebamcaaclubfxadlmueccxoimlema 0.002825 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds 0.946019
## channel_salesusilxuppasemubllopkaafesmlibmsdf 0.441995
## cons_gas_12m                                   0.010236 *
## cons_last_month                                0.009926 **
## forecast_discount_energy                       0.205092
## forecast_meter_rent_12m                        0.118041
## forecast_price_energy_p2                       0.014027 *
## forecast_price_pow_p1                          0.059096 .
## has_gas                                        0.451774
## imp_cons                                       0.488873
## margin_net_pow_ele                             2.65e-13 ***
## nb_prod_act                                    0.578718
## net_margin                                     0.016514 *
## num_years_antig                                0.000158 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce      0.984531
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws      0.534625
## origin_upldkssxwpmemidmecebumciepifcamkci      0.188870
## origin_uplxidpiddsbxsbosboudacockeimpuepw      0.069068 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      0.977598
## pow_max                                        0.004865 **
## p1_var                                         0.460662
## day_modif                                      0.127316
## day_renewal                                    0.483697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9626.4  on 15701  degrees of freedom
## AIC: 9746.4
##
## Number of Fisher Scoring iterations: 12
```

Update the most insignificant variable according to the p-values in summary as below:

```
# m12: remove nb_prod_act
m12 <- update(m11, . ~ . - nb_prod_act)
summary(m12)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
```

```
##        cons_last_month + forecast_discount_energy + forecast_meter_rent_12m +
##        forecast_price_energy_p2 + forecast_price_pow_p1 + has_gas +
##        imp_cons + margin_net_pow_ele + net_margin + num_years_antig +
##        origin_up + pow_max + p1_var + day_modif + day_renewal, family = binomial(link = logit),
##        data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0316  -0.5024  -0.4127  -0.3368   2.9655
##
## Coefficients:
##                                                Estimate Std. Error z value
## (Intercept)                                   -2.751e+00  6.128e-01  -4.490
## activity_newapdekpcbwosbxepsfxclislboipuxpop  -1.809e-02  1.460e-01  -0.124
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc  -1.368e-01  6.106e-01  -0.224
## activity_newbalskueexlmuccwdffilikwxasupasxf   3.413e-01  4.462e-01   0.765
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo   1.980e-01  4.486e-01   0.441
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf   5.966e-02  4.447e-01   0.134
## activity_newcfdsselwimsklimddecfifseabdkxfcs  -1.765e+00  1.013e+00  -1.743
## activity_newckadsdebplpkplelfspfoiucmxkeppus   2.997e-02  6.112e-01   0.049
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa  -4.159e-01  2.923e-01  -1.423
## activity_newcluecxlameloamldmasudocsbmaoamdw  -8.205e-02  3.157e-01  -0.260
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu  -3.044e-01  3.393e-01  -0.897
## activity_newdaobdssbkieoukwxbopxiiospudkopwl  -4.670e-01  5.228e-01  -0.893
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl  -9.570e-01  7.312e-01  -1.309
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk  -6.564e-02  6.149e-01  -0.107
## activity_newdupxuibdflmskeieweeofcaluuuiioix  -6.187e-01  7.314e-01  -0.846
## activity_newfkmblacmaapkaoauabpwpuweokkeiali   8.154e-02  5.508e-01   0.148
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb   2.065e-01  2.042e-01   1.011
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm   9.248e-02  4.403e-01   0.210
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi  -2.794e-01  5.337e-01  -0.524
## activity_newipdldckuswupeifllfbwccfpeafludfi   2.460e-01  3.860e-01   0.637
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce  -1.320e-02  1.769e-01  -0.075
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca  -1.323e+00  1.017e+00  -1.300
## activity_newkwuslieomapmswolewpobpplkaooaaew   7.300e-02  2.122e-01   0.344
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx  -9.440e-02  5.333e-01  -0.177
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil   1.668e-01  4.467e-01   0.373
## activity_newOthers                             9.875e-02  7.874e-02   1.254
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae   4.445e-01  4.252e-01   1.045
## activity_newsaxlifeumaobawxpemwuopbwwldlucff   4.387e-01  3.702e-01   1.185
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc  -2.631e-01  4.302e-01  -0.612
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu  -3.851e-01  4.658e-01  -0.827
## activity_newsumdxiaiudmaioicexmiwuudlblkissm  -3.156e-01  6.066e-01  -0.520
## activity_newsxublbwoeuckkocekklxkllcdxxaisop   4.947e-01  3.463e-01   1.429
## activity_newwxemiwkumpibllwklfbcooafckufkdlm   3.486e-01  2.682e-01   1.300
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.174e+01  2.669e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.017e-01  1.430e-01  -1.410
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  3.335e-02  9.147e-02   0.365
## channel_saleslmkebamcaaclubfxadlmueccxoimlema -3.664e-01  1.228e-01  -2.985
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds -1.145e+01  1.691e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf -9.441e-02  1.233e-01  -0.766
## cons_gas_12m                                  -1.167e-06  4.574e-07  -2.552
## cons_last_month                               -2.682e-06  1.029e-06  -2.606
## forecast_discount_energy                      -7.601e-03  5.605e-03  -1.356
```

```
## forecast_meter_rent_12m                                   8.137e-04  5.177e-04   1.572
## forecast_price_energy_p2                                   1.843e+00  7.523e-01   2.449
## forecast_price_pow_p1                                      1.225e-02  6.500e-03   1.885
## has_gas                                                   -1.300e-01  8.358e-02  -1.555
## imp_cons                                                   4.577e-05  6.629e-05   0.690
## margin_net_pow_ele                                         6.465e-03  8.813e-04   7.335
## net_margin                                                 2.407e-04  1.002e-04   2.404
## num_years_antig                                           -9.125e-02  2.420e-02  -3.770
## origin_upewxeelcelemmiwuafmddpobolfuxioce                -1.038e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws                 2.900e-01  4.688e-01   0.619
## origin_upldkssxwpmemidmecebumciepifcamkci                 6.136e-01  4.678e-01   1.312
## origin_uplxidpiddsbxsbosboudacockeimpuepw                 8.442e-01  4.652e-01   1.815
## origin_upusapbepcfoloekilkwsdiboslwaxobdp                -1.062e+01  3.782e+02  -0.028
## pow_max                                                   -5.495e-03  1.949e-03  -2.819
## p1_var                                                    -1.279e+00  1.725e+00  -0.741
## day_modif                                                 -5.709e-05  3.737e-05  -1.528
## day_renewal                                               -3.926e-04  5.448e-04  -0.721
##                                                           Pr(>|z|)
## (Intercept)                                               7.13e-06 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop              0.901440
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc              0.822751
## activity_newbalskueexlmuccwdffilikwxasupasxf              0.444314
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo              0.658976
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf              0.893293
## activity_newcfdsselwimsklimddecfifseabdkxfcs              0.081298 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus              0.960885
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa              0.154829
## activity_newcluecxlameloamldmasudocsbmaoamdw              0.794950
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu              0.369756
## activity_newdaobdssbkieoukwxbopxiiospudkopwl              0.371725
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl              0.190613
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk              0.914990
## activity_newdupxuibdflmskeieweeofcaluuuiioix              0.397586
## activity_newfkmblacmaapkaoauabpwpuweokkeiali              0.882304
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb              0.311901
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm              0.833635
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi              0.600620
## activity_newipdldckuswupeifllfbwccfpeafludfi              0.523884
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce              0.940528
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca              0.193516
## activity_newkwuslieomapmswolewpobpplkaooaaew              0.730765
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx              0.859494
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil              0.708793
## activity_newOthers                                        0.209810
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae              0.295823
## activity_newsaxlifeumaobawxpemwuopbwwldlucff              0.235966
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc              0.540845
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu              0.408322
## activity_newsumdxiaiudmaioicexmiwuudlblkissm              0.602829
## activity_newsxublbwoeuckkocekklxkllcdxxaisop              0.153134
## activity_newwxemiwkumpibllwklfbcooafckufkdlm              0.193571
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu             0.964912
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci             0.158395
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua             0.715392
```

```
## channel_saleslmkebamcaaclubfxadlmueccxoimlema 0.002836 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds 0.946030
## channel_salesusilxuppasemubllopkaafesmlibmsdf 0.443931
## cons_gas_12m                                  0.010709 *
## cons_last_month                               0.009173 **
## forecast_discount_energy                      0.175020
## forecast_meter_rent_12m                       0.116012
## forecast_price_energy_p2                      0.014312 *
## forecast_price_pow_p1                         0.059397 .
## has_gas                                       0.119887
## imp_cons                                      0.489910
## margin_net_pow_ele                            2.21e-13 ***
## net_margin                                    0.016237 *
## num_years_antig                               0.000163 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce     0.984532
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws     0.536244
## origin_upldkssxwpmemidmecebumciepifcamkci     0.189661
## origin_uplxidpiddsbxsbosboudacockeimpuepw     0.069554 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp     0.977597
## pow_max                                       0.004818 **
## p1_var                                        0.458534
## day_modif                                     0.126574
## day_renewal                                   0.471063
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9626.8  on 15702  degrees of freedom
## AIC: 9744.8
##
## Number of Fisher Scoring iterations: 12

# m13: remove imp_cons
m13 <- update(m12, . ~ . - imp_cons)
summary(m13)


##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_discount_energy + forecast_meter_rent_12m +
##     forecast_price_energy_p2 + forecast_price_pow_p1 + has_gas +
##     margin_net_pow_ele + net_margin + num_years_antig + origin_up +
##     pow_max + p1_var + day_modif + day_renewal, family = binomial(link = logit),
##     data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0189  -0.5023  -0.4124  -0.3368   2.9544
##
## Coefficients:
##                                               Estimate Std. Error z value
## (Intercept)                                  -2.768e+00  6.125e-01  -4.519
```

```
## activity_newapdekpcbwosbxepsfxclislboipuxpop    -2.140e-02  1.459e-01  -0.147
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    -1.367e-01  6.105e-01  -0.224
## activity_newbalskueexlmuccwdffilikwxasupasxf     3.395e-01  4.462e-01   0.761
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo     1.967e-01  4.486e-01   0.438
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf     6.479e-02  4.447e-01   0.146
## activity_newcfdsselwimsklimddecfifseabdkxfcs    -1.763e+00  1.013e+00  -1.741
## activity_newckadsdebplpkplelfspfoiucmxkeppus     2.911e-02  6.112e-01   0.048
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    -4.160e-01  2.923e-01  -1.423
## activity_newcluecxlameloamldmasudocsbmaoamdw    -8.307e-02  3.156e-01  -0.263
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    -3.053e-01  3.394e-01  -0.900
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    -4.660e-01  5.228e-01  -0.891
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    -9.553e-01  7.312e-01  -1.306
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    -6.826e-02  6.148e-01  -0.111
## activity_newdupxuibdflmskeieweeofcaluuuiioix    -6.214e-01  7.313e-01  -0.850
## activity_newfkmblacmaapkaoauabpwpuweokkeiali     8.143e-02  5.508e-01   0.148
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb     2.053e-01  2.044e-01   1.004
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm     9.174e-02  4.403e-01   0.208
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    -2.561e-01  5.322e-01  -0.481
## activity_newipdldckuswupeifllfbwccfpeafludfi     2.481e-01  3.861e-01   0.643
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    -1.393e-02  1.769e-01  -0.079
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    -1.322e+00  1.017e+00  -1.300
## activity_newkwuslieomapmswolewpobpplkaooaaew     7.843e-02  2.121e-01   0.370
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    -9.487e-02  5.333e-01  -0.178
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil     1.685e-01  4.467e-01   0.377
## activity_newOthers                               9.995e-02  7.872e-02   1.270
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae     4.421e-01  4.253e-01   1.040
## activity_newsaxlifeumaobawxpemwuopbwwldlucff     4.367e-01  3.701e-01   1.180
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    -2.632e-01  4.302e-01  -0.612
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    -3.858e-01  4.658e-01  -0.828
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    -3.201e-01  6.067e-01  -0.528
## activity_newsxublbwoeuckkocekklxkllcdxxaisop     4.946e-01  3.463e-01   1.428
## activity_newwxemiwkumpibllwklfbcooafckufkdlm     3.486e-01  2.682e-01   1.300
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu   -1.173e+01  2.668e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci   -2.018e-01  1.430e-01  -1.411
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua    3.342e-02  9.147e-02   0.365
## channel_saleslmkebamcaaclubfxadlmueccxoimlema   -3.661e-01  1.228e-01  -2.983
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds   -1.145e+01  1.691e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf   -9.446e-02  1.233e-01  -0.766
## cons_gas_12m                                    -1.180e-06  4.576e-07  -2.578
## cons_last_month                                 -2.569e-06  1.002e-06  -2.562
## forecast_discount_energy                        -7.633e-03  5.610e-03  -1.361
## forecast_meter_rent_12m                          8.098e-04  5.173e-04   1.565
## forecast_price_energy_p2                         1.846e+00  7.522e-01   2.453
## forecast_price_pow_p1                            1.257e-02  6.485e-03   1.938
## has_gas                                         -1.304e-01  8.359e-02  -1.560
## margin_net_pow_ele                               6.411e-03  8.773e-04   7.307
## net_margin                                       2.709e-04  9.366e-05   2.893
## num_years_antig                                 -9.119e-02  2.421e-02  -3.768
## origin_upewxeelcelemmiwuafmddpobolfuxioce       -1.038e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws        2.906e-01  4.689e-01   0.620
## origin_upldkssxwpmemidmecebumciepifcamkci        6.144e-01  4.679e-01   1.313
## origin_uplxidpiddsbxsbosboudacockeimpuepw        8.445e-01  4.652e-01   1.815
## origin_upusapbepcfoloekilkwsdiboslwaxobdp       -1.062e+01  3.782e+02  -0.028
## pow_max                                         -5.277e-03  1.919e-03  -2.750
```

```
## p1_var                                          -1.282e+00  1.724e+00  -0.743
## day_modif                                       -5.711e-05  3.737e-05  -1.528
## day_renewal                                     -3.902e-04  5.446e-04  -0.717
##                                                 Pr(>|z|)
## (Intercept)                                     6.22e-06 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.883384
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.822820
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.446648
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.661095
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.884159
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.081603 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.962006
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.154643
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.792393
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.368287
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.372780
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.191385
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.911594
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.395532
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.882469
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.315279
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.834947
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.630313
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.520411
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.937241
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.193612
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.711486
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.858798
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.706000
## activity_newOthers                             0.204197
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.298514
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    0.238101
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    0.540583
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    0.407540
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    0.597744
## activity_newsxublbwoeuckkocekklxkllcdxxaisop    0.153217
## activity_newwxemiwkumpibllwklfbcooafckufkdlm    0.193731
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu  0.964918
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci  0.158144
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  0.714867
## channel_saleslmkebamcaaclubfxadlmueccxoimlema  0.002859 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds  0.946035
## channel_salesusilxuppasemubllopkaafesmlibmsdf  0.443699
## cons_gas_12m                                   0.009932 **
## cons_last_month                                0.010406 *
## forecast_discount_energy                       0.173597
## forecast_meter_rent_12m                        0.117485
## forecast_price_energy_p2                       0.014150 *
## forecast_price_pow_p1                          0.052636 .
## has_gas                                        0.118664
## margin_net_pow_ele                             2.73e-13 ***
## net_margin                                     0.003820 **
## num_years_antig                                0.000165 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce      0.984530
```

```
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws         0.535441
## origin_upldkssxwpmemidmecebumciepifcamkci         0.189106
## origin_uplxidpiddsbxsbosboudacockeimpuepw         0.069468 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp         0.977609
## pow_max                                           0.005961 **
## p1_var                                            0.457202
## day_modif                                         0.126425
## day_renewal                                       0.473631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9627.2  on 15703  degrees of freedom
## AIC: 9743.2
##
## Number of Fisher Scoring iterations: 12
```

```
# m14: remove day_renewal
m14 <- update(m13, . ~ . - day_renewal)
summary(m14)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_discount_energy + forecast_meter_rent_12m +
##     forecast_price_energy_p2 + forecast_price_pow_p1 + has_gas +
##     margin_net_pow_ele + net_margin + num_years_antig + origin_up +
##     pow_max + p1_var + day_modif, family = binomial(link = logit),
##     data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0110  -0.5021  -0.4124  -0.3371   2.9562
##
## Coefficients:
##                                                  Estimate Std. Error z value
## (Intercept)                                     -2.909e+00  5.795e-01  -5.020
## activity_newapdekpcbwosbxepsfxclislboipuxpop    -2.176e-02  1.459e-01  -0.149
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    -1.374e-01  6.105e-01  -0.225
## activity_newbalskueexlmuccwdffilikwxasupasxf     3.347e-01  4.462e-01   0.750
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo     1.993e-01  4.486e-01   0.444
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf     5.169e-02  4.442e-01   0.116
## activity_newcfdsselwimsklimddecfifseabdkxfcs    -1.765e+00  1.013e+00  -1.744
## activity_newckadsdebplpkplelfspfoiucmxkeppus     3.045e-02  6.112e-01   0.050
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    -4.141e-01  2.923e-01  -1.417
## activity_newcluecxlameloamldmasudocsbmaoamdw    -8.474e-02  3.157e-01  -0.268
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    -3.012e-01  3.394e-01  -0.888
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    -4.693e-01  5.228e-01  -0.898
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    -9.594e-01  7.312e-01  -1.312
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    -7.015e-02  6.149e-01  -0.114
## activity_newdupxuibdflmskeieweeofcaluuuiioix    -6.185e-01  7.313e-01  -0.846
## activity_newfkmblacmaapkaoauabpwpuweokkeiali     8.611e-02  5.508e-01   0.156
```

20

```
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb      2.064e-01  2.044e-01   1.010
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm      9.149e-02  4.403e-01   0.208
## activity_newilkfsaapsxpkcpswblldddfmpamwelpxi    -2.537e-01  5.321e-01  -0.477
## activity_newipdldckuswupeifllfbwccfpeafludfi      2.448e-01  3.862e-01   0.634
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce     -1.261e-02  1.769e-01  -0.071
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca     -1.321e+00  1.017e+00  -1.299
## activity_newkwuslieomapmswolewpobpplkaooaaew      7.930e-02  2.120e-01   0.374
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx     -9.300e-02  5.333e-01  -0.174
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil      1.679e-01  4.466e-01   0.376
## activity_newOthers                                9.997e-02  7.872e-02   1.270
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae      4.391e-01  4.252e-01   1.033
## activity_newsaxlifeumaobawxpemwuopbwwldlucff      4.392e-01  3.701e-01   1.187
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc     -2.663e-01  4.303e-01  -0.619
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu     -3.872e-01  4.658e-01  -0.831
## activity_newsumdxiaiudmaioicexmiwuudlblkissm     -3.176e-01  6.066e-01  -0.524
## activity_newsxublbwoeuckkocekklxkllcdxxaisop      4.914e-01  3.464e-01   1.419
## activity_newwxemiwkumpibllwklfbcooafckufkdlm      3.487e-01  2.682e-01   1.300
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu    -1.176e+01  2.671e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci    -2.040e-01  1.430e-01  -1.427
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua     3.382e-02  9.146e-02   0.370
## channel_saleslmkebamcaaclubfxadlmueccxoimlema    -3.660e-01  1.228e-01  -2.982
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds    -1.146e+01  1.690e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf    -9.485e-02  1.233e-01  -0.769
## cons_gas_12m                                     -1.183e-06  4.580e-07  -2.583
## cons_last_month                                  -2.563e-06  1.003e-06  -2.556
## forecast_discount_energy                         -8.740e-03  5.394e-03  -1.620
## forecast_meter_rent_12m                           8.193e-04  5.167e-04   1.586
## forecast_price_energy_p2                          1.830e+00  7.515e-01   2.435
## forecast_price_pow_p1                             1.248e-02  6.482e-03   1.925
## has_gas                                          -1.306e-01  8.359e-02  -1.563
## margin_net_pow_ele                                6.429e-03  8.769e-04   7.332
## net_margin                                        2.712e-04  9.349e-05   2.901
## num_years_antig                                  -9.179e-02  2.420e-02  -3.793
## origin_upewxeelcelemmiwuafmddpobolfuxioce        -1.038e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws         2.904e-01  4.689e-01   0.619
## origin_upldkssxwpmemidmecebumciepifcamkci         6.136e-01  4.679e-01   1.312
## origin_uplxidpiddsbxsbosboudacockeimpuepw         8.445e-01  4.652e-01   1.815
## origin_upusapbepcfoloekilkwsdiboslwaxobdp        -1.061e+01  3.783e+02  -0.028
## pow_max                                          -5.288e-03  1.919e-03  -2.756
## p1_var                                           -1.283e+00  1.723e+00  -0.744
## day_modif                                        -5.375e-05  3.710e-05  -1.449
##                                                  Pr(>|z|)
## (Intercept)                                      5.16e-07 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop     0.881449
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc     0.821983
## activity_newbalskueexlmuccwdffilikwxasupasxf     0.453070
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo     0.656810
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf     0.907367
## activity_newcfdsselwimsklimddecfifseabdkxfcs     0.081237 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus     0.960258
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa     0.156494
## activity_newcluecxlameloamldmasudocsbmaoamdw     0.788372
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu     0.374736
## activity_newdaobdssbkieoukwxbopxiiospudkopwl     0.369362
```

21

```
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl 0.189534
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk 0.909167
## activity_newdupxuibdflmskeieweeofcaluuuiioix 0.397719
## activity_newfkmblacmaapkaoauabpwpuweokkeiali 0.875774
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb 0.312560
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm 0.835380
## activity_newilkfsaapsxpkcpswblldfmpamwelpxi 0.633488
## activity_newipdldckuswupeifllfbwccfpeafludfi 0.526189
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce 0.943180
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca 0.194043
## activity_newkwuslieomapmswolewpobpplkaooaaew 0.708384
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx 0.861557
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil 0.707004
## activity_newOthers                           0.204123
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae 0.301785
## activity_newsaxlifeumaobawxpemwuopbwwldlucff 0.235334
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc 0.536013
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu 0.405804
## activity_newsumdxiaiudmaioicexmiwuudlblkissm 0.600584
## activity_newsxublbwoeuckkocekklxkllcdxxaisop 0.156040
## activity_newwxemiwkumpibllwklfbcooafckufkdlm 0.193492
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu 0.964865
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci 0.153701
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua 0.711572
## channel_saleslmkebamcaaclubfxadlmueccxoimlema 0.002867 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds 0.945924
## channel_salesusilxuppasemubllopkaafesmlibmsdf 0.441814
## cons_gas_12m                                 0.009783 **
## cons_last_month                              0.010594 *
## forecast_discount_energy                     0.105187
## forecast_meter_rent_12m                      0.112823
## forecast_price_energy_p2                     0.014871 *
## forecast_price_pow_p1                        0.054267 .
## has_gas                                      0.118084
## margin_net_pow_ele                           2.27e-13 ***
## net_margin                                   0.003718 **
## num_years_antig                              0.000149 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce     0.984527
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws     0.535676
## origin_upldkssxwpmemidmecebumciepifcamkci     0.189676
## origin_uplxidpiddsbxsbosboudacockeimpuepw     0.069491 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp     0.977616
## pow_max                                      0.005849 **
## p1_var                                       0.456740
## day_modif                                    0.147369
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9627.7  on 15704  degrees of freedom
## AIC: 9741.7
##
```

```
## Number of Fisher Scoring iterations: 12
```

```
# m15: remove p1_var
m15 <- update(m14, . ~ . - p1_var)
summary(m15)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_discount_energy + forecast_meter_rent_12m +
##     forecast_price_energy_p2 + forecast_price_pow_p1 + has_gas +
##     margin_net_pow_ele + net_margin + num_years_antig + origin_up +
##     pow_max + day_modif, family = binomial(link = logit), data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9532  -0.5024  -0.4123  -0.3374   2.9550
##
## Coefficients:
##                                               Estimate Std. Error z value
## (Intercept)                                  -3.012e+00  5.646e-01  -5.335
## activity_newapdekpcbwosbxepsfxclislboipuxpop -2.079e-02  1.459e-01  -0.142
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc -1.341e-01  6.104e-01  -0.220
## activity_newbalskueexlmuccwdffilikwxasupasxf  3.332e-01  4.461e-01   0.747
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo  2.020e-01  4.490e-01   0.450
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf  5.130e-02  4.441e-01   0.116
## activity_newcfdsselwimsklimddecfifseabdkxfcs -1.759e+00  1.012e+00  -1.737
## activity_newckadsdebplpkplelfspfoiucmxkeppus  3.059e-02  6.111e-01   0.050
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa -4.114e-01  2.922e-01  -1.408
## activity_newcluecxlameloamldmasudocsbmaoamdw -8.074e-02  3.153e-01  -0.256
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu -3.002e-01  3.390e-01  -0.886
## activity_newdaobdssbkieoukwxbopxiiospudkopwl -4.690e-01  5.228e-01  -0.897
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl -9.512e-01  7.311e-01  -1.301
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk -6.598e-02  6.148e-01  -0.107
## activity_newdupxuibdflmskeieweeofcaluuuiioix -6.198e-01  7.312e-01  -0.848
## activity_newfkmblacmaapkaoauabpwpuweokkeiali  8.191e-02  5.511e-01   0.149
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb  2.069e-01  2.043e-01   1.012
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm  8.576e-02  4.402e-01   0.195
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi -2.536e-01  5.321e-01  -0.477
## activity_newipdldckuswupeifllfbwccfpeafludfi  2.512e-01  3.860e-01   0.651
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce -1.159e-02  1.769e-01  -0.066
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca -1.321e+00  1.017e+00  -1.299
## activity_newkwuslieomapmswolewpobpplkaooaaew  8.191e-02  2.120e-01   0.386
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx -9.094e-02  5.333e-01  -0.171
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil  1.665e-01  4.469e-01   0.373
## activity_newOthers                            1.026e-01  7.866e-02   1.305
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae  4.432e-01  4.252e-01   1.042
## activity_newsaxlifeumaobawxpemwuopbwwldlucff  4.368e-01  3.703e-01   1.180
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc -2.643e-01  4.303e-01  -0.614
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu -3.869e-01  4.658e-01  -0.831
## activity_newsumdxiaiudmaioicexmiwuudlblkissm -3.017e-01  6.063e-01  -0.498
## activity_newsxublbwoeuckkocekklxkllcdxxaisop  4.925e-01  3.464e-01   1.422
## activity_newwxemiwkumpibllwklfbcooafckufkdlm  3.477e-01  2.682e-01   1.296
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.176e+01  2.671e+02  -0.044
```

```
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.057e-01  1.429e-01  -1.439
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  3.176e-02  9.136e-02   0.348
## channel_saleslmkebamcaaclubfxadlmueccxoimlema -3.680e-01  1.227e-01  -3.000
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds -1.146e+01  1.690e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf -1.004e-01  1.231e-01  -0.816
## cons_gas_12m                                   -1.180e-06  4.573e-07  -2.579
## cons_last_month                                -2.580e-06  1.003e-06  -2.573
## forecast_discount_energy                       -9.704e-03  5.233e-03  -1.855
## forecast_meter_rent_12m                         9.520e-04  4.809e-04   1.980
## forecast_price_energy_p2                        1.858e+00  7.485e-01   2.482
## forecast_price_pow_p1                           1.048e-02  5.997e-03   1.748
## has_gas                                        -1.314e-01  8.358e-02  -1.573
## margin_net_pow_ele                              6.183e-03  8.090e-04   7.642
## net_margin                                      2.683e-04  9.315e-05   2.880
## num_years_antig                                -8.997e-02  2.405e-02  -3.741
## origin_upewxeelcelemmiwuafmddpobolfuxioce      -1.039e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       2.792e-01  4.686e-01   0.596
## origin_upldkssxwpmemidmecebumciepifcamkci       6.098e-01  4.678e-01   1.304
## origin_uplxidpiddsbxsbosboudacockeimpuepw       8.395e-01  4.651e-01   1.805
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      -1.063e+01  3.782e+02  -0.028
## pow_max                                        -5.086e-03  1.890e-03  -2.691
## day_modif                                      -5.356e-05  3.709e-05  -1.444
##                                                 Pr(>|z|)
## (Intercept)                                     9.57e-08 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.886703
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.826059
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.455127
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.652832
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.908047
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.082316 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.960072
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.159215
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.797924
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.375778
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.369711
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.193255
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.914548
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.396690
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.881840
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.311355
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.845530
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.633668
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.515262
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.947773
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.193948
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.699170
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.864592
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.709404
## activity_newOthers                              0.192010
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.297230
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    0.238103
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    0.538976
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    0.406144
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    0.618711
```

24

```
## activity_newsxublbwoeuckkocekklxkllcdxxaisop  0.155077
## activity_newwxemiwkumpibllwklfbcooafckufkdlm  0.194873
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu 0.964877
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci 0.150152
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua 0.728098
## channel_saleslmkebamcaaclubfxadlmueccxoimlema 0.002701 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds 0.945946
## channel_salesusilxuppasemubllopkaafesmlibmsdf 0.414472
## cons_gas_12m                                  0.009903 **
## cons_last_month                               0.010079 *
## forecast_discount_energy                      0.063656 .
## forecast_meter_rent_12m                       0.047728 *
## forecast_price_energy_p2                      0.013060 *
## forecast_price_pow_p1                         0.080515 .
## has_gas                                       0.115782
## margin_net_pow_ele                            2.14e-14 ***
## net_margin                                    0.003978 **
## num_years_antig                               0.000183 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce     0.984511
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws     0.551202
## origin_upldkssxwpmemidmecebumciepifcamkci     0.192377
## origin_uplxidpiddsbxsbosboudacockeimpuepw     0.071049 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp     0.977565
## pow_max                                       0.007131 **
## day_modif                                     0.148727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9628.3  on 15705  degrees of freedom
## AIC: 9740.3
##
## Number of Fisher Scoring iterations: 12
```

```r
# anova(update(m15, . ~ has_gas +. ), test = "Chisq")
# summary(update(m15, . ~ day_modif +. ))
```

```r
# m16: remove day_modif
m16 <- update(m15, . ~ . - day_modif)
summary(m16)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_discount_energy + forecast_meter_rent_12m +
##     forecast_price_energy_p2 + forecast_price_pow_p1 + has_gas +
##     margin_net_pow_ele + net_margin + num_years_antig + origin_up +
##     pow_max, family = binomial(link = logit), data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q    Median       3Q       Max
```

```
## -1.9414  -0.5012  -0.4123  -0.3388   2.9791
##
## Coefficients:
##                                             Estimate Std. Error z value
## (Intercept)                                -3.004e+00  5.634e-01  -5.333
## activity_newapdekpcbwosbxepsfxclislboipuxpop -4.297e-02  1.450e-01  -0.296
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc -1.155e-01  6.102e-01  -0.189
## activity_newbalskueexlmuccwdffilikwxasupasxf  3.465e-01  4.460e-01   0.777
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo  2.017e-01  4.488e-01   0.449
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf  5.673e-02  4.441e-01   0.128
## activity_newcfdsselwimsklimddecfifseabdkxfcs -1.744e+00  1.012e+00  -1.723
## activity_newckadsdebplpkplelfspfoiucmxkeppus  4.869e-02  6.109e-01   0.080
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa -4.011e-01  2.922e-01  -1.373
## activity_newcluecxlameloamldmasudocsbmaoamdw -8.211e-02  3.154e-01  -0.260
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu -2.960e-01  3.391e-01  -0.873
## activity_newdaobdssbkieoukwxbopxiiospudkopwl -4.614e-01  5.228e-01  -0.882
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl -9.454e-01  7.310e-01  -1.293
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk -6.161e-02  6.151e-01  -0.100
## activity_newdupxuibdflmskeieweeofcaluuuiioix -6.138e-01  7.311e-01  -0.840
## activity_newfkmblacmaapkaoauabpwpuweokkeiali  9.890e-02  5.505e-01   0.180
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb  2.142e-01  2.043e-01   1.049
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm  1.002e-01  4.402e-01   0.228
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi -2.454e-01  5.323e-01  -0.461
## activity_newipdldckuswupeifllfbwccfpeafludfi  2.506e-01  3.859e-01   0.649
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce -1.101e-02  1.769e-01  -0.062
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca -1.306e+00  1.017e+00  -1.284
## activity_newkwuslieomapmswolewpobpplkaooaaew  8.962e-02  2.119e-01   0.423
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx -9.081e-02  5.332e-01  -0.170
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil  1.727e-01  4.471e-01   0.386
## activity_newOthers                            1.098e-01  7.850e-02   1.398
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae  4.588e-01  4.252e-01   1.079
## activity_newsaxlifeumaobawxpemwuopbwwldlucff  4.467e-01  3.703e-01   1.206
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc -2.543e-01  4.302e-01  -0.591
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu -4.005e-01  4.658e-01  -0.860
## activity_newsumdxiaiudmaioicexmiwuudlblkissm -3.089e-01  6.062e-01  -0.509
## activity_newsxublbwoeuckkocekklxkllcdxxaisop  4.925e-01  3.463e-01   1.422
## activity_newwxemiwkumpibllwklfbcooafckufkdlm  3.538e-01  2.682e-01   1.319
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.174e+01  2.671e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.046e-01  1.429e-01  -1.431
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  3.504e-02  9.131e-02   0.384
## channel_saleslmkebamcaaclubfxadlmueccxoimlema -3.695e-01  1.227e-01  -3.012
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds -1.145e+01  1.690e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf -9.637e-02  1.230e-01  -0.783
## cons_gas_12m                                  -1.184e-06  4.575e-07  -2.587
## cons_last_month                               -2.571e-06  1.004e-06  -2.559
## forecast_discount_energy                      -8.418e-03  5.162e-03  -1.631
## forecast_meter_rent_12m                        9.687e-04  4.812e-04   2.013
## forecast_price_energy_p2                       1.833e+00  7.478e-01   2.451
## forecast_price_pow_p1                          1.028e-02  5.946e-03   1.728
## has_gas                                       -1.294e-01  8.356e-02  -1.549
## margin_net_pow_ele                             6.173e-03  8.089e-04   7.632
## net_margin                                     2.671e-04  9.309e-05   2.869
## num_years_antig                               -1.024e-01  2.254e-02  -4.545
## origin_upewxeelcelemmiwuafmddpobolfuxioce     -1.045e+01  5.354e+02  -0.020
```

26

```
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       2.720e-01  4.687e-01   0.580
## origin_upldkssxwpmemidmecebumciepifcamkci       6.025e-01  4.679e-01   1.288
## origin_uplxidpiddsbxsbosboudacockeimpuepw       8.341e-01  4.652e-01   1.793
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      -1.064e+01  3.782e+02  -0.028
## pow_max                                        -5.188e-03  1.893e-03  -2.740
##                                                 Pr(>|z|)
## (Intercept)                                     9.68e-08 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.76701
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.84984
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.43720
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.65317
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.89836
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.08493 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.93647
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.16980
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.79460
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.38269
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.37752
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.19592
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.92022
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.40116
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.85743
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.29431
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.82000
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.64473
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.51605
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.95037
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.19904
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.67240
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.86477
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.69933
## activity_newOthers                             0.16204
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.28060
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    0.22769
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    0.55444
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    0.38988
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    0.61041
## activity_newsxublbwoeuckkocekklxkllcdxxaisop    0.15499
## activity_newwxemiwkumpibllwklfbcooafckufkdlm    0.18703
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu   0.96494
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci   0.15230
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua   0.70115
## channel_saleslmkebamcaaclubfxadlmueccxoimlema   0.00259 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds   0.94600
## channel_salesusilxuppasemubllopkaafesmlibmsdf   0.43352
## cons_gas_12m                                    0.00968 **
## cons_last_month                                 0.01049 *
## forecast_discount_energy                        0.10296
## forecast_meter_rent_12m                         0.04411 *
## forecast_price_energy_p2                        0.01424 *
## forecast_price_pow_p1                           0.08392 .
## has_gas                                         0.12140
## margin_net_pow_ele                              2.32e-14 ***
## net_margin                                      0.00411 **
```

```
## num_years_antig                             5.49e-06 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce    0.98442
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws    0.56168
## origin_upldkssxwpmemidmecebumciepifcamkci    0.19784
## origin_uplxidpiddsbxsbosboudacockeimpuepw    0.07297 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp    0.97755
## pow_max                                      0.00615 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9630.4  on 15706  degrees of freedom
## AIC: 9740.4
##
## Number of Fisher Scoring iterations: 12
```

```
anova(m16, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: churn
##
## Terms added sequentially (first to last)
##
##
##                        Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
## NULL                                   15760    10020.9
## activity_new           32   72.248     15728     9948.6 6.059e-05 ***
## channel_sales           6   97.522     15722     9851.1 < 2.2e-16 ***
## cons_gas_12m            1   27.352     15721     9823.7 1.696e-07 ***
## cons_last_month         1    6.705     15720     9817.0  0.009612 **
## forecast_discount_energy 1    0.803     15719     9816.2  0.370102
## forecast_meter_rent_12m 1   19.082     15718     9797.1 1.252e-05 ***
## forecast_price_energy_p2 1    4.746     15717     9792.4  0.029361 *
## forecast_price_pow_p1   1    1.737     15716     9790.7  0.187543
## has_gas                 1    1.875     15715     9788.8  0.170928
## margin_net_pow_ele      1   54.439     15714     9734.3 1.604e-13 ***
## net_margin              1    9.097     15713     9725.2  0.002561 **
## num_years_antig         1   36.983     15712     9688.3 1.192e-09 ***
## origin_up               5   49.441     15707     9638.8 1.803e-09 ***
## pow_max                 1    8.442     15706     9630.4  0.003668 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# m17: By comparing the summary and ANOVA table of m16, we decide to remove forecast_discount_energy
m17 <- update(m16, . ~ . - forecast_discount_energy)
summary(m17)
```

```
##
```

```
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_meter_rent_12m + forecast_price_energy_p2 +
##     forecast_price_pow_p1 + has_gas + margin_net_pow_ele + net_margin +
##     num_years_antig + origin_up + pow_max, family = binomial(link = logit),
##     data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9653  -0.5018  -0.4119  -0.3390   2.9745
##
## Coefficients:
##                                                Estimate Std. Error z value
## (Intercept)                                   -2.992e+00  5.633e-01  -5.311
## activity_newapdekpcbwosbxepsfxclislboipuxpop  -3.326e-02  1.449e-01  -0.230
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc  -1.054e-01  6.101e-01  -0.173
## activity_newbalskueexlmuccwdffilikwxasupasxf   3.469e-01  4.459e-01   0.778
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo   2.116e-01  4.481e-01   0.472
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf   4.153e-02  4.441e-01   0.094
## activity_newcfdsselwimsklimddecfifseabdkxfcs  -1.748e+00  1.013e+00  -1.727
## activity_newckadsdebplpkplelfspfoiucmxkeppus   5.746e-02  6.109e-01   0.094
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa  -4.023e-01  2.921e-01  -1.377
## activity_newcluecxlameloamldmasudocsbmaoamdw  -7.129e-02  3.147e-01  -0.227
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu  -2.976e-01  3.387e-01  -0.879
## activity_newdaobdssbkieoukwxbopxiiospudkopwl  -4.543e-01  5.227e-01  -0.869
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl  -9.672e-01  7.312e-01  -1.323
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk  -6.282e-02  6.152e-01  -0.102
## activity_newdupxuibdflmskeieweeofcaluuuiioix  -6.102e-01  7.310e-01  -0.835
## activity_newfkmblacmaapkaoauabpwpuweokkeiali   1.103e-01  5.489e-01   0.201
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb   2.221e-01  2.042e-01   1.088
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm   1.040e-01  4.409e-01   0.236
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi  -2.487e-01  5.320e-01  -0.467
## activity_newipdldckuswupeiflllfbwccfpeafludfi  2.581e-01  3.859e-01   0.669
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce  -6.158e-03  1.768e-01  -0.035
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca  -1.325e+00  1.017e+00  -1.303
## activity_newkwuslieomapmswolewpobpplkaooaaew   9.191e-02  2.120e-01   0.434
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx  -8.077e-02  5.331e-01  -0.152
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil   1.792e-01  4.469e-01   0.401
## activity_newOthers                             1.121e-01  7.847e-02   1.428
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae   4.486e-01  4.258e-01   1.053
## activity_newsaxlifeumaobawxpemwuopbwwldlucff   4.421e-01  3.702e-01   1.194
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc  -2.739e-01  4.303e-01  -0.637
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu  -3.978e-01  4.658e-01  -0.854
## activity_newsumdxiaiudmaioicexmiwuudlblkissm  -3.046e-01  6.062e-01  -0.502
## activity_newsxublbwoeuckkocekklxkllcdxxaisop   4.862e-01  3.464e-01   1.404
## activity_newwxemiwkumpibllwklfbcooafckufkdlm   3.518e-01  2.682e-01   1.312
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.172e+01  2.671e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.033e-01  1.429e-01  -1.423
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua  3.467e-02  9.133e-02   0.380
## channel_saleslmkebamcaaclubfxadlmueccxoimlema -3.662e-01  1.226e-01  -2.986
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds -1.144e+01  1.690e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf -9.858e-02  1.230e-01  -0.801
## cons_gas_12m                                  -1.184e-06  4.565e-07  -2.594
## cons_last_month                               -2.545e-06  1.001e-06  -2.544
```

```
## forecast_meter_rent_12m                        1.012e-03  4.797e-04   2.109
## forecast_price_energy_p2                        1.705e+00  7.418e-01   2.299
## forecast_price_pow_p1                           9.824e-03  5.939e-03   1.654
## has_gas                                        -1.315e-01  8.354e-02  -1.574
## margin_net_pow_ele                              5.912e-03  7.953e-04   7.434
## net_margin                                      2.454e-04  8.961e-05   2.739
## num_years_antig                                -1.016e-01  2.255e-02  -4.504
## origin_upewxeelcelemmiwuafmddpobolfuxioce      -1.045e+01  5.354e+02  -0.020
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       2.772e-01  4.686e-01   0.591
## origin_upldkssxwpmemidmecebumciepifcamkci       6.063e-01  4.679e-01   1.296
## origin_uplxidpiddsbxsbosboudacockeimpuepw       8.343e-01  4.651e-01   1.794
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      -1.064e+01  3.783e+02  -0.028
## pow_max                                        -4.909e-03  1.872e-03  -2.622
##                                                 Pr(>|z|)
## (Intercept)                                     1.09e-07 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.81846
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.86280
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.43661
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.63672
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.92550
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.08426 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.92506
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.16855
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.82077
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.37958
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.38477
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.18590
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.91867
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.40387
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.84076
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.27678
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.81359
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.64017
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.50358
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.97221
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.19273
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.66465
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.87958
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.68847
## activity_newOthers                              0.15329
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.29214
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    0.23241
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    0.52434
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    0.39310
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    0.61533
## activity_newsxublbwoeuckkocekklxkllcdxxaisop    0.16042
## activity_newwxemiwkumpibllwklfbcooafckufkdlm    0.18960
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu   0.96499
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci   0.15488
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua   0.70421
## channel_saleslmkebamcaaclubfxadlmueccxoimlema   0.00283 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds   0.94603
## channel_salesusilxuppasemubllopkaafesmlibmsdf   0.42288
## cons_gas_12m                                    0.00949 **
```

```
## cons_last_month                                 0.01097 *
## forecast_meter_rent_12m                          0.03491 *
## forecast_price_energy_p2                         0.02149 *
## forecast_price_pow_p1                            0.09810 .
## has_gas                                          0.11556
## margin_net_pow_ele                               1.06e-13 ***
## net_margin                                       0.00617 **
## num_years_antig                                  6.67e-06 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce         0.98443
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws         0.55422
## origin_upldkssxwpmemidmecebumciepifcamkci         0.19500
## origin_uplxidpiddsbxsbosboudacockeimpuepw         0.07286 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp         0.97756
## pow_max                                          0.00875 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9633.2  on 15707  degrees of freedom
## AIC: 9741.2
##
## Number of Fisher Scoring iterations: 12
```

```
anova(m17, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: churn
##
## Terms added sequentially (first to last)
##
##
##                      Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
## NULL                                 15760    10020.9
## activity_new         32   72.248     15728     9948.6 6.059e-05 ***
## channel_sales         6   97.522     15722     9851.1 < 2.2e-16 ***
## cons_gas_12m          1   27.352     15721     9823.7 1.696e-07 ***
## cons_last_month       1    6.705     15720     9817.0 0.009612 **
## forecast_meter_rent_12m 1 18.933     15719     9798.1 1.354e-05 ***
## forecast_price_energy_p2 1 5.141     15718     9792.9 0.023362 *
## forecast_price_pow_p1  1    1.834     15717     9791.1 0.175705
## has_gas               1    1.848     15716     9789.3 0.174041
## margin_net_pow_ele    1   54.480     15715     9734.8 1.570e-13 ***
## net_margin            1    8.635     15714     9726.2 0.003297 **
## num_years_antig       1   36.483     15713     9689.7 1.540e-09 ***
## origin_up             5   48.802     15708     9640.9 2.436e-09 ***
## pow_max               1    7.699     15707     9633.2 0.005524 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# m18: remove has_gas
m18 <- update(m17, . ~ . - has_gas)
summary(m18)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_meter_rent_12m + forecast_price_energy_p2 +
##     forecast_price_pow_p1 + margin_net_pow_ele + net_margin +
##     num_years_antig + origin_up + pow_max, family = binomial(link = logit),
##     data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9660  -0.5020  -0.4125  -0.3402   2.9567
##
## Coefficients:
##                                            Estimate Std. Error z value
## (Intercept)                               -3.036e+00  5.633e-01  -5.390
## activity_newapdekpcbwosbxepsfxclislboipuxpop -2.307e-02  1.450e-01  -0.159
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc -1.174e-01  6.098e-01  -0.192
## activity_newbalskueexlmuccwdffilikwxasupasxf  3.491e-01  4.461e-01   0.783
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo  2.250e-01  4.480e-01   0.502
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf  5.399e-02  4.440e-01   0.122
## activity_newcfdsselwimsklimddecfifseabdkxfcs -1.740e+00  1.012e+00  -1.719
## activity_newckadsdebplpkplelfspfoiucmxkeppus  5.176e-02  6.106e-01   0.085
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa -4.193e-01  2.919e-01  -1.436
## activity_newcluecxlameloamldmasudocsbmaoamdw -6.446e-02  3.145e-01  -0.205
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu -3.074e-01  3.385e-01  -0.908
## activity_newdaobdssbkieoukwxbopxiiospudkopwl -4.407e-01  5.225e-01  -0.843
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl -9.717e-01  7.310e-01  -1.329
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk -4.267e-02  6.151e-01  -0.069
## activity_newdupxuibdflmskeieweeofcaluuuiioix -5.938e-01  7.309e-01  -0.812
## activity_newfkmblacmaapkaoauabpwpuweokkeiali  1.292e-01  5.488e-01   0.235
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb  2.031e-01  2.037e-01   0.997
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm  1.139e-01  4.407e-01   0.258
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi -2.535e-01  5.316e-01  -0.477
## activity_newipdldckuswupeifllfbwccfpeafludfi  2.712e-01  3.857e-01   0.703
## activity_newkkkklcdamwfafdcfwofuscwfwadblfmce  1.976e-03  1.767e-01   0.011
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca -1.315e+00  1.017e+00  -1.293
## activity_newkwuslieomapmswolewpobpplkaooaaew  5.774e-02  2.108e-01   0.274
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx -6.362e-02  5.330e-01  -0.119
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil  1.863e-01  4.469e-01   0.417
## activity_newOthers                          1.200e-01  7.828e-02   1.534
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae  4.645e-01  4.256e-01   1.091
## activity_newsaxlifeumaobawxpemwuopbwwldlucff  4.547e-01  3.700e-01   1.229
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc -2.772e-01  4.301e-01  -0.645
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu -3.927e-01  4.659e-01  -0.843
## activity_newsumdxiaiudmaioicexmiwuudlblkissm -3.013e-01  6.060e-01  -0.497
## activity_newsxublbwoeuckkocekklxkllcdxxaisop  4.901e-01  3.462e-01   1.416
## activity_newwxemiwkumpibllwklfbcooafckufkdlm  3.665e-01  2.681e-01   1.367
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu -1.170e+01  2.670e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci -2.035e-01  1.429e-01  -1.424
```
```
32
```

```
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua   4.014e-02  9.126e-02   0.440
## channel_saleslmkebamcaaclubfxadlmueccxoimlema  -3.641e-01  1.227e-01  -2.968
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds  -1.142e+01  1.690e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf  -9.767e-02  1.230e-01  -0.794
## cons_gas_12m                                   -1.438e-06  4.551e-07  -3.159
## cons_last_month                                -2.561e-06  1.002e-06  -2.556
## forecast_meter_rent_12m                         1.020e-03  4.797e-04   2.126
## forecast_price_energy_p2                        1.681e+00  7.418e-01   2.266
## forecast_price_pow_p1                           1.018e-02  5.961e-03   1.707
## margin_net_pow_ele                              5.944e-03  7.950e-04   7.476
## net_margin                                      2.407e-04  8.895e-05   2.706
## num_years_antig                                -1.015e-01  2.253e-02  -4.506
## origin_upewxeelcelemmiwuafmddpobolfuxioce      -1.042e+01  5.354e+02  -0.019
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       2.876e-01  4.687e-01   0.614
## origin_upldkssxwpmemidmecebumciepifcamkci       6.143e-01  4.679e-01   1.313
## origin_uplxidpiddsbxsbosboudacockeimpuepw       8.399e-01  4.652e-01   1.806
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      -1.061e+01  3.783e+02  -0.028
## pow_max                                        -4.903e-03  1.872e-03  -2.619
##                                                 Pr(>|z|)
## (Intercept)                                     7.06e-08 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.87358
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.84740
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.43381
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.61553
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.90322
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.08570 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.93245
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.15089
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.83763
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.36380
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.39905
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.18376
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.94470
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.41659
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.81392
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.31867
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.79611
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.63343
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.48195
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.99108
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.19614
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.78417
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.90498
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.67675
## activity_newOthers                              0.12514
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.27506
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    0.21915
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    0.51924
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    0.39931
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    0.61902
## activity_newsxublbwoeuckkocekklxkllcdxxaisop    0.15687
## activity_newwxemiwkumpibllwklfbcooafckufkdlm    0.17168
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu   0.96504
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci   0.15441
```

```
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua   0.66005
## channel_saleslmkebamcaaclubfxadlmueccxoimlema   0.00300 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds   0.94613
## channel_salesusilxuppasemubllopkaafesmlibmsdf   0.42697
## cons_gas_12m                                    0.00158 **
## cons_last_month                                 0.01058 *
## forecast_meter_rent_12m                         0.03349 *
## forecast_price_energy_p2                        0.02348 *
## forecast_price_pow_p1                           0.08774 .
## margin_net_pow_ele                             7.67e-14 ***
## net_margin                                      0.00680 **
## num_years_antig                                6.61e-06 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce       0.98447
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       0.53944
## origin_upldkssxwpmemidmecebumciepifcamkci       0.18925
## origin_uplxidpiddsbxsbosboudacockeimpuepw       0.07099 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp       0.97762
## pow_max                                         0.00882 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9635.7  on 15708  degrees of freedom
## AIC: 9741.7
##
## Number of Fisher Scoring iterations: 12
```

```
# m19: remove forecast_price_pow_p1
m19 <- update(m18, . ~ . - forecast_price_pow_p1)
summary(m19)
```

```
##
## Call:
## glm(formula = churn ~ activity_new + channel_sales + cons_gas_12m +
##     cons_last_month + forecast_meter_rent_12m + forecast_price_energy_p2 +
##     margin_net_pow_ele + net_margin + num_years_antig + origin_up +
##     pow_max, family = binomial(link = logit), data = data_combo_new)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9848  -0.5016  -0.4118  -0.3417   2.9617
##
## Coefficients:
##                                               Estimate Std. Error z value
## (Intercept)                                 -2.555e+00  4.881e-01  -5.234
## activity_newapdekpcbwosbxepsfxclislboipuxpop -1.821e-02  1.451e-01  -0.125
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc -1.365e-01  6.097e-01  -0.224
## activity_newbalskueexlmuccwdffilikwxasupasxf  3.408e-01  4.461e-01   0.764
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo  2.342e-01  4.482e-01   0.523
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf  4.250e-02  4.439e-01   0.096
## activity_newcfdsselwimsklimddecfifseabdkxfcs -1.736e+00  1.012e+00  -1.714
## activity_newckadsdebplpkplelfspfoiucmxkeppus  4.498e-02  6.104e-01   0.074
```

34

```
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa   -4.274e-01  2.919e-01  -1.464
## activity_newcluecxlameloamldmasudocsbmaoamdw   -7.035e-02  3.145e-01  -0.224
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu   -3.117e-01  3.385e-01  -0.921
## activity_newdaobdssbkieoukwxbopxiiospudkopwl   -4.530e-01  5.226e-01  -0.867
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl   -9.845e-01  7.310e-01  -1.347
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk   -5.253e-02  6.151e-01  -0.085
## activity_newdupxuibdflmskeieweeofcaluuuiioix   -5.959e-01  7.310e-01  -0.815
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    1.315e-01  5.471e-01   0.240
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    1.896e-01  2.035e-01   0.932
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    1.042e-01  4.409e-01   0.236
## activity_newilkfsaapsxpkcpswblldfmpamwelpxi    -2.408e-01  5.314e-01  -0.453
## activity_newipdldckuswupeifllfbwccfpeafludfi    2.667e-01  3.857e-01   0.692
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce   -8.084e-03  1.766e-01  -0.046
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca   -1.317e+00  1.017e+00  -1.295
## activity_newkwuslieomapmswolewpobpplkaooaaew    4.415e-02  2.107e-01   0.210
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx   -7.320e-02  5.330e-01  -0.137
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    1.758e-01  4.462e-01   0.394
## activity_newOthers                              1.181e-01  7.825e-02   1.510
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    4.517e-01  4.255e-01   1.062
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    4.397e-01  3.700e-01   1.188
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc   -2.874e-01  4.301e-01  -0.668
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu   -3.984e-01  4.659e-01  -0.855
## activity_newsumdxiaiudmaioicexmiwuudlblkissm   -2.334e-01  6.046e-01  -0.386
## activity_newsxublbwoeuckkocekklxkllcdxxaisop    4.765e-01  3.461e-01   1.377
## activity_newwxemiwkumpibllwklfbcooafckufkdlm    3.531e-01  2.681e-01   1.317
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu  -1.173e+01  2.670e+02  -0.044
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci  -2.152e-01  1.427e-01  -1.509
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua   3.497e-02  9.111e-02   0.384
## channel_saleslmkebamcaaclubfxadlmueccxoimlema  -3.724e-01  1.226e-01  -3.038
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds  -1.144e+01  1.690e+02  -0.068
## channel_salesusilxuppasemubllopkaafesmlibmsdf  -9.886e-02  1.228e-01  -0.805
## cons_gas_12m                                   -1.455e-06  4.568e-07  -3.186
## cons_last_month                                -2.562e-06  1.001e-06  -2.559
## forecast_meter_rent_12m                         1.105e-03  4.797e-04   2.304
## forecast_price_energy_p2                        1.372e+00  7.212e-01   1.903
## margin_net_pow_ele                              5.948e-03  7.994e-04   7.441
## net_margin                                      2.402e-04  8.815e-05   2.725
## num_years_antig                                -1.037e-01  2.242e-02  -4.626
## origin_upewxeelcelemmiwuafmddpobolfuxioce      -1.044e+01  5.354e+02  -0.020
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws       2.767e-01  4.688e-01   0.590
## origin_upldkssxwpmemidmecebumciepifcamkci       6.108e-01  4.681e-01   1.305
## origin_uplxidpiddsbxsbosboudacockeimpuepw       8.213e-01  4.651e-01   1.766
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      -1.061e+01  3.784e+02  -0.028
## pow_max                                        -4.764e-03  1.881e-03  -2.533
##                                                Pr(>|z|)
## (Intercept)                                    1.66e-07 ***
## activity_newapdekpcbwosbxepsfxclislboipuxpop    0.90014
## activity_newaxsupumdipebmlbiwolspmkdouoiddbc    0.82291
## activity_newbalskueexlmuccwdffilikwxasupasxf    0.44491
## activity_newbwpaswkpcilmlklklcapcwwumwaodaoo    0.60126
## activity_newbxopwkbwdewxssbmkwcummkaakbwafxf    0.92373
## activity_newcfdsselwimsklimddecfifseabdkxfcs    0.08645 .
## activity_newckadsdebplpkplelfspfoiucmxkeppus    0.94126
## activity_newckfxocssowaeipxueikxcmaxdmcduxsa    0.14312
```

```
## activity_newcluecxlameloamldmasudocsbmaoamdw    0.82297
## activity_newcwofmuicebbcmiaaxufmfimpowpacobu    0.35710
## activity_newdaobdssbkieoukwxbopxiiospudkopwl    0.38598
## activity_newddkpdekmbfdffwdmabkiiilolsxswccl    0.17804
## activity_newdfcsaaowsemmabpepocaeaaecfwppxxk    0.93194
## activity_newdupxuibdflmskeieweeofcaluuuiioix    0.41498
## activity_newfkmblacmaapkaoauabpwpuweokkeiali    0.80999
## activity_newfmwdwsxillemwbbwelxsampiuwwpcdcb    0.35155
## activity_newibkiiwcxiccxpoedpweiuxwbxbuewbxm    0.81312
## activity_newilkfsaapsxpkcpswbllddfmpamwelpxi    0.65042
## activity_newipdldckuswupeifllfbwccfpeafludfi    0.48921
## activity_newkkklcdamwfafdcfwofuscwfwadblfmce    0.96349
## activity_newkmxccaddbdpaaolkbidlobeefsbbcxca    0.19522
## activity_newkwuslieomapmswolewpobpplkaooaaew    0.83401
## activity_newlkeudbeowbapkpfodoxacpwdpaeuwxcx    0.89076
## activity_newmpicaaibskkfmxoblmwwwuuwpkecacil    0.69352
## activity_newOthers                             0.13115
## activity_newppcxfxbffsxaakxamcdpexdoxulfwwae    0.28842
## activity_newsaxlifeumaobawxpemwuopbwwldlucff    0.23467
## activity_newsffadmsbuamddwapeumdfibkmpkdicmc    0.50391
## activity_newsfisfxfcocfpcmckuekokxuseixdaoeu    0.39239
## activity_newsumdxiaiudmaioicexmiwuudlblkissm    0.69953
## activity_newsxublbwoeuckkocekklxkllcdxxaisop    0.16860
## activity_newwxemiwkumpibllwklfbcooafckufkdlm    0.18787
## channel_salesepumfxlbckeskwekxbiuasklxalciiuu   0.96496
## channel_salesewpakwlliwisiwduibdlfmalxowmwpci   0.13139
## channel_salesfoosdfpfkusacimwkcsosbicdxkicaua   0.70113
## channel_saleslmkebamcaaclubfxadlmueccxoimlema   0.00238 **
## channel_salessddiedcslfslkckwlfkdpoeeailfpeds   0.94605
## channel_salesusilxuppasemubllopkaafesmlibmsdf   0.42086
## cons_gas_12m                                   0.00144 **
## cons_last_month                                0.01050 *
## forecast_meter_rent_12m                        0.02122 *
## forecast_price_energy_p2                       0.05708 .
## margin_net_pow_ele                             1.00e-13 ***
## net_margin                                     0.00644 **
## num_years_antig                                3.73e-06 ***
## origin_upewxeelcelemmiwuafmddpobolfuxioce      0.98444
## origin_upkamkkxfxxuwbdslkwifmmcsiusiuosws      0.55496
## origin_upldkssxwpmemidmecebumciepifcamkci      0.19189
## origin_uplxidpiddsbxsbosboudacockeimpuepw      0.07744 .
## origin_upusapbepcfoloekilkwsdiboslwaxobdp      0.97762
## pow_max                                        0.01132 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 10020.9  on 15760  degrees of freedom
## Residual deviance:  9638.7  on 15709  degrees of freedom
## AIC: 9742.7
##
## Number of Fisher Scoring iterations: 12
```

```
anova(m19, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: churn
##
## Terms added sequentially (first to last)
##
##
##                         Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
## NULL                                   15760    10020.9
## activity_new            32   72.248      15728     9948.6 6.059e-05 ***
## channel_sales            6   97.522      15722     9851.1 < 2.2e-16 ***
## cons_gas_12m             1   27.352      15721     9823.7 1.696e-07 ***
## cons_last_month          1    6.705      15720     9817.0  0.009612 **
## forecast_meter_rent_12m  1   18.933      15719     9798.1 1.354e-05 ***
## forecast_price_energy_p2 1    5.141      15718     9792.9  0.023362 *
## margin_net_pow_ele       1   54.343      15717     9738.6 1.683e-13 ***
## net_margin               1    8.701      15716     9729.9  0.003180 **
## num_years_antig          1   36.935      15715     9693.0 1.222e-09 ***
## origin_up                5   47.093      15710     9645.9 5.440e-09 ***
## pow_max                  1    7.188      15709     9638.7  0.007338 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
churn_final <- m19
```

Now all the remaining variables are at least marginally significant and the ANOVA table past Chi-squared test. In our final model, the churn rate will depend on category of the company's activity, code of the sales channel, gas consumption of the past 12 months, electricity consumption of the last month, forecasted bill of meter rental for the next 12 months, forecasted energy price for 2nd period, net margin on power subscription, total net margin, antiquity of the client (in number of years), code of the electricity campaign the customer first subscribed to and subscribed power.

A negative estimate indicates that a client with the corresponding feature is less likely to churn, and a smaller p-value indicates a stronger impact of that feature. For example, a client from the sales channel lmkebamcaaclubfxadlmueccxoimlema is much less likely to churn and will be a loyal customer.

During the constuction of our final model, we delete many variables due to multicollinearity or insignificance thus we may not be ablt to adjust the removed variable in our model. We notice that the p-value for `forecast_price_energy_p2` is only 0.057 and positive estimate means a larger forecasted energy price for 2nd period will lead to larger churn rate, which confirm the hypothesis that clients are price sensitive.

And the churn probability and status can be predicted using the final model above as follows:

```
churn_rate <- round(predict(churn_final, data_combo_new, type = "response"), 4)
set.seed(2021)
predict_churn <- rbinom(length(churn_rate), 1, prob = churn_rate)

new.df <- data.frame(churn_rate, predict_churn)

data_combo_predict <- cbind(data_combo_new, new.df)
```

Thus, the prediction accuracy is 82.60% as shown below:

```
sum(data_combo_predict$churn == data_combo_predict$predict_churn) / nrow(data_combo_predict)
```

```
## [1] 0.8260263
```

We then use a holdout set to evaluate our model. We devide the whole dataset into two: the first 14000 items for training and the remaining for testing.

```
train_set <- data_combo_new[1:14000, ]
test_set <- data_combo_new[14001:15761, ]
churn_test <- update(churn_final, data = train_set)

churn_rate <- round(predict(churn_test, test_set, type = "response"), 4)
set.seed(2021)
predict_churn <- rbinom(length(churn_rate), 1, prob = churn_rate)

new.df <- data.frame(churn_rate, predict_churn)

test_predict <- cbind(test_set, new.df)
sum(test_predict$churn == test_predict$predict_churn) / nrow(test_predict)
```

```
## [1] 0.8336173
```

And the prediction accuracy is 83.36%.

Now we apply a 20% discount to `forecast_price_energy_p2` in our `churn_test` model and here is our prediction afterwards:

```
test_set_2 <- test_set
test_set_2$forecast_price_energy_p2 <- test_set$forecast_price_energy_p2 * 0.8

churn_rate_2 <- round(predict(churn_test, test_set_2, type = "response"), 4)
set.seed(2021)
predict_churn_2 <- rbinom(length(churn_rate_2), 1, prob = churn_rate_2)

sum(predict_churn)
```

```
## [1] 175
```

```
sum(predict_churn_2)
```

```
## [1] 174
```

By comparing the two predicted churn status, there are only 1 out of 175 customer stop churning given a 20% discount, so we can conclude that there is no need to provide such discount.