Additional MLR on nonzero dataset

Gulzina Kuttubekova

12/6/2019

```r
games <- read.csv('games.csv')
sales <- read.csv('sales.csv')
```

```r
# filter zero sales out
sales <- sales %>% select_all() %>% filter(Sales != 0)
```

```r
set.seed(13)
sets <- stratified(sales, c('Platform', 'Genre', 'Publisher',
                            'Developer', 'Rating', 'Decade',
                            'Platform_Company', 'Platform_Gen',
                            'Family_Platform', 'Main_Developer',
                            'Developer_Country', 'Main_Publisher',
                            'Region'), 0.7, bothSets = TRUE)


train_sales <- sets$SAMP1
test_sales <- sets$SAMP2
```

MLR 3 - on the dataset without zeros, interactions are kept

```r
mlr3_nonzero <- lm(log(Sales) ~ Region + Genre + Rating +
               Critic_Score + Platform_Gen +
               Region:Genre, data = train_sales)
mlr3_nonzero %>% summary()
```

```
##
## Call:
## lm(formula = log(Sales) ~ Region + Genre + Rating + Critic_Score +
##     Platform_Gen + Region:Genre, data = train_sales)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6277 -0.8499 -0.0231  0.8060  5.3956
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -4.4400441  0.6964844  -6.375 1.88e-10
## RegionJP_Sales          -0.6149401  0.0660793  -9.306  < 2e-16
## RegionNA_Sales           0.6679604  0.0463250  14.419  < 2e-16
## RegionOther_Sales       -0.8451266  0.0476266 -17.745  < 2e-16
## GenreAdventure          -0.4499721  0.0914119  -4.922 8.62e-07
## GenreFighting           -0.1146056  0.0792772  -1.446 0.148299
## GenreMisc                0.0774425  0.0795933   0.973 0.330578
## GenrePlatform           -0.0320778  0.0760597  -0.422 0.673217
## GenrePuzzle             -0.6385461  0.1525679  -4.185 2.86e-05
## GenreRacing              0.0648065  0.0662593   0.978 0.328051
## GenreRole-Playing       -0.3077583  0.0630717  -4.880 1.07e-06
```

```
## GenreShooter                          0.0808435  0.0565646   1.429 0.152957
## GenreSimulation                      -0.1293950  0.0861957  -1.501 0.133328
## GenreSports                          -0.2450235  0.0616640  -3.974 7.11e-05
## GenreStrategy                        -0.6030057  0.0902822  -6.679 2.48e-11
## RatingE                              -0.7087034  0.6927934  -1.023 0.306339
## RatingE10+                           -0.9387095  0.6929253  -1.355 0.175529
## RatingK-A                            -0.3211333  0.9172848  -0.350 0.726275
## RatingM                              -0.7787717  0.6926118  -1.124 0.260860
## RatingRP                             -0.7612237  1.0964946  -0.694 0.487545
## RatingT                              -0.9527897  0.6925487  -1.376 0.168910
## Critic_Score                          0.0348766  0.0006845  50.954  < 2e-16
## Platform_Gen6th Gen                   0.3652004  0.0354397  10.305  < 2e-16
## Platform_Gen7th Gen                   0.7134911  0.0346951  20.565  < 2e-16
## Platform_Gen8th Gen                   0.5079627  0.0413375  12.288  < 2e-16
## RegionJP_Sales:GenreAdventure         0.5043131  0.1917350   2.630 0.008539
## RegionNA_Sales:GenreAdventure        -0.0796208  0.1278214  -0.623 0.533354
## RegionOther_Sales:GenreAdventure     -0.0306850  0.1333085  -0.230 0.817954
## RegionJP_Sales:GenreFighting          0.3220268  0.1390767   2.315 0.020599
## RegionNA_Sales:GenreFighting          0.1698350  0.1086001   1.564 0.117870
## RegionOther_Sales:GenreFighting       0.0263919  0.1123271   0.235 0.814246
## RegionJP_Sales:GenreMisc              0.2705429  0.1670046   1.620 0.105256
## RegionNA_Sales:GenreMisc              0.1122286  0.1072580   1.046 0.295417
## RegionOther_Sales:GenreMisc           0.0019144  0.1103593   0.017 0.986160
## RegionJP_Sales:GenrePlatform          0.2865898  0.1512153   1.895 0.058077
## RegionNA_Sales:GenrePlatform          0.0358029  0.1033567   0.346 0.729045
## RegionOther_Sales:GenrePlatform      -0.1145549  0.1077541  -1.063 0.287745
## RegionJP_Sales:GenrePuzzle            0.8588600  0.2508119   3.424 0.000618
## RegionNA_Sales:GenrePuzzle            0.0336626  0.1952603   0.172 0.863126
## RegionOther_Sales:GenrePuzzle        -0.0986393  0.2038105  -0.484 0.628409
## RegionJP_Sales:GenreRacing           -0.5593531  0.1432012  -3.906 9.42e-05
## RegionNA_Sales:GenreRacing           -0.1706397  0.0911496  -1.872 0.061212
## RegionOther_Sales:GenreRacing        -0.0628239  0.0936382  -0.671 0.502279
## RegionJP_Sales:GenreRole-Playing      1.0152895  0.1058270   9.594  < 2e-16
## RegionNA_Sales:GenreRole-Playing      0.0235496  0.0866898   0.272 0.785891
## RegionOther_Sales:GenreRole-Playing  -0.0327451  0.0890359  -0.368 0.713047
## RegionJP_Sales:GenreShooter          -0.8277268  0.1133686  -7.301 2.98e-13
## RegionNA_Sales:GenreShooter          -0.0770460  0.0784096  -0.983 0.325814
## RegionOther_Sales:GenreShooter       -0.0379093  0.0809700  -0.468 0.639655
## RegionJP_Sales:GenreSimulation        0.5898957  0.1716264   3.437 0.000589
## RegionNA_Sales:GenreSimulation        0.0427230  0.1229585   0.347 0.728251
## RegionOther_Sales:GenreSimulation    -0.0274382  0.1243426  -0.221 0.825355
## RegionJP_Sales:GenreSports           -0.1332638  0.1359470  -0.980 0.326970
## RegionNA_Sales:GenreSports            0.1710847  0.0808016   2.117 0.034245
## RegionOther_Sales:GenreSports         0.0118723  0.0828041   0.143 0.885994
## RegionJP_Sales:GenreStrategy          0.3305179  0.2315213   1.428 0.153427
## RegionNA_Sales:GenreStrategy         -0.2385242  0.1322419  -1.804 0.071296
## RegionOther_Sales:GenreStrategy      -0.0525043  0.1343041  -0.391 0.695850
##
## (Intercept)                          ***
## RegionJP_Sales                       ***
## RegionNA_Sales                       ***
## RegionOther_Sales                    ***
## GenreAdventure                       ***
## GenreFighting
```

```
## GenreMisc
## GenrePlatform
## GenrePuzzle                        ***
## GenreRacing
## GenreRole-Playing                  ***
## GenreShooter
## GenreSimulation
## GenreSports                        ***
## GenreStrategy                      ***
## RatingE
## RatingE10+
## RatingK-A
## RatingM
## RatingRP
## RatingT
## Critic_Score                       ***
## Platform_Gen6th Gen                ***
## Platform_Gen7th Gen                ***
## Platform_Gen8th Gen                ***
## RegionJP_Sales:GenreAdventure      **
## RegionNA_Sales:GenreAdventure
## RegionOther_Sales:GenreAdventure
## RegionJP_Sales:GenreFighting       *
## RegionNA_Sales:GenreFighting
## RegionOther_Sales:GenreFighting
## RegionJP_Sales:GenreMisc
## RegionNA_Sales:GenreMisc
## RegionOther_Sales:GenreMisc
## RegionJP_Sales:GenrePlatform       .
## RegionNA_Sales:GenrePlatform
## RegionOther_Sales:GenrePlatform
## RegionJP_Sales:GenrePuzzle         ***
## RegionNA_Sales:GenrePuzzle
## RegionOther_Sales:GenrePuzzle
## RegionJP_Sales:GenreRacing         ***
## RegionNA_Sales:GenreRacing         .
## RegionOther_Sales:GenreRacing
## RegionJP_Sales:GenreRole-Playing   ***
## RegionNA_Sales:GenreRole-Playing
## RegionOther_Sales:GenreRole-Playing
## RegionJP_Sales:GenreShooter        ***
## RegionNA_Sales:GenreShooter
## RegionOther_Sales:GenreShooter
## RegionJP_Sales:GenreSimulation     ***
## RegionNA_Sales:GenreSimulation
## RegionOther_Sales:GenreSimulation
## RegionJP_Sales:GenreSports
## RegionNA_Sales:GenreSports         *
## RegionOther_Sales:GenreSports
## RegionJP_Sales:GenreStrategy
## RegionNA_Sales:GenreStrategy       .
## RegionOther_Sales:GenreStrategy
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## Residual standard error: 1.198 on 17390 degrees of freedom
## Multiple R-squared:  0.3119, Adjusted R-squared:  0.3097
## F-statistic: 138.3 on 57 and 17390 DF,  p-value: < 2.2e-16
```

```
anova(mlr3_nonzero)
```

```
## Analysis of Variance Table
## 
## Response: log(Sales)
##                 Df  Sum Sq Mean Sq  F value    Pr(>F)
## Region           3  5990.9  1997.0 1390.840 < 2.2e-16 ***
## Genre           11   523.6    47.6   33.152 < 2.2e-16 ***
## Rating           6   211.1    35.2   24.508 < 2.2e-16 ***
## Critic_Score     1  3295.6  3295.6 2295.294 < 2.2e-16 ***
## Platform_Gen     3   754.9   251.6  175.257 < 2.2e-16 ***
## Region:Genre    33   543.4    16.5   11.468 < 2.2e-16 ***
## Residuals    17390 24968.3     1.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We observed the same trend on assumptions violations! We'll check all possible model assessments at the end.

MLR 4 - on the dataset without zeros, interactions are kept

```
mlr4_nonzero <- lm(log(Sales) ~ Region + Genre + Rating +
                Main_Publisher + Platform_Gen +
                Region:Genre + Region:Main_Publisher, data = train_sales)
mlr4_nonzero %>% summary()
```

```
## 
## Call:
## lm(formula = log(Sales) ~ Region + Genre + Rating + Main_Publisher +
##     Platform_Gen + Region:Genre + Region:Main_Publisher, data = train_sales)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3274 -0.8817 -0.0795  0.8062  5.3006
## 
## Coefficients: (1 not defined because of singularities)
##                                   Estimate Std. Error t value
## (Intercept)                     -8.996e-01  7.076e-01  -1.271
## RegionJP_Sales                  -1.131e+00  1.854e-01  -6.099
## RegionNA_Sales                   7.728e-01  9.386e-02   8.233
## RegionOther_Sales               -9.043e-01  9.654e-02  -9.367
## GenreAdventure                  -5.063e-01  9.337e-02  -5.422
## GenreFighting                    9.656e-02  8.207e-02   1.177
## GenreMisc                        7.783e-02  8.142e-02   0.956
## GenrePlatform                    5.765e-02  7.789e-02   0.740
## GenrePuzzle                     -4.290e-01  1.562e-01  -2.747
## GenreRacing                      1.976e-01  6.775e-02   2.916
## GenreRole-Playing               -1.545e-01  6.619e-02  -2.334
## GenreShooter                     5.106e-02  5.762e-02   0.886
```

```
## GenreSimulation                        -5.857e-02  8.818e-02  -0.664
## GenreSports                            -1.277e-01  6.502e-02  -1.964
## GenreStrategy                          -4.310e-01  9.238e-02  -4.666
## RatingE                                -1.542e+00  7.043e-01  -2.189
## RatingE10+                             -1.698e+00  7.045e-01  -2.411
## RatingK-A                              -8.503e-01  9.333e-01  -0.911
## RatingM                                -1.294e+00  7.041e-01  -1.838
## RatingRP                               -1.371e+00  1.173e+00  -1.169
## RatingT                                -1.632e+00  7.041e-01  -2.318
## Main_PublisherEA                        1.252e-01  7.956e-02   1.574
## Main_PublisherMicrosoft                 3.097e-02  1.270e-01   0.244
## Main_PublisherNamco                    -5.618e-01  9.374e-02  -5.993
## Main_PublisherNintendo                  5.983e-01  1.028e-01   5.819
## Main_PublisherOther                    -6.293e-01  6.873e-02  -9.156
## Main_PublisherSega                     -1.229e-01  8.582e-02  -1.432
## Main_PublisherSONY                      3.648e-01  9.930e-02   3.674
## Main_PublisherSquare Enix              -1.053e-01  1.322e-01  -0.796
## Main_PublisherTencent                  -7.978e-01  4.271e-01  -1.868
## Main_PublisherUbisoft                  -3.762e-01  8.289e-02  -4.539
## Platform_Gen6th Gen                     2.028e-01  3.606e-02   5.623
## Platform_Gen7th Gen                     4.854e-01  3.518e-02  13.796
## Platform_Gen8th Gen                     3.128e-01  4.240e-02   7.376
## RegionJP_Sales:GenreAdventure           4.098e-01  1.954e-01   2.098
## RegionNA_Sales:GenreAdventure          -3.529e-02  1.306e-01  -0.270
## RegionOther_Sales:GenreAdventure       -2.088e-02  1.362e-01  -0.153
## RegionJP_Sales:GenreFighting            9.396e-02  1.458e-01   0.645
## RegionNA_Sales:GenreFighting            1.902e-01  1.125e-01   1.691
## RegionOther_Sales:GenreFighting         6.737e-03  1.166e-01   0.058
## RegionJP_Sales:GenreMisc               -2.209e-01  1.729e-01  -1.278
## RegionNA_Sales:GenreMisc                9.960e-02  1.097e-01   0.908
## RegionOther_Sales:GenreMisc             2.291e-02  1.130e-01   0.203
## RegionJP_Sales:GenrePlatform           -2.477e-02  1.569e-01  -0.158
## RegionNA_Sales:GenrePlatform            6.450e-02  1.059e-01   0.609
## RegionOther_Sales:GenrePlatform        -9.463e-02  1.105e-01  -0.857
## RegionJP_Sales:GenrePuzzle             -3.449e-03  2.594e-01  -0.013
## RegionNA_Sales:GenrePuzzle              1.020e-02  2.001e-01   0.051
## RegionOther_Sales:GenrePuzzle          -4.466e-02  2.091e-01  -0.214
## RegionJP_Sales:GenreRacing             -3.860e-01  1.483e-01  -2.602
## RegionNA_Sales:GenreRacing             -1.921e-01  9.324e-02  -2.060
## RegionOther_Sales:GenreRacing          -8.883e-02  9.584e-02  -0.927
## RegionJP_Sales:GenreRole-Playing        6.794e-01  1.110e-01   6.119
## RegionNA_Sales:GenreRole-Playing        4.747e-02  9.102e-02   0.522
## RegionOther_Sales:GenreRole-Playing    -3.085e-02  9.366e-02  -0.329
## RegionJP_Sales:GenreShooter            -5.694e-01  1.178e-01  -4.833
## RegionNA_Sales:GenreShooter            -7.979e-02  7.986e-02  -0.999
## RegionOther_Sales:GenreShooter         -3.300e-02  8.251e-02  -0.400
## RegionJP_Sales:GenreSimulation          4.104e-01  1.754e-01   2.340
## RegionNA_Sales:GenreSimulation          2.808e-02  1.258e-01   0.223
## RegionOther_Sales:GenreSimulation      -8.865e-03  1.273e-01  -0.070
## RegionJP_Sales:GenreSports              1.254e-01  1.433e-01   0.875
## RegionNA_Sales:GenreSports              1.204e-01  8.596e-02   1.401
## RegionOther_Sales:GenreSports           4.720e-03  8.804e-02   0.054
## RegionJP_Sales:GenreStrategy            3.319e-02  2.361e-01   0.141
## RegionNA_Sales:GenreStrategy           -1.749e-01  1.354e-01  -1.292
```

```
## RegionOther_Sales:GenreStrategy                9.811e-04  1.373e-01   0.007
## RegionJP_Sales:Main_PublisherEA               -5.022e-01  2.066e-01  -2.431
## RegionNA_Sales:Main_PublisherEA                1.373e-02  1.103e-01   0.124
## RegionOther_Sales:Main_PublisherEA             2.444e-02  1.126e-01   0.217
## RegionJP_Sales:Main_PublisherMicrosoft        -1.691e-01  2.734e-01  -0.619
## RegionNA_Sales:Main_PublisherMicrosoft         2.029e-01  1.815e-01   1.118
## RegionOther_Sales:Main_PublisherMicrosoft      2.496e-05  1.866e-01   0.000
## RegionJP_Sales:Main_PublisherNamco             1.369e+00  2.171e-01   6.309
## RegionNA_Sales:Main_PublisherNamco            -1.273e-01  1.300e-01  -0.979
## RegionOther_Sales:Main_PublisherNamco          1.147e-01  1.354e-01   0.847
## RegionJP_Sales:Main_PublisherNintendo          1.405e+00  2.118e-01   6.633
## RegionNA_Sales:Main_PublisherNintendo          1.300e-01  1.423e-01   0.914
## RegionOther_Sales:Main_PublisherNintendo      -2.657e-01  1.445e-01  -1.839
## RegionJP_Sales:Main_PublisherOther             7.798e-01  1.851e-01   4.212
## RegionNA_Sales:Main_PublisherOther            -1.946e-01  9.466e-02  -2.056
## RegionOther_Sales:Main_PublisherOther          9.893e-02  9.757e-02   1.014
## RegionJP_Sales:Main_PublisherSega              3.852e-01  2.099e-01   1.835
## RegionNA_Sales:Main_PublisherSega             -1.557e-01  1.188e-01  -1.311
## RegionOther_Sales:Main_PublisherSega           7.532e-02  1.216e-01   0.620
## RegionJP_Sales:Main_PublisherSONY              6.001e-01  2.196e-01   2.732
## RegionNA_Sales:Main_PublisherSONY             -3.383e-01  1.367e-01  -2.475
## RegionOther_Sales:Main_PublisherSONY           2.039e-02  1.384e-01   0.147
## RegionJP_Sales:Main_PublisherSquare Enix       1.177e+00  2.502e-01   4.705
## RegionNA_Sales:Main_PublisherSquare Enix      -1.081e-01  1.853e-01  -0.583
## RegionOther_Sales:Main_PublisherSquare Enix    5.145e-02  1.877e-01   0.274
## RegionJP_Sales:Main_PublisherTencent                  NA         NA        NA
## RegionNA_Sales:Main_PublisherTencent          -1.223e+00  7.492e-01  -1.633
## RegionOther_Sales:Main_PublisherTencent       -2.104e-01  6.938e-01  -0.303
## RegionJP_Sales:Main_PublisherUbisoft           3.408e-01  2.193e-01   1.554
## RegionNA_Sales:Main_PublisherUbisoft          -1.107e-01  1.154e-01  -0.959
## RegionOther_Sales:Main_PublisherUbisoft        5.143e-02  1.185e-01   0.434
##                                               Pr(>|t|)
## (Intercept)                                    0.20363
## RegionJP_Sales                                 1.09e-09 ***
## RegionNA_Sales                                 < 2e-16 ***
## RegionOther_Sales                              < 2e-16 ***
## GenreAdventure                                 5.96e-08 ***
## GenreFighting                                  0.23937
## GenreMisc                                      0.33909
## GenrePlatform                                  0.45918
## GenrePuzzle                                    0.00602 **
## GenreRacing                                    0.00355 **
## GenreRole-Playing                              0.01959 *
## GenreShooter                                   0.37557
## GenreSimulation                                0.50660
## GenreSports                                    0.04951 *
## GenreStrategy                                  3.10e-06 ***
## RatingE                                        0.02858 *
## RatingE10+                                     0.01592 *
## RatingK-A                                      0.36225
## RatingM                                        0.06615 .
## RatingRP                                       0.24252
## RatingT                                        0.02047 *
## Main_PublisherEA                               0.11554
```

```
## Main_PublisherMicrosoft                          0.80728
## Main_PublisherNamco                               2.10e-09 ***
## Main_PublisherNintendo                            6.04e-09 ***
## Main_PublisherOther                               < 2e-16 ***
## Main_PublisherSega                                0.15209
## Main_PublisherSONY                                0.00024 ***
## Main_PublisherSquare Enix                         0.42579
## Main_PublisherTencent                             0.06179 .
## Main_PublisherUbisoft                             5.70e-06 ***
## Platform_Gen6th Gen                               1.90e-08 ***
## Platform_Gen7th Gen                               < 2e-16 ***
## Platform_Gen8th Gen                               1.70e-13 ***
## RegionJP_Sales:GenreAdventure                     0.03594 *
## RegionNA_Sales:GenreAdventure                     0.78694
## RegionOther_Sales:GenreAdventure                  0.87814
## RegionJP_Sales:GenreFighting                      0.51917
## RegionNA_Sales:GenreFighting                      0.09088 .
## RegionOther_Sales:GenreFighting                   0.95392
## RegionJP_Sales:GenreMisc                          0.20124
## RegionNA_Sales:GenreMisc                          0.36396
## RegionOther_Sales:GenreMisc                       0.83926
## RegionJP_Sales:GenrePlatform                      0.87454
## RegionNA_Sales:GenrePlatform                      0.54240
## RegionOther_Sales:GenrePlatform                   0.39172
## RegionJP_Sales:GenrePuzzle                        0.98939
## RegionNA_Sales:GenrePuzzle                        0.95937
## RegionOther_Sales:GenrePuzzle                     0.83085
## RegionJP_Sales:GenreRacing                        0.00928 **
## RegionNA_Sales:GenreRacing                        0.03940 *
## RegionOther_Sales:GenreRacing                     0.35405
## RegionJP_Sales:GenreRole-Playing                  9.59e-10 ***
## RegionNA_Sales:GenreRole-Playing                  0.60197
## RegionOther_Sales:GenreRole-Playing               0.74188
## RegionJP_Sales:GenreShooter                       1.35e-06 ***
## RegionNA_Sales:GenreShooter                       0.31779
## RegionOther_Sales:GenreShooter                    0.68921
## RegionJP_Sales:GenreSimulation                    0.01928 *
## RegionNA_Sales:GenreSimulation                    0.82339
## RegionOther_Sales:GenreSimulation                 0.94449
## RegionJP_Sales:GenreSports                        0.38154
## RegionNA_Sales:GenreSports                        0.16130
## RegionOther_Sales:GenreSports                     0.95724
## RegionJP_Sales:GenreStrategy                      0.88823
## RegionNA_Sales:GenreStrategy                      0.19639
## RegionOther_Sales:GenreStrategy                   0.99430
## RegionJP_Sales:Main_PublisherEA                   0.01506 *
## RegionNA_Sales:Main_PublisherEA                   0.90094
## RegionOther_Sales:Main_PublisherEA                0.82821
## RegionJP_Sales:Main_PublisherMicrosoft            0.53623
## RegionNA_Sales:Main_PublisherMicrosoft            0.26366
## RegionOther_Sales:Main_PublisherMicrosoft         0.99989
## RegionJP_Sales:Main_PublisherNamco                2.88e-10 ***
## RegionNA_Sales:Main_PublisherNamco                0.32771
## RegionOther_Sales:Main_PublisherNamco             0.39698
```

```
## RegionJP_Sales:Main_PublisherNintendo         3.38e-11 ***
## RegionNA_Sales:Main_PublisherNintendo          0.36093
## RegionOther_Sales:Main_PublisherNintendo       0.06588 .
## RegionJP_Sales:Main_PublisherOther            2.54e-05 ***
## RegionNA_Sales:Main_PublisherOther             0.03982 *
## RegionOther_Sales:Main_PublisherOther          0.31063
## RegionJP_Sales:Main_PublisherSega             0.06658 .
## RegionNA_Sales:Main_PublisherSega              0.18990
## RegionOther_Sales:Main_PublisherSega           0.53557
## RegionJP_Sales:Main_PublisherSONY             0.00629 **
## RegionNA_Sales:Main_PublisherSONY             0.01334 *
## RegionOther_Sales:Main_PublisherSONY           0.88287
## RegionJP_Sales:Main_PublisherSquare Enix     2.56e-06 ***
## RegionNA_Sales:Main_PublisherSquare Enix      0.55973
## RegionOther_Sales:Main_PublisherSquare Enix  0.78405
## RegionJP_Sales:Main_PublisherTencent               NA
## RegionNA_Sales:Main_PublisherTencent          0.10252
## RegionOther_Sales:Main_PublisherTencent       0.76174
## RegionJP_Sales:Main_PublisherUbisoft          0.12013
## RegionNA_Sales:Main_PublisherUbisoft          0.33738
## RegionOther_Sales:Main_PublisherUbisoft       0.66426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.217 on 17352 degrees of freedom
## Multiple R-squared:  0.2915, Adjusted R-squared:  0.2877
## F-statistic: 75.16 on 95 and 17352 DF,  p-value: < 2.2e-16
```

```r
anova(mlr4_nonzero)
```

```
## Analysis of Variance Table
##
## Response: log(Sales)
##                       Df  Sum Sq Mean Sq    F value    Pr(>F)
## Region                 3  5990.9 1996.95 1347.8386 < 2.2e-16 ***
## Genre                 11   523.6   47.60   32.1271 < 2.2e-16 ***
## Rating                 6   211.1   35.19   23.7500 < 2.2e-16 ***
## Main_Publisher        10  2480.3  248.03  167.4083 < 2.2e-16 ***
## Platform_Gen           3   381.3  127.10   85.7827 < 2.2e-16 ***
## Region:Genre          33   487.6   14.77    9.9719 < 2.2e-16 ***
## Region:Main_Publisher 29   504.4   17.39   11.7385 < 2.2e-16 ***
## Residuals          17352 25708.6    1.48
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

AdjR6 dropped significantly!!!

We observed the same trend on assumptions violations! We'll check all possible model assessments at the end.

MLR 5

```r
mlr5_nonzero <- lm(log(Sales) ~ Region + Genre + Rating +
            Main_Developer + Platform_Company +
```

```
              Region:Genre + Region:Main_Developer, data = train_sales)
mlr5_nonzero %>% summary()
```

```
##
## Call:
## lm(formula = log(Sales) ~ Region + Genre + Rating + Main_Developer +
##     Platform_Company + Region:Genre + Region:Main_Developer,
##     data = train_sales)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9621 -0.8901 -0.0755  0.8275  5.2965
##
## Coefficients:
##                                  Estimate Std. Error t value
## (Intercept)                     -0.0893959  0.7154339  -0.125
## RegionJP_Sales                  -1.4777765  0.2943356  -5.021
## RegionNA_Sales                   0.6672261  0.1552447   4.298
## RegionOther_Sales               -1.0385381  0.1601638  -6.484
## GenreAdventure                  -0.5271555  0.0935192  -5.637
## GenreFighting                    0.0425827  0.0817935   0.521
## GenreMisc                        0.1543569  0.0816876   1.890
## GenrePlatform                    0.0046951  0.0782240   0.060
## GenrePuzzle                     -0.3311955  0.1565199  -2.116
## GenreRacing                      0.1524249  0.0681807   2.236
## GenreRole-Playing               -0.2938273  0.0665380  -4.416
## GenreShooter                     0.0831811  0.0579952   1.434
## GenreSimulation                 -0.0596474  0.0883337  -0.675
## GenreSports                     -0.1367271  0.0672204  -2.034
## GenreStrategy                   -0.4948191  0.0917430  -5.394
## RatingE                         -2.0249015  0.7077292  -2.861
## RatingE10+                      -2.0822630  0.7076897  -2.942
## RatingK-A                       -1.2909593  0.9367018  -1.378
## RatingM                         -1.7459602  0.7073570  -2.468
## RatingRP                        -2.3235595  1.1195368  -2.075
## RatingT                         -2.1293727  0.7073600  -3.010
## Main_DeveloperEA                -0.0708213  0.1231487  -0.575
## Main_DeveloperMicrosoft          0.3267870  0.1631768   2.003
## Main_DeveloperNamco             -0.5244444  0.1641987  -3.194
## Main_DeveloperNintendo           0.4715449  0.1425396   3.308
## Main_DeveloperOther             -0.7430983  0.1104502  -6.728
## Main_DeveloperSega              -0.3856475  0.1499372  -2.572
## Main_DeveloperSONY               0.0299626  0.1399204   0.214
## Main_DeveloperSquare Enix       -0.0709756  0.1915349  -0.371
## Main_DeveloperTencent           -0.4484610  0.3100017  -1.447
## Main_DeveloperUbisoft           -0.3084366  0.1296324  -2.379
## Platform_CompanyNintendo         0.1179343  0.0284221   4.149
## Platform_CompanySega             1.1210481  0.2865395   3.912
## Platform_CompanySONY             0.4519979  0.0232440  19.446
## RegionJP_Sales:GenreAdventure    0.5261245  0.1963428   2.680
## RegionNA_Sales:GenreAdventure   -0.0477944  0.1307882  -0.365
## RegionOther_Sales:GenreAdventure -0.0346095  0.1364207  -0.254
## RegionJP_Sales:GenreFighting     0.1746385  0.1450552   1.204
## RegionNA_Sales:GenreFighting     0.1690989  0.1120608   1.509
```

```
## RegionOther_Sales:GenreFighting              -0.0083404  0.1159519  -0.072
## RegionJP_Sales:GenreMisc                      0.1796033  0.1725110   1.041
## RegionNA_Sales:GenreMisc                      0.1011181  0.1100998   0.918
## RegionOther_Sales:GenreMisc                   0.0113254  0.1133204   0.100
## RegionJP_Sales:GenrePlatform                  0.1783972  0.1587912   1.123
## RegionNA_Sales:GenrePlatform                  0.0432712  0.1065055   0.406
## RegionOther_Sales:GenrePlatform              -0.1032967  0.1111656  -0.929
## RegionJP_Sales:GenrePuzzle                    0.3958836  0.2592499   1.527
## RegionNA_Sales:GenrePuzzle                   -0.0019222  0.2004928  -0.010
## RegionOther_Sales:GenrePuzzle                -0.0618457  0.2093599  -0.295
## RegionJP_Sales:GenreRacing                   -0.4037382  0.1490520  -2.709
## RegionNA_Sales:GenreRacing                   -0.1852917  0.0936493  -1.979
## RegionOther_Sales:GenreRacing                -0.0934333  0.0962797  -0.970
## RegionJP_Sales:GenreRole-Playing              0.7537091  0.1118032   6.741
## RegionNA_Sales:GenreRole-Playing              0.0575255  0.0915005   0.629
## RegionOther_Sales:GenreRole-Playing          -0.0066350  0.0941618  -0.070
## RegionJP_Sales:GenreShooter                  -0.5779815  0.1174468  -4.921
## RegionNA_Sales:GenreShooter                  -0.0695608  0.0803756  -0.865
## RegionOther_Sales:GenreShooter               -0.0243462  0.0830428  -0.293
## RegionJP_Sales:GenreSimulation                0.4336201  0.1764034   2.458
## RegionNA_Sales:GenreSimulation                0.0496052  0.1262683   0.393
## RegionOther_Sales:GenreSimulation            -0.0036576  0.1277367  -0.029
## RegionJP_Sales:GenreSports                    0.1612140  0.1472961   1.094
## RegionNA_Sales:GenreSports                    0.0876727  0.0893503   0.981
## RegionOther_Sales:GenreSports                 0.0040837  0.0914409   0.045
## RegionJP_Sales:GenreStrategy                 -0.0331618  0.2374033  -0.140
## RegionNA_Sales:GenreStrategy                 -0.1962713  0.1352420  -1.451
## RegionOther_Sales:GenreStrategy               0.0020173  0.1374071   0.015
## RegionJP_Sales:Main_DeveloperEA              -0.3035966  0.3209273  -0.946
## RegionNA_Sales:Main_DeveloperEA               0.1883452  0.1716410   1.097
## RegionOther_Sales:Main_DeveloperEA            0.2027471  0.1762335   1.150
## RegionJP_Sales:Main_DeveloperMicrosoft       -0.0234890  0.3646858  -0.064
## RegionNA_Sales:Main_DeveloperMicrosoft        0.1038195  0.2303096   0.451
## RegionOther_Sales:Main_DeveloperMicrosoft    -0.0005943  0.2349842  -0.003
## RegionJP_Sales:Main_DeveloperNamco            1.8412337  0.3529364   5.217
## RegionNA_Sales:Main_DeveloperNamco           -0.0722983  0.2283818  -0.317
## RegionOther_Sales:Main_DeveloperNamco         0.2178597  0.2362839   0.922
## RegionJP_Sales:Main_DeveloperNintendo         1.7546749  0.3211627   5.464
## RegionNA_Sales:Main_DeveloperNintendo         0.3191842  0.1977369   1.614
## RegionOther_Sales:Main_DeveloperNintendo     -0.0516393  0.2019303  -0.256
## RegionJP_Sales:Main_DeveloperOther            1.1199954  0.2934692   3.816
## RegionNA_Sales:Main_DeveloperOther           -0.0288737  0.1537179  -0.188
## RegionOther_Sales:Main_DeveloperOther         0.2265381  0.1586362   1.428
## RegionJP_Sales:Main_DeveloperSega             0.4116833  0.3468028   1.187
## RegionNA_Sales:Main_DeveloperSega             0.2883006  0.2069127   1.393
## RegionOther_Sales:Main_DeveloperSega          0.2278167  0.2116343   1.076
## RegionJP_Sales:Main_DeveloperSONY             0.7686718  0.3248557   2.366
## RegionNA_Sales:Main_DeveloperSONY            -0.2137059  0.1927584  -1.109
## RegionOther_Sales:Main_DeveloperSONY          0.1497919  0.1965084   0.762
## RegionJP_Sales:Main_DeveloperSquare Enix      1.7974323  0.3713815   4.840
## RegionNA_Sales:Main_DeveloperSquare Enix     -0.0043903  0.2628907  -0.017
## RegionOther_Sales:Main_DeveloperSquare Enix   0.2568321  0.2731129   0.940
## RegionJP_Sales:Main_DeveloperTencent          0.3391882  0.6856396   0.495
## RegionNA_Sales:Main_DeveloperTencent          0.1992526  0.4636178   0.430
```

```
## RegionOther_Sales:Main_DeveloperTencent       0.1048020  0.4575501   0.229
## RegionJP_Sales:Main_DeveloperUbisoft          0.1311644  0.3361779   0.390
## RegionNA_Sales:Main_DeveloperUbisoft         -0.1054690  0.1813388  -0.582
## RegionOther_Sales:Main_DeveloperUbisoft       0.1602694  0.1863802   0.860
##                                               Pr(>|t|)
## (Intercept)                                   0.900562
## RegionJP_Sales                                5.20e-07 ***
## RegionNA_Sales                                1.73e-05 ***
## RegionOther_Sales                             9.16e-11 ***
## GenreAdventure                                1.76e-08 ***
## GenreFighting                                 0.602644
## GenreMisc                                     0.058828 .
## GenrePlatform                                 0.952139
## GenrePuzzle                                   0.034359 *
## GenreRacing                                   0.025390 *
## GenreRole-Playing                             1.01e-05 ***
## GenreShooter                                  0.151511
## GenreSimulation                               0.499526
## GenreSports                                   0.041966 *
## GenreStrategy                                 7.00e-08 ***
## RatingE                                       0.004226 **
## RatingE10+                                    0.003262 **
## RatingK-A                                     0.168160
## RatingM                                       0.013586 *
## RatingRP                                      0.037958 *
## RatingT                                       0.002614 **
## Main_DeveloperEA                              0.565239
## Main_DeveloperMicrosoft                       0.045230 *
## Main_DeveloperNamco                           0.001406 **
## Main_DeveloperNintendo                        0.000941 ***
## Main_DeveloperOther                           1.78e-11 ***
## Main_DeveloperSega                            0.010118 *
## Main_DeveloperSONY                            0.830440
## Main_DeveloperSquare Enix                     0.710968
## Main_DeveloperTencent                         0.148016
## Main_DeveloperUbisoft                         0.017356 *
## Platform_CompanyNintendo                      3.35e-05 ***
## Platform_CompanySega                          9.17e-05 ***
## Platform_CompanySONY                           < 2e-16 ***
## RegionJP_Sales:GenreAdventure                 0.007377 **
## RegionNA_Sales:GenreAdventure                 0.714792
## RegionOther_Sales:GenreAdventure              0.799733
## RegionJP_Sales:GenreFighting                  0.228627
## RegionNA_Sales:GenreFighting                  0.131319
## RegionOther_Sales:GenreFighting               0.942659
## RegionJP_Sales:GenreMisc                      0.297838
## RegionNA_Sales:GenreMisc                      0.358411
## RegionOther_Sales:GenreMisc                   0.920392
## RegionJP_Sales:GenrePlatform                  0.261253
## RegionNA_Sales:GenrePlatform                  0.684541
## RegionOther_Sales:GenrePlatform               0.352791
## RegionJP_Sales:GenrePuzzle                    0.126771
## RegionNA_Sales:GenrePuzzle                    0.992351
## RegionOther_Sales:GenrePuzzle                 0.767689
```

```
## RegionJP_Sales:GenreRacing                     0.006761 **
## RegionNA_Sales:GenreRacing                     0.047880 *
## RegionOther_Sales:GenreRacing                  0.331842
## RegionJP_Sales:GenreRole-Playing               1.62e-11 ***
## RegionNA_Sales:GenreRole-Playing               0.529560
## RegionOther_Sales:GenreRole-Playing            0.943825
## RegionJP_Sales:GenreShooter                    8.68e-07 ***
## RegionNA_Sales:GenreShooter                    0.386806
## RegionOther_Sales:GenreShooter                 0.769391
## RegionJP_Sales:GenreSimulation                 0.013976 *
## RegionNA_Sales:GenreSimulation                 0.694431
## RegionOther_Sales:GenreSimulation              0.977157
## RegionJP_Sales:GenreSports                     0.273756
## RegionNA_Sales:GenreSports                     0.326496
## RegionOther_Sales:GenreSports                  0.964379
## RegionJP_Sales:GenreStrategy                   0.888910
## RegionNA_Sales:GenreStrategy                   0.146726
## RegionOther_Sales:GenreStrategy               0.988286
## RegionJP_Sales:Main_DeveloperEA                0.344163
## RegionNA_Sales:Main_DeveloperEA                0.272517
## RegionOther_Sales:Main_DeveloperEA             0.249976
## RegionJP_Sales:Main_DeveloperMicrosoft         0.948645
## RegionNA_Sales:Main_DeveloperMicrosoft         0.652152
## RegionOther_Sales:Main_DeveloperMicrosoft      0.997982
## RegionJP_Sales:Main_DeveloperNamco             1.84e-07 ***
## RegionNA_Sales:Main_DeveloperNamco             0.751575
## RegionOther_Sales:Main_DeveloperNamco          0.356528
## RegionJP_Sales:Main_DeveloperNintendo          4.73e-08 ***
## RegionNA_Sales:Main_DeveloperNintendo          0.106505
## RegionOther_Sales:Main_DeveloperNintendo       0.798163
## RegionJP_Sales:Main_DeveloperOther             0.000136 ***
## RegionNA_Sales:Main_DeveloperOther             0.851008
## RegionOther_Sales:Main_DeveloperOther          0.153300
## RegionJP_Sales:Main_DeveloperSega              0.235212
## RegionNA_Sales:Main_DeveloperSega              0.163534
## RegionOther_Sales:Main_DeveloperSega           0.281735
## RegionJP_Sales:Main_DeveloperSONY              0.017983 *
## RegionNA_Sales:Main_DeveloperSONY              0.267587
## RegionOther_Sales:Main_DeveloperSONY           0.445911
## RegionJP_Sales:Main_DeveloperSquare Enix       1.31e-06 ***
## RegionNA_Sales:Main_DeveloperSquare Enix       0.986676
## RegionOther_Sales:Main_DeveloperSquare Enix 0.347032
## RegionJP_Sales:Main_DeveloperTencent           0.620816
## RegionNA_Sales:Main_DeveloperTencent           0.667363
## RegionOther_Sales:Main_DeveloperTencent        0.818833
## RegionJP_Sales:Main_DeveloperUbisoft           0.696420
## RegionNA_Sales:Main_DeveloperUbisoft           0.560835
## RegionOther_Sales:Main_DeveloperUbisoft        0.389853
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.224 on 17351 degrees of freedom
## Multiple R-squared:  0.2835, Adjusted R-squared:  0.2795
## F-statistic: 71.51 on 96 and 17351 DF,  p-value: < 2.2e-16
```

```
anova(mlr5_nonzero)
```

```
## Analysis of Variance Table
##
## Response: log(Sales)
##                       Df  Sum Sq Mean Sq  F value    Pr(>F)
## Region                 3  5990.9 1996.95 1332.6390 < 2.2e-16 ***
## Genre                 11   523.6   47.60   31.7648 < 2.2e-16 ***
## Rating                 6   211.1   35.19   23.4822 < 2.2e-16 ***
## Main_Developer        10  1979.8  197.98  132.1164 < 2.2e-16 ***
## Platform_Company       3   612.8  204.27  136.3199 < 2.2e-16 ***
## Region:Genre          33   483.7   14.66    9.7825 < 2.2e-16 ***
## Region:Main_Developer 30   485.5   16.18   10.7989 < 2.2e-16 ***
## Residuals          17351 26000.4    1.50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Still all 5 variables including the interactions are significant. Adj-R^2 is 0.4094, dropped from mlr3 and mlr4.
Model violations are the same.

MLR 6

```
mlr6_nonzero <- lm(log(Sales) ~ Region + Genre + Rating + Region:Genre,
          data = train_sales)
mlr6_nonzero %>% summary()
```

```
##
## Call:
## lm(formula = log(Sales) ~ Region + Genre + Rating + Region:Genre,
##     data = train_sales)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1932 -0.9525 -0.0864  0.8635  5.7079
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -0.838898   0.747676  -1.122  0.26187
## RegionJP_Sales     -0.352351   0.071066  -4.958 7.19e-07
## RegionNA_Sales      0.654450   0.050005  13.088  < 2e-16
## RegionOther_Sales  -0.808886   0.051407 -15.735  < 2e-16
## GenreAdventure     -0.553074   0.098678  -5.605 2.12e-08
## GenreFighting      -0.003957   0.085568  -0.046  0.96312
## GenreMisc           0.195551   0.085896   2.277  0.02282
## GenrePlatform       0.050192   0.081996   0.612  0.54046
## GenrePuzzle        -0.315613   0.164616  -1.917  0.05522
## GenreRacing         0.128901   0.071445   1.804  0.07122
## GenreRole-Playing  -0.151063   0.067919  -2.224  0.02615
## GenreShooter        0.118988   0.061009   1.950  0.05115
## GenreSimulation    -0.100595   0.092764  -1.084  0.27820
## GenreSports         0.002621   0.066288   0.040  0.96846
## GenreStrategy      -0.623415   0.096528  -6.458 1.09e-10
```

```
## RatingE                               -1.505741   0.747790  -2.014   0.04407
## RatingE10+                            -1.652882   0.747804  -2.210   0.02710
## RatingK-A                             -0.853730   0.989820  -0.863   0.38842
## RatingM                               -1.437637   0.747603  -1.923   0.05450
## RatingRP                              -2.214359   1.183431  -1.871   0.06134
## RatingT                               -1.718124   0.747537  -2.298   0.02155
## RegionJP_Sales:GenreAdventure          0.613424   0.207045   2.963   0.00305
## RegionNA_Sales:GenreAdventure         -0.034081   0.138010  -0.247   0.80495
## RegionOther_Sales:GenreAdventure      -0.019248   0.143958  -0.134   0.89364
## RegionJP_Sales:GenreFighting           0.286449   0.150164   1.908   0.05646
## RegionNA_Sales:GenreFighting           0.172402   0.117278   1.470   0.14157
## RegionOther_Sales:GenreFighting        0.018520   0.121304   0.153   0.87866
## RegionJP_Sales:GenreMisc               0.137438   0.180314   0.762   0.44594
## RegionNA_Sales:GenreMisc               0.082681   0.115821   0.714   0.47532
## RegionOther_Sales:GenreMisc           -0.032798   0.119173  -0.275   0.78315
## RegionJP_Sales:GenrePlatform           0.343641   0.163281   2.105   0.03534
## RegionNA_Sales:GenrePlatform           0.046826   0.111615   0.420   0.67483
## RegionOther_Sales:GenrePlatform       -0.116752   0.116362  -1.003   0.31570
## RegionJP_Sales:GenrePuzzle             0.735731   0.270848   2.716   0.00661
## RegionNA_Sales:GenrePuzzle            -0.065514   0.210849  -0.311   0.75602
## RegionOther_Sales:GenrePuzzle         -0.153786   0.220097  -0.699   0.48474
## RegionJP_Sales:GenreRacing            -0.483730   0.154636  -3.128   0.00176
## RegionNA_Sales:GenreRacing            -0.171557   0.098426  -1.743   0.08135
## RegionOther_Sales:GenreRacing         -0.075409   0.101122  -0.746   0.45584
## RegionJP_Sales:GenreRole-Playing       0.873901   0.114215   7.651 2.09e-14
## RegionNA_Sales:GenreRole-Playing       0.053586   0.093605   0.572   0.56701
## RegionOther_Sales:GenreRole-Playing   -0.039290   0.096143  -0.409   0.68279
## RegionJP_Sales:GenreShooter           -0.723846   0.122359  -5.916 3.37e-09
## RegionNA_Sales:GenreShooter           -0.070129   0.084674  -0.828   0.40755
## RegionOther_Sales:GenreShooter        -0.025542   0.087440  -0.292   0.77021
## RegionJP_Sales:GenreSimulation         0.532396   0.185212   2.875   0.00405
## RegionNA_Sales:GenreSimulation         0.093747   0.132730   0.706   0.48001
## RegionOther_Sales:GenreSimulation     -0.003924   0.134271  -0.029   0.97668
## RegionJP_Sales:GenreSports            -0.204810   0.146774  -1.395   0.16291
## RegionNA_Sales:GenreSports             0.200537   0.087247   2.298   0.02155
## RegionOther_Sales:GenreSports          0.012123   0.089408   0.136   0.89214
## RegionJP_Sales:GenreStrategy           0.394444   0.249668   1.580   0.11415
## RegionNA_Sales:GenreStrategy          -0.109164   0.142642  -0.765   0.44410
## RegionOther_Sales:GenreStrategy        0.050504   0.144989   0.348   0.72760
##
## (Intercept)
## RegionJP_Sales                      ***
## RegionNA_Sales                      ***
## RegionOther_Sales                   ***
## GenreAdventure                      ***
## GenreFighting
## GenreMisc                           *
## GenrePlatform
## GenrePuzzle                         .
## GenreRacing                         .
## GenreRole-Playing                   *
## GenreShooter                        .
## GenreSimulation
## GenreSports
```

```
## GenreStrategy                        ***
## RatingE                              *
## RatingE10+                           *
## RatingK-A
## RatingM                              .
## RatingRP                             .
## RatingT                              *
## RegionJP_Sales:GenreAdventure        **
## RegionNA_Sales:GenreAdventure
## RegionOther_Sales:GenreAdventure
## RegionJP_Sales:GenreFighting         .
## RegionNA_Sales:GenreFighting
## RegionOther_Sales:GenreFighting
## RegionJP_Sales:GenreMisc
## RegionNA_Sales:GenreMisc
## RegionOther_Sales:GenreMisc
## RegionJP_Sales:GenrePlatform         *
## RegionNA_Sales:GenrePlatform
## RegionOther_Sales:GenrePlatform
## RegionJP_Sales:GenrePuzzle           **
## RegionNA_Sales:GenrePuzzle
## RegionOther_Sales:GenrePuzzle
## RegionJP_Sales:GenreRacing           **
## RegionNA_Sales:GenreRacing           .
## RegionOther_Sales:GenreRacing
## RegionJP_Sales:GenreRole-Playing     ***
## RegionNA_Sales:GenreRole-Playing
## RegionOther_Sales:GenreRole-Playing
## RegionJP_Sales:GenreShooter          ***
## RegionNA_Sales:GenreShooter
## RegionOther_Sales:GenreShooter
## RegionJP_Sales:GenreSimulation       **
## RegionNA_Sales:GenreSimulation
## RegionOther_Sales:GenreSimulation
## RegionJP_Sales:GenreSports
## RegionNA_Sales:GenreSports           *
## RegionOther_Sales:GenreSports
## RegionJP_Sales:GenreStrategy
## RegionNA_Sales:GenreStrategy
## RegionOther_Sales:GenreStrategy
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.294 on 17394 degrees of freedom
## Multiple R-squared:  0.1974, Adjusted R-squared:  0.1949
## F-statistic: 80.69 on 53 and 17394 DF,  p-value: < 2.2e-16
```

```
anova(mlr6_nonzero)
```

```
## Analysis of Variance Table
##
## Response: log(Sales)
##                 Df  Sum Sq Mean Sq   F value     Pr(>F)
## Region           3  5990.9 1996.95 1192.5684 < 2.2e-16 ***
## Genre           11   523.6   47.60   28.4261 < 2.2e-16 ***
```

```
## Rating              6    211.1    35.19    21.0141 < 2.2e-16 ***
## Region:Genre       33    436.0    13.21     7.8896 < 2.2e-16 ***
## Residuals       17394 29126.2     1.67
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Still all 3 variables including the interactions are significant. Adj-R^2 is really badddd, decreased from all previous cases. Model violations are the same.

We observed the same trend on assumptions violations! We'll check all possible model assessments at the end.

Model Comparison

Predictions and RMSE:

```r
# rmse on train data
sqrt(sum((mlr3_nonzero$fitted.values - train_sales$Sales)^2) / length(mlr3_nonzero$fitted.values))
```

```
## [1] 2.890188
```

```r
predicted_sales <- predict(mlr3_nonzero, test_sales[,c(3,5,10,14,19)])

# rmse on test data
sqrt(sum((predicted_sales - test_sales$Sales)^2) / length(predicted_sales))
```

```
## [1] 2.837222
```

```r
# rmse on train data
sqrt(sum((mlr4_nonzero$fitted.values - train_sales$Sales)^2) / length(mlr4_nonzero$fitted.values))
```

```
## [1] 2.886228
```

```r
predicted_sales <- predict(mlr4_nonzero, test_sales[,c(3,10,14,18,19)])
```

```
## Warning in predict.lm(mlr4_nonzero, test_sales[, c(3, 10, 14, 18, 19)]):
## prediction from a rank-deficient fit may be misleading
```

```r
# rmse on test data
sqrt(sum((predicted_sales - test_sales$Sales)^2) / length(predicted_sales))
```

```
## [1] 2.951015
```

```r
# rmse on train data
sqrt(sum((mlr5_nonzero$fitted.values - train_sales$Sales)^2) / length(mlr5_nonzero$fitted.values))
```

```
## [1] 2.886106
```

```r
predicted_sales <- predict(mlr5_nonzero, test_sales[,c(3, 10, 16, 13, 19)])

# rmse on test data
sqrt(sum((predicted_sales - test_sales$Sales)^2) / length(predicted_sales))
```

```
## [1] 2.929614
```

```r
# rmse on train data
sqrt(sum((mlr6_nonzero$fitted.values - train_sales$Sales)^2) / length(mlr6_nonzero$fitted.values))
```

```
## [1] 2.875492
```

```
predicted_sales <- predict(mlr6_nonzero, test_sales[,c(3, 10, 19)])

# rmse on test data
sqrt(sum((predicted_sales - test_sales$Sales)^2) / length(predicted_sales))
```

```
## [1] 2.986754
```

Create dataframe for model assessment:

```
df <- data.frame(AIC = c(AIC(mlr3_nonzero), AIC(mlr4_nonzero),
                         AIC(mlr5_nonzero), AIC(mlr6_nonzero)),
                 BIC = c(BIC(mlr3_nonzero), BIC(mlr4_nonzero),
                         BIC(mlr5_nonzero), BIC(mlr6_nonzero)),
                 AdjRsq = c(0.3097, 0.2877, 0.2795, 0.1949),
                 trainRMSE = c(2.890188, 2.886228, 2.886106, 2.875492),
                 testRMSE = c(2.837222, 2.951015, 2.929614, 2.986754))
row.names(df) <- c('MLR3_nonzero', 'MLR4_nonzero',
                   'MLR5_nonzero', 'MLR6_nonzero')
df
```

```
##                   AIC      BIC AdjRsq trainRMSE testRMSE
## MLR3_nonzero 55886.36 56344.61 0.3097  2.890188 2.837222
## MLR4_nonzero 56472.16 57225.55 0.2877  2.886228 2.951015
## MLR5_nonzero 56671.03 57432.19 0.2795  2.886106 2.929614
## MLR6_nonzero 58565.85 58993.03 0.1949  2.875492 2.986754
```