SAP projekt - Analiza poadataka o igračima NHL lige

Zvonimir Petar Rezo, Dominik Ćurić i Ivan Sabolić

Uvod

Sportska industrija jedna je od najvećih industrija na svijetu sa procjenom vrijednosti od oko 600 milijardi dolara. Jedan od sportova koji pripadaaju sportskoj industriji te jedan od četiri glavna sporta u Sjedinjenim Američkim Državama je i hokej.Nacionalna Hokejaška Liga(NHL) najbolja je heokejaška liga na svijetu sa procjenom vrijednosti oko 18.4 milijarde dolara. NHL je osnovana 1917. godine i obuhvaća 24 američke momčadi i 7 kanadskih momčadi koje se svake sezone natječu za prestižni Stanley kup.

Podatkovni skup : NHL 2016-2017

Podatkovni skup koji imamo na raspolaganju vezan nam je uz sezonu 2016-2017.Sadrži nam podatke za sve igrače koji su nastupili u toj sezoni te njihove karakteristike poput:visine,težine,pozicije na terenu,godišnje plaće itd.

Učitavanje podatkovnog skupa

```
AllSits <- read.csv2(file.choose())
fiveOnFive <- read.csv2(file.choose())
fiveOnFour <- read.csv2(file.choose())
fourOnFive <- read.csv2(file.choose())
Goalies <- read.csv2(file.choose())
fiveVsfive <- read.csv2(file.choose())
```

Prikat ćemo naše podatake i upoznati se s nekim karakteristikama naših podataka

```
head(AllSits)
```

```
##         Born       City Pr.St Cntry Nat Ht  Wt DftYr DftRd Ovrl Hand   NHLid
## 1 1988-04-30   Hamilton    ON   CAN CAN 69 170    NA    NA   NA    R 8476805
## 2 1987-02-25   Muskegon    MI   USA USA 74 218  2005     2   42    L 8471716
## 3 1993-09-23  Stockholm        SWE SWE 71 196  2012     2   37    R 8476857
## 4 1991-12-01   Johnston    RI   USA USA 70 208    NA    NA   NA    R 8478569
## 5 1992-04-30 Morristown    NJ   USA USA 72 202  2010     5  140    L 8475844
## 6 1997-07-26      Rauma         FIN FIN 71 172  2015     2   35    L 8478427
##    Last.Name First.Name Position Team GP  G  A A1 A2 PTS X... E... PIM Shifts
## 1     Abbott    Spencer       LW  CHI  1  0  0  0  0   0    0  0.1   0     12
## 2 Abdelkader     Justin    LW/RW  DET 64  7 14  9  5  21  -20 -3.7  50   1397
## 3      Aberg     Pontus       LW  NSH 15  1  1  0  1   2   -2 -1.2   4    256
## 4    Acciari       Noel        C  BOS 29  2  3  1  2   5    3  0.5  16    431
## 5   Agostino      Kenny       LW  STL  7  1  2  2  0   3    0  1.9   2    117
## 6        Aho  Sebastian    RW/LW  CAR 82 24 25 12 13  49   -1  5.1  26   1814
##     TOI  TOI.1 TOI.GP TOI.GP.1  TOI.  IPP.  SH.   SV.  PDO F.60 A.60  Pct. Diff
## 1   514    8.6   8.57     8.57 16,7%  0,0% 0,0% 1,000 1000 0,00 0,00  0,0%    0
## 2 63969 1040.9  16.65    16.52 28,3% 61,8% 6,8%  ,909  977 1,96 2,59 43,0%  -11
## 3 11102  185.0  12.33    12.33 24,9% 40,0% 6,5%  ,921  986 1,62 2,27 41,7%   -2
```

```
## 4 18047  300.5  10.23    10.32 20,8% 55,6% 7,6%  ,959 1035 1,80 1,20 60,0%    3
## 5  5366   89.4  12.78    12.78 24,8% 75,0% 7,0%  ,912  982 2,69 2,01 57,1%    1
## 6 82591 1357.5  16.78    16.71 29,4% 69,0% 9,4%  ,911 1005 3,14 2,21 58,7%   21
##   Diff.60 iCF iCF.1 iFF iSF iSF.1 iSF.2  ixG iSCF iRB iRS     iDS sDist sDist.1
## 1    0,00   2     2   1   1     1     1  0.1    0   0   0       0  15.0    14.1
## 2   -0,63 158   160 138 102   104   104 13.0   49   7  10      17  26.8    25.0
## 3   -0,65  24    24  15  12    12    12  0.5    0   0   1       1  32.9    31.8
## 4    0,60  NA    43  NA  NA    24    24  2.6   10  NA   2 #VALUE!  23.8      NA
## 5    0,67  26    26  24  17    17    17  1.4    4   1   1       2  31.4    32.7
## 6    0,93 360   362 285 213   214   214 24.4   77  25  26      51  26.0    25.7
##    Pass iHF iHF.1 iHA iHDf iMiss iGVA iTKA iBLK iGVA.1 iTKA.1 iBLK.1 BLK. iFOW
## 1   0,0   0     0   3   -3     0    0    1    0      0      1       0 0,0%    0
## 2 237,7 144   140  74   66    36   18   15   27     18     14      25 2,8%    5
## 3   0,0   9     9  17   -8     3    2    5    2      2      5       2 1,3%    1
## 4  25,6  80    80  59   21     8    4   10    5      4     10       5 2,0%   11
## 5  38,5   9     9   7    2     7    4    1    3      4      1       3 4,6%    0
## 6 215,3  51    49  91  -42    71   57   40   14     56     40      14 1,4%   17
##   iFOL iFOW.1 iFOL.1    FO. X.FOT dzFOW dzFOL nzFOW nzFOL ozFOW ozFOL FOW.Up
## 1    0      0      0   0,0%  0,0%     0     0     0     0     0     0      0
## 2   12      5     12  29,4%  1,8%     0     0     1     4     4     8      1
## 3    0      1      0 100,0%  0,5%     1     0     0     0     0     0      1
## 4   18     11     18  37,9% 10,5%     3     6     2     3     6     9      7
## 5    0      0      0   0,0%  0,0%     0     0     0     0     0     0      0
## 6   18     17     17  48,6%  2,7%     1     0    11    10     5     8      2
##   FOL.Up FOW.Down FOL.Down FOW.Close FOL.Close OTG X1G GWG ENG PSG PSA G.Bkhd
## 1      0        0        0         0         0   0   0   0   0   0   0      0
## 2      2        2        3         2        10   1   5   1   0   0   0      1
## 3      0        0        0         1         0   0   0   0   0   0   0      0
## 4      3        1        6         6        12   0   1   0   0   0   0      0
## 5      0        0        0         0         0   0   0   0   0   0   0      0
## 6      3       10        2         9        14   1   5   4   2   0   0      1
##   G.Dflct G.Slap G.Snap G.Tip G.Wrap G.Wrst CBar Post Over Wide S.Bkhd S.Dflct
## 1       0      0      0     0      0      0    0    0    0    0      0       0
## 2       1      1      1     0      0      3    0    0    1   35      6       6
## 3       0      0      1     0      0      0    0    1    0    2      0       1
## 4       0      0      0     0      0      2    0    0    0    8      1       0
## 5       0      0      1     0      0      0    0    0    1    6      4       0
## 6       2      4      7     1      0      9    0    2    3   66     26       8
##   S.Slap S.Snap S.Tip S.Wrap S.Wrst iPenT iPenD iPENT iPEND iPenDf  NPD Min Maj
## 1      0      0     1      0      0     0     0     0     0      0  0.0   0   0
## 2     16     11    19      1     45    22    32    20    28     10  8.6  20   2
## 3      0      1     1      1      8     2     3     2     2      1  0.0   2   0
## 4      0      5     2      1     14     4     5     4     2      1 -2.0   3   0
## 5      2      3     0      0      8     1     1     1     1      0  0.1   1   0
## 6     17     47    12      7     97    13    22    12    20      9  7.2  13   0
##   Match Misc Game   CF   CA   FF  FA  SF  SA  xGF  xGA SCF SCA GF GA RBF RBA
## 1     0    0    0   11    7    6   6   5   4  0.2  0.4   0   1  0  0   0   0
## 2     0    0    0  916  895  697 671 498 495 48.5 44.3 168 151 34 45  24  31
## 3     0    0    0  142  159  106 117  77  89  5.6  6.0  18  20  5  7   4   4
## 4     0    1    0  232  251  172 193 118 146 10.3  9.3  31  24  9  6   8   3
## 5     0    0    0   95   65   77  48  57  34  5.2  3.3  20  10  4  3   3   4
## 6     0    0    0 1443 1018 1070 798 757 561 80.1 50.4 283 167 71 50  82  43
##   RSF RSA DSF DSA FOW FOL  HF  HA GVA TKA PENT PEND  OPS DPS   PS SOS SOG SOGDG
## 1   0   1   0   1   4   2   2   9   1   1    0    1  0.0 0.0  0.0  NA  NA    NA
```

```
## 2  37  45  61  76 500 471 361 367 143  81   64   74  0.1 0.5  0.7  NA  NA    NA
## 3  10   9  14  13  97  92  68  72  28  23   13   13 -0.2 0.1 -0.1  NA  NA    NA
## 4  10  13  18  16 144 132 184 177  28  45   29   23  0.0 0.5  0.5  NA  NA    NA
## 5   3   4   6   8  29  38  30  39  14   8    5    3  0.2 0.1  0.3  NA  NA    NA
## 6  68  53 150  96 705 558 373 493 236 203   53   66  4.2 1.8  5.9   5   2     0
##      OTOI Grit DAP  Pace   GS GS.G X1st X2nd X3rd MGL              Injuries
## 1   42.63    0   0 126,0  0.3 0.28   NA   NA   NA  NA
## 2 2641.98  223   8 104,4 16.1 0.26    0    0    1  18       Lower body, Knee
## 3  558.17   15   7  97,6  1.0 0.06    0    0    1  NA
## 4 1141.29  101  30  96,4  4.8 0.17    0    0    1  15 Lower body, Upper body
## 5  271.29   14  10 107,4  4.6 0.66    0    1    0  NA
## 6 3267.64   91   7 108,8 57.6 0.71    5    1    4  NA
##          CHIP NMC Status        Salary       Cap.Hit
## 1                    UFA   $575.000,00   $575.000,00
## 2 $932.926,83 NTC    UFA $5.500.000,00 $4.250.000,00
## 3                    RFA   $842.500,00   $780.833,00
## 4 $144.969,51        RFA   $892.500,00   $792.500,00
## 5                    UFA   $625.000,00   $625.000,00
## 6                    RFA   $925.000,00   $925.000,00
```

```
head(fiveOnFive)
```

```
##    Last.Name First.Name Position Team GP  G  A A1 PTS PTS.60  IPP.   TOI  TOI.1
## 1     Abbott    Spencer       LW  CHI  1  0  0  0   0   0.00  0,0%   514    8.6
## 2 Abdelkader     Justin    LW/RW  DET 64  2 11  6  13   0.96 78,8% 50402  778.2
## 3      Aberg     Pontus       LW  NSH 15  1  1  0   2   0.67 40,0% 10863  177.5
## 4    Acciari       Noel        C  BOS 29  2  3  1   5   1.03 55,6% 17648  286.0
## 5   Agostino      Kenny       LW  STL  7  1  1  1   2   1.50 66,7%  4830   79.6
## 6        Aho  Sebastian    RW/LW  CAR 82 17 13  6  30   1.59 73,2% 70606 1094.8
##    TOI.2  TOI.3  TOI.4 TOI.GP TOI.GP.1  TOI.    ZS. RelZS. TOI.QoT TOI.QoC
## 1    8.6    8.6    8.6   8.57     8.57 16,7% 100,0%  47,4%   30,6%   29,9%
## 2  803.5  794.4  803.8  13.12    12.70 25,9%  44,8%  -7,1%   30,1%   29,1%
## 3  179.5  179.5  179.5  12.08    11.98 25,8%  47,9%  -1,3%   29,9%   28,9%
## 4  292.9  292.4  292.9  10.02    10.04 22,1%  41,2% -13,7%   28,6%   29,0%
## 5   80.2   80.2   80.2  11.50    11.46 23,9%  60,4%  16,7%   30,0%   29,3%
## 6 1098.6 1108.5 1109.3  14.25    13.78 27,7%  53,6%  -1,8%   30,8%   29,5%
##   CF.QoC  SH.   SV.  PDO F.60 A.60  Pct. Diff Diff.60 RelF.60 RelA.60 RelPct.
## 1  46,6% 0,0% 1.000 1000 77,0 49,0 61,1%    4    28,0    -3,2     8,2    3,9%
## 2  50,5% 4,7% 0.916  963 46,9 55,2 46,0% -111    -8,3    -3,1     1,2    0,5%
## 3  49,5% 6,7% 0.921  988 44,8 52,5 46,0%  -23    -7,7   -17,0    -1,6   -4,7%
## 4  50,0% 7,7% 0.957 1035 46,8 49,3 48,7%  -12    -2,5   -16,9    -3,7   -2,8%
## 5  49,5% 5,9% 0.910  969 63,2 47,6 57,0%   21    15,6    15,6   -11,4   12,8%
## 6  50,3% 7,3% 0.917  990 56,1 48,9 53,4%  133     7,2    -1,3    -7,1    4,5%
##   RelDf.60 RelC. RelC.60  RelF. RelF.60.1 QRelCF60 QRelCA60 QRelDFF60 QRelDFA60
## 1     27,8 -5,2%   28.02 -14,9%      0.00     -6.1      7.5     -33.4       3.6
## 2     -2,1 -2,3%   -6.47  -1,9%     -4.50     -1.3      2.1      -0.9       1.8
## 3     -8,0 -7,0%   -6.81  -6,0%     -4.54    -11.1     -0.1      -8.0      -2.4
## 4     -3,6 -5,8%   -2.39  -7,3%     -2.79    -11.7     -1.4      -3.5      -4.9
## 5     16,4 12,4%   14.09  10,4%     14.09     11.1     -7.8       7.8       0.9
## 6      9,2  2,7%    5.62   2,4%      2.88      0.1     -3.8       0.3      -2.2
##   iCF iFF iSF  ixG iSCF iCF.1 iFF.1 iSF.1 iRB iRS     iDS sDist  Pass iHF iHA
## 1   2   1   1  0.1    0     2     1     1   0   0       0  14.1   0.0   0   3
## 2 123 104  74  7.1   21   124   105    75   3   9      12  27.9 230.2 129  60
## 3  22  15  12  0.5    0    22    15    12   0   1       1  31.8   0.0   9  16
## 4  NA  NA  NA  2.6   10    43    32    24  NA   2 #VALUE!    NA  25.6  79  59
```

```
## 5  23  21  15  1.3    4    23    21    15   1   1       2  32.3  28.5   9   7
## 6 275 219 167 19.2   59   273   216   166  23  25      48  24.2 152.2  46  79
##   iHDf iGVA iTKA iBLK BLK. iFOW iFOL iFOW.1 iFOL.1    FO. X.FOT iPENT iPEND
## 1   -3    0    1    0 0,0%    0    0      0      0   0,0%  0,0%     0     0
## 2   69   13   12   18 2,5%    4    7      4      7  36,4%  1,5%     7    22
## 3   -7    2    5    2 1,3%    1    0      1      0 100,0%  0,5%     2     2
## 4   20    3    9    5 2,1%   10   18     10     17  37,0% 10,2%     4     2
## 5    2    4    1    3 4,7%    0    0      0      0   0,0%  0,0%     1     1
## 6  -33   45   35   14 1,5%   17   15     16     13  55,2%  2,8%    11    16
##   iPenDf OZS DZS OZS.1 DZS.1 NZS OTF OZS.2 DZS.2 NZS.1 OZF DZF NZF   CF  CA  FF
## 1      0   3   0     3     0   3   8     3     0     3   4   2   1   11   7   6
## 2     15 215 262   210   261 259 602   215   262   264 291 281 231  637 752 487
## 3      0  57  62    57    62  64 133    57    62    64  71  66  52  137 158 103
## 4     -2  70 100    70   100  96 251    70   100    97 124 127  67  228 240 168
## 5      0  29  19    29    19  18  73    29    19    18  27  30  21   85  64  68
## 6      5 343 297   341   293 397 794   335   292   395 421 338 264 1055 926 789
##    FA CF.1 CA.1 FF.1 FA.1  SF  SA SCF SCA  xGF  xGA GF GA RBF RBA RSF RSA DSF
## 1   6   11    7    6    6   5   4   0   1  0.2  0.4  0  0   0   0   0   1   0
## 2 567  626  733  480  552 348 412  89 117 28.0 34.9 16 35  12  27  31  38  43
## 3 117  137  158  103  117  75  89  17  20  5.4  6.0  5  7   4   4  10   9  14
## 4 182  230  242  170  184 116 141  31  23 10.2  8.6  9  6   8   3  10  11  18
## 5  47   85   64   68   47  51  33  18  10  4.5  3.2  3  3   2   4   3   4   5
## 6 723 1039  913  779  713 563 507 179 147 52.5 44.1 41 42  58  39  60  47 118
##   DSA CF.2 CA.2 FF.2 FA.2 SF.1 SA.1 GF.1 GA.1 CF.3 CA.3 FF.3 FA.3 CF.4 CA.4
## 1   1   12    7    7    6    6    4    0    0   11    7    6    6   11    7
## 2  65  640  751  490  566  353  422   17   35  685  790  523  593  619  727
## 3  13  134  156  100  116   74   89    5    7  138  159  104  117  126  155
## 4  14  231  239  171  182  117  140    9    6  230  246  170  188  223  237
## 5   8   86   65   68   48   51   34    3    3   86   64   69   47   84   61
## 6  86 1041  910  781  710  564  504   41   42 1126  990  842  774 1043  893
##   SF.2 SA.2 xGF.1 xGA.1 GF.2 GA.2 RBF.1 RBA.1 CF.5 CA.5   DFF   DFA C.Up C.Down
## 1    5    4   0.3   0.2    0    0     2     2   10    7   2.9   4.3    0      0
## 2  335  401  27.3  34.4   17   34    87   112  622  735 462.5 557.5  -67    -25
## 3   67   89   5.4   6.6    5    7    14    22  136  158  95.5 101.1  -33     -1
## 4  113  138   9.9  10.3    9    6    43    26  226  241 167.3 140.5  -42     34
## 5   50   30   3.6   2.7    3    3    19    15   82   64  63.5  43.7   14      2
## 6  557  494  46.4  40.0   41   42   188   149 1044  914 786.9 722.4   13     79
##   C.Tied C.Close F.Up F.Down F.Tied F.Close FOW FOL  HF  HA GVA TKA PENT PEND
## 1      4       5    0     -1      1       1   4   2   2   9   1   1    0    1
## 2    -23     -83  -46    -24    -10     -61 366 364 325 327 105  65   41   56
## 3     13     -16  -27      2     11      -8  94  89  68  69  25  23   11   12
## 4     -4       5  -32     23     -5       2 141 125 181 177  27  44   28   16
## 5      5      21   14      2      5      19  29  37  30  37  12   8    4    3
## 6     37      92    5     50     11      42 570 461 348 452 186 177   43   59
##     OTOI  OCF  OCA  OFF  OFA  OSF  OSA   OxGF   OxGA OSCF OSCA OGF OGA OOZS
## 1   42.6   57   29   37   20   29   12   3.13   0.92   12    2   3   1   10
## 2 2222.1 1854 1999 1421 1542 1009 1096  84.91  89.78  278  281  83  79  723
## 3  511.2  527  461  391  351  279  272  22.96  20.04   74   64  25  13  152
## 4 1005.2 1068  888  791  639  567  459  44.17  39.22  133  135  34  34  368
## 5  253.3  201  249  157  163  112  115  10.11   7.94   38   28  11   4   66
## 6 2854.5 2729 2666 1977 1996 1380 1408 125.73 125.87  422  435 100 123  932
##   ODZS ONZS TMCF60 TMCA60 TMFF60 TMFA60 TMSF60 TMSA60 TMGF60 TMGA60 TMGF TMGA
## 1    9   11   55.4   54.6   39.6   40.1   29.5   30.4   2.36   2.01  0.3  0.1
## 2  668  749   50.5   53.3   38.7   40.8   27.8   28.9   2.36   2.40 33.1 31.1
```

```
## 3  157  199   58.3   54.4   43.0   40.9   31.4   29.0   2.51   2.10  6.9  2.6
## 4  303  337   57.7   51.7   43.3   37.7   31.0   26.8   1.87   2.35  6.2 11.0
## 5   85   88   51.7   51.9   38.2   38.4   27.5   27.6   2.39   2.01  5.3  1.8
## 6  749  876   58.3   53.6   42.4   40.3   29.6   28.7   2.29   2.59 43.2 58.0
##   TmxGF TmxGA TOI.QoT.1 CF.QoT xGF.QoT OppCF60 OppCA60 OppFF60 OppFA60 OppSF60
## 1   0.7   0.2      30.6   50.5    49.1    51.4    59.0    39.4    44.7    28.7
## 2  31.4  30.9      30.1   48.9    49.1    55.7    54.6    41.4    40.7    29.8
## 3   9.2   7.4      29.9   51.3    52.4    53.8    54.9    39.8    40.6    28.7
## 4   9.9  16.9      28.6   52.6    53.1    55.3    55.2    41.3    41.4    29.7
## 5   3.0   2.5      30.0   49.6    50.8    55.3    56.3    40.9    42.1    29.4
## 6  45.9  42.2      30.8   51.9    51.7    55.9    55.2    41.6    40.9    30.0
##   OppSA60 OppGF60 OppGA60 CF.QoC.1 DFF.QoC TOI.QoC.1 CF.QoC.2 xGF.QoC NSPF
## 1    33.1    1.90    2.24     -5.9   -26.7      29.9     46.9    48.1  0.0
## 2    29.4    2.30    2.27     -1.6    -1.7      29.1     50.3    50.2 -2.9
## 3    29.3    2.11    2.32     -5.4    -3.8      28.9     49.6    49.4 -0.8
## 4    29.9    2.33    2.27     -4.9     1.6      29.0     50.1    50.0 -0.9
## 5    30.4    2.17    2.32      8.8     3.7      29.3     49.5    49.3  0.0
## 6    29.5    2.34    2.25      1.9     1.5      29.5     50.1    50.2 -1.2
##    NGPF ozFO ozSFPF ozSAPF ozGFPF ozGAPF Exp.ozNSPF Exp.ozNGPF nzFO nzSFPF
## 1  0.00    0      0      0      0      0       0.00       0.00    0      0
## 2 -0.06    8      0      1      0      0       1.94       0.06    3      0
## 3  0.01    0      0      0      0      0       0.00       0.00    0      0
## 4 -0.06   15      2      0      0      0       3.64       0.11    5      1
## 5  0.00    0      0      0      0      0       0.00       0.00    0      0
## 6 -0.06   10      3      0      0      0       2.43       0.07   18      0
##   nzSAPF nzGFPF nzGAPF dzFO dzSFPF dzSAPF dzGFPF dzGAPF Exp.dzNSPF Exp.dzNGPF
## 1      0      0      0    0      0      0      0      0       0.00       0.00
## 2      0      0      0    0      0      0      0      0       0.00       0.00
## 3      0      0      0    1      0      1      0      0      -0.24      -0.01
## 4      0      0      0    7      0      2      0      0      -1.70      -0.05
## 5      0      0      0    0      0      0      0      0       0.00       0.00
## 6      2      0      0    1      0      0      0      0      -0.24      -0.01
##   FOvsL FO.vsL FOvsR FO.vsR OpFO OpFOW HopFO HopFOW RopFO RopFOW  Pace
## 1    NA     NA    NA     NA   NA    NA    NA     NA    NA     NA 129,5
## 2    10   50.0     7    0.0   NA    NA    NA     NA    NA     NA 104,3
## 3     1  100.0    NA     NA   NA    NA    NA     NA    NA     NA  98,3
## 4    24   41.7     5   20.0   NA    NA    NA     NA    NA     NA  97,7
## 5    NA     NA    NA     NA   NA    NA    NA     NA    NA     NA 112,7
## 6    26   50.0     9   44.4    1     1     1      1     0      0 106,8
```

```
head(fiveOnFour)
```

```
##    Last.Name First.Name Position Team GP G  A A1 PTS PTS.60   IPP.   TOI TOI.1
## 1     Abbott    Spencer       LW  CHI  1 0  0 NA   0   0.00   0,0%     0    NA
## 2 Abdelkader     Justin    LW/RW  DET 64 5  2  0   7   2.93  56,0%  8902 141.3
## 3      Aberg     Pontus       LW  NSH 15 0  0  0   0   0.00   0,0%   233   5.1
## 4    Acciari       Noel        C  BOS 29 0  0 NA   0   0.00   0,0%    49   5.4
## 5   Agostino      Kenny       LW  STL  7 0  1  1   1   6.60 100,0%   536   9.4
## 6        Aho  Sebastian    RW/LW  CAR 82 6 11  4  17   5.72  82,9% 10911 179.4
##   TOI.2 TOI.GP  TOI.    ZS. RelZS. TOI.QoT TOI.QoC CF.QoC   SH.   SV.  PDO
## 1    NA   0.00  0,0%   0,0%   0,0%    0,0%    0,0%   0,0%  0,0% 0.000    0
## 2 139.7   2.24 45,5%  96,5%   4,9%   29,9%   30,2%  11,8% 11,3% 0.829  942
## 3   3.7   0.28 20,1% 100,0%   8,3%   30,2%   31,8%  11,5%  0,0% 0.000    0
## 4    NA   0.11  8,3% 100,0%   4,7%   28,8%   30,1%   0,0%  0,0% 0.000    0
## 5   8.9   1.30 37,3% 100,0%   0,0%   32,2%   29,5%  11,8% 16,7% 1.000 1167
```

```
## 6 174.0   2.18 47,7%  95,9%  -2,2%   30,7%   29,8%  11,4% 13,9% 0.909 1048
##    F.60 A.60   Pct. Diff Diff.60 RelF.60 RelA.60 RelPct. RelDf.60 iCF iFF iSF
## 1   0.0  0.0   0,0%    0     0.0     0.0     0.0    0,0%      0.0  NA  NA  NA
## 2  94.2  9.0  91,3%  200    85.2    25.0    -2.7    5,8%     27.8  25  25  22
## 3  28.1  0.0 100,0%    2    28.1   -46.6    -6.0    7,4%    -40.7   2   0   0
## 4   0.0  0.0   0,0%    0     0.0   -91.1   -15.0  -85,8%    -76.1   0   0   0
## 5  65.4  6.5  90,9%    9    58.9   -14.2     2.8   -4,5%    -16.9   3   3   2
## 6 109.9  8.5  92,8%  299   101.4    36.1     0.9    2,2%     35.2  62  46  33
##   iSF.1 ixG iSCF iCF.1 iFF.1 iSF.2 iRB iRS iDS sDist Pass iHF iHF.1 iHA iHDf
## 1     0  NA   NA    NA    NA    NA  NA  NA   0    NA  0.0   0    NA  NA    0
## 2    24 4.7   22    26    26    23   3   0   3  13.4  0.0   7     7  10   -3
## 3     0 0.0    0     2     0     0   0   0   0    NA  0.0   0     0   1   -1
## 4     0 0.0    0    NA    NA    NA   0   0   0    NA  0.0   0     0   0    0
## 5     2 0.1    0     3     3     2   0   0   0  35.8 10.0   0     0   0    0
## 6    34 3.3   11    62    47    34   1   0   1  32.7 63.1   2     2   9   -7
##   iMiss iGVA iTKA iGVA.1 iTKA.1 iBLK BLK. iFOW iFOL iFOW.1 iFOL.1   FO. X.FOT
## 1     0    0    0     NA     NA   NA 0,0%    0    0     NA     NA  0,0%  0,0%
## 2     5    3    0      3      0    0 0,0%    1    3      1      3 25,0%  2,9%
## 3     0    0    0      0      0    0 0,0%    0    0      0      0  0,0%  0,0%
## 4     0    0    0      0      0    0 0,0%    0    0      0      0  0,0%  0,0%
## 5     1    0    0      0      0    0 0,0%    0    0      0      0  0,0%  0,0%
## 6    13   10    2      9      2    0 0,0%    0    2      0      2  0,0%  1,1%
##   iPENT iPEND iPenDf OZS DZS NZS OTF OZS.1 DZS.1 NZS.1 OZF DZF NZF  CF CA  FF
## 1    NA    NA      0  NA  NA  NA  NA    NA    NA    NA  NA  NA  NA  NA NA  NA
## 2    10     2     -8 110   4  23  43   113     4    25  59  15  27 219 20 163
## 3     0     0      0   3   0   2   4     3     0     2   1   1   2   2  0   0
## 4     0     0      0   1   0   3  14    NA    NA    NA   3   0   0   0  0   0
## 5     0     0      0   1   0   0  11     1     0     0   3   0   3  10  1   9
## 6     0     0      0 141   6  29  49   141     6    29  60  11  43 321 25 227
##   FA  SF SA  xGF xGA SCF SCA GF GA RBF RBA RSF RSA DSF DSA CF.1 CA.1 FF.1 FA.1
## 1 NA  NA NA   NA  NA  NA  NA NA NA  NA  NA  NA  NA   0   0    0    0    0    0
## 2 17 109 14 16.1 2.0  62   8 12  2   8   0   4   5  12   5  222   22  166   18
## 3  0   0  0  0.0 0.0   0   0  0  0   0   0   0   0   0   0    2    0    0    0
## 4  0   0  0  0.0 0.0   0   0  0  0   0   0   0   0   0   0    0    0    0    0
## 5  1   6  1  0.7 0.0   2   0  1  0   1   0   0   0   1   0   10    1    9    1
## 6 20 146 11 22.8 2.2  88  10 21  1  20   1   3   2  23   3  326   25  233   20
##   SF.1 SA.1 GF.1 GA.1 FOW FOL HF HA GVA TKA PENT PEND  OTOI OCF OCA OFF OFA OSF
## 1    0    0    0    0  NA  NA NA NA  NA  NA   NA   NA    NA  NA  NA  NA  NA  NA
## 2  112   15   13    3  85  52 13 26  28   7   15    7 169.2 195  33 134  28  89
## 3    0    0    0    0   2   3  0  2   2   0    1    0  20.1  25   2  20   2  14
## 4    0    0    0    0   2   2  0  0   0   0    1    1  60.0  91  15  71  13  50
## 5    6    1    1    0   0   1  0  2   2   0    1    0  15.8  21   1  15   1  10
## 6  150   11   20    1 110  66 15 34  35  12    5    1 196.7 242  25 184  21 135
##   OSA  OxGF OxGA OSCF OSCA OGF OGA OOZS ODZS ONZS TMCF60 TMCA60 TMFF60 TMFA60
## 1  NA    NA   NA   NA   NA  NA  NA   NA   NA   NA     NA     NA     NA     NA
## 2  24 12.06 1.82   48    5  11   5  110   10   28   82.8   13.3   59.4   11.3
## 3   2  2.03 0.31    5    1   1   0   11    1    3   94.9   11.4   69.3   10.0
## 4   9  5.15 0.95   18    2   2   0   41    2   14     NA     NA     NA     NA
## 5   0  0.99 0.01    5    0   0   0   12    0    2   89.8   10.9   67.5    9.2
## 6  16 17.13 1.75   57    6  11   3  153    3   36  107.2    8.0   77.7    6.3
##   TMSF60 TMSA60 TMGF60 TMGA60 TOI.QoT.1 CF.QoT xGF.QoT OppCF60 OppCA60 OppFF60
## 1     NA     NA     NA     NA        NA     NA      NA      NA      NA      NA
## 2   40.5    9.4   4.48   1.36      29.9   49.8    50.4    12.7    94.8    11.0
## 3   49.1    7.8   6.13   0.57      30.2   52.4    52.6    12.4    94.7    10.9
```

```
## 4     NA     NA     NA     NA      28.8   51.8    52.6      NA      NA      NA
## 5   50.9    6.6   6.60   0.56      32.2   50.8    51.4    13.1    98.2    11.1
## 6   50.6    4.0   6.46   0.51      30.7   53.2    52.9    12.6    97.9    10.8
##   OppFA60 OppSF60 OppSA60 OppGF60 OppGA60 TOI.QoC.1 CF.QoC.1 xGF.QoC  Pace
## 1      NA      NA      NA      NA      NA        NA       NA      NA   0.0
## 2    71.2     8.9    51.0    0.82    6.25      30.2     49.7    49.9 103.1
## 3    70.1     8.5    50.7    0.82    5.80      31.8     52.1    52.0  28.1
## 4      NA      NA      NA      NA      NA      30.1     49.3    49.6   0.0
## 5    74.3     8.7    53.2    0.78    6.51      29.5     49.1    49.5  72.0
## 6    73.4     8.6    52.3    0.87    6.62      29.8     49.3    49.8 118.4
```

```
head(fourOnFive)
```

```
##    Last.Name First.Name Position Team GP G A A1 PTS  TOI TOI.1 TOI.2 TOI.GP
## 1     Abbott    Spencer       LW  CHI  1 0 0 NA   0    0    NA    NA   0.00
## 2 Abdelkader     Justin    LW/RW  DET 64 0 1  1   1 4665  81.8  77.8   1.24
## 3      Aberg     Pontus       LW  NSH 15 0 0 NA   0    6   0.7    NA   0.03
## 4    Acciari       Noel        C  BOS 29 0 0  0   0  350   7.3   5.2   0.21
## 5   Agostino      Kenny       LW  STL  7 0 0 NA   0    0   0.3    NA   0.02
## 6        Aho  Sebastian    RW/LW  CAR 82 1 1  0   2 1074  25.0  17.0   0.24
##    TOI.   ZS. RelZS. TOI.QoT TOI.QoC CF.QoC   SH.   SV.  PDO F.60 A.60  Pct.
## 1  0,0%  0,0%   0,0%    0,0%    0,0%   0,0%  0,0% 0.000    0  0.0  0.0  0,0%
## 2 31,6%  8,1%  -3,4%   30,6%   29,4%  88,1% 22,2% 0.915 1137 11.3 86.6 11,5%
## 3  3,5%  0,0% -11,1% #VALUE! #VALUE!   0,0%  0,0% 0.000    0  0.0  0.0  0,0%
## 4 10,4% 25,0%  21,0%   29,2%   30,8%  87,5%  0,0% 1.000 1000  9.9 49.4 16,7%
## 5 13,1%  0,0%   0,0%   25,9%   30,0%   0,0%  0,0% 0.000    0  0.0  0.0  0,0%
## 6 17,0%  0,0%  -4,7%   31,4%   29,2%  87,7% 33,3% 0.826 1159 17.8 55.2 24,4%
##   Diff Diff.60 RelF.60 RelA.60 RelPct. RelDf.60 iCF iFF iSF iSF.1 ixG iSCF
## 1    0     0.0     0.0     0.0    0,0%      0.0  NA  NA   0    NA  NA   NA
## 2 -100   -75.3    -2.9    -7.8   -1,6%      4.9   3   2   2     2 0.3    1
## 3    0     0.0   -12.7   -92.1  -12,1%     79.4   0   0   0     0 0.0    0
## 4   -4   -39.5     1.3   -38.7    7,8%     39.9   0   0   0     0 0.0    0
## 5    0     0.0     0.0   -82.2    0,0%     82.2   0   0   0     0 0.0    0
## 6  -13   -37.4     8.5   -16.3   12,9%     24.8   3   3   3     3 0.4    2
##   iCF.1 iFF.1 iSF.2 iRB iRS iDS sDist Pass iHF iHF.1 iHA iHDf iMiss iGVA iTKA
## 1    NA    NA    NA  NA  NA   0    NA  0,0   0    NA  NA    0     0    0    0
## 2     3     2     2   0   1   1  32.7  7,5   3     3   1    2     0    2    2
## 3    NA    NA    NA   0   0   0    NA  0,0   0     0   0    0     0    0    0
## 4     0     0     0   0   0   0    NA  0,0   0     0   0    0     0    1    1
## 5    NA    NA    NA   0   0   0    NA  0,0   0     0   0    0     0    0    0
## 6     3     3     3   0   1   1  21.2  0,0   0     0   0    0     0    0    0
##   iBLK iGVA.1 iTKA.1 iBLK.1 BLK. iFOW iFOL iFOW.1 iFOL.1    FO. X.FOT iPENT
## 1    0     NA     NA     NA 0,0%    0    0     NA     NA   0,0%  0,0%    NA
## 2    5      2      2      5 4,6%    0    2      0      2   0,0%  3,0%     1
## 3    0      0      0      0 0,0%    0    0      0      0   0,0%  0,0%     0
## 4    0      1      1      0 0,0%    1    0      1      0 100,0% 25,0%     0
## 5    0      0      0      0 0,0%    0    0      0      0   0,0%  0,0%     0
## 6    0      0      0      0 0,0%    0    1      0      1   0,0% 20,0%     1
##   iPEND iPenDf OZS DZS NZS OTF OZS.1 DZS.1 NZS.1 OZF DZF NZF CF  CA FF FA SF SA
## 1    NA      0  NA  NA  NA  NA    NA    NA    NA  NA  NA  NA NA  NA NA NA NA NA
## 2     1      0   4  43  19  86     4    48    20   5  33  20 15 109 10 79  9 55
## 3     0      0   0   0   0   3    NA    NA    NA   0   1   0  0   0  0  0  0  0
## 4     0      0   1   3   0  13     1     3     0   3   7   2  1   5  1  5  1  3
## 5     0      0   0   0   0   1    NA    NA    NA   0   0   0  0   0  0  0  0  0
## 6     1      0   0   3   2  52     0     3     2   1   4   7  6  18  6 16  6 11
```

```
##   xGF xGA SCF SCA GF GA RBF RBA RSF RSA DSF DSA CF.1 CA.1 FF.1 FA.1 SF.1 SA.1
## 1  NA  NA  NA  NA NA NA  NA  NA  NA  NA   0   0    0    0    0    0    0    0
## 2 0.9 6.1   3  22  2  5   1   4   1   0   2   4   15  121   10   88    9   63
## 3 0.0 0.0   0   0  0  0   0   0   0   0   0   0    0    0    0    0    0    0
## 4 0.0 0.6   0   1  0  0   0   0   0   1   0   1    1    5    1    5    1    3
## 5 0.0 0.0   0   0  0  0   0   0   0   0   0   0    0    0    0    0    0    0
## 6 0.8 1.0   3   3  2  2   1   0   2   1   3   1    6   19    6   17    6   12
##   GF.1 GA.1 FOW FOL HF HA GVA TKA PENT PEND  OTOI OCF OCA OFF OFA OSF OSA OxGF
## 1    0    0  NA  NA NA NA  NA  NA   NA   NA    NA  NA  NA  NA  NA  NA  NA   NA
## 2    2    5  30  36 16  9   8   5    2    6 177.2  42 279  37 209  26 152 2.05
## 3    0    0   0   0  0  0   0   0    0    0  18.9   4  29   3  23   3  18 0.15
## 4    0    0   1   3  2  0   1   1    0    5  62.7   9  92   9  71   8  44 0.84
## 5    0    0   0   0  0  0   0   0    0    0   2.2   0   3   0   1   0   1 0.00
## 6    2    2   3   2  2  1   0   4    1    2 122.5  19 146  17 100  14  60 0.97
##    OxGA OSCF OSCA OGF OGA OOZS ODZS ONZS TMCF60 TMCA60 TMFF60 TMFA60 TMSF60
## 1    NA   NA   NA  NA  NA   NA   NA   NA     NA     NA     NA     NA     NA
## 2 18.12    4   66   0  10   19  146   30   14.8  102.1   12.0   77.3    7.7
## 3  1.64    1    6   1   0    2   16    5     NA     NA     NA     NA     NA
## 4  5.87    4   19   1   4    2   48    7   16.4   89.8   15.0   67.2   12.7
## 5  0.20    0    1   0   0    0    2    0     NA     NA     NA     NA     NA
## 6  9.78    4   36   1   5    4   81   22   17.1   74.0   14.7   53.3   12.3
##   TMSA60 TMGF60 TMGA60 TOI.QoT.1 CF.QoT xGF.QoT OppCF60 OppCA60 OppFF60 OppFA60
## 1     NA     NA     NA        NA     NA      NA      NA      NA      NA      NA
## 2   57.1   0.42   7.25      30.6   49.0    49.1    95.3    12.8    71.8    11.1
## 3     NA     NA     NA        NA     NA      NA      NA      NA      NA      NA
## 4   45.1   0.77   5.10      29.2   52.3    53.3    91.0    13.1    69.9    10.7
## 5     NA     NA     NA      25.9   50.9    55.3      NA      NA      NA      NA
## 6   36.1   1.56   5.02      31.4   51.3    50.7    92.2    13.0    69.6    10.7
##   OppSF60 OppSA60 OppGF60 OppGA60 TOI.QoC.1 CF.QoC.1 xGF.QoC Pace
## 1      NA      NA      NA      NA        NA       NA      NA  0.0
## 2    50.7     8.8    6.81    0.89      29.4     50.7    50.1 97.9
## 3      NA      NA      NA      NA        NA       NA      NA  0.0
## 4    50.6     9.0    6.15    0.68      30.8     50.8    52.6 59.2
## 5      NA      NA      NA      NA      30.0     47.6    46.8  0.0
## 6    49.2     8.7    5.77    0.79      29.2     49.8    49.9 73.0
```

```
head(Goalies)
```

```
##   Last.Name First.Name Team.s.        DOB     Birth.City S.P Cntry Nat Ht  Wt
## 1     Allen       Jake     STL 1990-08-07    Fredericton  NB   CAN CAN 74 203
## 2     Alves      Jorge     CAR 1979-01-30         Boston  MA   USA USA 69 185
## 3  Andersen   Frederik     TOR 1989-10-02        Herning       DNK DNK 76 230
## 4  Anderson      Craig     OTT 1981-05-21     Park Ridge  IL   USA USA 74 187
## 5   Bachman    Richard     VAN 1987-07-25 Salt Lake City  UT   USA USA 70 183
## 6   Bernier   Jonathan     ANA 1988-08-07          Laval  QC   CAN CAN 72 184
##   Sh Dft.Yr Rd Ovrl GP GS GR  W  L T OTL   SA   SV  GA   SV.  GAA GSAA    MIN
## 1  L   2008  2   34 61 60  1 33 20 0   5 1620 1482 138 0.915 2.42  2.2 3418.3
## 2  L     NA NA   NA  1  0  1  0  0 0   0    0    0   0 0.000 0.00  0.0    0.1
## 3  L   2012  3   87 66 66  0 33 16 0  14 2052 1883 169 0.918 2.67  8.6 3799.3
## 4  L   2001  3   73 40 40  0 25 11 0   4 1247 1155  92 0.926 2.28 15.9 2421.2
## 5  L   2006  4  120  5  5  0  2  3 0   0  162  149  13 0.920 2.64  1.0  295.2
## 6  L   2006  1   11 39 33  6 21  7 0   4  982  899  83 0.915 2.50  2.0 1993.2
##   SO G A PTS PIM PenT PenD PenDf Supp   QoC  StMin StSV StGA StSV. StGAA QS
## 1  4 0 1   1   4    2    2     0 2.42 8,50% 3400.5 1479  138 0.915  2.43 33
## 2  0 0 0   0   0    0    0     0 0.00          0.0    0    0 0.000  0.00  0
```

```
## 3  4 0 1   1  16    3    9     6 3.04 8,58% 3799.3 1883  169 0.918  2.67 38
## 4  5 0 1   1   0    0    1     1 2.51 8,52% 2421.2 1155   92 0.926  2.28 25
## 5  0 0 0   0   0    0    0     0 2.28 9,07%  295.2  149   13 0.920  2.64  2
## 6  2 0 0   0   4    2    2     0 3.00 8,76% 1899.0  870   78 0.918  2.46 19
##     QS. RBS Pull ReMin ReSV ReGA ReSV. ReGAA Dist Reb Rush  DS HighS HighG
## 1 55,0%  11    8  17.8    3    0 1.000  0.00 32.0 100  101 201   352    66
## 2  0,0%   0   NA   0.1    0    0 0.000  0.00   NA   0    0   0     0     0
## 3 57,6%  11    6   0.0    0    0 0.000  0.00 29.4 118  184 302   481    86
## 4 62,5%   4    1   0.0    0    0 0.000  0.00 34.1  99  130 229   276    49
## 5 40,0%   0    0   0.0    0    0 0.000  0.00 27.3   8   16  24    48     8
## 6 57,6%   4    3  94.2   29    5 0.853  3.18 27.8  65   96 161   216    43
##   HighSV. MedS MedG LowS LowG AdjSV. EVSA EVGA EV.SV. PP.SA PP.GA PP.SV. SH.SA
## 1   0.813  593   55  675   17  0.910 1350  110  0.919   241    24  0.900    29
## 2   0.000    0    0    0    0  0.000    0    0  0.000     0     0  0.000     0
## 3   0.821  842   73  730   10  0.919 1704  133  0.922   293    31  0.894    55
## 4   0.822  406   32  565   11  0.920 1051   62  0.941   167    27  0.838    29
## 5   0.833   65    4   49    1  0.929  141   11  0.922    20     2  0.900     1
## 6   0.801  414   30  352   10  0.914  782   60  0.923   183    20  0.891    17
##   SH.GA SH.SV. AdjSV..1   CA   FA SA.1   xGA GA.1 TrueSV. ExpSV. xGSAA    ZS.
## 1     4  0.862    0.915 3022 2241 1620 136.0  138   0.954  0.916  -2.0  46,6%
## 2     0  0.000    0.000    0    0    0   0.0    0      NA     NA   0.0 100,0%
## 3     5  0.909    0.918 3842 2923 2053 188.1  169   0.956  0.908  19.1  48,1%
## 4     3  0.897    0.925 2435 1739 1247 105.0   92   0.962  0.916  13.0  51,0%
## 5     0  1.000    0.921  305  236  162  15.2   13   0.957  0.906   2.2  47,9%
## 6     3  0.824    0.916 1985 1489  982  92.2   83   0.958  0.906   9.2  44,2%
##    OZS  DZS  NZS   CF   FF   SF   xGF  GF RebF RushF SOS SOG SO.SV. SOW SOL NMC
## 1 1041 1195 1050 2905 2190 1611 137.7 162  100   111  15   4  0.733   1   2
## 2    1    0    0    0    0    0   0.0   0    0     0   0   0  0.000   0   0
## 3 1246 1345 1275 3689 2776 1949 192.8 204  107   209  23  11  0.522   1   7
## 4  837  803  748 2200 1632 1213 101.5 104  111   132  24   5  0.792   3   3
## 5   93  101   92  253  185  127  11.2   9    4    19   0   0  0.000   0   0
## 6  623  786  625 1800 1362  979  99.6  95   78    95  33  11  0.667   1   3
##   Status        Salary       Cap.Hit  Pace X1st X2nd X3rd Star  GPS Ginj
## 1    UFA $2.500.000,00 $2.350.000,00 104.0    9    7    6   22  9.8   NA
## 2                                      0.0   NA   NA   NA   NA  0.0   NA
## 3    UFA $5.000.000,00 $5.000.000,00 118.9    9    6    6   21 13.1    2
## 4    UFA $4.750.000,00 $4.200.000,00 114.9    8    2    5   15  9.2    2
## 5    UFA   $575.000,00   $575.000,00 113.4    1    0    1    2  1.1   NA
## 6    UFA $4.150.000,00 $4.150.000,00 113.9    5    0    2    7  6.0    3
##     Injuries        CHIP
## 1
## 2
## 3 Upper body $121.951,22
## 4 Lower body $102.439,02
## 5
## 6 Upper body $151.829,27
```

```
head(fiveVsfive)
```

```
##   First.Name Last.Name Team    TOI  GAA GSAA   SV. TrueSV. ExpSV. xGSAA   CA
## 1      Allen      Jake  STL 2692.8 2.14  1.7 0.925   0.960  0.924   1.3 2378
## 2      Alves     Jorge  CAR    0.1 0.00  0.0 0.000   0.000  0.000   0.0    0
## 3   Andersen  Frederik  TOR 2997.2 2.34  6.5 0.927   0.961  0.917  16.9 3006
## 4   Anderson     Craig  OTT 1908.4 1.85 17.0 0.940   0.969  0.923  17.5 1919
## 5    Bachman   Richard  VAN  244.8 2.70 -0.3 0.921   0.958  0.908   1.8  262
```

```
## 6    Bernier  Jonathan  ANA 1535.9 2.23 -0.4 0.923   0.963  0.917   4.3 1521
##     FA   SA   xGA  GA HighS HighG HighSV. MedS MedG LowS LowG AdjSV.  Dist Reb
## 1 1752 1273  97.3  96   239    45   0.812  594   14  440   37  0.919 32.69  78
## 2    0    0   0.0   0     0     0   0.000    0    0    0    0  0.000    NA   0
## 3 2280 1609 133.9 117   336    58   0.827  642    8  631   51  0.929 30.30  91
## 4 1374  990  76.5  59   199    34   0.829  501    8  290   17  0.935 35.34  79
## 5  204  139  12.8  11    38     8   0.789   44    1   57    2  0.933 26.77   8
## 6 1128  737  61.3  57   144    28   0.806  301    8  292   21  0.921 28.52  46
##   Rush  DS PenT PenD PenDf   CF   FF   SF    xGF  GF RebF RushF Supp    ZS. OZS
## 1   83 161    2    1    -1 2371 1774 1307 104.22 113   78    97 2.52  47,6% 789
## 2    0   0    0    0     0    0    0    0   0.00   0    0     0 0.00 100,0%   1
## 3  156 247    2    7     5 2962 2202 1531 140.48 139   74   181 2.78  47,8% 922
## 4  117 196    0    0     0 1787 1314  986  78.14  74   95   119 2.33  52,5% 635
## 5   16  24    0    0     0  218  155  107   8.77   8    4    18 1.96  47,7%  72
## 6   79 125    2    2     0 1393 1058  760  70.45  67   51    80 2.62  46,0% 450
##    DZS  NZS
## 1  867  919
## 2    0    0
## 3 1006 1100
## 4  574  647
## 5   79   80
## 6  529  529
```

Pregled dimenzija naših podatkvnih okvira

```
dim(AllSits)
```

```
## [1] 888 167
```

```
dim(fiveOnFive)
```

```
## [1] 888 225
```

```
dim(fiveOnFour)
```

```
## [1] 888 150
```

```
dim(fourOnFive)
```

```
## [1] 888 149
```

```
dim(Goalies)
```

```
## [1]  95 111
```

```
dim(fiveVsfive)
```

```
## [1] 95 42
```

Pregled tipova podataka u našim podatkvnim okvirima

```
sapply(AllSits, class)
```

```
##        Born        City       Pr.St       Cntry         Nat          Ht
## "character" "character" "character" "character" "character"   "integer"
##          Wt       DftYr       DftRd        Ovrl        Hand       NHLid
##   "integer"   "integer"   "integer"   "integer" "character"   "integer"
##   Last.Name  First.Name    Position        Team          GP           G
## "character" "character" "character" "character"   "integer"   "integer"
##           A          A1          A2         PTS        X...        E...
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
```

```
##         PIM      Shifts         TOI       TOI.1      TOI.GP    TOI.GP.1
##   "integer"   "integer"   "integer"   "numeric"   "numeric"   "numeric"
##        TOI.        IPP.         SH.         SV.         PDO        F.60
## "character" "character" "character" "character" "character" "character"
##        A.60        Pct.        Diff     Diff.60         iCF       iCF.1
## "character" "character"   "integer" "character"   "integer"   "integer"
##         iFF         iSF       iSF.1       iSF.2         ixG        iSCF
##   "integer"   "integer"   "integer"   "integer"   "numeric"   "integer"
##         iRB         iRS         iDS       sDist     sDist.1        Pass
##   "integer"   "integer" "character"   "numeric"   "numeric" "character"
##         iHF       iHF.1         iHA        iHDf       iMiss        iGVA
##   "integer"   "integer"   "integer" "character"   "integer"   "integer"
##        iTKA        iBLK      iGVA.1      iTKA.1      iBLK.1        BLK.
##   "integer"   "integer"   "integer"   "integer"   "integer" "character"
##        iFOW        iFOL      iFOW.1      iFOL.1         FO.        X.FOT
##   "integer"   "integer"   "integer"   "integer" "character" "character"
##       dzFOW       dzFOL       nzFOW       nzFOL       ozFOW       ozFOL
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      FOW.Up      FOL.Up    FOW.Down    FOL.Down   FOW.Close   FOL.Close
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         OTG         X1G         GWG         ENG         PSG         PSA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      G.Bkhd      G.Dflct      G.Slap      G.Snap       G.Tip      G.Wrap
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      G.Wrst        CBar        Post        Over        Wide      S.Bkhd
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##     S.Dflct      S.Slap      S.Snap       S.Tip      S.Wrap      S.Wrst
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##       iPenT       iPenD       iPENT       iPEND      iPenDf         NPD
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##         Min         Maj       Match        Misc        Game          CF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##          CA          FF          FA          SF          SA         xGF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##         xGA         SCF         SCA          GF          GA         RBF
##   "numeric"   "integer"   "integer"   "integer"   "integer"   "integer"
##         RBA         RSF         RSA         DSF         DSA         FOW
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         FOL          HF          HA         GVA         TKA        PENT
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        PEND         OPS         DPS          PS         SOS         SOG
##   "integer"   "numeric"   "numeric"   "numeric"   "integer"   "integer"
##       SOGDG        OTOI        Grit         DAP        Pace          GS
##   "integer"   "numeric"   "integer"   "numeric" "character"   "numeric"
##        GS.G        X1st        X2nd        X3rd         MGL    Injuries
##   "numeric"   "integer"   "integer"   "integer"   "integer" "character"
##        CHIP         NMC      Status      Salary     Cap.Hit
## "character" "character" "character" "character" "character"
```

```
sapply(fiveOnFive, class)
```

```
##   Last.Name  First.Name    Position        Team          GP           G
## "character" "character" "character" "character"   "integer"   "integer"
##           A          A1         PTS      PTS.60        IPP.         TOI
##   "integer"   "integer"   "integer"   "numeric" "character"   "integer"
```

```
##       TOI.1       TOI.2       TOI.3       TOI.4      TOI.GP    TOI.GP.1
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##        TOI.         ZS.      RelZS.     TOI.QoT     TOI.QoC      CF.QoC
## "character" "character" "character" "character" "character" "character"
##         SH.         SV.         PDO        F.60        A.60        Pct.
## "character"   "numeric"   "integer" "character" "character" "character"
##        Diff     Diff.60     RelF.60     RelA.60     RelPct.    RelDf.60
##   "integer" "character" "character" "character" "character" "character"
##       RelC.     RelC.60       RelF.   RelF.60.1    QRelCF60    QRelCA60
## "character"   "numeric" "character"   "numeric"   "numeric"   "numeric"
##   QRelDFF60   QRelDFA60         iCF         iFF         iSF         ixG
##   "numeric"   "numeric"   "integer"   "integer"   "integer"   "numeric"
##        iSCF       iCF.1       iFF.1       iSF.1         iRB         iRS
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         iDS       sDist        Pass         iHF         iHA        iHDf
## "character"   "numeric"   "numeric"   "integer"   "integer" "character"
##        iGVA        iTKA        iBLK        BLK.        iFOW        iFOL
##   "integer"   "integer"   "integer" "character"   "integer"   "integer"
##      iFOW.1      iFOL.1         FO.       X.FOT       iPENT       iPEND
##   "integer"   "integer" "character" "character"   "integer"   "integer"
##      iPenDf         OZS         DZS       OZS.1       DZS.1         NZS
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         OTF       OZS.2       DZS.2       NZS.1         OZF         DZF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         NZF          CF          CA          FF          FA        CF.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        CA.1        FF.1        FA.1          SF          SA         SCF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         SCA         xGF         xGA          GF          GA         RBF
##   "integer"   "numeric"   "numeric"   "integer"   "integer"   "integer"
##         RBA         RSF         RSA         DSF         DSA        CF.2
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        CA.2        FF.2        FA.2        SF.1        SA.1        GF.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        GA.1        CF.3        CA.3        FF.3        FA.3        CF.4
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        CA.4        SF.2        SA.2       xGF.1       xGA.1        GF.2
##   "integer"   "integer"   "integer"   "numeric"   "numeric"   "integer"
##        GA.2       RBF.1       RBA.1        CF.5        CA.5         DFF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##         DFA         C.Up       C.Down      C.Tied     C.Close         F.Up
##   "numeric"   "integer"   "integer"   "integer"   "integer"   "integer"
##       F.Down      F.Tied     F.Close         FOW         FOL          HF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##          HA         GVA         TKA        PENT        PEND        OTOI
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##         OCF         OCA         OFF         OFA         OSF         OSA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        OxGF        OxGA        OSCF        OSCA         OGF         OGA
##   "numeric"   "numeric"   "integer"   "integer"   "integer"   "integer"
##        OOZS        ODZS        ONZS       TMCF60      TMCA60      TMFF60
##   "integer"   "integer"   "integer"   "numeric"   "numeric"   "numeric"
##       TMFA60      TMSF60      TMSA60      TMGF60      TMGA60        TMGF
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
```

```
##        TMGA       TmxGF       TmxGA   TOI.QoT.1      CF.QoT     xGF.QoT
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     OppCF60     OppCA60     OppFF60     OppFA60     OppSF60     OppSA60
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     OppGF60     OppGA60    CF.QoC.1     DFF.QoC   TOI.QoC.1    CF.QoC.2
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     xGF.QoC        NSPF        NGPF        ozFO     ozSFPF      ozSAPF
##   "numeric"   "numeric"   "numeric"   "integer"   "integer"   "integer"
##      ozGFPF      ozGAPF  Exp.ozNSPF  Exp.ozNGPF        nzFO      nzSFPF
##   "integer"   "integer"   "numeric"   "numeric"   "integer"   "integer"
##      nzSAPF      nzGFPF      nzGAPF        dzFO      dzSFPF      dzSAPF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      dzGFPF      dzGAPF  Exp.dzNSPF  Exp.dzNGPF       FOvsL      FO.vsL
##   "integer"   "integer"   "numeric"   "numeric"   "integer"   "numeric"
##       FOvsR      FO.vsR        OpFO       OpFOW       HopFO      HopFOW
##   "integer"   "numeric"   "integer"   "integer"   "integer"   "integer"
##       RopFO      RopFOW        Pace
##   "integer"   "integer" "character"
```

```
sapply(fiveOnFour, class)
```

```
##   Last.Name  First.Name    Position        Team          GP           G
## "character" "character" "character" "character"   "integer"   "integer"
##           A          A1         PTS      PTS.60        IPP.         TOI
##   "integer"   "integer"   "integer"   "numeric" "character"   "integer"
##       TOI.1       TOI.2      TOI.GP        TOI.         ZS.      RelZS.
##   "numeric"   "numeric"   "numeric" "character" "character" "character"
##     TOI.QoT     TOI.QoC      CF.QoC         SH.         SV.         PDO
## "character" "character" "character" "character"   "numeric"   "integer"
##        F.60        A.60        Pct.        Diff    Diff.60    RelF.60
##   "numeric"   "numeric" "character"   "integer"   "numeric"   "numeric"
##     RelA.60     RelPct.    RelDf.60         iCF         iFF         iSF
##   "numeric" "character"   "numeric"   "integer"   "integer"   "integer"
##       iSF.1         ixG        iSCF       iCF.1       iFF.1       iSF.2
##   "integer"   "numeric"   "integer"   "integer"   "integer"   "integer"
##         iRB         iRS         iDS       sDist        Pass         iHF
##   "integer"   "integer"   "integer"   "numeric"   "numeric"   "integer"
##       iHF.1         iHA        iHDf       iMiss        iGVA        iTKA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      iGVA.1      iTKA.1        iBLK        BLK.        iFOW        iFOL
##   "integer"   "integer"   "integer" "character"   "integer"   "integer"
##      iFOW.1      iFOL.1         FO.       X.FOT       iPENT       iPEND
##   "integer"   "integer" "character" "character"   "integer"   "integer"
##      iPenDf         OZS         DZS         NZS         OTF       OZS.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##       DZS.1       NZS.1         OZF         DZF         NZF          CF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##          CA          FF          FA          SF          SA         xGF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##         xGA         SCF         SCA          GF          GA         RBF
##   "numeric"   "integer"   "integer"   "integer"   "integer"   "integer"
##         RBA         RSF         RSA         DSF         DSA        CF.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        CA.1        FF.1        FA.1        SF.1        SA.1        GF.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
```

```
##        GA.1         FOW         FOL          HF          HA         GVA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         TKA        PENT        PEND        OTOI         OCF         OCA
##   "integer"   "integer"   "integer"   "numeric"   "integer"   "integer"
##         OFF         OFA         OSF         OSA        OxGF        OxGA
##   "integer"   "integer"   "integer"   "integer"   "numeric"   "numeric"
##        OSCF        OSCA         OGF         OGA        OOZS        ODZS
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        ONZS      TMCF60      TMCA60      TMFF60      TMFA60      TMSF60
##   "integer"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##      TMSA60      TMGF60      TMGA60   TOI.QoT.1      CF.QoT     xGF.QoT
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     OppCF60     OppCA60     OppFF60     OppFA60     OppSF60     OppSA60
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     OppGF60     OppGA60   TOI.QoC.1    CF.QoC.1     xGF.QoC        Pace
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
```

```
sapply(fourOnFive, class)
```

```
##   Last.Name  First.Name    Position        Team          GP           G
## "character" "character" "character" "character"   "integer"   "integer"
##           A          A1         PTS         TOI       TOI.1       TOI.2
##   "integer"   "integer"   "integer"   "integer"   "numeric"   "numeric"
##      TOI.GP        TOI.         ZS.      RelZS.     TOI.QoT     TOI.QoC
##   "numeric" "character" "character" "character" "character" "character"
##      CF.QoC         SH.         SV.         PDO        F.60        A.60
## "character" "character"   "numeric"   "integer"   "numeric"   "numeric"
##        Pct.        Diff     Diff.60     RelF.60     RelA.60     RelPct.
## "character"   "integer"   "numeric"   "numeric"   "numeric" "character"
##    RelDf.60         iCF         iFF         iSF       iSF.1         ixG
##   "numeric"   "integer"   "integer"   "integer"   "integer"   "numeric"
##        iSCF       iCF.1       iFF.1       iSF.2         iRB         iRS
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         iDS       sDist        Pass         iHF       iHF.1         iHA
##   "integer"   "numeric" "character"   "integer"   "integer"   "integer"
##        iHDf       iMiss        iGVA        iTKA        iBLK      iGVA.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      iTKA.1      iBLK.1        BLK.        iFOW        iFOL      iFOW.1
##   "integer"   "integer" "character"   "integer"   "integer"   "integer"
##      iFOL.1         FO.       X.FOT       iPENT       iPEND      iPenDf
##   "integer" "character" "character"   "integer"   "integer"   "integer"
##         OZS         DZS         NZS         OTF       OZS.1       DZS.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##       NZS.1         OZF         DZF         NZF          CF          CA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##          FF          FA          SF          SA         xGF         xGA
##   "integer"   "integer"   "integer"   "integer"   "numeric"   "numeric"
##         SCF         SCA          GF          GA         RBF         RBA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         RSF         RSA         DSF         DSA        CF.1        CA.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        FF.1        FA.1        SF.1        SA.1        GF.1        GA.1
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         FOW         FOL          HF          HA         GVA         TKA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
```

```
##        PENT        PEND        OTOI         OCF         OCA         OFF
##   "integer"   "integer"   "numeric"   "integer"   "integer"   "integer"
##         OFA         OSF         OSA        OxGF        OxGA        OSCF
##   "integer"   "integer"   "integer"   "numeric"   "numeric"   "integer"
##        OSCA         OGF         OGA        OOZS        ODZS        ONZS
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##      TMCF60      TMCA60      TMFF60      TMFA60      TMSF60      TMSA60
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##      TMGF60      TMGA60   TOI.QoT.1      CF.QoT     xGF.QoT     OppCF60
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     OppCA60     OppFF60     OppFA60     OppSF60     OppSA60     OppGF60
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
##     OppGA60   TOI.QoC.1    CF.QoC.1     xGF.QoC        Pace
##   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
```

```
sapply(Goalies, class)
```

```
##   Last.Name  First.Name     Team.s.         DOB  Birth.City         S.P
## "character" "character" "character" "character" "character" "character"
##       Cntry         Nat          Ht          Wt          Sh      Dft.Yr
## "character" "character"   "integer"   "integer" "character"   "integer"
##          Rd        Ovrl          GP          GS          GR           W
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##           L           T         OTL          SA          SV          GA
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##         SV.         GAA        GSAA         MIN          SO           G
##   "numeric"   "numeric"   "numeric"   "numeric"   "integer"   "integer"
##           A         PTS         PIM        PenT        PenD       PenDf
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##        Supp         QoC       StMin        StSV        StGA       StSV.
##   "numeric" "character"   "numeric"   "integer"   "integer"   "numeric"
##       StGAA          QS         QS.         RBS        Pull       ReMin
##   "numeric"   "integer" "character"   "integer"   "integer"   "numeric"
##        ReSV        ReGA       ReSV.      ReGAA        Dist         Reb
##   "integer"   "integer"   "numeric"   "numeric"   "numeric"   "integer"
##        Rush          DS       HighS      HighG     HighSV.        MedS
##   "integer"   "integer"   "integer"   "integer"   "numeric"   "integer"
##        MedG        LowS        LowG      AdjSV.        EVSA        EVGA
##   "integer"   "integer"   "integer"   "numeric"   "integer"   "integer"
##      EV.SV.       PP.SA       PP.GA      PP.SV.       SH.SA       SH.GA
##   "numeric"   "integer"   "integer"   "numeric"   "integer"   "integer"
##      SH.SV.    AdjSV..1          CA          FA        SA.1         xGA
##   "numeric"   "numeric"   "integer"   "integer"   "integer"   "numeric"
##        GA.1     TrueSV.      ExpSV.       xGSAA          ZS.         OZS
##   "integer"   "numeric"   "numeric"   "numeric" "character"   "integer"
##         DZS         NZS          CF          FF          SF         xGF
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##          GF        RebF       RushF         SOS         SOG      SO.SV.
##   "integer"   "integer"   "integer"   "integer"   "integer"   "numeric"
##         SOW         SOL         NMC      Status      Salary     Cap.Hit
##   "integer"   "integer" "character" "character" "character" "character"
##        Pace        X1st        X2nd        X3rd        Star         GPS
##   "numeric"   "integer"   "integer"   "integer"   "integer"   "numeric"
##        Ginj    Injuries        CHIP
##   "integer" "character" "character"
```

```
sapply(fiveVsfive, class)
```

```
##  First.Name   Last.Name        Team         TOI         GAA        GSAA
## "character" "character" "character"   "numeric"   "numeric"   "numeric"
##         SV.      TrueSV.      ExpSV.       xGSAA          CA          FA
##   "numeric"   "numeric"   "numeric"   "numeric"   "integer"   "integer"
##          SA         xGA          GA       HighS       HighG     HighSV.
##   "integer"   "numeric"   "integer"   "integer"   "integer"   "numeric"
##        MedS        MedG        LowS        LowG      AdjSV.        Dist
##   "integer"   "integer"   "integer"   "integer"   "numeric"   "numeric"
##         Reb        Rush          DS        PenT        PenD       PenDf
##   "integer"   "integer"   "integer"   "integer"   "integer"   "integer"
##          CF          FF          SF         xGF          GF        RebF
##   "integer"   "integer"   "integer"   "numeric"   "integer"   "integer"
##       RushF        Supp         ZS.         OZS         DZS         NZS
##   "integer"   "numeric" "character"   "integer"   "integer"   "integer"
```

Plaće NHL igrača

Nakon što smo se upoznali s podacima, možemo se prebaciti na analizu.Jedna od najvažnijih stavki svakog zaposlenog čovjeka, iznos novca kojeg dobije za svoj rad.Stoga ćemo se u ovom dijelu koncentrirati na zadatke vezane uz plaće.Stoga će naš prvi zadatak bit: Možemo li uočiti nekakvu distribuciju plaća NHL igrača?

###ZADATAK 1 ###DISTRIBUCIJA PLAĆE

```
require(stringr)
### IZDVAJANJE U ZASEBNE ZABLICE PLAĆE IGRAČA
### 3 TABLICE:IGRAČI+GOLMANI, IGRAČI I GOLMANI
playersSalary <- as.data.frame(AllSits[, c("Last.Name", "First.Name","Hand","Position","Salary")])
goaliesSalary <- as.data.frame(Goalies[, c("Last.Name", "First.Name", "Salary")])
goaliesSalary$Position <- "GK"
goaliesSalary$Hand <- NA

salary <- rbind(playersSalary, goaliesSalary)

salary$Salary <- str_remove(salary$Salary, "[$]")
playersSalary$Salary <- str_remove(playersSalary$Salary, "[$]")
goaliesSalary$Salary <- str_remove(goaliesSalary$Salary, "[$]")

salary$Salary <- str_replace_all(salary$Salary, "[.]", "")
playersSalary$Salary <- str_replace_all(playersSalary$Salary, "[.]", "")
goaliesSalary$Salary <- str_replace_all(goaliesSalary$Salary, "[.]", "")

salary$Salary <- str_replace_all(salary$Salary, "[,]", ".")
playersSalary$Salary <- str_replace_all(playersSalary$Salary, "[,]", ".")
goaliesSalary$Salary <- str_replace_all(goaliesSalary$Salary, "[,]", ".")

salary$Salary <- as.numeric(salary$Salary)
playersSalary$Salary <- as.numeric(playersSalary$Salary)
goaliesSalary$Salary <- as.numeric(goaliesSalary$Salary)

### Plaće podijeljene s milijun
salary$Salary <- salary$Salary / 1e6
playersSalary$Salary <- playersSalary$Salary / 1e6
goaliesSalary$Salary <- goaliesSalary$Salary / 1e6
```

```
### postotak redaka kojima je salary NA
sum(is.na(salary$Salary)) / nrow(salary)
```

```
## [1] 0.02034588
```

```
sum(is.na(playersSalary$Salary)) / nrow(playersSalary)
```

```
## [1] 0.01576577
```

```
sum(is.na(goaliesSalary$Salary)) / nrow(goaliesSalary)
```

```
## [1] 0.06315789
```

```
###  odbacujemo one retke u kojima je salary postavljena na NA zbog prethodno dobivenih malih postotaka
salary_pure <- as.data.frame(salary[complete.cases(salary$Salary), ])
playersSalary_pure <- as.data.frame(playersSalary[complete.cases(playersSalary$Salary), ])
goaliesSalary_pure <- as.data.frame(goaliesSalary[complete.cases(goaliesSalary$Salary), ])
```

Mjere centralne tendencije vezane uz plaću

Izračunat ćemo neke standradne mjere poput max,min, srednje vrijednosti, medijana, kvartila

```
###max i min
max(salary_pure$Salary)
```

```
## [1] 14
```

```
min(salary_pure$Salary)
```

```
## [1] 0.575
```

```
max(playersSalary_pure$Salary)
```

```
## [1] 14
```

```
min(playersSalary_pure$Salary)
```

```
## [1] 0.575
```

```
max(goaliesSalary_pure$Salary)
```

```
## [1] 9.5
```

```
min(goaliesSalary_pure$Salary)
```

```
## [1] 0.575
```

```
### Aritmeticka sredina
mean(salary_pure$Salary)
```

```
## [1] 2.349043
```

```
mean(playersSalary_pure$Salary)
```

```
## [1] 2.325289
```

```
mean(goaliesSalary_pure$Salary)
```

```
## [1] 2.582316
```

```
### Medijan
median(salary_pure$Salary)
```

```
## [1] 0.925
```

```
median(playersSalary_pure$Salary)
```

```
## [1] 0.925
```

```
median(goaliesSalary_pure$Salary)
```

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

```
### 1., 2. i 3. kvartil
quantile(salary_pure$Salary, probs = c(0.25, 0.5, 0.75))
```

```
##     25%     50%     75%
## 0.74125 0.92500 3.75000
```

```
quantile(playersSalary_pure$Salary, probs = c(0.25, 0.5, 0.75))
```

```
##    25%    50%    75%
## 0.7425 0.9250 3.7000
```

```
quantile(goaliesSalary_pure$Salary, probs = c(0.25, 0.5, 0.75))
```

```
##     25%     50%     75%
## 0.70875 1.00000 4.50000
```

Mjere raspršenosti

Ovdje ćemo računati mjere raspršenosti plaća poput Ranga,IQRanga,varijance, standardne devijacije

```
### Rang
max(salary_pure$Salary) - min(salary_pure$Salary)
```

```
## [1] 13.425
```

```
max(playersSalary_pure$Salary) - min(playersSalary_pure$Salary)
```

```
## [1] 13.425
```

```
max(goaliesSalary_pure$Salary) - min(goaliesSalary_pure$Salary)
```

```
## [1] 8.925
```

```
### IQR - interkvartilni rang
IQR(salary_pure$Salary)
```

```
## [1] 3.00875
```

```
IQR(playersSalary_pure$Salary)
```

```
## [1] 2.9575
```

```
IQR(goaliesSalary_pure$Salary)
```

```
## [1] 3.79125
```

```
### Varijanca
var(salary_pure$Salary)
```

```
## [1] 5.331964
```

```
var(playersSalary_pure$Salary)
```

```
## [1] 5.281965
```

```r
var(goaliesSalary_pure$Salary)
```

```
## [1] 5.827925
```

```r
### Standardna devijacija
sd(salary_pure$Salary)
```

```
## [1] 2.309104
```

```r
sd(playersSalary_pure$Salary)
```

```
## [1] 2.298253
```

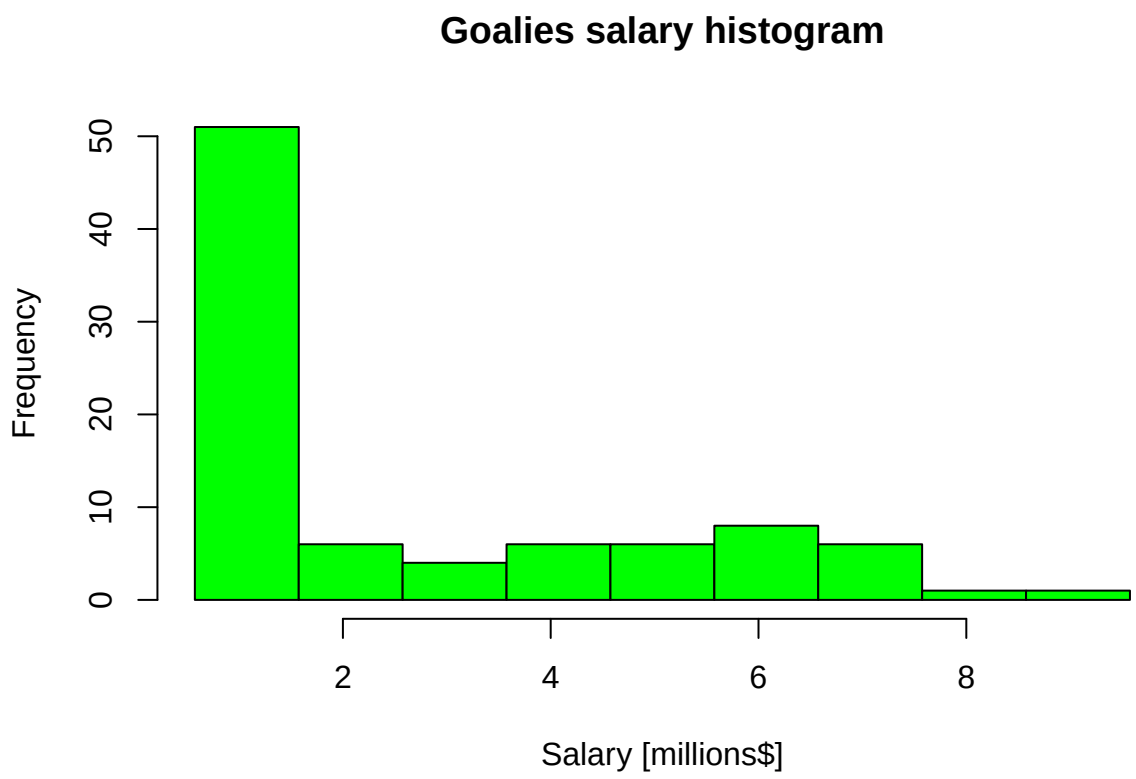
```r
sd(goaliesSalary_pure$Salary)
```

```
## [1] 2.41411
```

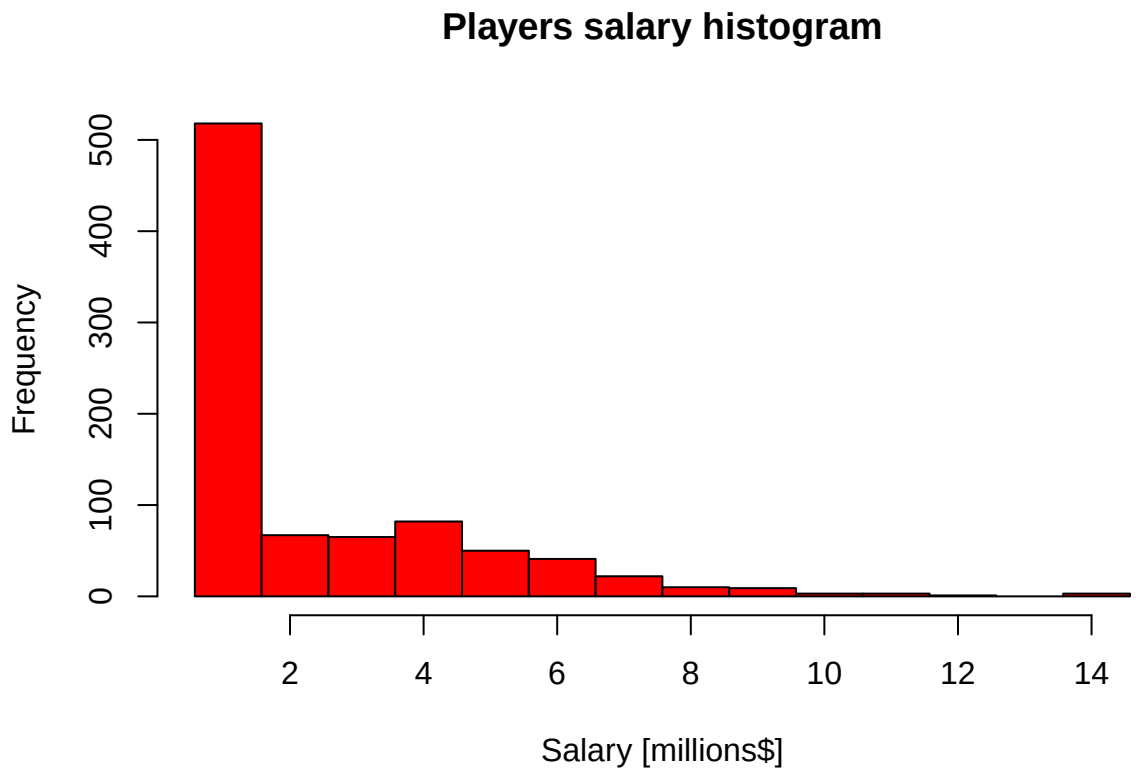
PRIKAZ plaća pomoću histograma

Sada nakon što smo izračunali neke korisne mjere i još više se približili varijabli Salary,prikazat ćemo Salary pomoću histograma.Isto tako probat ćemo pronaći u prikazu oblik normalne distribucije plaća.

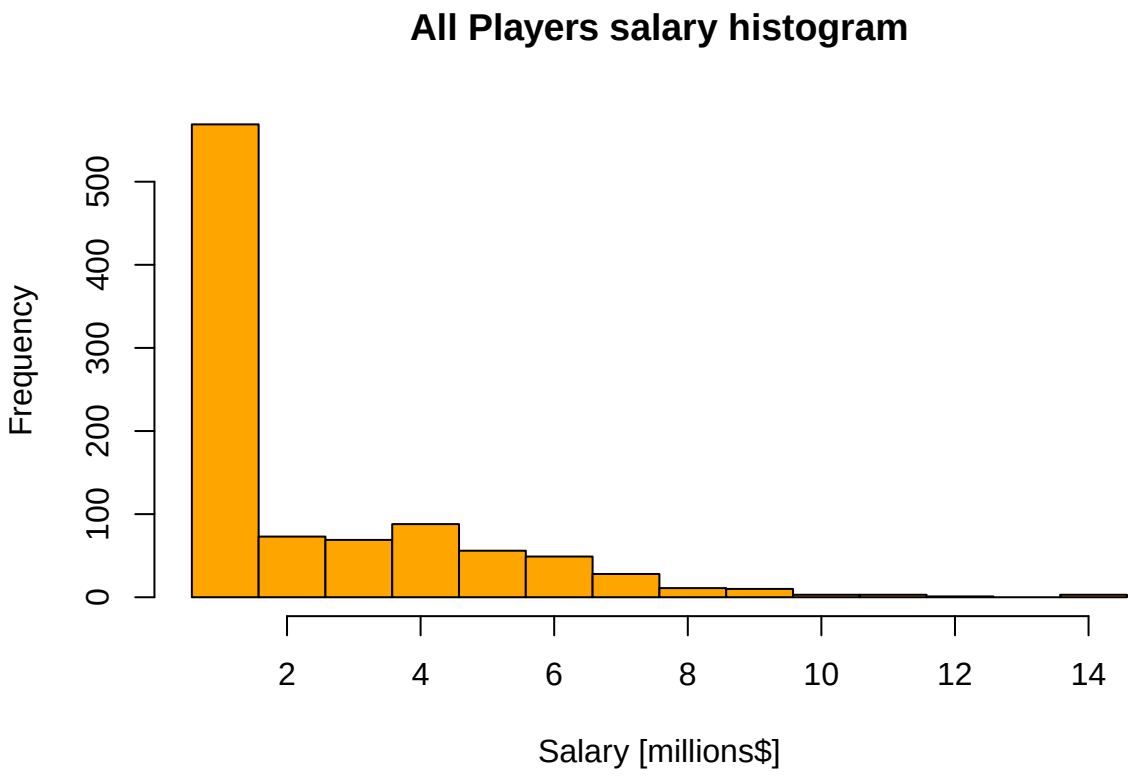
```r
histGoaliesSalary <- hist(goaliesSalary_pure$Salary,
         breaks = seq(min(goaliesSalary_pure$Salary),max(goaliesSalary_pure$Salary)+1, 1),
         main="Goalies salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="green"
         )
```



```
histPlayersSalary <- hist(playersSalary_pure$Salary,
         breaks = seq(min(playersSalary_pure$Salary),max(playersSalary_pure$Salary)+1, 1),
         main="Players salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="red",
         )
```



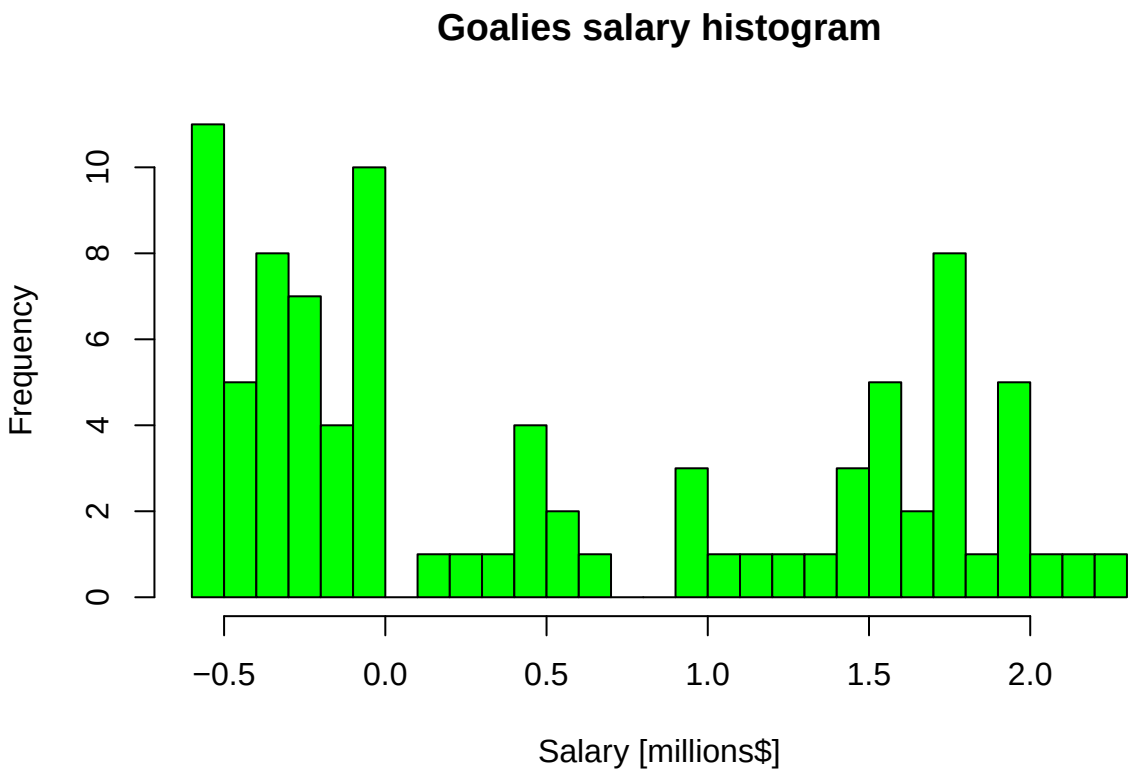

```
histSalary <- hist(salary_pure$Salary,
         breaks = seq(min(salary_pure$Salary),max(salary_pure$Salary)+1, 1),
         main="All Players salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="orange",
         )
```



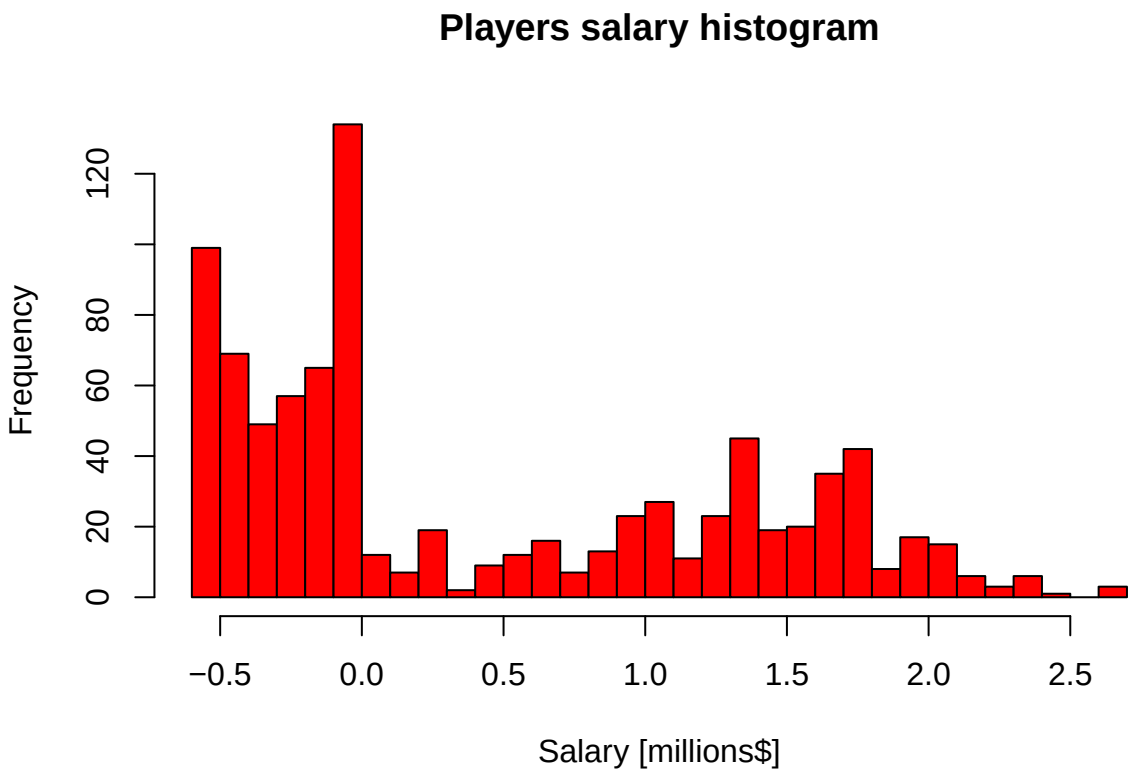
Provjera log distribucije

S obzirom da iz histograma ne možemo uočiti normalnu distribuciju poslužiti ćemo se log distribucijom te tako transformirati podatake sve u svrhu donekle prikaza normalne distribucije.

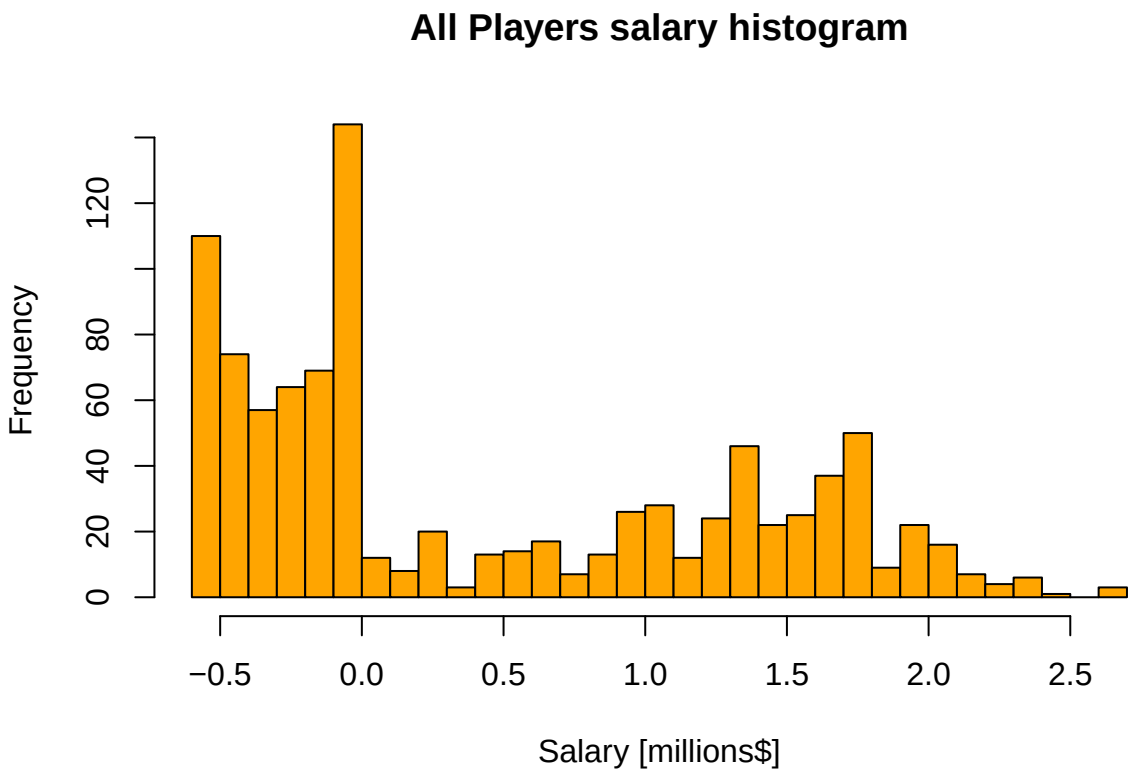
```
hist(log(goaliesSalary_pure$Salary),
         breaks = 30,
         main="Goalies salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="green"
         )
```




```
hist(log(playersSalary_pure$Salary),
         breaks = 30,
         main="Players salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="red"
         )
```



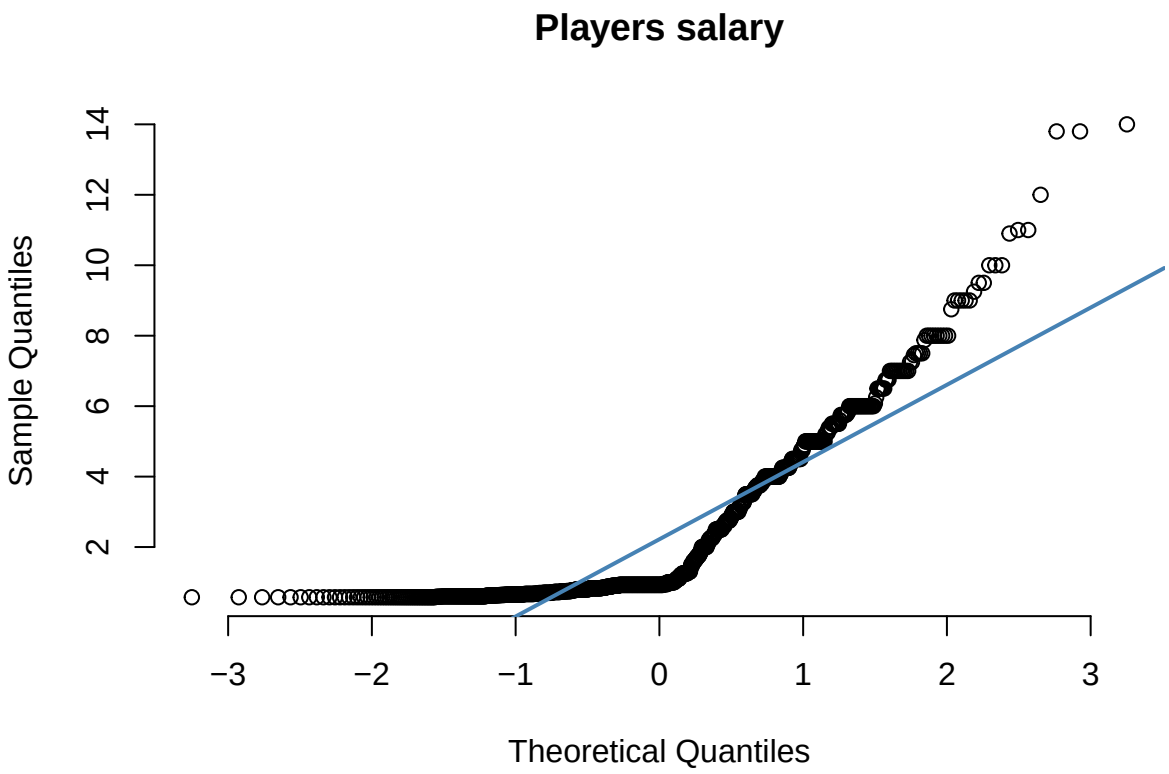

```
hist(log(salary_pure$Salary),
         breaks = 30,
         main="All Players salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="orange",
         )
```



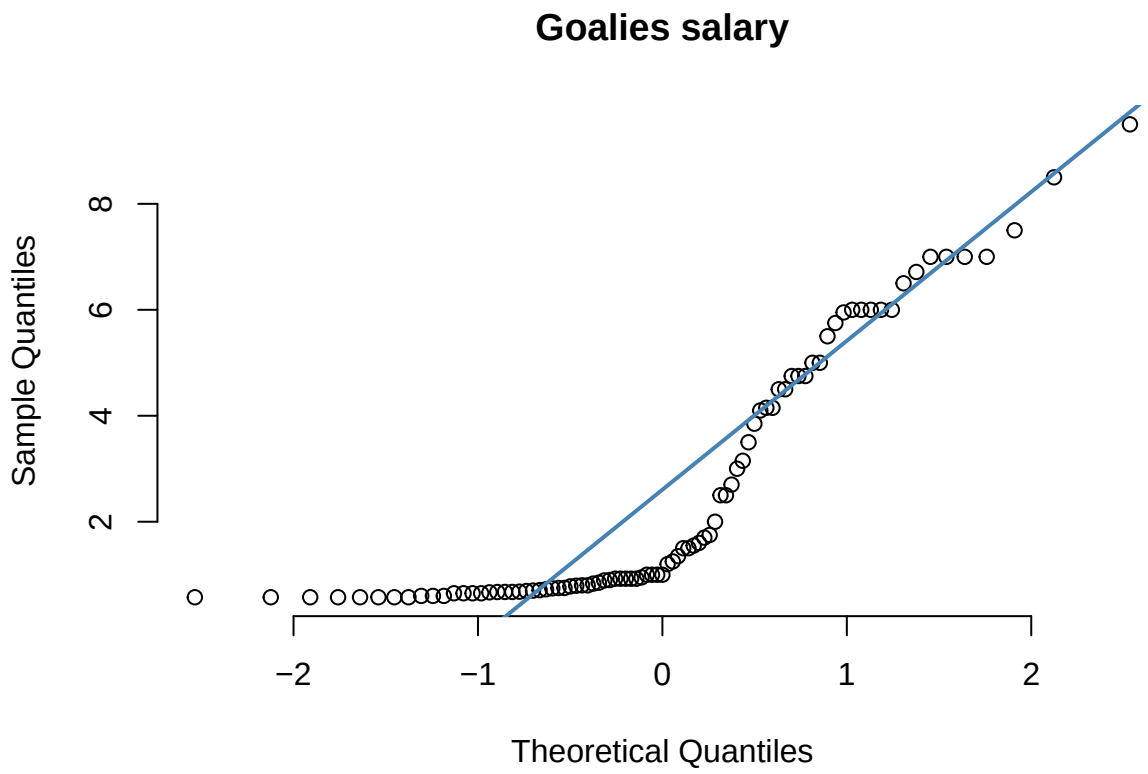
qq plotovi

Iz histograma se ne može baš zaključiti dolazi li uzorak iz normalne distribucije te ćemo pokušati koristiti qq plotove.

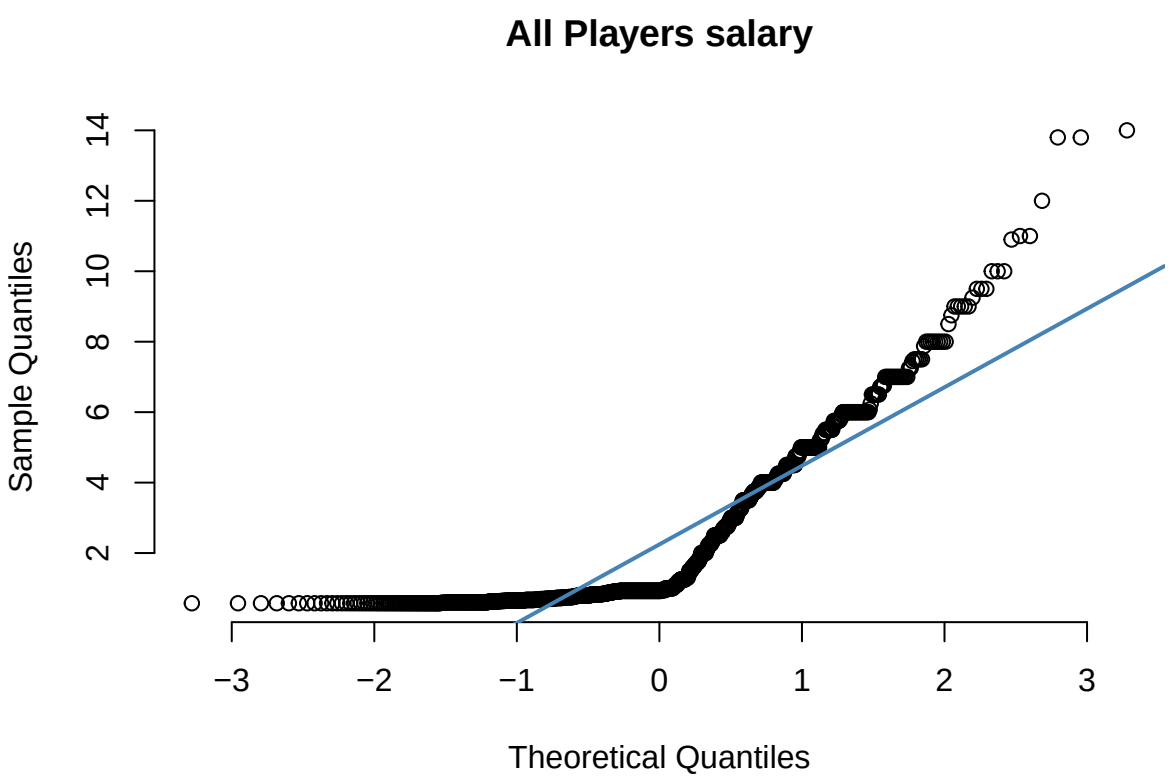
```
qqnorm(playersSalary_pure$Salary, pch = 1, frame = FALSE,main='Players salary')
qqline(playersSalary_pure$Salary, col = "steelblue", lwd = 2)
```




```
qqnorm(goaliesSalary_pure$Salary, pch = 1, frame = FALSE,main='Goalies salary')
qqline(goaliesSalary_pure$Salary, col = "steelblue", lwd = 2)
```

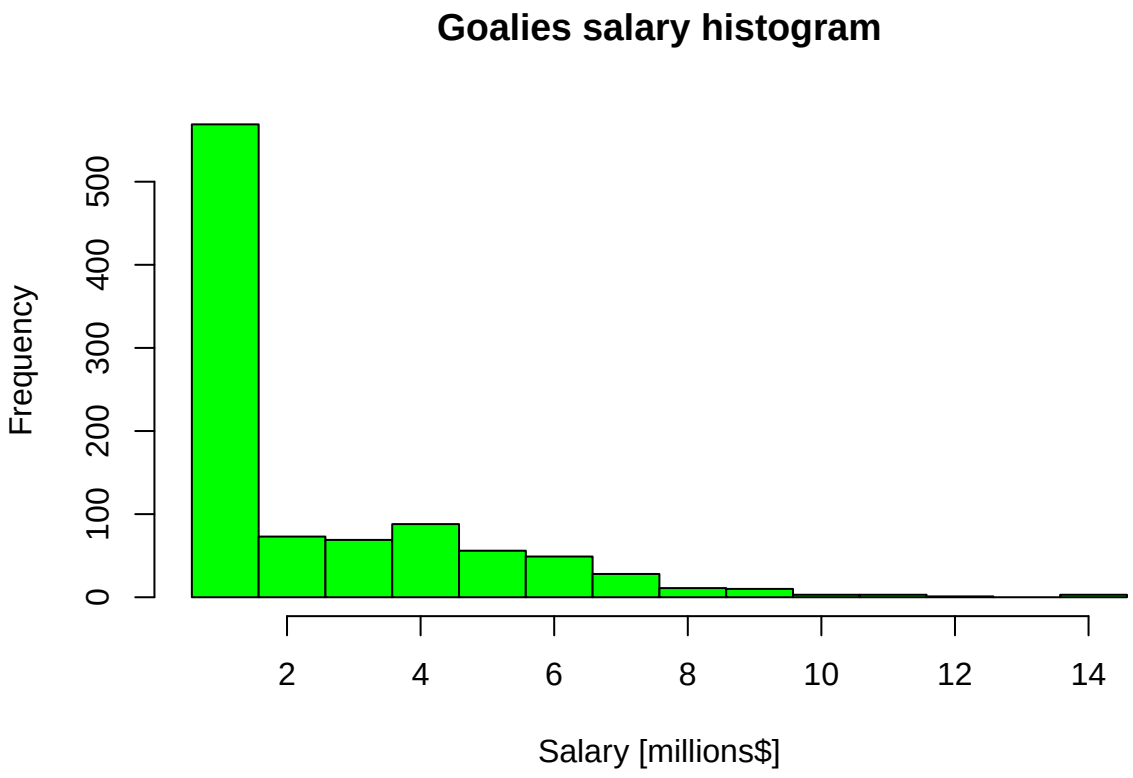



```
qqnorm(salary_pure$Salary, pch = 1, frame = FALSE,main='All Players salary')
qqline(salary_pure$Salary, col = "steelblue", lwd = 2)
```



###ZAKLJUČAK
QQ plotovi su pokazali istu stvar, ne možemo zaključiti da uzorak dolazi iz normalne distribucije. No ako ponovno malo bolje pogledamo u histogram mogli bi uočiti da uzorak dolazi iz eksponencijalne distribucije.

```
hist(salary_pure$Salary,
         breaks = seq(min(salary_pure$Salary),max(salary_pure$Salary)+1, 1),
         main="Goalies salary histogram",
         xlab="Salary [millions$]",
         ylab = "Frequency",
         col="green",
         )
```



###ZADATAK 2 ## Provjerit ćemo jesu li igrači na centarskoj poziciji plaćeni više od obrambenih igrača.

Vrlo često u sportovima koje možemo podijeliti po mjestima na kojima igrači igraju na terenu(obrambeni,srednji i napadački dio) događa se da oni igrači koji su u napadu su većinom više plaćeni od drugi.Stoga će naš zadatak biti ispitati jesu li igači na centarskoj poziciji plaćeni više od obrambenih igrača.

```
att = playersSalary_pure[playersSalary_pure$Position == "C" | playersSalary_pure$Position == "C/LW"| pla
def = playersSalary_pure[playersSalary_pure$Position == "D",]
```

Probat cemo prvo usporediti srednje vrijednosti:

```
mean(att$Salary)
```

```
## [1] 2.32232
```

```
mean(def$Salary)
```

```
## [1] 2.368636
```

Srednje vrijednosti su jako blizu, što nas navodi da su plaće centralnih i obrambenih igrača slične.

Iako iz histograma nismo mogli zaključiti da uzorak dolazi iz normalne distribucije, primjenom CGT možemo to zaključiti. Naime, promatramo stotine igrača što znači da je naš uzorak dovoljno velik, a i jasno je da su igrači međusobno nezavisni.

Provjerit ćemo jesu li varijance jednake:

```
var(att$Salary)
```

```
## [1] 6.282969
```

```r
var(def$Salary)
```

```
## [1] 4.512176
```

Razlika nije nezamjeriva. Provodimo var.test()

```r
var.test(att$Salary, def$Salary)
```

```
##
##  F test to compare two variances
##
## data:  att$Salary and def$Salary
## F = 1.3924, num df = 256, denom df = 291, p-value = 0.00623
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.098524 1.768258
## sample estimates:
## ratio of variances
##           1.392448
```

p vrijednost iznosi 0.00623. Na temelju dobivene p vrijednosti odbacujemo H0 pretpostavku da su varijance jednake. Nastavljamo s računom uzimajući u obzir da nam varijance nisu jednake.

```r
t.test(att$Salary, def$Salary, alt="greater", var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  att$Salary and def$Salary
## t = -0.23187, df = 504.56, p-value = 0.5916
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.37548      Inf
## sample estimates:
## mean of x mean of y
##  2.322320  2.368636
```

###ZAKLJUČAK p-vrijednost iznosi 0.5916. To znači da ne odbacujemo pretpostavku da centralni igrači zarađuju jednako kao i obrambeni.

Rezultati su pomalo iznenađujući, uzimajući u obzir da u većini ostalih mainstream sportova napadači i igrači sredine zarađuju najviše novaca.

Većina timski sportova definira formacije koje se sastoje od pozicija na kojima pojedini igrači tima igraju te tako svojim sposobnostima doprinose timu.Isto tako ako uzmemo primjer pozicije bočnog obrambenog igrača iz nogometa.Često možemo vidjeti da lijevog bočnu poziciju igra igrač kojem je preferirana noga za šutiranje lopte lijeva, isti je slučaj i sa desnom bočnom pozicijom.Stoga je naš sljedeći zadatak ispitati: Jesu li pozicije igrača na terenu nezavisne od preferirane ruke s kojom igrač puca.

###ZADATAK 3 Postavljanje hipoteza: H0 - pozicije su nezavisne H1 - pozicije su zavisne

```r
knitr::include_graphics("hockey_positions.png")
```

Pazimo na to koje su moguće vrijednosti naše varijable, a koje se stvarno pojavljuju u podatcima

```r
levels(AllSits$Position)
```

```
## NULL
```

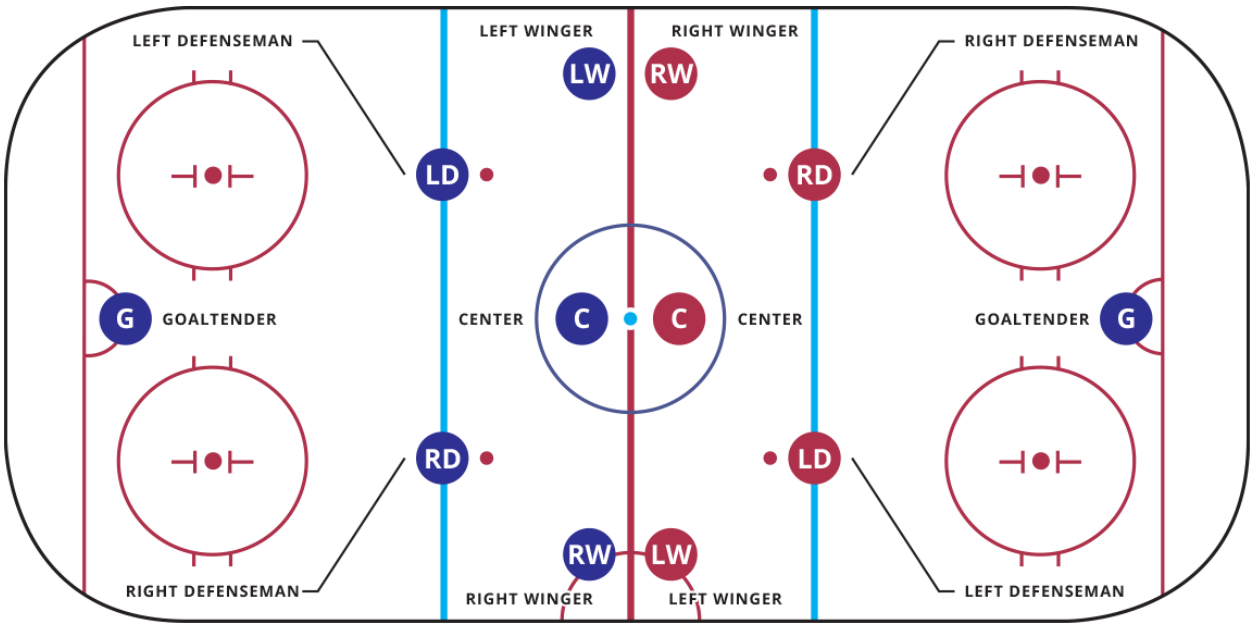


Figure 1: Klasifikacija pozicije igraca na terenu u hokeju

```
levels(AllSits$Hand)
```

```
## NULL
```

```
table(AllSits$Position)
```

```
##
##       C     C/D    C/LW  C/LW/C C/LW/RW    C/RW C/RW/LW       D    D/LW    D/RW
##     144       1      56       1       9      42       7     296       1       2
##      LW    LW/C LW/C/RW   LW/RW LW/RW/C      RW    RW/C RW/C/LW   RW/LW RW/LW/C
##      79      47      10      34       5      91      23       5      31       4
```

```
table(AllSits$Hand)
```

```
##
##   L   R
## 537 351
```

Kopirajmo najprije podatke u novi data.frame kako ne bi promijenili prave vrijednosti.

```
AllSits_copy = data.frame(AllSits)
tracemem(AllSits)==tracemem(AllSits_copy)
```

```
## [1] FALSE
```

```
untracemem(AllSits_copy)
untracemem(AllSits_copy)
```

Pretpostavili smo da je D (defender) centralna pozicija

```
require(tidyr)
```

```
## Loading required package: tidyr
```

```
AllSits_copy['Position'] <- sapply(AllSits_copy['Position'], as.character)

for (column_name in c("C","C/D","C/LW","C/LW/C","C/LW/RW","C/RW", "C/RW/LW", "D", "D/LW", "D/RW")){
  AllSits_copy$Position[AllSits_copy$Position == column_name] = "Central_positions";
}
for (column_name in c("LW","LW/C", "LW/C/RW","LW/RW", "LW/RW/C")){
  AllSits_copy$Position[AllSits_copy$Position == column_name] = "Left_positions";
}
for (column_name in c("RW", "RW/C", "RW/C/LW", "RW/LW", "RW/LW/C")){
  AllSits_copy$Position[AllSits_copy$Position == column_name] = "Right_positions";
}

tbl = table(AllSits_copy$Position)
print(tbl)
```

```
##
## Central_positions    Left_positions   Right_positions
##               559               175               154
```

```
tbl = table(AllSits_copy$Position,
            AllSits_copy$Hand)
tbl = addmargins(tbl)
tbl
```

```
##
##                       L   R Sum
##   Central_positions 341 218 559
##   Left_positions    150  25 175
##   Right_positions    46 108 154
##   Sum               537 351 888
```

Da pokažemo nezavisnost među razredima koristit ćemo chi^2 test koji je implementiran u funkciji `chisq.test()`. Da bi koristili taj test, moramo provjeriti je li očekivana vrijednost svakog od razreda barem 5

```
for (col_names in colnames(tbl)){
  for (row_names in rownames(tbl)){
    if (!(row_names == 'Sum' | col_names == 'Sum') ){
      cat('Očekivane frekvencije za razred ',col_names,'-',row_names,': ',(tbl[row_names,'Sum'] * tbl['
    }
  }
}
```

```
## Ocekivane frekvencije za razred  L - Central_positions :  338.0439
## Ocekivane frekvencije za razred  L - Left_positions :  105.8277
## Ocekivane frekvencije za razred  L - Right_positions :  93.12838
## Ocekivane frekvencije za razred  R - Central_positions :  220.9561
## Ocekivane frekvencije za razred  R - Left_positions :  69.1723
## Ocekivane frekvencije za razred  R - Right_positions :  60.87162
```

Kao što sam ispis pokazuje očekivane vrijednosti su iznad 5 stoga je uvjet za chi^2 je ispunjen.

```
### Izvođenje hi kvadrat testa
chisq.test(tbl,correct=F)
```

```
##
##  Pearson's Chi-squared test
```

```
##
## data:  tbl
## X-squared = 107.05, df = 6, p-value < 2.2e-16
```

###ZAKLJUČAK

S obzirom da je p vrijednost iznosi 2.2e-16, možemo odbaciti h0 u korist h1, odnosno možemo reći da su pozicije igrača na terenu i preferirana ruka zavisne.

Jedna od najbitnijih stavki kod profesionalnih sportaša,ali i općenito jest fizička pripremljenost.Mnogo parametara se mjeri i ispituje da bi se odredila fizička spremnost igrača.Jedan od najvažnijih parametara koji se provjerava jest težina.Stoga će naš sljedeći zadatak biti ispitati: jesu li igrači iz neke države teži od igrača iz druge države?Konkretno u našem slučaju ispitivat ćemo težine ruskih i švedskih igrača.

###ZADATAK 4 ###TJELESNA TEŽINA

```
###  DVIJE TABLICE KOJE SADRŽE RUSKE I ŠVEDSKE IGRAČE
russian_players <- AllSits[AllSits$Nat == "RUS", ]
swedish_players <- AllSits[AllSits$Nat == "SWE", ]
```

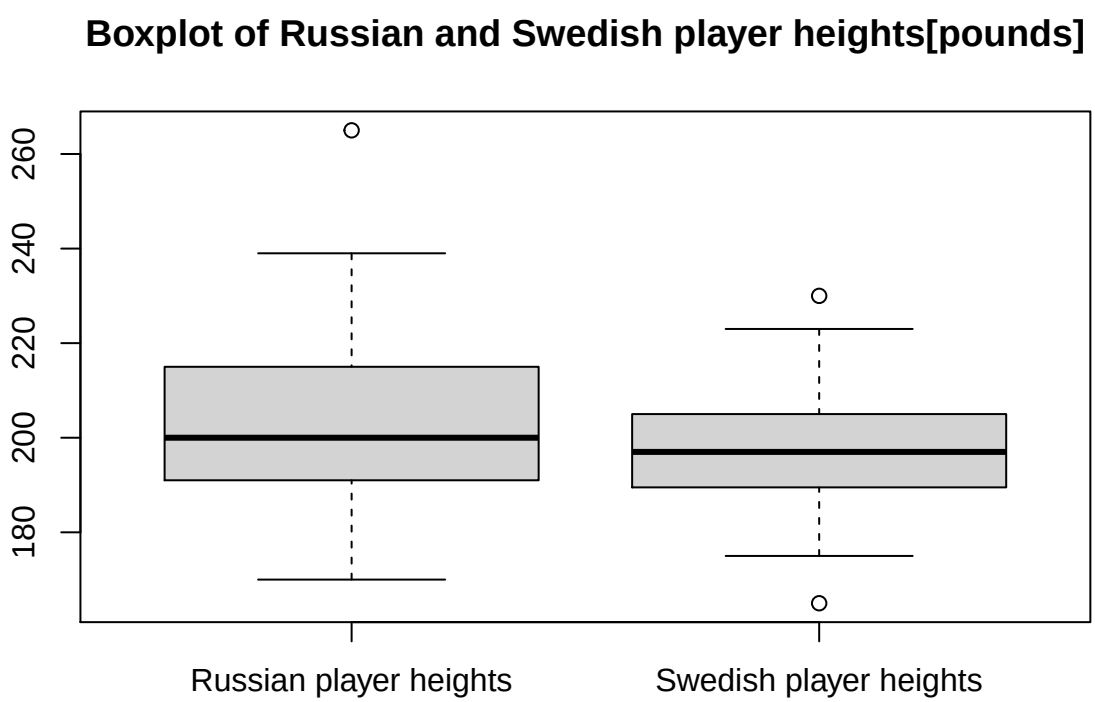
```
cat('Prosječna visina kanadskih igrača iznosi ', mean(russian_players$Wt),'\n')
```

```
## Prosjecna visina kanadskih igraca iznosi  202.7105
```

```
cat('Prosječna visina španjolskih igrača iznosi ', mean(swedish_players$Wt), '\n')
```

```
## Prosjecna visina španjolskih igraca iznosi  198.0127
```

```
boxplot(russian_players$Wt, swedish_players$Wt,
        names = c('Russian player heights','Swedish player heights'),
        main='Boxplot of Russian and Swedish player heights[pounds]')
```



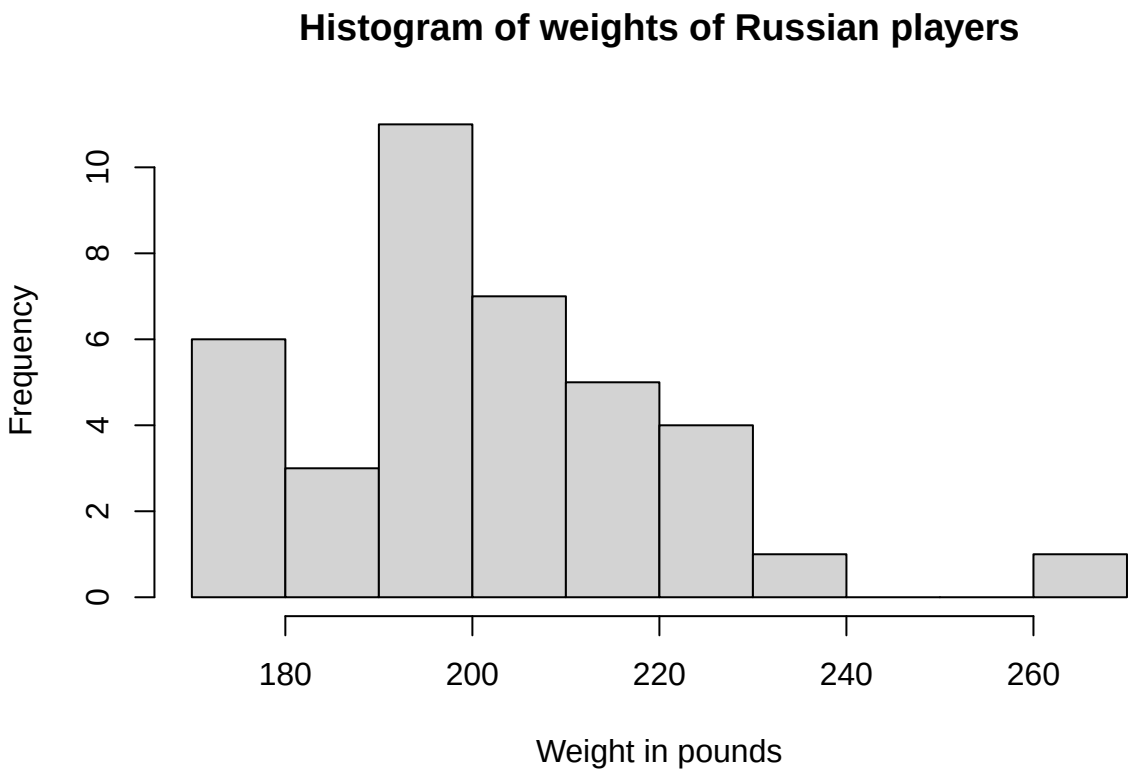
Postoje indikacije da su ruskih igrači teži od švedskih.

Ovakvo ispitivanje možemo provesti t-testom.

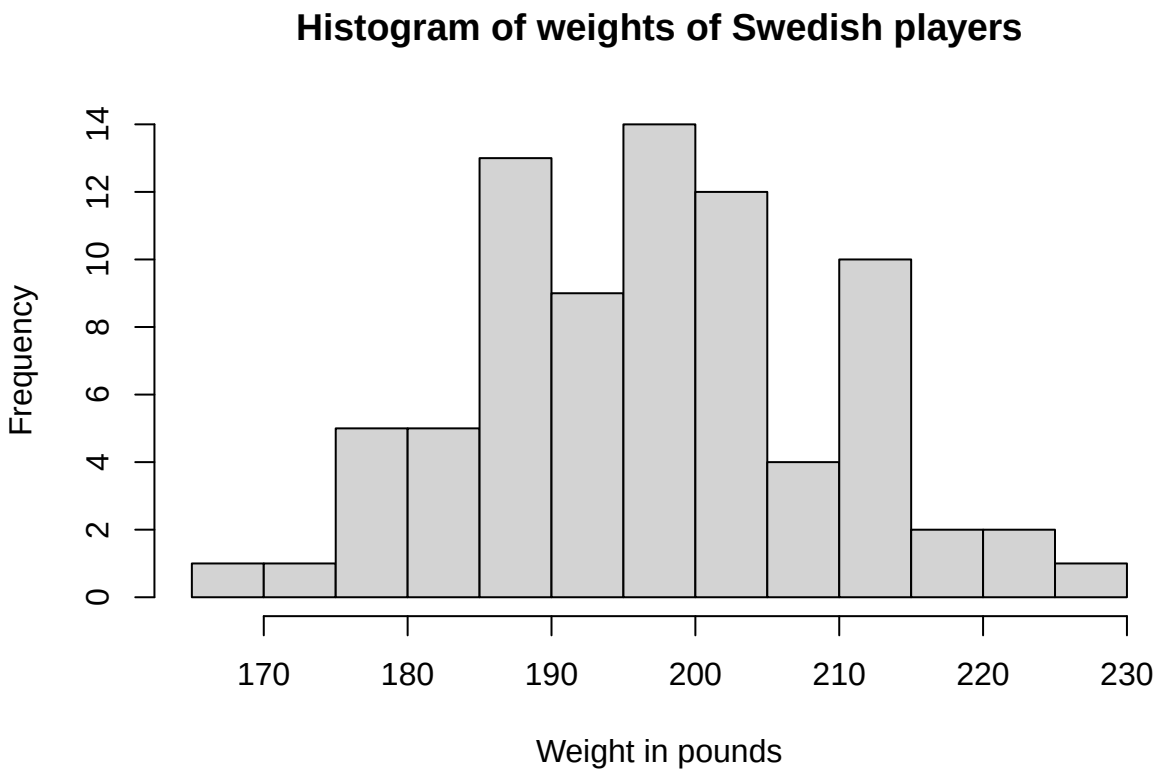
Kako bi mogli provesti test, moramo najprije provjeriti pretpostavke normalnosti i nezavisnosti uzorka. Obzirom da razmatramo dva uzoraka iz dvije različite zemlje, možemo pretpostaviti njihovu nezavisnost.Sljedeći korak je provjeriti normalnost podataka.

Pokušat ćemo preko histograma pokazati normalnosti

```
hist(russian_players$Wt,
     breaks= 10,
     main='Histogram of weights of Russian players',
     xlab='Weight in pounds')
```

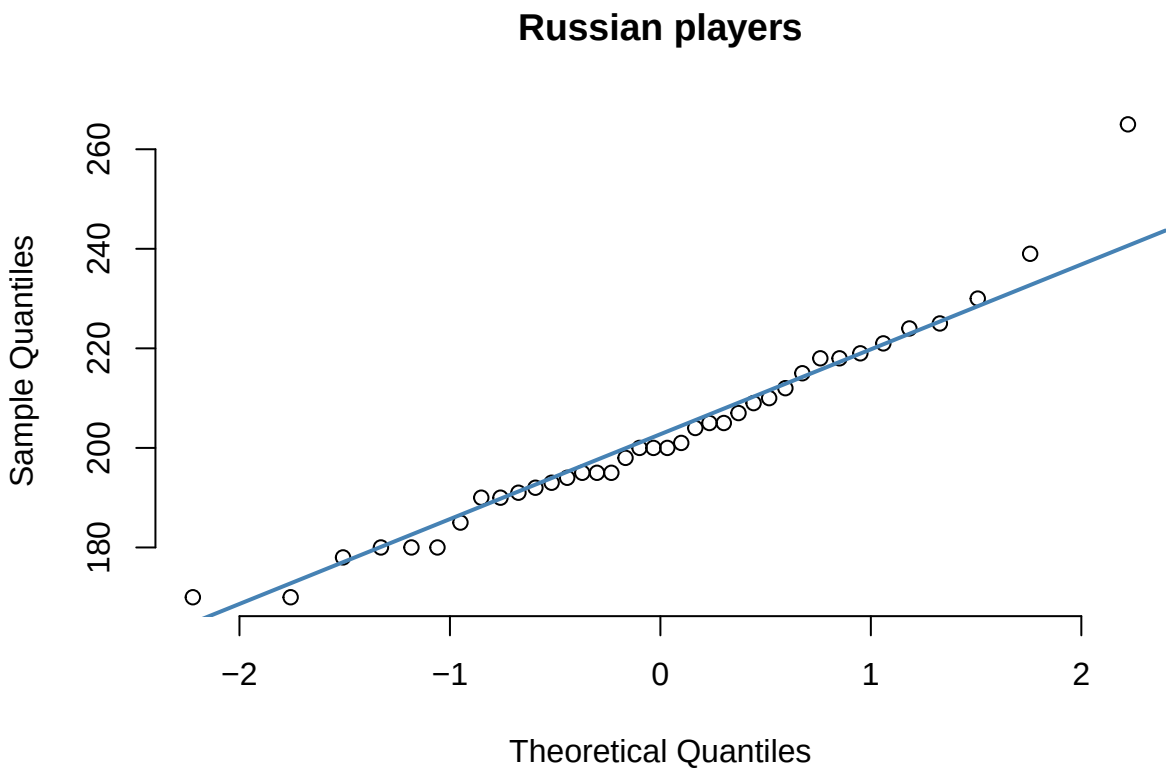



```
hist(swedish_players$Wt ,
     breaks=10,
     main='Histogram of weights of Swedish players',
     xlab='Weight in pounds')
```

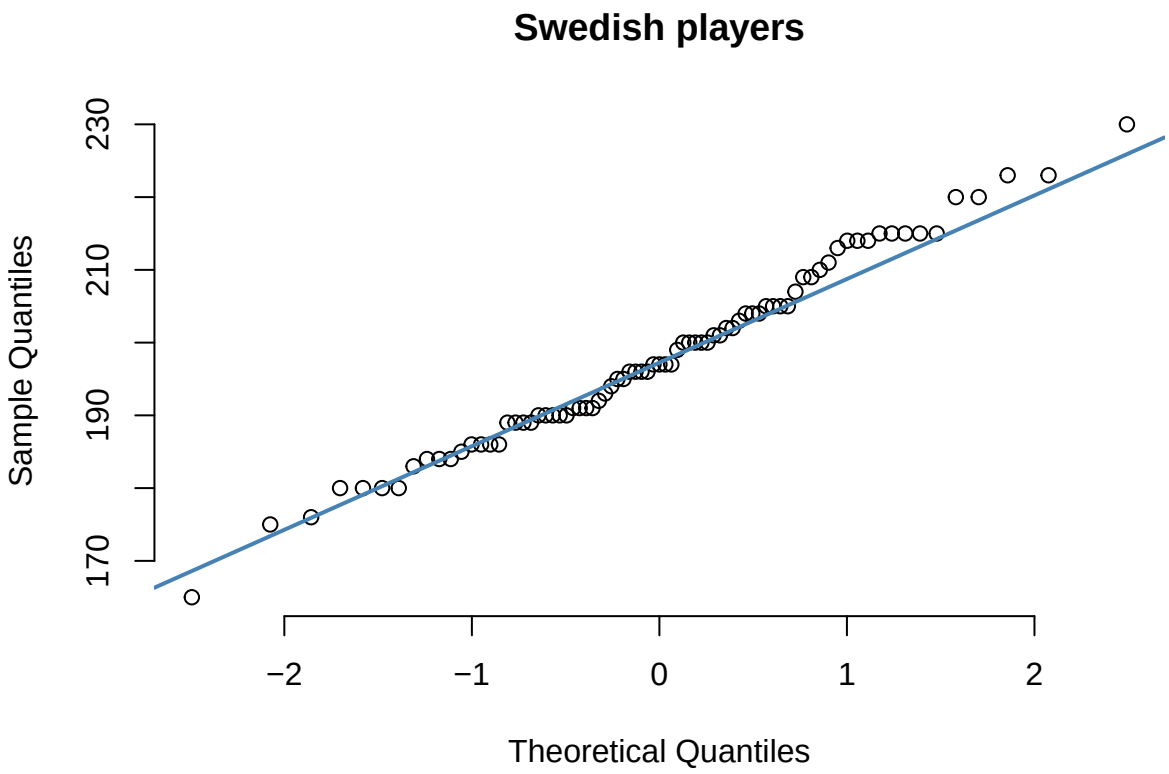


Iz histograma je možda teško vidljiva normalnost, pa ćemo probati s qqplot-ovima

```
qqnorm(russian_players$Wt, pch = 1, frame = FALSE,main='Russian players')
qqline(russian_players$Wt, col = "steelblue", lwd = 2)
```




```
qqnorm(swedish_players$Wt, pch = 1, frame = FALSE,main='Swedish players')
qqline(swedish_players$Wt, col = "steelblue", lwd = 2)
```



qqplotova daju bolju potvrdu da težine dolaze iz normalne distribucije stoga možemo pretpostavit normalnost te nastaviti sa t-testom.

Da bi smo proveli t-test trebamo još izračunati i varijance populacija.

```
var(russian_players$Wt)
```

```
## [1] 375.7248
```

```
var(swedish_players$Wt)
```

```
## [1] 161.8845
```

Razlika između varijanci je dosta velika,ali ipak ćemo provesti testiranje.Pretpostavit ćemo da su varijance jednake i to će biti H0,dok će alternativna hipoteza biti da nisu jednake.

Provest ćemo test nad varijancama kako bi provjerili naše pretpostavke.

```
var.test(russian_players$Wt, swedish_players$Wt)
```

```
##
##  F test to compare two variances
##
## data:  russian_players$Wt and swedish_players$Wt
## F = 2.3209, num df = 37, denom df = 78, p-value = 0.001846
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.363353 4.180389
## sample estimates:
## ratio of variances
```

```
##           2.320944
```

p-vrijednost od 0.001846 nam govori da možemo odbaciti hipotezu H_0 te da su varijance naša dva uzorka nejednaka.

Provedimo sada t-test uz pretpostavku nejednakosti varijanci.

```
t.test(russian_players$Wt, swedish_players$Wt, alt = "greater", var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  russian_players$Wt and swedish_players$Wt
## t = 1.3598, df = 52.849, p-value = 0.08984
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -1.08641      Inf
## sample estimates:
## mean of x mean of y
##  202.7105  198.0127
```

###ZAKLJUČAK

Zbog jako male p-vrijednost koja iznosi 0.08984 ne možemo odbaciti H_0 hipotezu o jednakosti prosječnih vrijednosti u korist H_1, odnosno ne možemo reći da su ruski igrači u prosjeku značajno teži od švedskih igrača.

S obzirom da se naše društvo svakog dana sve više mijenja i napreduje u svim aspektima, ista je stvar i sa sportovim.Grane sporta poput nogometa,košarke,hokeja sve se više šire na globalnoj razini i postaju dostupnije mnogo većem broju ljudi nego prije.U domaćim ligama sve je više stranih igrača, a iznimka nije niti NHL.Tijekom godina sve više europskih igrača nastupa u NHL,stoga je naš sljedeći zadataka:Koliki udio NHL igrača čine igrači iz Europe,ako se pretpostavlja da oko 30 posto NHL igrača čine europski igrači?

###ZADATAK 5 ### UDIO EUROPSKIH IGRAČA U NHL Za analizu koristit ćemo test o jednoj proporciji H0 - p0 = 0.3 H1 - p0 < 0.3

```
AllPlayers <- merge(AllSits,Goalies, all = TRUE)
```

```
## tracemem[0x00000000154bb3d8 -> 0x0000000012f85058]: merge.data.frame merge eval eval withVisible wit
## tracemem[0x0000000012f85058 -> 0x0000000013cd42e0]: .rowNamesDF<- row.names<-.data.frame row.names<-
```

```
nationality <- table(AllPlayers$Nat)
print(nationality)
```

```
##
## AUT CAN CHE CZE DEU DNK FIN FRA GBR HRV LVA NOR RUS SVK SVN SWE USA
##   3 450  14  38   7  10  39   4   1   1   2   2  42  13   1  89 267
```

Iz tablice vidimo da imamo 983 podataka, te da su svi igrači iz Europe ili Sjeverne Amerike

```
n <- nrow(AllPlayers)
k <- nrow(AllPlayers[!AllPlayers$Nat %in% c("CAN", "USA"), ])

prop.test(x = k,n = n, p = 0.3 , alternative = "less")
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  k out of n, null probability 0.3
## X-squared = 3.9072, df = 1, p-value = 0.02404
```

```
## alternative hypothesis: true p is less than 0.3
## 95 percent confidence interval:
##  0.0000000 0.2950349
## sample estimates:
##         p
## 0.2706002
```

###ZAKLJUČAK

p-vrijednost iznosi 0.02404 možemo odbaciti H_0 hipotezu u korist H_1, odnosno možemo reći da manje od 30 posto svih NHL igrača čine igrači iz Europe.

Isto kao što smo analizirali udio eruopski igrača u NHL,tako možemo povezati pozicije na kojima igraju igrači iz Sj.Amerike i Europe te probati odgovorit Je li udio obrambenih igrača u odnosu na ostale igrače u Sj.Americi veći od udjela obrambenih igrača u odnosu na ostale igrače iz Europe.

###ZADATAK 6 ### POZICIJE I BODOVI

```
require(dplyr)
PointsPoistion <- AllSits[,c("First.Name", "Last.Name","GP","Position","G","A","PTS")]

### UZIMAMO SAMO JEDNU POZICIJU U OBZIR
PointsPoistion$OnePosition <- word(PointsPoistion$Position, start = 1, sep = fixed("/"))

### PROVJERA VRIJEDNOSTI
unique(PointsPoistion$OnePosition)
```

```
## [1] "LW" "C"  "RW" "D"
```

```
### Tablice koje sadrže statistiku vezanu za poena,golove i asistencije
PointsStat <- PointsPoistion %>% group_by(OnePosition) %>% summarise(Pospts = sum(PTS))%>%unique()
GoalsStat <- PointsPoistion %>% group_by(OnePosition) %>% summarise(Pospts = sum(G))%>%unique()
AsistsStat <- PointsPoistion %>% group_by(OnePosition) %>% summarise(Pospts = sum(A))%>%unique()
```

PRIKAZ PODATAKA

```
print(PointsStat)
```

```
## # A tibble: 4 x 2
##   OnePosition Pospts
##   <chr>        <int>
## 1 C             6196
## 2 D             4575
## 3 LW            3789
## 4 RW            3383
```

```
print(GoalsStat)
```

```
## # A tibble: 4 x 2
##   OnePosition Pospts
##   <chr>        <int>
## 1 C             2468
## 2 D             1032
## 3 LW            1704
## 4 RW            1500
```

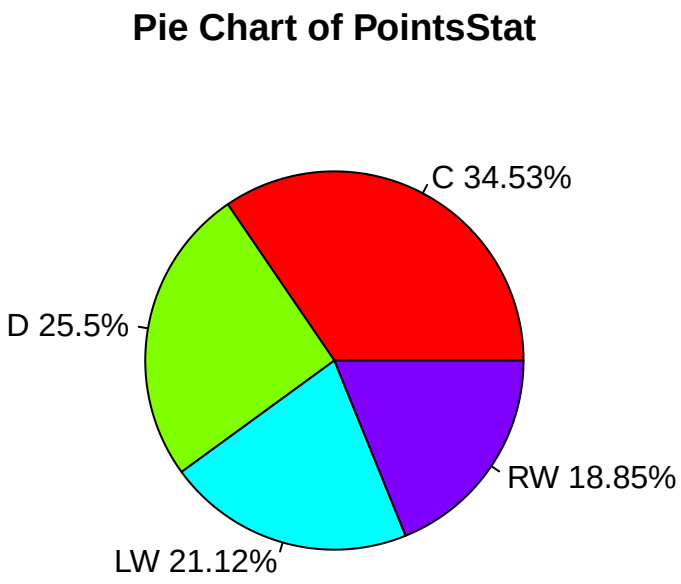
```
print(AsistsStat)
```

```
## # A tibble: 4 x 2
```

```
##   OnePosition Pospts
##   <chr>        <int>
## 1 C             3729
## 2 D             3542
## 3 LW            2085
## 4 RW            1883
```

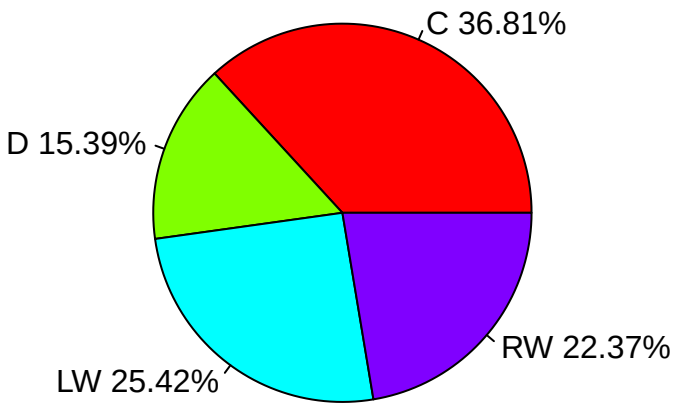
VIZUALIZACIJA PODATAKA

```
library(ggplot2)
per <- round(PointsStat$Pospts / sum(PointsStat$Pospts)*100, 2)
lbls <- paste(PointsStat$OnePosition,per)
lbls <- paste(lbls,"%", sep = "")
pie(PointsStat$Pospts,labels = lbls, col=rainbow(length(lbls)),
   main="Pie Chart of PointsStat")
```

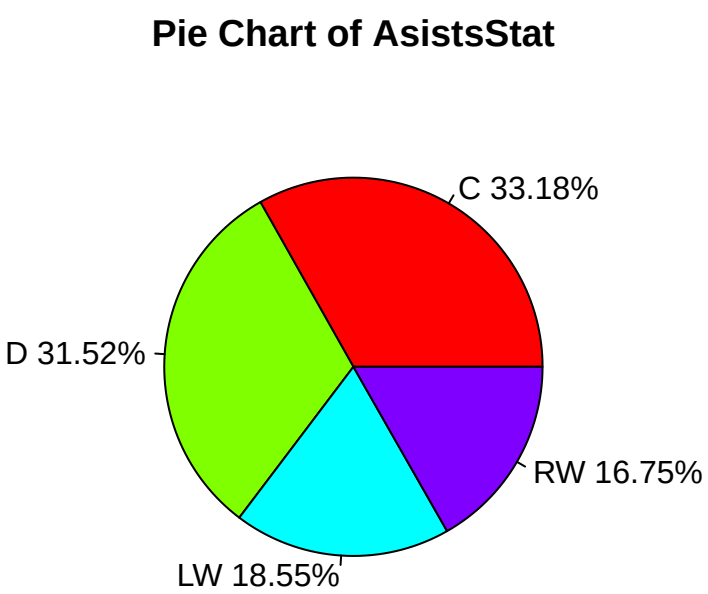


###VIZUALIZACIJA PODATAKA

```
per <- round(GoalsStat$Pospts / sum(GoalsStat$Pospts)*100, 2)
lbls <- paste(GoalsStat$OnePosition,per)
lbls <- paste(lbls,"%", sep = "")
pie(GoalsStat$Pospts,labels = lbls, col=rainbow(length(lbls)),
   main="Pie Chart of GoalsStat")
```




```
per <- round(AsistsStat$Pospts / sum(AsistsStat$Pospts)*100, 2)
lbls <- paste(AsistsStat$OnePosition,per)
lbls <- paste(lbls,"%", sep = "")
pie(AsistsStat$Pospts,labels = lbls, col=rainbow(length(lbls)),
   main="Pie Chart of AsistsStat")
```



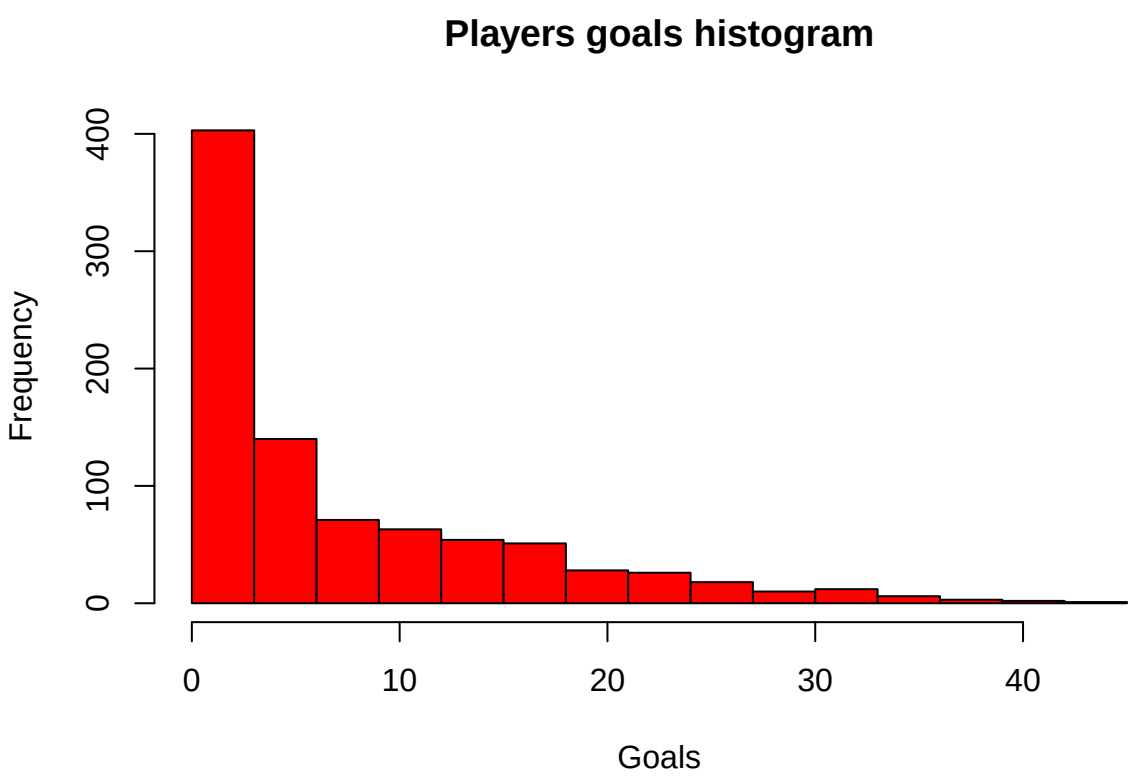
###ZAKLJUČAK

Naša analiza zadatak nam otkriva da najviše bodova osvaja centarska pozicija s udjelom od 34,53% svih ostvarenih bodova.Da smo prije analize pokušali dati odgovor na pitanje koja pozicija donosi najviše bodova, vrlo vjerovatno bi naš odgovor bio centarska pozicija i temeljio bi se na tome da igrači na centarskoj poziciji igraju bliže protivničkom golu te zbog toga imaju veću šansu za postizanjem pogotka ili asistencije.Stoga ova analiza nema nekog prevelikog smisla jer smo i sami bez detaljne analize mogli vrlo vjerovatno točno pretpostaviti odgovor na naše pitanje.

Prethodno smo analizirali koje pozicije ostvaruje PTS, i nismo se iznenadili kad je analiza pokazala da centarska pozicija ostvaruje najviše PTS.No sada možemo i igru uvesti malo geografije i podijeliti igrače iz NHL na one iz Sj.Amerike i Europe te pokušati odgovoriti na pitanje: Zabijaju li Američki igrači više od Europskih?

###ZADATAK 7

```
histPlayersGoals <- hist(AllSits$G,
         breaks = seq(0, max(AllSits$G) + 3, 3),
         main="Players goals histogram",
         xlab="Goals",
         ylab = "Frequency",
         col="red"
         )
```



Plotanjem histograma ne možemo pokazati da uzorak broja golova dolazi iz normalne distribucije.

```
americans = AllSits[AllSits$Nat == "USA" | AllSits$Nat == "CAN",]
europeans = AllSits[!(AllSits$Nat == "USA" | AllSits$Nat == "CAN"),]
```

Probat cemo prvo usporediti srednje vrijednosti:

```
mean(americans$G)
```

```
## [1] 7.16311
```

```
mean(europeans$G)
```

```
## [1] 8.642241
```

Pomalo iznenađujuće, ako gledamo srednje vrijednosti europljani zabijaju više golova od amerikanaca i kanađana. Provest ćemo test da pokažemo vrijedi li ta pretpstavka.

Iako iz histograma nismo mogli zaključiti da uzorak dolazi iz normalne distribucije, primjenom CGT možemo to zaključiti. Naime, promatramo stotine igrača što znači da je naš uzorak dovoljno velik, a i jasno je da su igrači međusobno nezavisni.

Provjerit ćemo jesu li varijance jednake:

```
var(americans$G)
```

```
## [1] 69.92297
```

```
var(europeans$G)
```

```
## [1] 84.67232
```

Razlika od 15ak je nezamjeriva. Provodimo var.test()

```
var.test(americans$G, europeans$G)
```

```
##
##  F test to compare two variances
##
## data:  americans$G and europeans$G
## F = 0.82581, num df = 655, denom df = 231, p-value = 0.07014
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.6640578 1.0157761
## sample estimates:
## ratio of variances
##          0.8258067
```

p vrijednost iznosi 0.07014. Mi smo uzeli razinu značajnosti od 5%, te na temelju rezultata ovo testa ne odbacujemo pretpostavku da su varijance jednake.

Stvorili su se uvjeti da provedemo t.test() (Mogli bi provesti i z test s obzirom na veličinu populaciju, ali z.test() nije dio običnog r paketa)

```
t.test(europeans$G, americans$G, alt="greater", var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  europeans$G and americans$G
## t = 2.2546, df = 886, p-value = 0.0122
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.3988736       Inf
## sample estimates:
## mean of x mean of y
##  8.642241  7.163110
```

###ZAKLJUČAK

p-vrijednost iznosi 0.0122. To znači da odbacujemo pretpostavku da amerikanci i europljani zabijaju jednako golova, u korist da europljani zabijaju više.

```
americansT <- americans[americans$TOI.GP > 5,]
europeansT <- europeans[europeans$TOI.GP > 5,]
mean(americansT$G)
```

```
## [1] 7.185015
```

```
mean(europeansT$G)
```

```
## [1] 8.679654
```

###DODATAK ZAKLJUČKU

Rezultati su bili malo iznenađujući, pa smo usporedili aritmetičku sredinu postignutih golova amerikanaca i europljana koji su igrali prosječno više od 5 minuta. Rezultati nagovještaju naš prijašnji zaključak.

Osim dosta dobrih stvari koje profesionalni sport pruža sportašima.Postoje i neke loše stvari,a jedna od takvih stvari jesu i ozlijede.Igrači koji se ozlijede i prođu proces rehabilitacije,često nakon povratka na sportske terene ne igraju na razini prije ozlijede i minutaža im se smanji.Stoga će naš sljedeći zadataka biti:Provjeriti ovisnost između minuta provedenih na ledu i ozlijeda koje su pretprjeli pojedini igrači

###ZADATAK 8 ###UDIO OBRAMBENIH IGRAČA U AMERICI I EUROPI U ovom primjeru možemo se poslužiti testom o dvije proporcije.

```
table(americans$Position)
```

```
##
##       C     C/D    C/LW  C/LW/C C/LW/RW    C/RW C/RW/LW       D    D/RW      LW
##     113       1      44       1       6      33       5     224       2      58
##    LW/C LW/C/RW   LW/RW LW/RW/C      RW    RW/C RW/C/LW   RW/LW RW/LW/C
##      33       5      24       4      58      20       3      18       4
```

```
table(europeans$Position)
```

```
##
##       C    C/LW C/LW/RW    C/RW C/RW/LW       D    D/LW      LW    LW/C LW/C/RW
##      31      12       3       9       2      72       1      21      14       5
##   LW/RW LW/RW/C      RW    RW/C RW/C/LW   RW/LW
##      10       1      33       3       2      13
```

```
NumOfAmericans <- nrow(americans)
NumOfEuropeans <- nrow(europeans)
americansDefenders <- nrow(americans[grepl("D",americans$Position), ])
europeansDefenders <- nrow(europeans[grepl("D",europeans$Position), ])

prop.test(x=c(americansDefenders, europeansDefenders), n = c(NumOfAmericans, NumOfEuropeans),
          alternative = "greater")
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(americansDefenders, europeansDefenders) out of c(NumOfAmericans, NumOfEuropeans)
## X-squared = 0.62073, df = 1, p-value = 0.2154
## alternative hypothesis: greater
## 95 percent confidence interval:
##  -0.03025692  1.00000000
## sample estimates:
##    prop 1    prop 2
## 0.3460366 0.3146552
```

###ZAKLJUČAK

S obzirom da je p-vrijednost jednaka 0.2154, nemožemo odbaciti H0 u korist H1 i tvrditi da je broj braniča u odnosu na ostale igrače veći u Sj.Americi nego u Europi.

Kao i u većini sportova,cilj svakog sporta je postići više golova/koševa/poena/... od svog protivnika i tako ga pobijediti.Također neki igrači koji igraju na određenim pozicijama postižu više golova/koševa/poena/... od drugih igrača.Stoga je naš treći zadataka istražiti: Koje pozicije u hokeju ostvaruju više bodova(PTS = golovi + asistencije)u odigranim utakmicama te imali uopće smisla razmatrati pozicije i bodove.

###ZADATAK 9 #Provjera ovisnosti između minuta provedenih na terenu po utakmici i ozlijeđivanja Provesti ćemo test nezavinosti. Podatke ćemo prikazati kao kategorijske na način da minute provedene na terenu podijelimo u dvije kategorije jednake veličine (odnosno manje ili veće od medijana). Ozlijede dijelimo na one koje nisu imali ozlijedu te koji su imali jednu ili više ozlijeda.

H0: ozlijeđivanje i minute provedene na terenu su nezavisne H1: ozlijeđivanje i minute provedene na terenu nisu nezavisne

Koristimo kopiju naše prvotne tablice.

```
AllSits_copy$Injuries[nchar(AllSits$Injuries) > 0] = "injury"
AllSits_copy$Injuries[nchar(AllSits$Injuries) == 0] = "no_injury"
```

Svim igračima koji su bili ozlijeđeni retke smo postavili u 1, ostalima u 0.

```
AllSits$Injuries
```

```
##   [1] ""
##   [2] "Lower body, Knee"
##   [3] ""
##   [4] "Lower body, Upper body"
##   [5] ""
##   [6] ""
##   [7] ""
##   [8] "Flu, Lower body"
##   [9] ""
##  [10] "Upper body"
##  [11] "Shoulder"
##  [12] "Upper body, Lower body"
##  [13] ""
##  [14] "Lower body"
##  [15] "Upper body"
##  [16] "Knee, Upper body, Undisclosed"
##  [17] ""
##  [18] ""
##  [19] "Upper body, Lower body"
##  [20] "Elbow, Concussion"
##  [21] ""
##  [22] ""
##  [23] "Foot, Concussion, Upper body"
##  [24] ""
##  [25] ""
##  [26] "Foot"
##  [27] ""
##  [28] ""
##  [29] ""
##  [30] ""
##  [31] "Back, Upper body"
##  [32] "Lower body"
##  [33] ""
##  [34] "Illness"
##  [35] "Upper body"
##  [36] "Illness"
##  [37] "Illness"
##  [38] "Neck"
##  [39] "Foot, Illness"
##  [40] ""
##  [41] "Knee"
##  [42] ""
##  [43] "Foot, Upper body"
##  [44] "Undisclosed"
##  [45] "Leg, Lower body"
##  [46] ""
##  [47] "Undisclosed"
```

```
##  [48] "Lower body"
##  [49] ""
##  [50] ""
##  [51] "Leg"
##  [52] "Illness"
##  [53] ""
##  [54] "Facial"
##  [55] "Upper body, Hand"
##  [56] "Undisclosed"
##  [57] "Hand, Lower body"
##  [58] "Illness"
##  [59] ""
##  [60] ""
##  [61] ""
##  [62] ""
##  [63] "Knee, Ribs"
##  [64] ""
##  [65] "Illness"
##  [66] "Flu, Upper body"
##  [67] "Lower body, Upper body"
##  [68] ""
##  [69] "Shoulder"
##  [70] "Arm"
##  [71] ""
##  [72] "Groin, Lower body, Head"
##  [73] "Lower body"
##  [74] "Flu, Upper body, Foot"
##  [75] "Upper body, Lower body"
##  [76] "Lower body, Illness"
##  [77] ""
##  [78] ""
##  [79] ""
##  [80] "Finger"
##  [81] "Upper body, Foot"
##  [82] ""
##  [83] "Finger"
##  [84] ""
##  [85] "Upper body"
##  [86] "Upper body, Undisclosed"
##  [87] ""
##  [88] "Back"
##  [89] "Hand"
##  [90] ""
##  [91] "Upper body"
##  [92] ""
##  [93] "Neck, Eye, Illness, Upper body"
##  [94] ""
##  [95] "Lower body"
##  [96] "Upper body"
##  [97] "Hip"
##  [98] "Undisclosed"
##  [99] "Hip, Lower body"
## [100] "Head, Illness, Oblique"
## [101] "Upper body"
```

```
## [102] ""
## [103] ""
## [104] "Illness"
## [105] ""
## [106] ""
## [107] "Lower body"
## [108] ""
## [109] ""
## [110] ""
## [111] "Upper body"
## [112] ""
## [113] "Hand, Flu, Knee"
## [114] ""
## [115] "Lower body"
## [116] ""
## [117] ""
## [118] "Illness"
## [119] "Lower body, Illness"
## [120] "Upper body"
## [121] ""
## [122] ""
## [123] ""
## [124] "Concussion"
## [125] "Upper body, Hand"
## [126] "Lower body"
## [127] ""
## [128] "Upper body"
## [129] "Upper body, Lower body"
## [130] "Upper body"
## [131] ""
## [132] "Lower body"
## [133] ""
## [134] "Upper body"
## [135] "Groin, Leg"
## [136] ""
## [137] ""
## [138] ""
## [139] "Lower body, Upper body"
## [140] "Illness"
## [141] "Upper body"
## [142] "Upper body"
## [143] ""
## [144] "Upper body"
## [145] "Illness, Upper body, Mouth"
## [146] "Knee"
## [147] ""
## [148] "Lower body"
## [149] "Foot, Lower body"
## [150] "Concussion"
## [151] "Upper body, Lower body"
## [152] "Foot, Undisclosed"
## [153] "Lower body"
## [154] ""
## [155] "Upper body, Knee"
```

```
## [156] "Knee"
## [157] ""
## [158] "Leg, Undisclosed"
## [159] ""
## [160] "Upper body, Illness"
## [161] ""
## [162] ""
## [163] ""
## [164] ""
## [165] ""
## [166] "Lower body"
## [167] "Wrist"
## [168] ""
## [169] "Knee"
## [170] "Lower body"
## [171] "Knee"
## [172] "Concussion"
## [173] "Upper body"
## [174] ""
## [175] ""
## [176] ""
## [177] "Lower body"
## [178] "Hand"
## [179] "Upper body"
## [180] "Shoulder, Neck"
## [181] ""
## [182] ""
## [183] "Lower body"
## [184] "Upper body"
## [185] ""
## [186] "Upper body, Undisclosed, Flu"
## [187] "Undisclosed"
## [188] "Head, Illness"
## [189] ""
## [190] "Lower body, Illness"
## [191] ""
## [192] "Jaw"
## [193] ""
## [194] ""
## [195] "Knee"
## [196] "Lower body"
## [197] ""
## [198] "Upper body, Finger"
## [199] "Upper body"
## [200] ""
## [201] "Ankle"
## [202] "Concussion, Neck"
## [203] ""
## [204] ""
## [205] ""
## [206] "Upper body"
## [207] "Upper body, Lower body"
## [208] "Lower body"
## [209] ""
```

```
## [210] ""
## [211] "Groin"
## [212] "Lower body, Concussion, Knee"
## [213] "Back, Wrist"
## [214] "Leg"
## [215] ""
## [216] ""
## [217] ""
## [218] "Knee"
## [219] "Lower body"
## [220] "Calf"
## [221] ""
## [222] "Groin"
## [223] "Upper body, Shoulder, Undisclosed"
## [224] "Upper body, Lower body"
## [225] "Undisclosed, Groin"
## [226] "Shoulder, Foot"
## [227] "Upper body"
## [228] "Lower body, Illness"
## [229] ""
## [230] "Lower body, Upper body"
## [231] "Lower body"
## [232] ""
## [233] "Illness, Upper body, Lower body"
## [234] "Neck, Upper body"
## [235] "Knee"
## [236] "Illness"
## [237] "Knee, Upper body"
## [238] ""
## [239] ""
## [240] ""
## [241] "Upper body"
## [242] "Knee"
## [243] "Lower body, Foot, Undisclosed"
## [244] ""
## [245] ""
## [246] ""
## [247] "Foot"
## [248] ""
## [249] ""
## [250] "Knee"
## [251] "Hand"
## [252] ""
## [253] ""
## [254] "Lower body, Illness"
## [255] ""
## [256] ""
## [257] "Finger"
## [258] "Upper body"
## [259] ""
## [260] ""
## [261] "Foot"
## [262] ""
## [263] "Upper body, Lower body"
```

```
## [264] ""
## [265] ""
## [266] ""
## [267] ""
## [268] "Groin, Lower body, Ankle"
## [269] "Mid-body"
## [270] ""
## [271] ""
## [272] "Foot"
## [273] "Illness"
## [274] ""
## [275] "Oblique"
## [276] ""
## [277] "Upper body"
## [278] "Foot, Hip, Abdominal"
## [279] "Hand"
## [280] ""
## [281] "Hip"
## [282] "Undisclosed"
## [283] ""
## [284] "Illness, Wrist"
## [285] "Shoulder"
## [286] ""
## [287] ""
## [288] "Lower body, Illness"
## [289] "Wrist"
## [290] "Back"
## [291] ""
## [292] ""
## [293] "Concussion"
## [294] ""
## [295] ""
## [296] ""
## [297] "Undisclosed"
## [298] ""
## [299] "Illness, Upper body"
## [300] "Wrist"
## [301] "Concussion"
## [302] ""
## [303] "Upper body, Lower body"
## [304] ""
## [305] "Concussion, Lower body"
## [306] ""
## [307] "Upper body"
## [308] "Upper body"
## [309] "Knee, Lower body"
## [310] "Undisclosed"
## [311] "Illness"
## [312] ""
## [313] "Thumb, Lower body, Upper body"
## [314] ""
## [315] ""
## [316] ""
## [317] "Ribs, Knee, Upper body"
```

```
## [318] "Lower body, Upper body, Illness"
## [319] ""
## [320] ""
## [321] ""
## [322] "Lower body"
## [323] "Lower body"
## [324] "Upper body"
## [325] "Lower body"
## [326] ""
## [327] "Lower body"
## [328] "Lower body"
## [329] "Illness"
## [330] ""
## [331] ""
## [332] ""
## [333] "Shoulder, Knee"
## [334] "Groin, Hip"
## [335] "Lower body"
## [336] ""
## [337] ""
## [338] "Knee"
## [339] "Lower body"
## [340] "Upper body"
## [341] "Upper body"
## [342] ""
## [343] "Undisclosed, Illness, Groin"
## [344] "Undisclosed"
## [345] ""
## [346] ""
## [347] "Upper body"
## [348] "Concussion"
## [349] "Lower body"
## [350] "Lower body, Upper body"
## [351] ""
## [352] ""
## [353] "Achilles"
## [354] "Illness, Lower body"
## [355] "Chest"
## [356] ""
## [357] "Lower body"
## [358] "Hand"
## [359] ""
## [360] ""
## [361] ""
## [362] ""
## [363] ""
## [364] "Lower body"
## [365] "Ankle, Upper body"
## [366] ""
## [367] ""
## [368] ""
## [369] ""
## [370] ""
## [371] "Lower body, Illness"
```

```
## [372] "Leg"
## [373] ""
## [374] "Undisclosed, Lower body"
## [375] ""
## [376] "Knee, Illness, Undisclosed"
## [377] ""
## [378] ""
## [379] "Foot"
## [380] "Upper body"
## [381] "Concussion, Illness, Upper body"
## [382] "Upper body, Lower body"
## [383] ""
## [384] "Back"
## [385] ""
## [386] "Illness, Lower body"
## [387] ""
## [388] ""
## [389] "Ribs, Illness"
## [390] ""
## [391] ""
## [392] ""
## [393] "Lower body, Foot"
## [394] "Ankle, Lower body"
## [395] "Illness"
## [396] ""
## [397] "Lower body"
## [398] ""
## [399] ""
## [400] ""
## [401] ""
## [402] ""
## [403] ""
## [404] ""
## [405] ""
## [406] ""
## [407] ""
## [408] "Illness, Wrist"
## [409] "Undisclosed"
## [410] ""
## [411] ""
## [412] ""
## [413] "Back, Illness"
## [414] ""
## [415] ""
## [416] "Illness"
## [417] ""
## [418] "Ankle"
## [419] "Upper body, Illness"
## [420] ""
## [421] ""
## [422] "Neck, Upper body"
## [423] ""
## [424] "Knee, Undisclosed, Groin, Lower body"
## [425] "Lower body"
```

```
## [426] "Upper body, Illness"
## [427] "Upper body, Lower body, Illness"
## [428] "Lower body"
## [429] "Illness"
## [430] "Undisclosed, Upper body"
## [431] "Lower body, Back, Upper body"
## [432] "Lower body"
## [433] ""
## [434] ""
## [435] ""
## [436] ""
## [437] "Upper body"
## [438] "Upper body"
## [439] "Concussion, Illness"
## [440] ""
## [441] "Lower body"
## [442] ""
## [443] "Upper body"
## [444] "Illness, Upper body"
## [445] "Lower body"
## [446] ""
## [447] "Wrist"
## [448] ""
## [449] "Lower body"
## [450] "Concussion, Upper body"
## [451] "Upper body"
## [452] "Illness"
## [453] "Upper body"
## [454] "Upper body, Illness"
## [455] ""
## [456] "Lower body, Undisclosed"
## [457] ""
## [458] "Upper body, Lower body, Neck"
## [459] ""
## [460] ""
## [461] ""
## [462] "Concussion"
## [463] "Hip"
## [464] ""
## [465] ""
## [466] "Lower body"
## [467] "Lower body, Upper body"
## [468] "Upper body"
## [469] "Lower body"
## [470] ""
## [471] ""
## [472] ""
## [473] ""
## [474] ""
## [475] "Illness, Hand"
## [476] "Concussion"
## [477] "Lower body"
## [478] ""
## [479] ""
```

```
## [480] "Upper body, Concussion"
## [481] "Lower body, Shoulder"
## [482] ""
## [483] "Concussion, Upper body"
## [484] ""
## [485] "Finger"
## [486] ""
## [487] "Upper body, Shoulder"
## [488] "Lower body"
## [489] "Lower body, Flu"
## [490] "Lower body"
## [491] ""
## [492] "Upper body"
## [493] ""
## [494] ""
## [495] "Lower body"
## [496] ""
## [497] "Upper body"
## [498] ""
## [499] ""
## [500] ""
## [501] "Lower body, Shoulder"
## [502] ""
## [503] ""
## [504] "Shoulder"
## [505] ""
## [506] ""
## [507] "Lower body"
## [508] ""
## [509] ""
## [510] ""
## [511] "Flu, Undisclosed"
## [512] ""
## [513] ""
## [514] "Lower body, Undisclosed, Upper body"
## [515] "Upper body"
## [516] ""
## [517] "Upper body"
## [518] "Upper body, Facial"
## [519] "Lower body"
## [520] "Arm"
## [521] ""
## [522] "Upper body"
## [523] ""
## [524] ""
## [525] "Shoulder, Illness, Upper body"
## [526] "Upper body"
## [527] ""
## [528] "Finger, Illness, Upper body"
## [529] "Lower body, Illness, Finger"
## [530] ""
## [531] ""
## [532] ""
## [533] "Lower body"
```

```
## [534] "Flu"
## [535] ""
## [536] "Hand, Illness, Concussion"
## [537] ""
## [538] "Hip"
## [539] "Flu"
## [540] ""
## [541] ""
## [542] ""
## [543] ""
## [544] ""
## [545] "Concussion"
## [546] ""
## [547] ""
## [548] ""
## [549] "Lower body"
## [550] ""
## [551] ""
## [552] "Upper body"
## [553] "Lower body"
## [554] "Lower body, Upper body, Hand"
## [555] ""
## [556] "Lower body"
## [557] "Undisclosed"
## [558] "Undisclosed, Upper body, Lower body"
## [559] "Groin, Upper body, Undisclosed"
## [560] ""
## [561] "Upper body"
## [562] "Upper body, Finger"
## [563] "Upper body"
## [564] ""
## [565] ""
## [566] ""
## [567] "Undisclosed"
## [568] ""
## [569] ""
## [570] "Upper body"
## [571] "Knee"
## [572] ""
## [573] "Upper body, Lower body"
## [574] ""
## [575] ""
## [576] "Lower body, Leg"
## [577] ""
## [578] ""
## [579] ""
## [580] "Ankle"
## [581] "Undisclosed"
## [582] ""
## [583] "Upper body"
## [584] ""
## [585] ""
## [586] ""
## [587] ""
```

```
## [588] ""
## [589] "Undisclosed, Upper body, Appendectomy"
## [590] "Lower body, Ankle"
## [591] ""
## [592] "Knee, Ribs, Illness"
## [593] "Lower body, Hand, Undisclosed"
## [594] ""
## [595] "Upper body"
## [596] ""
## [597] "Lower body"
## [598] "Upper body"
## [599] "Shoulder"
## [600] ""
## [601] ""
## [602] ""
## [603] ""
## [604] ""
## [605] "Knee"
## [606] "Undisclosed, Lower body"
## [607] "Upper body"
## [608] ""
## [609] ""
## [610] "Upper body, Lower body"
## [611] "Undisclosed"
## [612] ""
## [613] "Finger"
## [614] "Lower body, Illness, Upper body"
## [615] "Undisclosed, Elbow"
## [616] "Ankle"
## [617] ""
## [618] ""
## [619] ""
## [620] ""
## [621] "Upper body"
## [622] ""
## [623] "Upper body"
## [624] ""
## [625] ""
## [626] ""
## [627] "Lower body"
## [628] "Ankle"
## [629] "Leg, Flu"
## [630] ""
## [631] ""
## [632] "Illness"
## [633] ""
## [634] "Knee"
## [635] "Upper body, Undisclosed"
## [636] "Upper body"
## [637] ""
## [638] "Illness"
## [639] ""
## [640] "Hernia"
## [641] ""
```

```
## [642] "Undisclosed"
## [643] "Undisclosed"
## [644] "Lower body, Upper body, Leg"
## [645] ""
## [646] "Back"
## [647] ""
## [648] "Concussion"
## [649] "Lower body, Undisclosed"
## [650] ""
## [651] "Foot"
## [652] "Lower body"
## [653] ""
## [654] ""
## [655] ""
## [656] "Upper body"
## [657] "Upper body"
## [658] "Illness, Lower body"
## [659] "Abdominal, Lower body"
## [660] "Abdominal, Lower body"
## [661] "Lower body"
## [662] ""
## [663] "Illness"
## [664] ""
## [665] ""
## [666] ""
## [667] "Oblique, Lower body, Upper body"
## [668] ""
## [669] "Foot"
## [670] ""
## [671] "Flu"
## [672] ""
## [673] ""
## [674] ""
## [675] ""
## [676] "Leg"
## [677] "Lower body"
## [678] "Knee"
## [679] ""
## [680] "Lower body"
## [681] "Upper body, Neck"
## [682] ""
## [683] "Knee"
## [684] ""
## [685] ""
## [686] ""
## [687] "Upper body, Hand"
## [688] ""
## [689] "Lower body, Upper body"
## [690] "Illness, Upper body"
## [691] "Lower body, Undisclosed, Upper body"
## [692] ""
## [693] "Finger, Upper body"
## [694] "Finger, Hand, Undisclosed, Upper body"
## [695] ""
```

```
## [696] ""
## [697] "Upper body, Lower body"
## [698] ""
## [699] ""
## [700] "Undisclosed"
## [701] ""
## [702] "Illness, Lower body"
## [703] "Shoulder"
## [704] "Lower body"
## [705] "Lower body"
## [706] ""
## [707] ""
## [708] ""
## [709] "Illness, Undisclosed, Upper body, Lower body"
## [710] ""
## [711] "Illness"
## [712] ""
## [713] ""
## [714] ""
## [715] "Concussion"
## [716] ""
## [717] "Elbow"
## [718] "Upper body"
## [719] ""
## [720] ""
## [721] "Lower body, Upper body"
## [722] ""
## [723] "Jaw"
## [724] "Illness"
## [725] ""
## [726] "Upper body"
## [727] ""
## [728] "Upper body"
## [729] "Foot, Lower body"
## [730] "Concussion, Hip"
## [731] ""
## [732] "Concussion"
## [733] ""
## [734] "Upper body"
## [735] "Eye, Upper body"
## [736] ""
## [737] ""
## [738] ""
## [739] "Upper body"
## [740] ""
## [741] "Upper body"
## [742] ""
## [743] ""
## [744] ""
## [745] "Illness, Upper body"
## [746] "Upper body, Undisclosed, Ankle"
## [747] "Lower body, Neck"
## [748] "Undisclosed"
## [749] ""
```

```
## [750] ""
## [751] "Upper body"
## [752] "Knee"
## [753] ""
## [754] "Illness, Undisclosed, Upper body"
## [755] ""
## [756] "Concussion"
## [757] ""
## [758] "Upper body"
## [759] "Lower body"
## [760] ""
## [761] ""
## [762] ""
## [763] ""
## [764] "Upper body, Undisclosed"
## [765] ""
## [766] "Undisclosed, Lower body, Upper body, Back"
## [767] "Concussion"
## [768] "Knee"
## [769] "Upper body, Lower body"
## [770] ""
## [771] "Concussion"
## [772] "Concussion"
## [773] "Upper body, Lower body, Illness"
## [774] ""
## [775] "Upper body"
## [776] "Knee"
## [777] "Upper body, Lower body, Illness"
## [778] "Illness, Upper body"
## [779] "Upper body"
## [780] ""
## [781] ""
## [782] ""
## [783] "Illness"
## [784] ""
## [785] "Neck, Lower body"
## [786] "Knee, Upper body"
## [787] "Abdominal"
## [788] ""
## [789] "Upper body"
## [790] ""
## [791] "Shoulder"
## [792] ""
## [793] "Illness, Wrist"
## [794] "Lower body"
## [795] "Upper body"
## [796] ""
## [797] ""
## [798] ""
## [799] "Undisclosed"
## [800] ""
## [801] "Undisclosed"
## [802] "Lower body, Upper body, Illness"
## [803] ""
```

```
## [804] ""
## [805] "Hamstring"
## [806] ""
## [807] ""
## [808] ""
## [809] ""
## [810] "Achilles, Upper body"
## [811] ""
## [812] ""
## [813] "Undisclosed"
## [814] "Lower body"
## [815] ""
## [816] "Flu, Upper body"
## [817] "Hand"
## [818] "Upper body"
## [819] "Lower body"
## [820] ""
## [821] "Lower body, Illness"
## [822] ""
## [823] ""
## [824] "Upper body"
## [825] "Illness"
## [826] ""
## [827] "Finger"
## [828] ""
## [829] "Groin, Lower body"
## [830] ""
## [831] ""
## [832] "Upper body"
## [833] ""
## [834] "Hip, Lower body, Ankle"
## [835] ""
## [836] "Illness, Undisclosed, Lower body, Upper body"
## [837] "Foot, Upper body"
## [838] ""
## [839] ""
## [840] ""
## [841] "Groin, Undisclosed, Soreness"
## [842] "Upper body"
## [843] "Undisclosed"
## [844] ""
## [845] "Lower body, Facial, Flu, Illness"
## [846] ""
## [847] ""
## [848] "Lower body"
## [849] ""
## [850] ""
## [851] ""
## [852] ""
## [853] ""
## [854] "Upper body, Lower body, Flu"
## [855] "Lower body"
## [856] "Lower body"
## [857] ""
```

```
## [858] ""
## [859] "Neck"
## [860] "Shoulder"
## [861] ""
## [862] ""
## [863] "Lower body, Flu"
## [864] "Upper body, Lower body"
## [865] ""
## [866] ""
## [867] "Lower body"
## [868] "Upper body, Lower body"
## [869] "Lower body"
## [870] ""
## [871] "Undisclosed"
## [872] ""
## [873] ""
## [874] "Lower body, Undisclosed, Shoulder"
## [875] ""
## [876] "Illness, Undisclosed"
## [877] ""
## [878] "Facial, Concussion"
## [879] "Ankle"
## [880] ""
## [881] "Illness"
## [882] ""
## [883] ""
## [884] "Leg, Hand"
## [885] "Undisclosed"
## [886] "Undisclosed"
## [887] "Lower body"
## [888] "Upper body"
```

```
AllSits_copy$Injuries
```

```
##   [1] "no_injury" "injury"    "no_injury" "injury"    "no_injury" "no_injury"
##   [7] "no_injury" "injury"    "no_injury" "injury"    "injury"    "injury"
##  [13] "no_injury" "injury"    "injury"    "injury"    "no_injury" "no_injury"
##  [19] "injury"    "injury"    "no_injury" "no_injury" "injury"    "no_injury"
##  [25] "no_injury" "injury"    "no_injury" "no_injury" "no_injury" "no_injury"
##  [31] "injury"    "injury"    "no_injury" "injury"    "injury"    "injury"
##  [37] "injury"    "injury"    "injury"    "no_injury" "injury"    "no_injury"
##  [43] "injury"    "injury"    "injury"    "no_injury" "injury"    "injury"
##  [49] "no_injury" "no_injury" "injury"    "injury"    "no_injury" "injury"
##  [55] "injury"    "injury"    "injury"    "injury"    "no_injury" "no_injury"
##  [61] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "injury"
##  [67] "injury"    "no_injury" "injury"    "injury"    "no_injury" "injury"
##  [73] "injury"    "injury"    "injury"    "injury"    "no_injury" "no_injury"
##  [79] "no_injury" "injury"    "injury"    "no_injury" "injury"    "no_injury"
##  [85] "injury"    "injury"    "no_injury" "injury"    "injury"    "no_injury"
##  [91] "injury"    "no_injury" "injury"    "no_injury" "injury"    "injury"
##  [97] "injury"    "injury"    "injury"    "injury"    "injury"    "no_injury"
## [103] "no_injury" "injury"    "no_injury" "no_injury" "injury"    "no_injury"
## [109] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [115] "injury"    "no_injury" "no_injury" "injury"    "injury"    "injury"
## [121] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "injury"
```

```
## [127] "no_injury" "injury"    "injury"    "injury"    "no_injury" "injury"
## [133] "no_injury" "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [139] "injury"    "injury"    "injury"    "injury"    "no_injury" "injury"
## [145] "injury"    "injury"    "no_injury" "injury"    "injury"    "injury"
## [151] "injury"    "injury"    "injury"    "no_injury" "injury"    "injury"
## [157] "no_injury" "injury"    "no_injury" "injury"    "no_injury" "no_injury"
## [163] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [169] "injury"    "injury"    "injury"    "injury"    "injury"    "no_injury"
## [175] "no_injury" "no_injury" "injury"    "injury"    "injury"    "injury"
## [181] "no_injury" "no_injury" "injury"    "injury"    "no_injury" "injury"
## [187] "injury"    "injury"    "no_injury" "injury"    "no_injury" "injury"
## [193] "no_injury" "no_injury" "injury"    "injury"    "no_injury" "injury"
## [199] "injury"    "no_injury" "injury"    "injury"    "no_injury" "no_injury"
## [205] "no_injury" "injury"    "injury"    "injury"    "no_injury" "no_injury"
## [211] "injury"    "injury"    "injury"    "injury"    "no_injury" "no_injury"
## [217] "no_injury" "injury"    "injury"    "injury"    "no_injury" "injury"
## [223] "injury"    "injury"    "injury"    "injury"    "injury"    "injury"
## [229] "no_injury" "injury"    "injury"    "no_injury" "injury"    "injury"
## [235] "injury"    "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [241] "injury"    "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [247] "injury"    "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [253] "no_injury" "injury"    "no_injury" "no_injury" "injury"    "injury"
## [259] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [265] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [271] "no_injury" "injury"    "injury"    "no_injury" "injury"    "no_injury"
## [277] "injury"    "injury"    "injury"    "no_injury" "injury"    "injury"
## [283] "no_injury" "injury"    "injury"    "no_injury" "no_injury" "injury"
## [289] "injury"    "injury"    "no_injury" "no_injury" "injury"    "no_injury"
## [295] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "injury"
## [301] "injury"    "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [307] "injury"    "injury"    "injury"    "injury"    "injury"    "no_injury"
## [313] "injury"    "no_injury" "no_injury" "no_injury" "injury"    "injury"
## [319] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "injury"
## [325] "injury"    "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [331] "no_injury" "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [337] "no_injury" "injury"    "injury"    "injury"    "injury"    "no_injury"
## [343] "injury"    "injury"    "no_injury" "no_injury" "injury"    "injury"
## [349] "injury"    "injury"    "no_injury" "no_injury" "injury"    "injury"
## [355] "injury"    "no_injury" "injury"    "injury"    "no_injury" "no_injury"
## [361] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [367] "no_injury" "no_injury" "no_injury" "no_injury" "injury"    "injury"
## [373] "no_injury" "injury"    "no_injury" "injury"    "no_injury" "no_injury"
## [379] "injury"    "injury"    "injury"    "injury"    "no_injury" "injury"
## [385] "no_injury" "injury"    "no_injury" "no_injury" "injury"    "no_injury"
## [391] "no_injury" "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [397] "injury"    "no_injury" "no_injury" "no_injury" "no_injury" "no_injury"
## [403] "no_injury" "no_injury" "no_injury" "no_injury" "no_injury" "injury"
## [409] "injury"    "no_injury" "no_injury" "no_injury" "injury"    "no_injury"
## [415] "no_injury" "injury"    "no_injury" "injury"    "injury"    "no_injury"
## [421] "no_injury" "injury"    "no_injury" "injury"    "injury"    "injury"
## [427] "injury"    "injury"    "injury"    "injury"    "injury"    "injury"
## [433] "no_injury" "no_injury" "no_injury" "no_injury" "injury"    "injury"
## [439] "injury"    "no_injury" "injury"    "no_injury" "injury"    "injury"
## [445] "injury"    "no_injury" "injury"    "no_injury" "injury"    "injury"
```

```
## [451] "injury"    "injury"    "injury"    "injury"    "no_injury" "injury"
## [457] "no_injury" "injury"    "no_injury" "no_injury" "no_injury" "injury"
## [463] "injury"    "no_injury" "no_injury" "injury"    "injury"    "injury"
## [469] "injury"    "no_injury" "no_injury" "no_injury" "no_injury" "no_injury"
## [475] "injury"    "injury"    "injury"    "no_injury" "no_injury" "injury"
## [481] "injury"    "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [487] "injury"    "injury"    "injury"    "injury"    "no_injury" "injury"
## [493] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [499] "no_injury" "no_injury" "injury"    "no_injury" "no_injury" "injury"
## [505] "no_injury" "no_injury" "injury"    "no_injury" "no_injury" "no_injury"
## [511] "injury"    "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [517] "injury"    "injury"    "injury"    "injury"    "no_injury" "injury"
## [523] "no_injury" "no_injury" "injury"    "injury"    "no_injury" "injury"
## [529] "injury"    "no_injury" "no_injury" "no_injury" "injury"    "injury"
## [535] "no_injury" "injury"    "no_injury" "injury"    "injury"    "no_injury"
## [541] "no_injury" "no_injury" "no_injury" "no_injury" "injury"    "no_injury"
## [547] "no_injury" "no_injury" "injury"    "no_injury" "no_injury" "injury"
## [553] "injury"    "injury"    "no_injury" "injury"    "injury"    "injury"
## [559] "injury"    "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [565] "no_injury" "no_injury" "injury"    "no_injury" "no_injury" "injury"
## [571] "injury"    "no_injury" "injury"    "no_injury" "no_injury" "injury"
## [577] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [583] "injury"    "no_injury" "no_injury" "no_injury" "no_injury" "no_injury"
## [589] "injury"    "injury"    "no_injury" "injury"    "injury"    "no_injury"
## [595] "injury"    "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [601] "no_injury" "no_injury" "no_injury" "no_injury" "injury"    "injury"
## [607] "injury"    "no_injury" "no_injury" "injury"    "injury"    "no_injury"
## [613] "injury"    "injury"    "injury"    "injury"    "no_injury" "no_injury"
## [619] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [625] "no_injury" "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [631] "no_injury" "injury"    "no_injury" "injury"    "injury"    "injury"
## [637] "no_injury" "injury"    "no_injury" "injury"    "no_injury" "injury"
## [643] "injury"    "injury"    "no_injury" "injury"    "no_injury" "injury"
## [649] "injury"    "no_injury" "injury"    "injury"    "no_injury" "no_injury"
## [655] "no_injury" "injury"    "injury"    "injury"    "injury"    "injury"
## [661] "injury"    "no_injury" "injury"    "no_injury" "no_injury" "no_injury"
## [667] "injury"    "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [673] "no_injury" "no_injury" "no_injury" "injury"    "injury"    "injury"
## [679] "no_injury" "injury"    "injury"    "no_injury" "injury"    "no_injury"
## [685] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "injury"
## [691] "injury"    "no_injury" "injury"    "injury"    "no_injury" "no_injury"
## [697] "injury"    "no_injury" "no_injury" "injury"    "no_injury" "injury"
## [703] "injury"    "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [709] "injury"    "no_injury" "injury"    "no_injury" "no_injury" "no_injury"
## [715] "injury"    "no_injury" "injury"    "injury"    "no_injury" "no_injury"
## [721] "injury"    "no_injury" "injury"    "injury"    "no_injury" "injury"
## [727] "no_injury" "injury"    "injury"    "injury"    "no_injury" "injury"
## [733] "no_injury" "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [739] "injury"    "no_injury" "injury"    "no_injury" "no_injury" "no_injury"
## [745] "injury"    "injury"    "injury"    "injury"    "no_injury" "no_injury"
## [751] "injury"    "injury"    "no_injury" "injury"    "no_injury" "injury"
## [757] "no_injury" "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [763] "no_injury" "injury"    "no_injury" "injury"    "injury"    "injury"
## [769] "injury"    "no_injury" "injury"    "injury"    "injury"    "no_injury"
```

```
## [775] "injury"    "injury"    "injury"    "injury"    "injury"    "no_injury"
## [781] "no_injury" "no_injury" "injury"    "no_injury" "injury"    "injury"
## [787] "injury"    "no_injury" "injury"    "no_injury" "injury"    "no_injury"
## [793] "injury"    "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [799] "injury"    "no_injury" "injury"    "injury"    "no_injury" "no_injury"
## [805] "injury"    "no_injury" "no_injury" "no_injury" "no_injury" "injury"
## [811] "no_injury" "no_injury" "injury"    "injury"    "no_injury" "injury"
## [817] "injury"    "injury"    "injury"    "no_injury" "injury"    "no_injury"
## [823] "no_injury" "injury"    "injury"    "no_injury" "injury"    "no_injury"
## [829] "injury"    "no_injury" "no_injury" "injury"    "no_injury" "injury"
## [835] "no_injury" "injury"    "injury"    "no_injury" "no_injury" "no_injury"
## [841] "injury"    "injury"    "injury"    "no_injury" "injury"    "no_injury"
## [847] "no_injury" "injury"    "no_injury" "no_injury" "no_injury" "no_injury"
## [853] "no_injury" "injury"    "injury"    "injury"    "no_injury" "no_injury"
## [859] "injury"    "injury"    "no_injury" "no_injury" "injury"    "injury"
## [865] "no_injury" "no_injury" "injury"    "injury"    "injury"    "no_injury"
## [871] "injury"    "no_injury" "no_injury" "injury"    "no_injury" "injury"
## [877] "no_injury" "injury"    "injury"    "no_injury" "injury"    "no_injury"
## [883] "no_injury" "injury"    "injury"    "injury"    "injury"    "injury"
```

```
AllSits_copy$TOI.GP[AllSits$TOI.GP >= median(AllSits$TOI.GP)] = "big minutes"
AllSits_copy$TOI.GP[AllSits$TOI.GP < median(AllSits$TOI.GP)] = "small minutes"
```

Igračima koji su u gornjoj polovici kod minutaže postavili smo minute u big minutes, ostalima u small minutes

```
table_inj_min = table(AllSits_copy$Injuries, AllSits_copy$TOI.GP)
table_inj_min <- addmargins(table_inj_min)
table_inj_min
```

```
##
##             big minutes small minutes Sum
##   injury            268           209 477
##   no_injury         177           234 411
##   Sum               445           443 888
```

Moramo provjeriti još jednu stvar prije provedbe samog chisq.test() testa: je li frekvencija svakog razreda veća ili jednaka 5.

```
for (col_names in colnames(table_inj_min)){
  for (row_names in rownames(table_inj_min)){
    if (!(row_names == 'Sum' | col_names == 'Sum') ){
      cat('Očekivane frekvencije za razred ',col_names,'-',row_names,': ',(table_inj_min[row_names,'Sum
    }
  }
}
```

```
## Ocekivane frekvencije za razred  big minutes - injury :  239.0372
## Ocekivane frekvencije za razred  big minutes - no_injury :  205.9628
## Ocekivane frekvencije za razred  small minutes - injury :  237.9628
## Ocekivane frekvencije za razred  small minutes - no_injury :  205.0372
```

Kao što vidimo, uvjet je zadovoljen. Ostalo je provesti sami test.

```
chisq.test(tbl)
```

```
##
##  Pearson's Chi-squared test
##
```

```
## data:  tbl
## X-squared = 107.05, df = 6, p-value < 2.2e-16
```

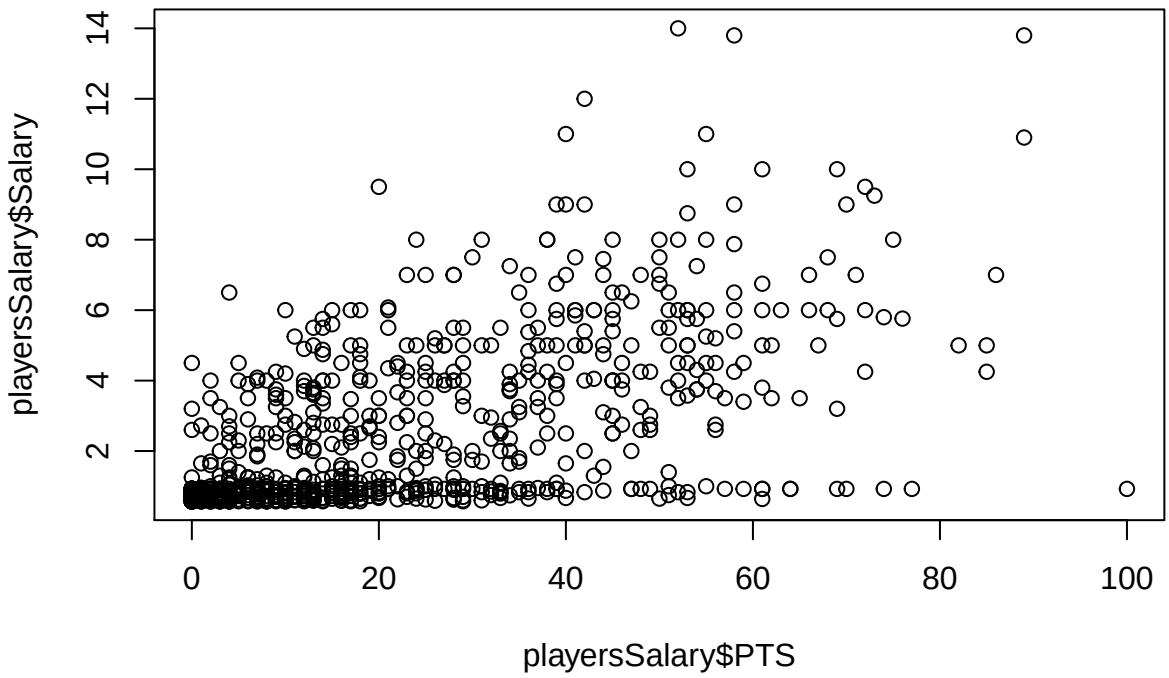
###ZAKLJUČAK

S obzirom na jako malu p vrijednost odbacujemo H0 u korist H1, odnosno pokazali smo da minute igranja i ozlijeđivanje nisu nezavisne varijable

U školi/fakultetu često pišemo ispite iz kojih dobivamo različite ocjene.Na našu ocjenu utječe dosta faktora, od težine zadataka, gradiva koje pišemo, vremena utrošenog na savladavanje tog gradiva i slično.Istu stvar možmo primjeniti i na plaće igrača u našem podatkovom skupu.Dosta faktora može utjecati na plaću igrača, stoga će naš sljedeći zadataka biti:Ispitati vezu između plaće igrača s jedne strane i određenih varijabla s druge strane.U našem konkretnom slučaju uzet ćemo tri varijable koje bi mogle utjecati na plaću, a to su:PTS,minute i izbori u prvu,drugu ili treću najbolju šestorku lige.

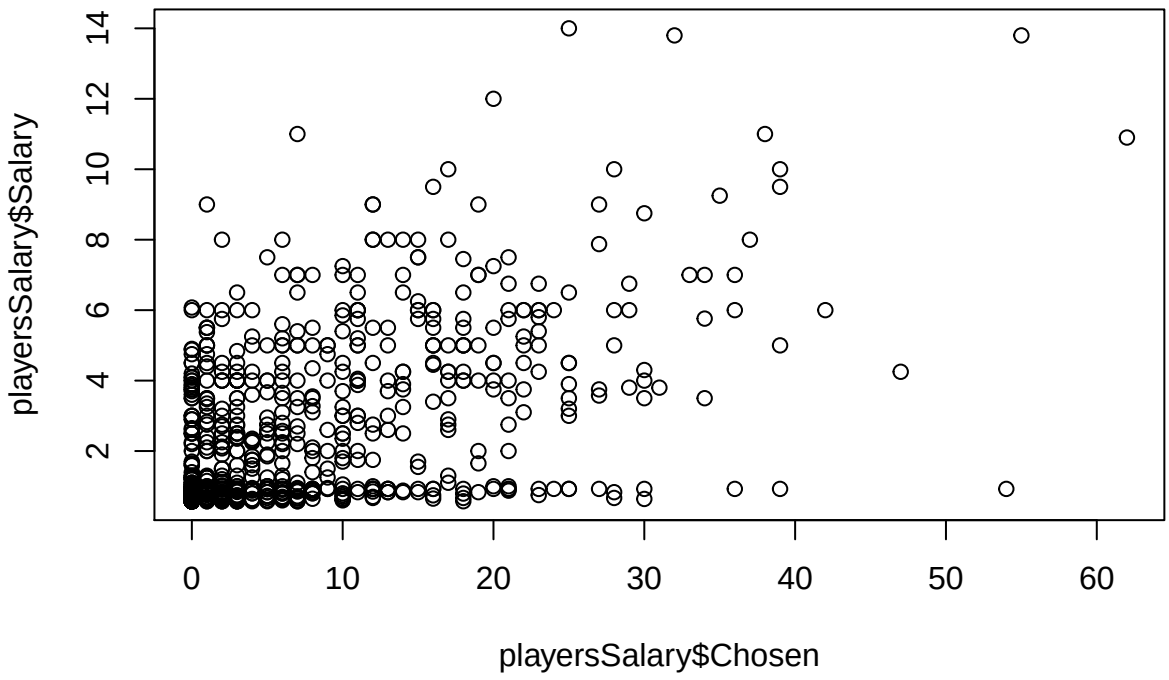
###ZADATAK 10 ###PLAĆE I PTS, PLAĆE I IZBOR U STARS TE PLAĆE I MINUTE IGRANJA
Ovdje ćemo koristiti model linearne regresije

```
require(stringr)
playersSalary$PTS <- AllSits[ ,c("PTS")]
playersSalary$TOI.GP <- AllSits[ ,c("TOI.GP")]
playersSalary$X1st <- AllSits[ ,c("X1st")]
playersSalary$X2nd <- AllSits[ ,c("X2nd")]
playersSalary$X3rd <- AllSits[ ,c("X3rd")]
playersSalary$PIM <- AllSits[ ,c("PIM")]
playersSalary$SCA <- AllSits[ ,c("SCA")]

playersSalary <- playersSalary[!is.na(playersSalary$Salary),]


#umjesto NA stavljamo 0 jer da je dobar igrac ne bi bio NA
playersSalary$X1st <- ifelse(is.na(playersSalary$X1st), 0, playersSalary$X1st)
playersSalary$X2nd <- ifelse(is.na(playersSalary$X2nd), 0, playersSalary$X2nd)
playersSalary$X3rd <- ifelse(is.na(playersSalary$X3rd), 0, playersSalary$X3rd)

playersSalary$Chosen <- playersSalary$X1st * 3 + playersSalary$X2nd * 2 + playersSalary$X3rd


plot(playersSalary$PTS, playersSalary$Salary)
```

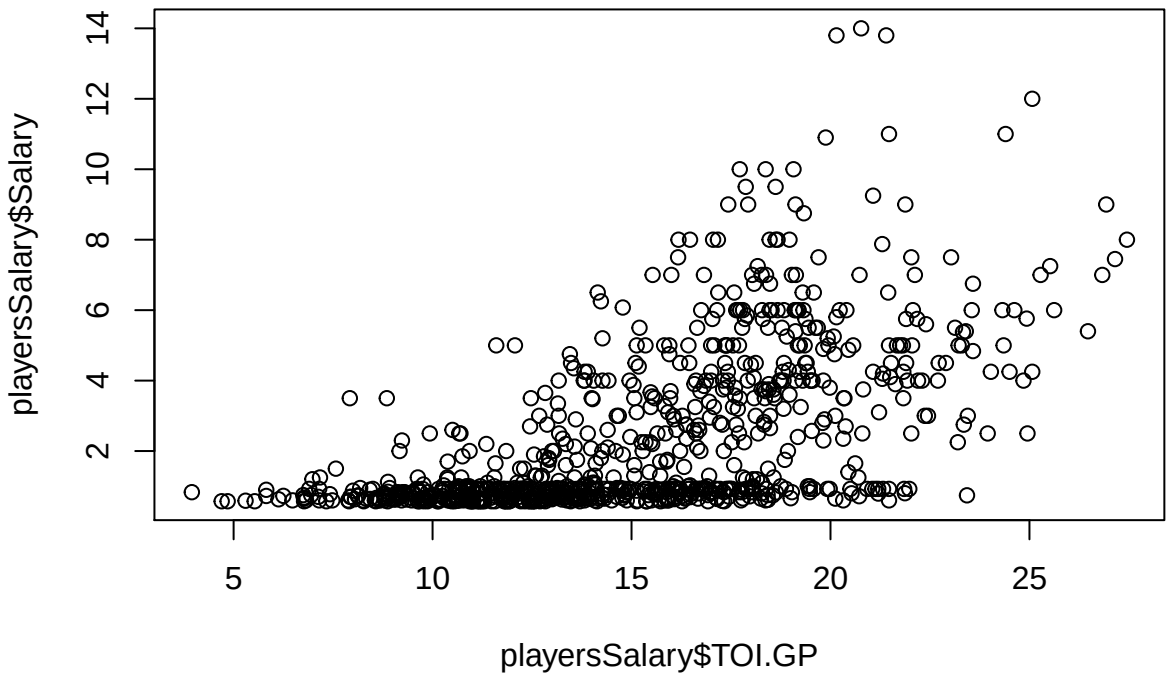



```
plot(playersSalary$SCA, playersSalary$Salary)
```


```
plot(playersSalary$Chosen, playersSalary$Salary)
```



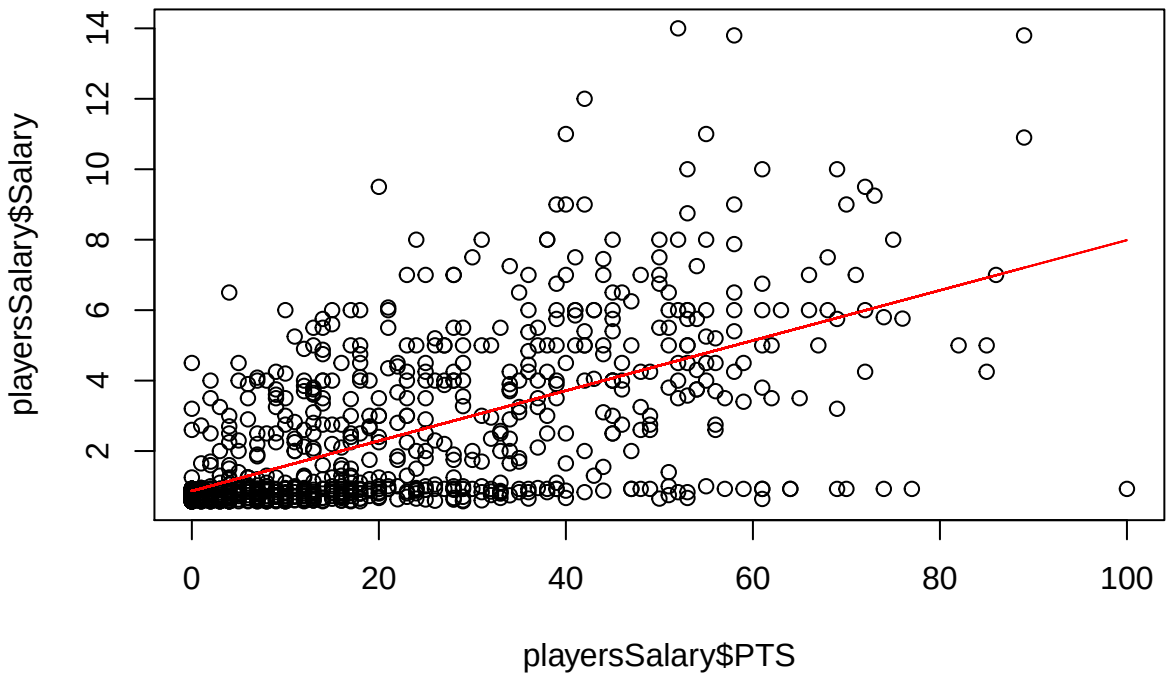

```
plot(playersSalary$TOI.GP, playersSalary$Salary)
```



Na osnovu scatter plotova možemo pretpostaviti da varijable koje promatramo pozitivno utječu na izlaznu varijablu(plaću igrača). Kako bismo ispitali utjecaje varijabli na izlaznu varijablu, procijenjujemo model jednostavne regresije za svaku nezavisnu varijablu. Zavisna varijabla regresijskih modela bit će plaća(salary).

Linearna regresija

```
#linearni model plaće igrača (Salary) i poena (golovi + asistencije = PTS)
fit.PTS = lm(playersSalary$Salary~playersSalary$PTS,data=playersSalary)

plot(playersSalary$PTS,playersSalary$Salary)+ #grafički prikaz podataka
lines(playersSalary$PTS,fit.PTS$fitted.values,col='red') #grafički prikaz procijenjenih vrijednosti iz
```

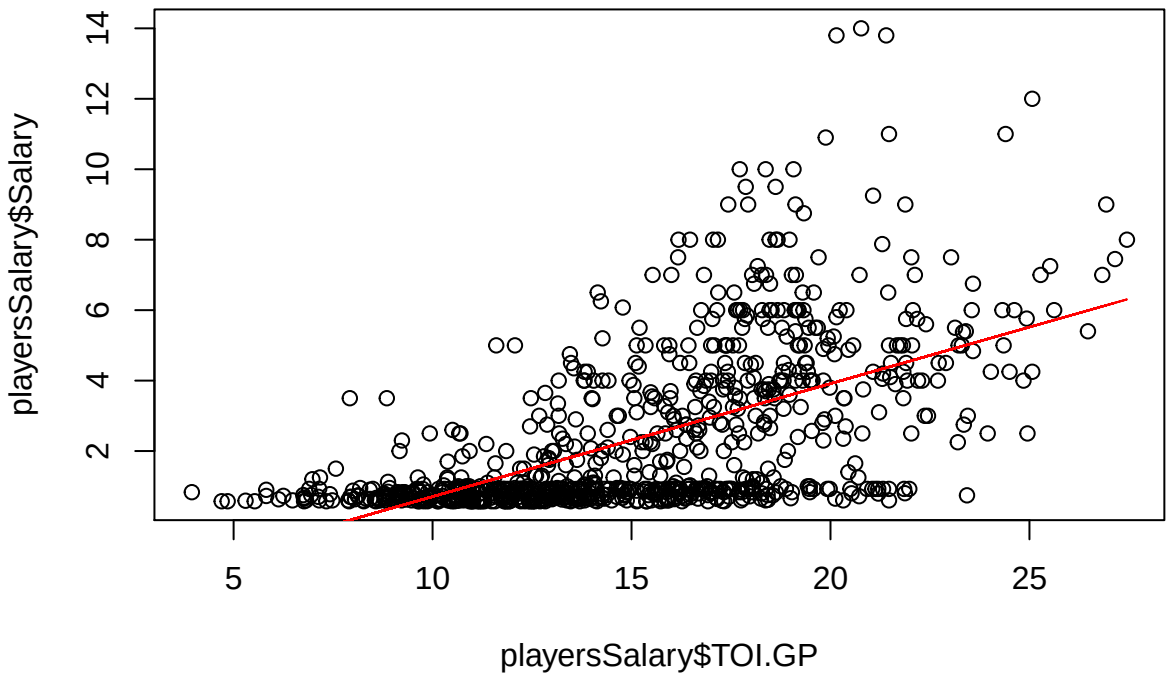



```
## integer(0)
```

```
#linearni model plaće igrača (Salary) i odabira u tri najbolja igrača
fit.chosen = lm(playersSalary$Salary~playersSalary$Chosen,data=playersSalary)

plot(playersSalary$Chosen,playersSalary$Salary)+ #graficki prikaz podataka
lines(playersSalary$Chosen,fit.chosen$fitted.values,col='red') #graficki prikaz procijenjenih vrijednos
```


```
## integer(0)
```

```
#linearni model plaće igrača (Salary) i minuta po utakmici
fit.minutes = lm(playersSalary$Salary~playersSalary$TOI.GP,data=playersSalary)

plot(playersSalary$TOI.GP,playersSalary$Salary)+ #graficki prikaz podataka
lines(playersSalary$TOI.GP,fit.minutes$fitted.values,col='red') #graficki prikaz procijenjenih vrijedno
```




```
## integer(0)
```

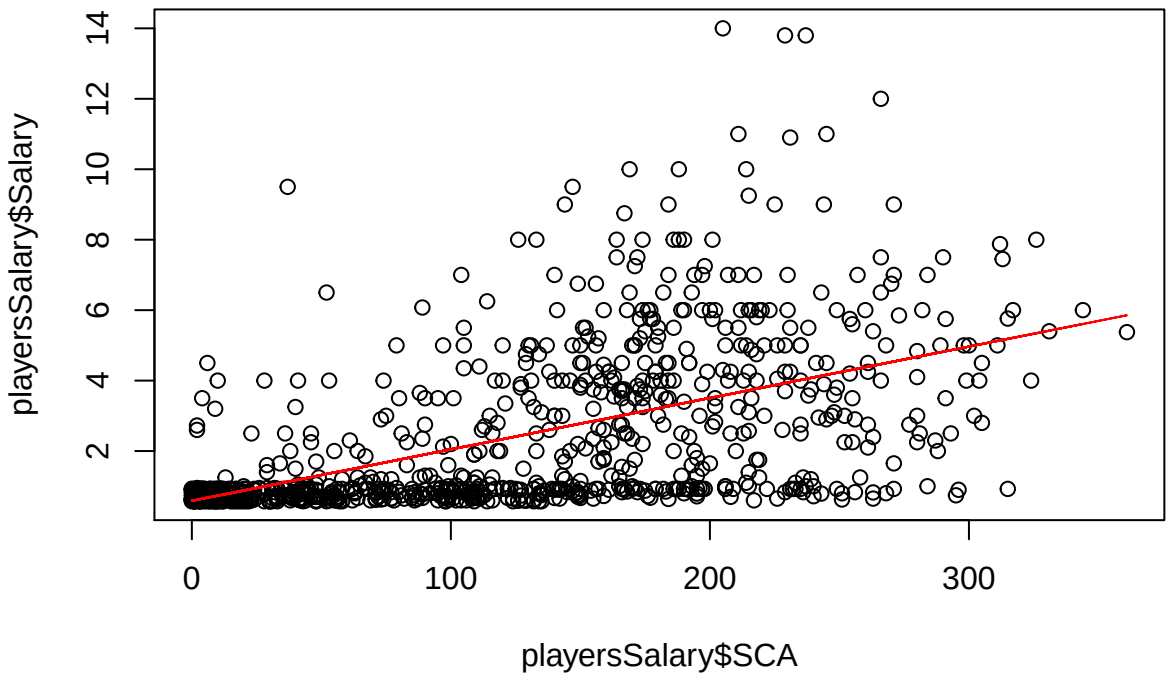
```
#linearni model plaće igrača (Salary) i šansi za gol kreiranih dok je igrač bio u igri
fit.SCA = lm(playersSalary$Salary~playersSalary$SCA,data=playersSalary)

length(playersSalary$SCA)
```

```
## [1] 874
```

```
length(fit.SCA$fitted.values)
```

```
## [1] 873
```

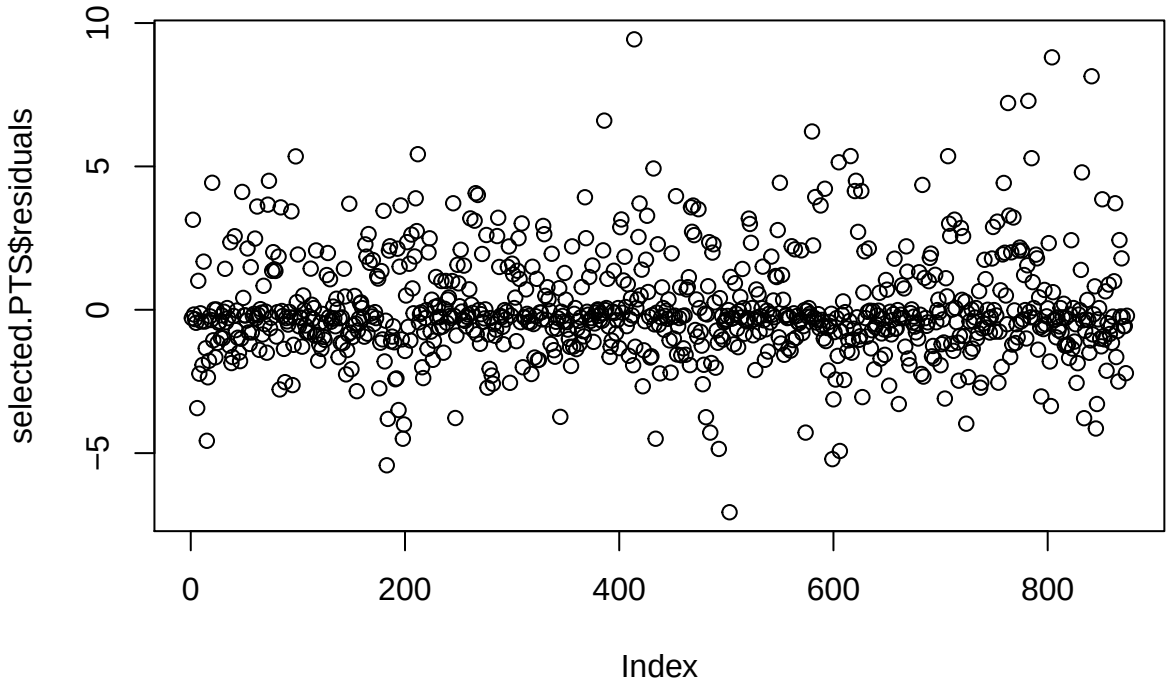
```
plot(playersSalary$SCA,playersSalary$Salary)+ #grafički prikaz podataka
lines(playersSalary$SCA[!is.na(playersSalary$SCA)],fit.SCA$fitted.values,col='red') #grafički prikaz pr
```



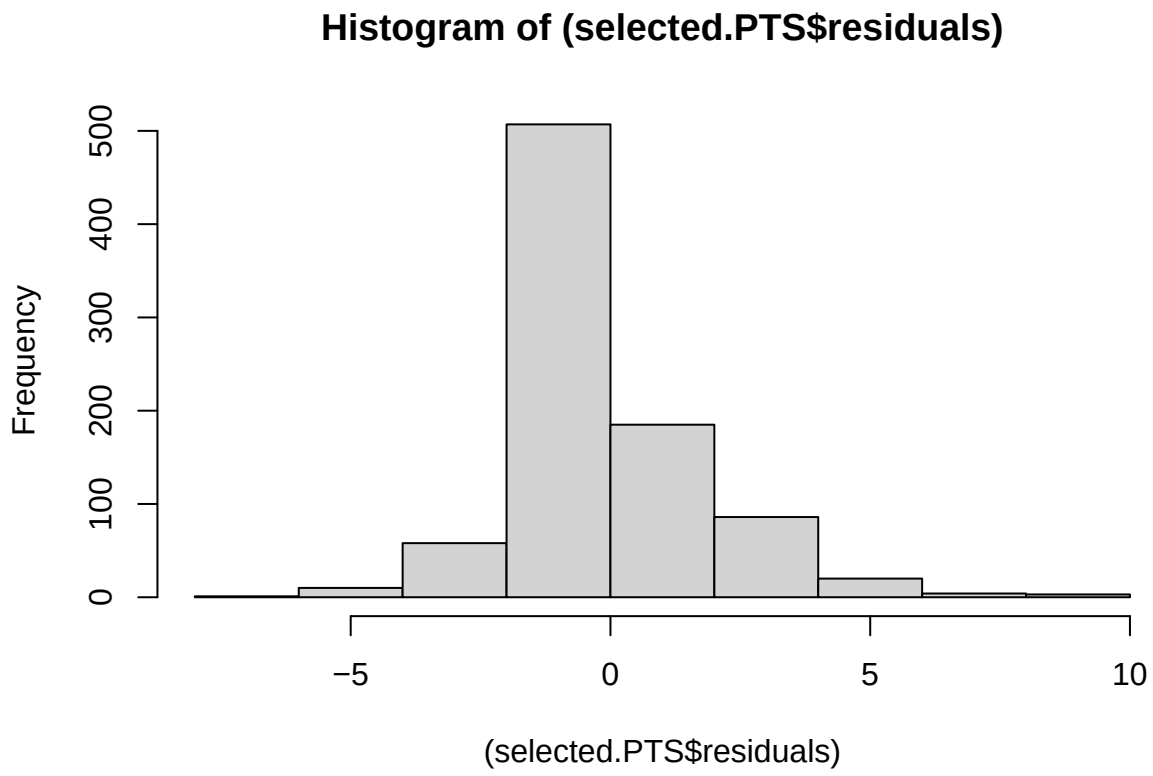

```
## integer(0)
```

Iz nagiba pravaca linearne regresije možemo vidjeti da su naše pretpostavke o efektima promatranih varijabli bile opravdane. Kako bi dalje analizirali dobivene modele, moramo provjeriti da naše pretpostavke normalnosti reziduala i homogenosti varijance nisu pogrešne.

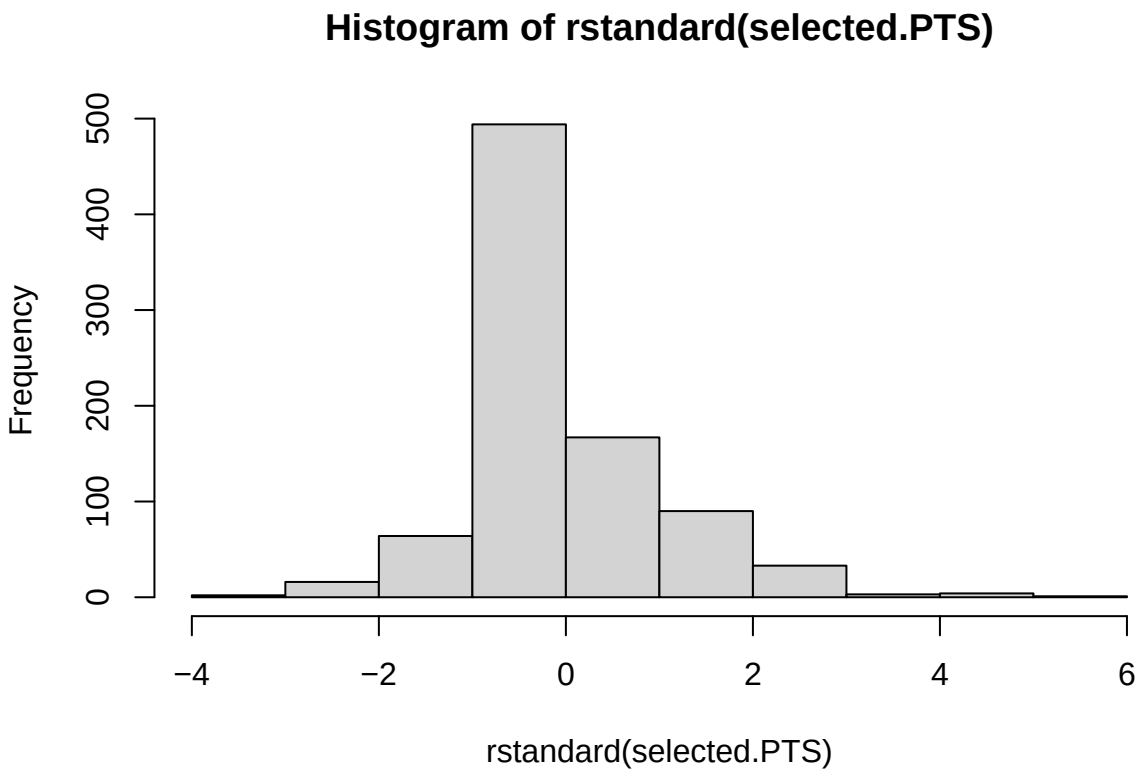
```
####  Bodovi
selected.PTS = fit.PTS

plot(selected.PTS$residuals) #gledajuci reziduale na ovaj nacin tesko je suditi o normalnosti
```



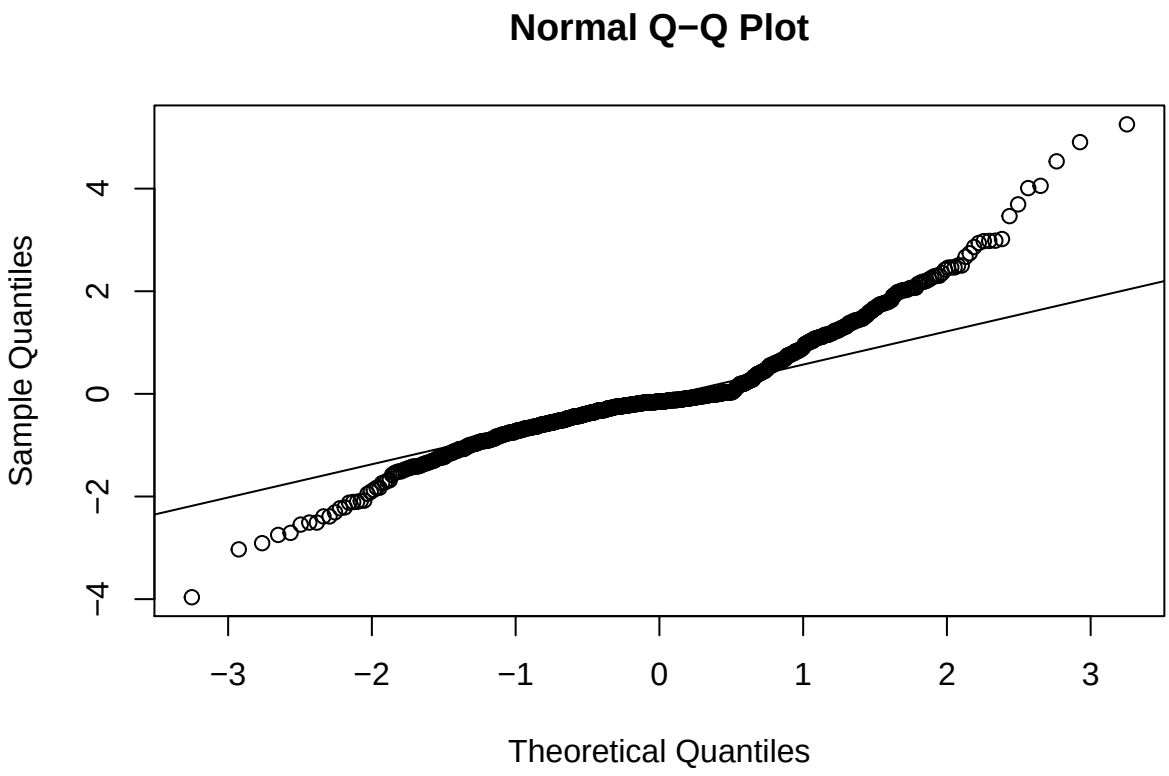

```
#histogram je vrlo interpretativan
hist((selected.PTS$residuals))
```



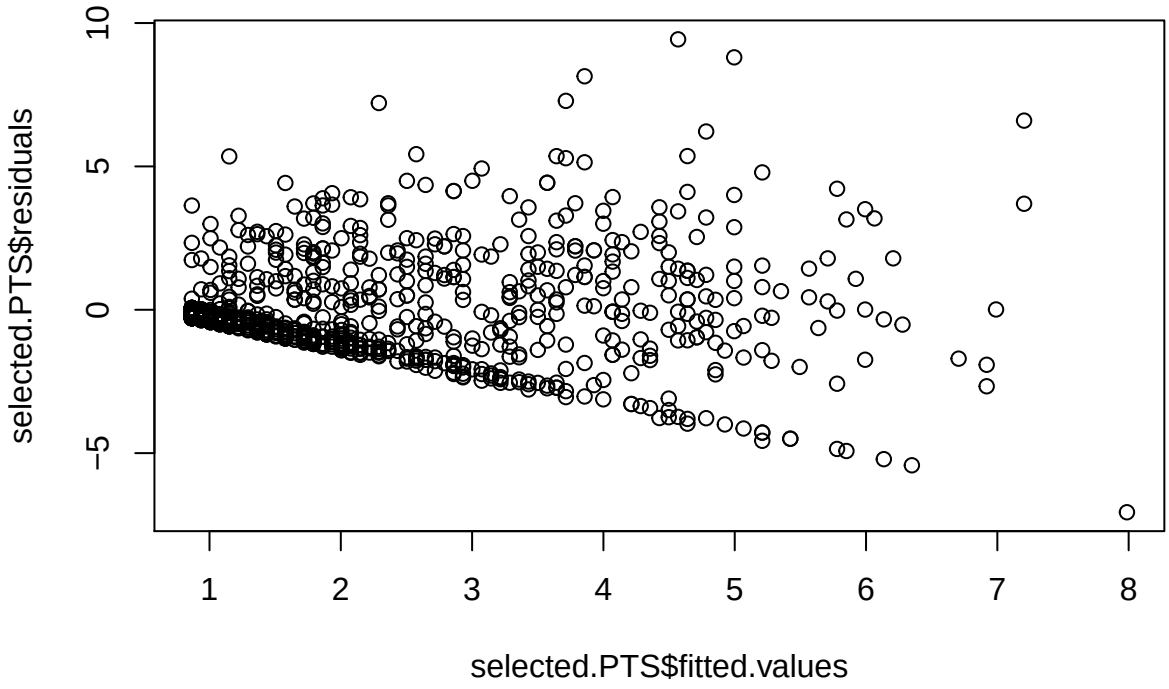

```
hist(rstandard(selected.PTS))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.PTS))
qqline(rstandard(selected.PTS))
```




```
plot(selected.PTS$fitted.values,selected.PTS$residuals) #reziduale je dobro prikazati u ovisnosti o pro
```




```
#KS test na normalnost
ks.test(rstandard(fit.PTS),'pnorm')
```

```
## Warning in ks.test(rstandard(fit.PTS), "pnorm"): ties should not be present for
## the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.PTS)
## D = 0.18129, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

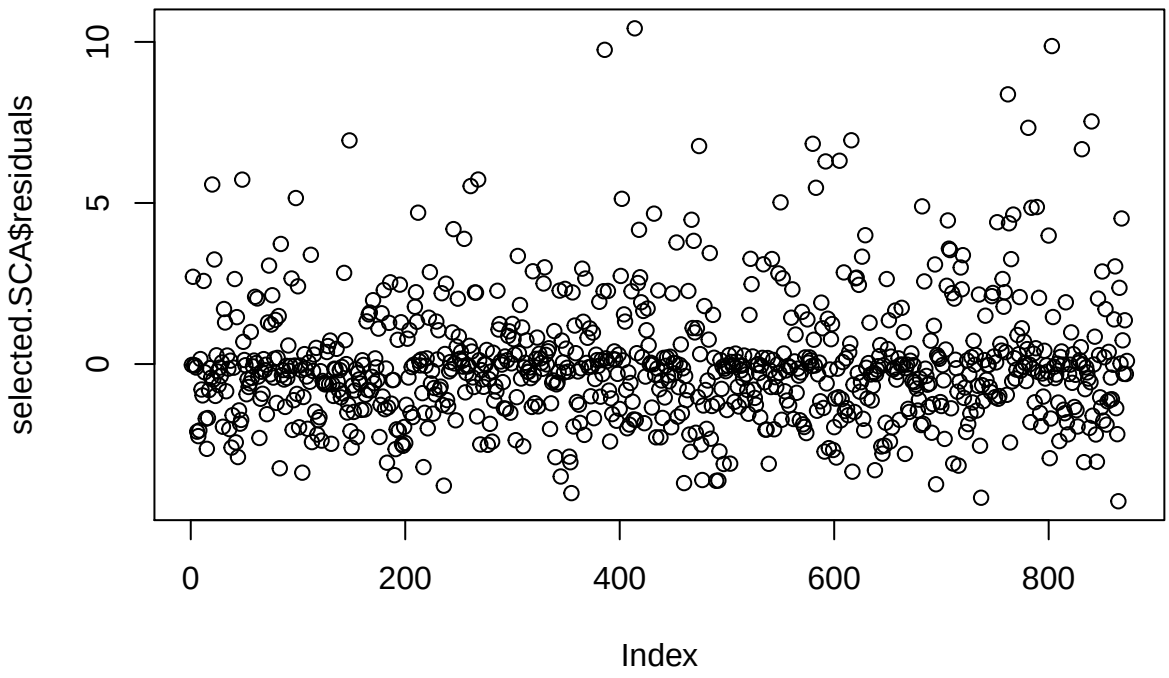
```
require(nortest)
```

```
## Loading required package: nortest
```

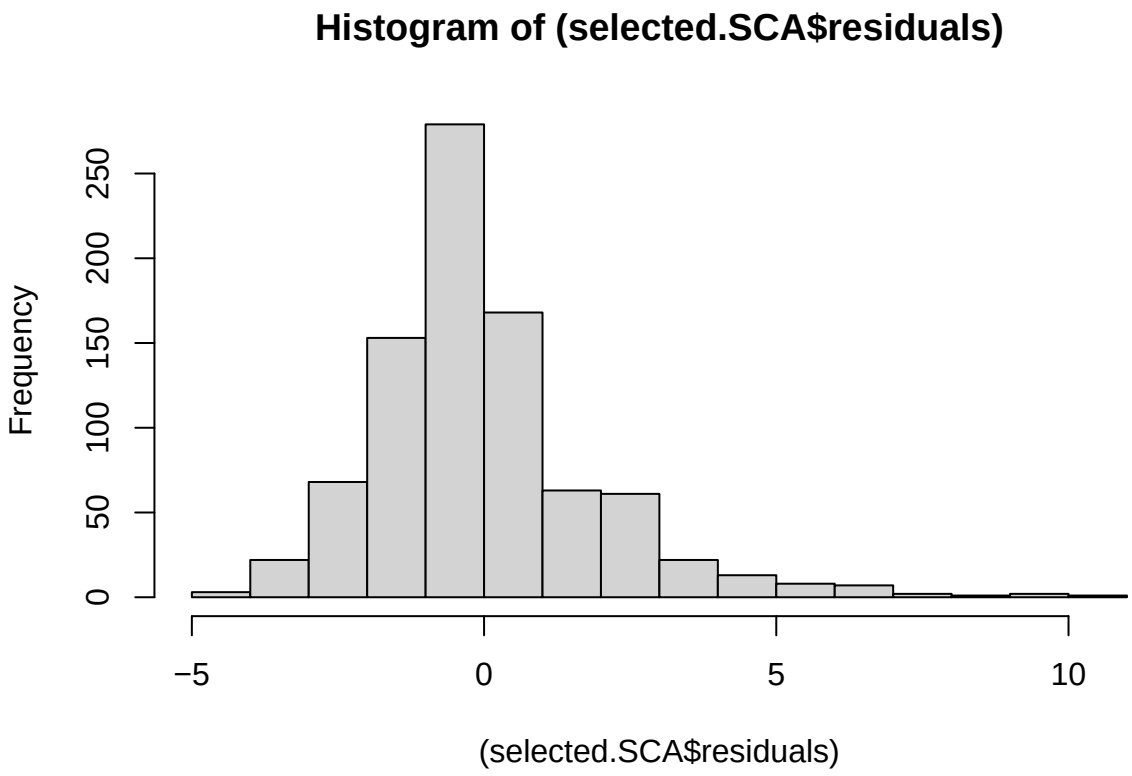
```
lillie.test(rstandard(fit.PTS))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.PTS)
## D = 0.18129, p-value < 2.2e-16
```

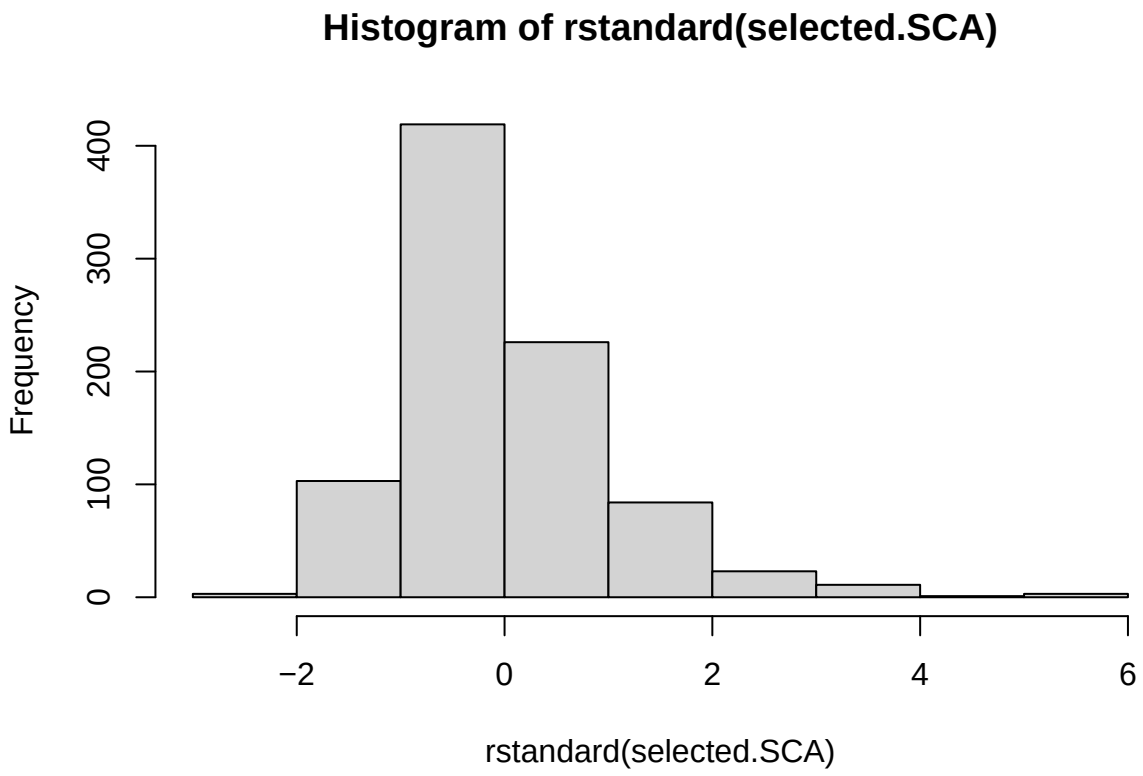
```
#### Bodovi
selected.SCA = fit.SCA

plot(selected.SCA$residuals)
```



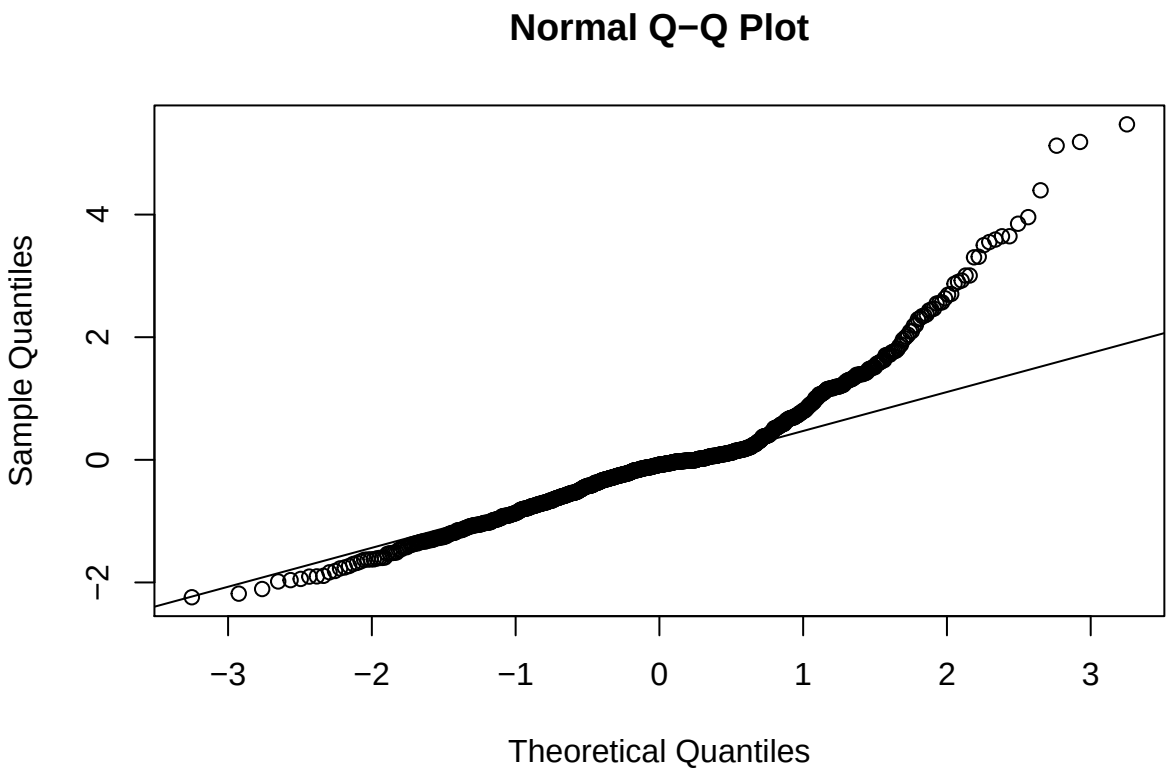

```
#histogram je vrlo interpretativan
hist((selected.SCA$residuals))
```



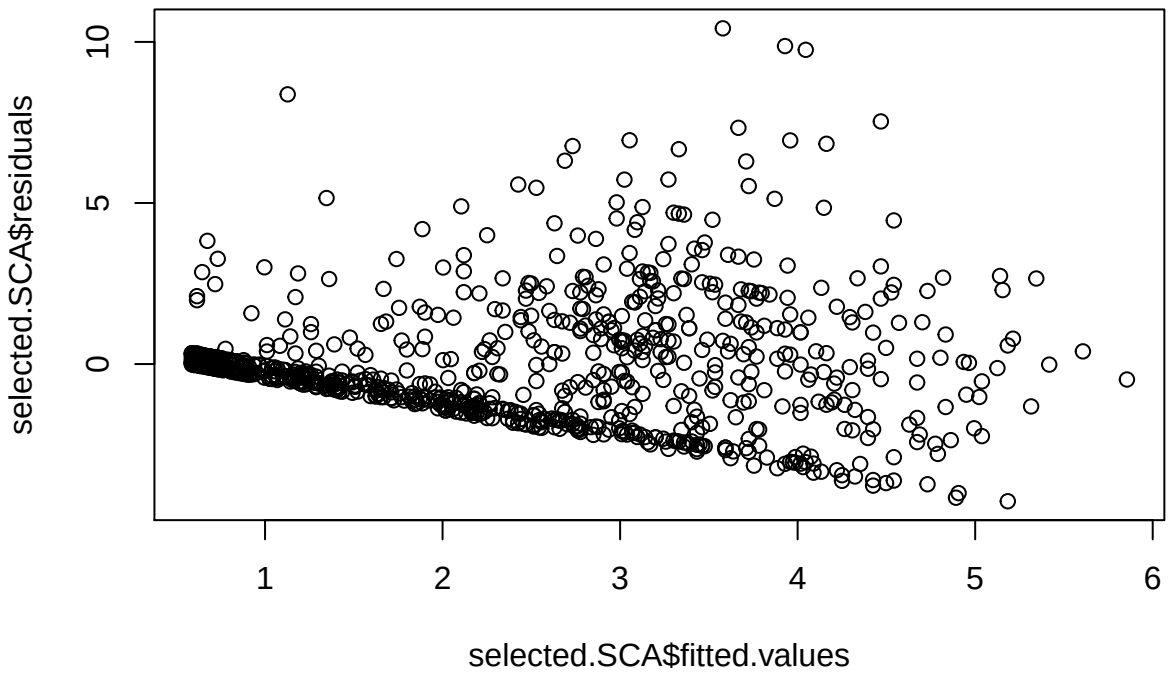

```
hist(rstandard(selected.SCA))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.SCA))
qqline(rstandard(selected.SCA))
```




```
plot(selected.SCA$fitted.values,selected.SCA$residuals) #reziduale je dobro prikazati u ovisnosti o pro
```




```
#KS test na normalnost
ks.test(rstandard(fit.SCA),'pnorm')
```

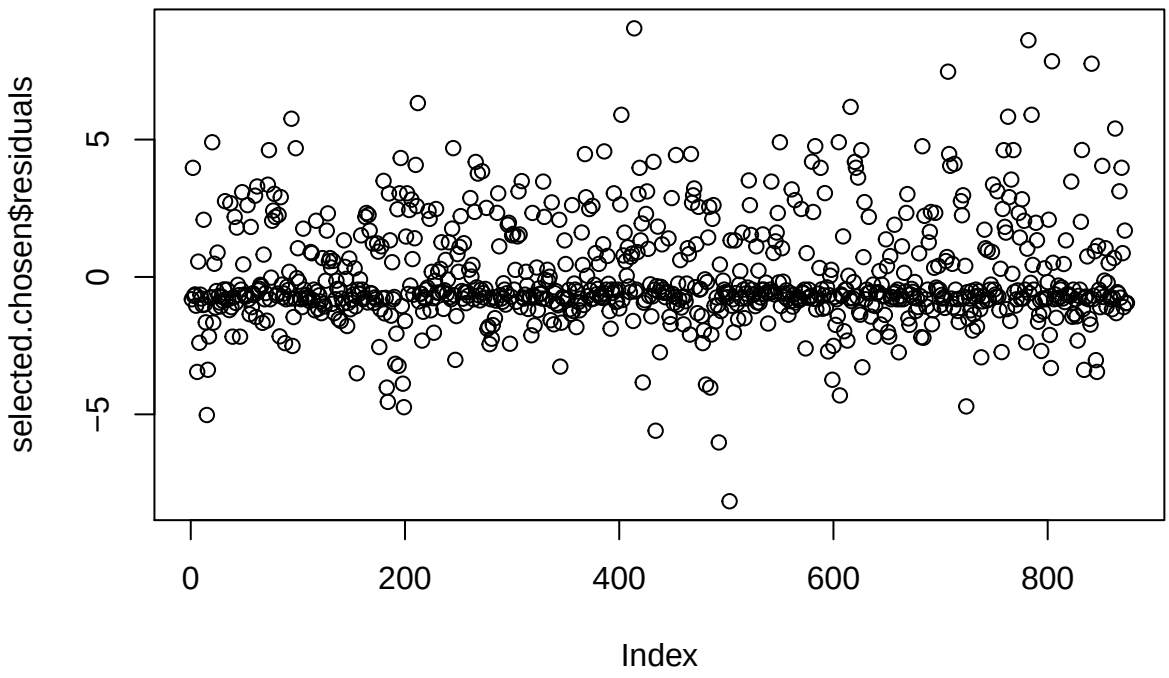
```
## Warning in ks.test(rstandard(fit.SCA), "pnorm"): ties should not be present for
## the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.SCA)
## D = 0.15584, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

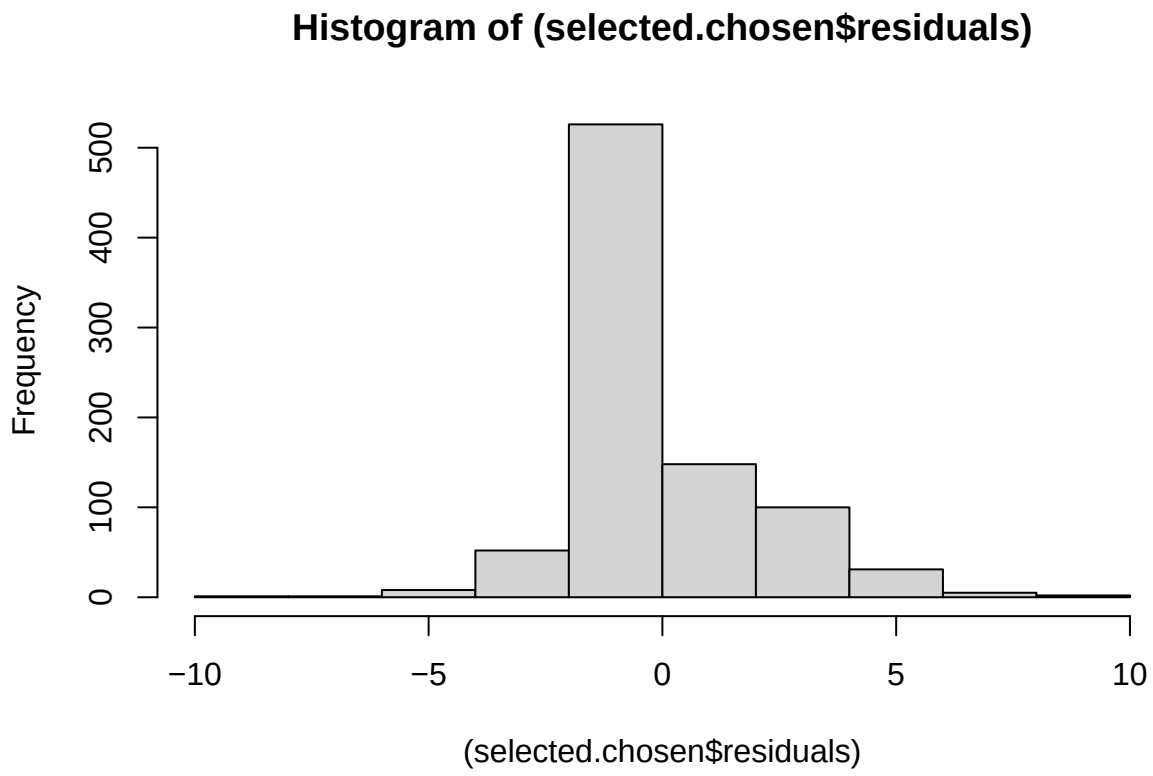
```
require(nortest)
lillie.test(rstandard(fit.SCA))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.SCA)
## D = 0.15588, p-value < 2.2e-16
```

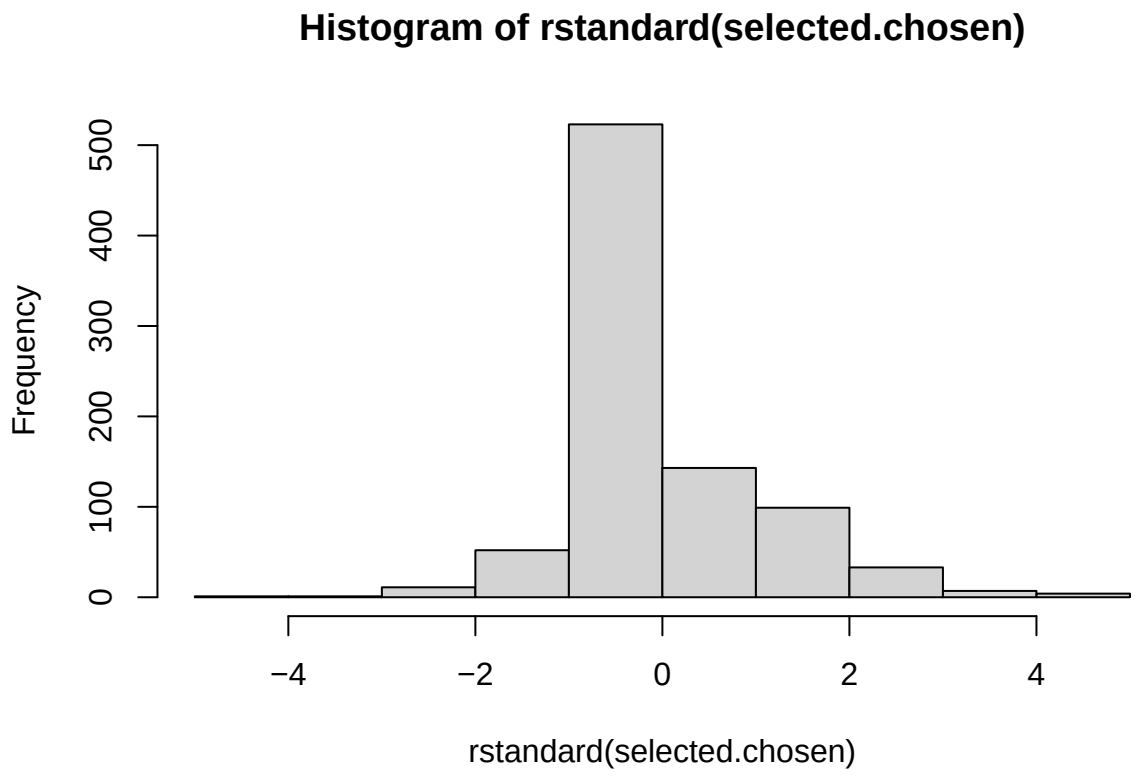
```
####  Izabranost u stars
selected.chosen = fit.chosen

plot(selected.chosen$residuals) #gledajuci reziduale na ovaj nacin tesko je suditi o normalnosti
```



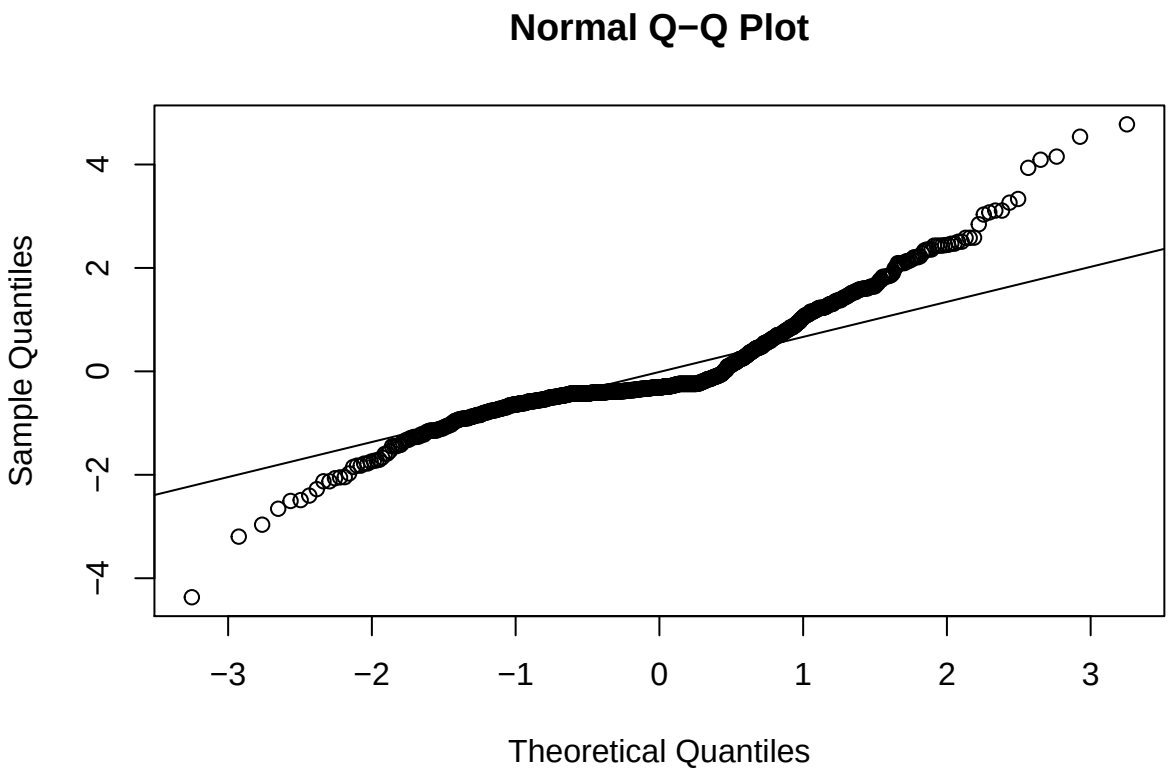

```
#histogram je vrlo interpretativan
hist((selected.chosen$residuals))
```



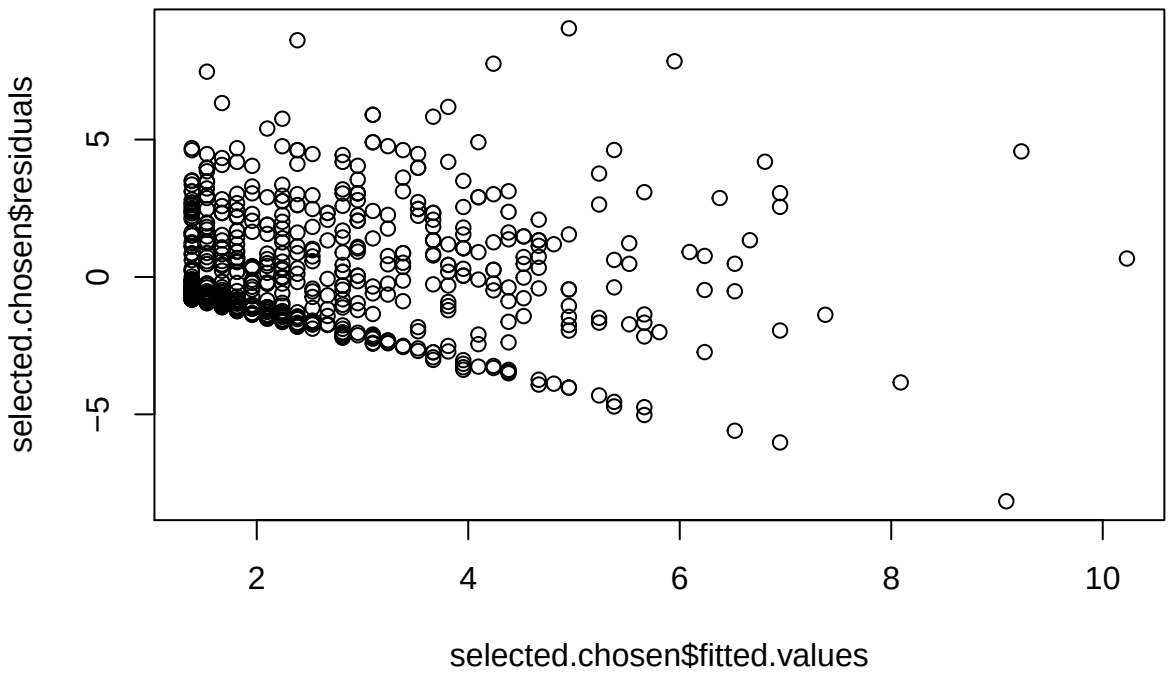

```
hist(rstandard(selected.chosen))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.chosen))
qqline(rstandard(selected.chosen))
```




```
plot(selected.chosen$fitted.values,selected.chosen$residuals) #reziduale je dobro prikazati u ovisnosti
```




```
#KS test na normalnost
ks.test(rstandard(fit.chosen),'pnorm')
```

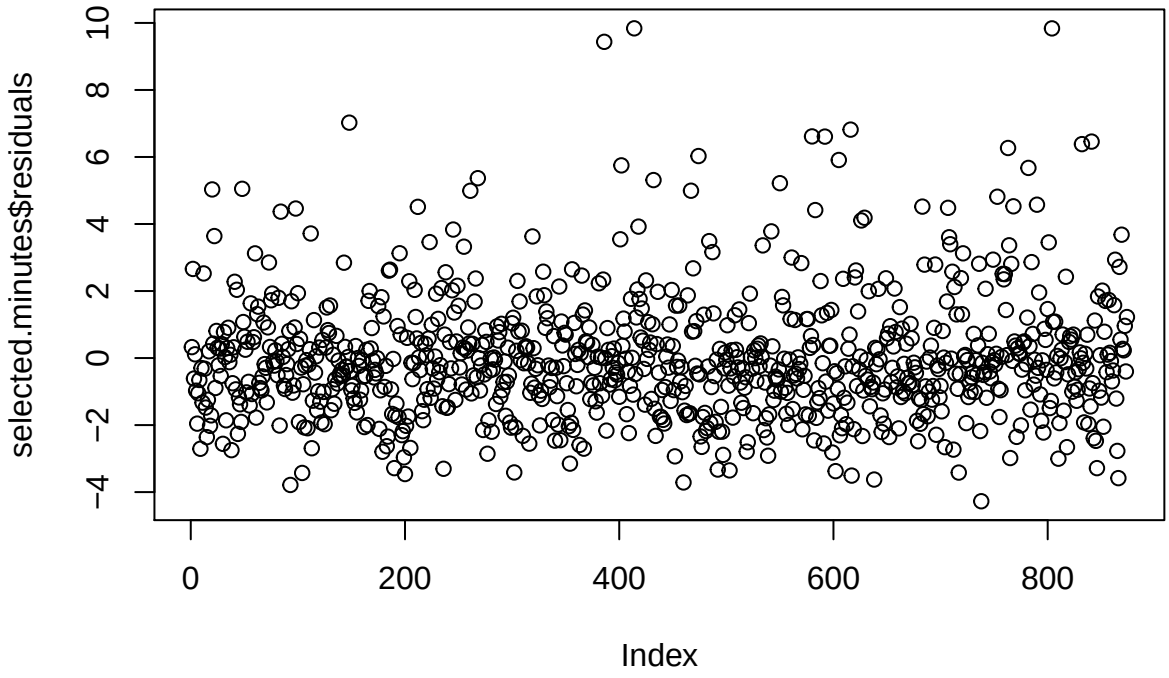
```
## Warning in ks.test(rstandard(fit.chosen), "pnorm"): ties should not be present
## for the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.chosen)
## D = 0.20184, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

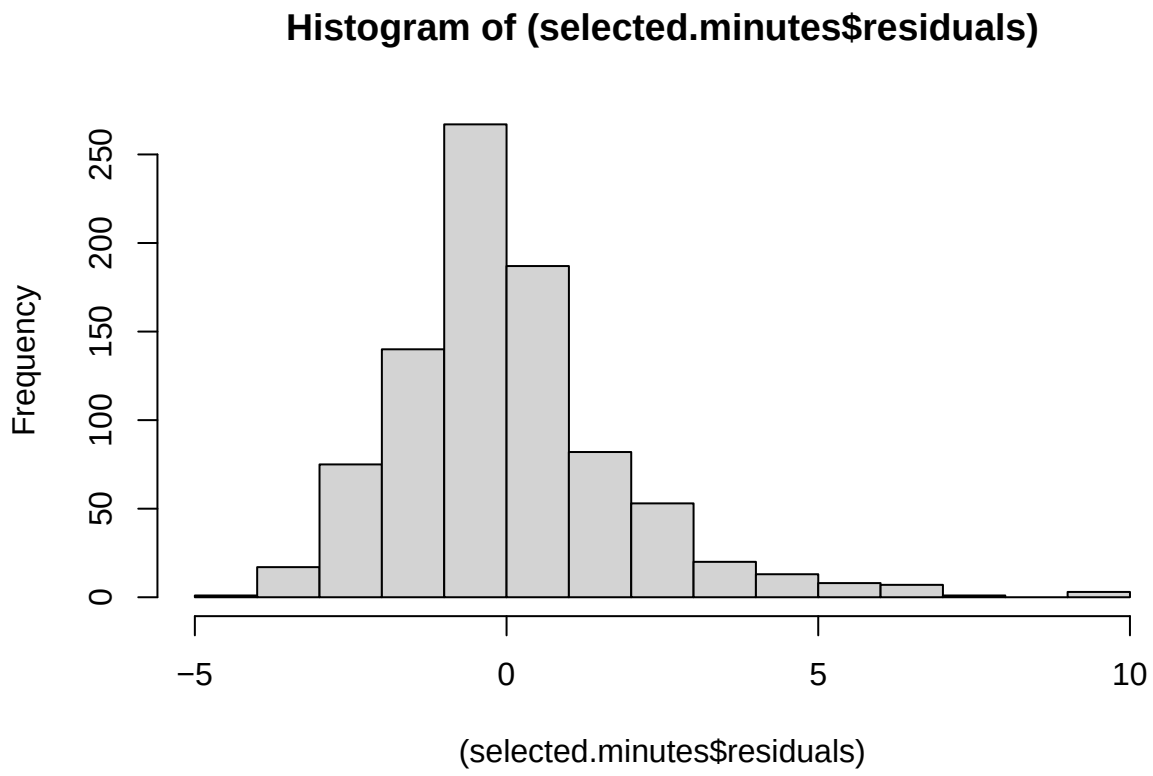
```
lillie.test(rstandard(fit.chosen))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.chosen)
## D = 0.20168, p-value < 2.2e-16
```

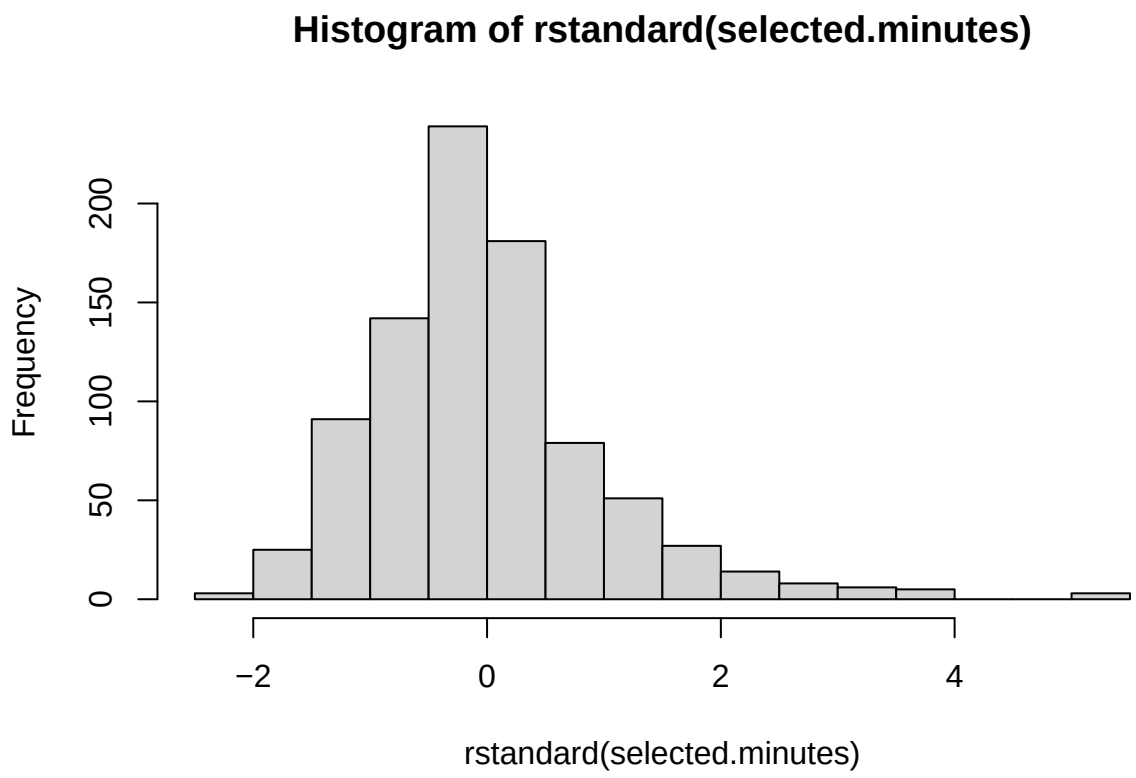
```
####  Vrijeme igranja po utakmici
selected.minutes = fit.minutes

plot(selected.minutes$residuals) #gledajuci reziduale na ovaj nacin tesko je suditi o normalnosti
```



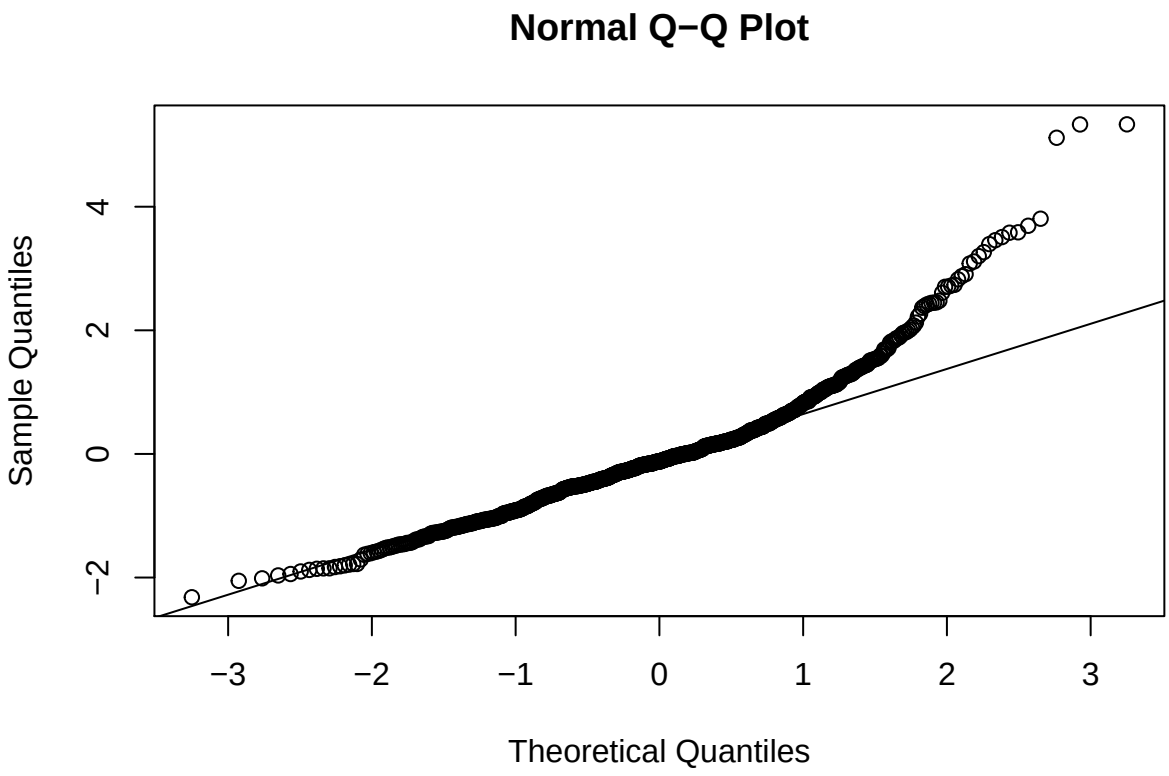

```
#histogram je vrlo interpretativan
hist((selected.minutes$residuals))
```



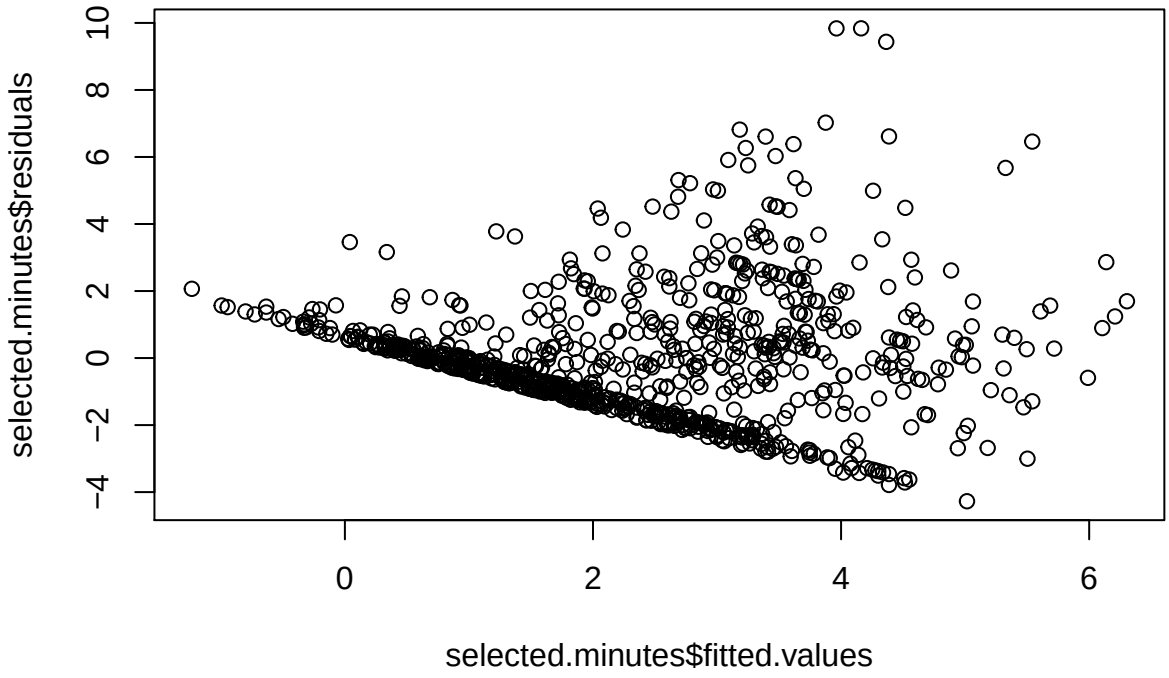

```
hist(rstandard(selected.minutes))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.minutes))
qqline(rstandard(selected.minutes))
```




```
plot(selected.minutes$fitted.values,selected.minutes$residuals) #reziduale je dobro prikazati u ovisnos
```




```
#KS test na normalnost
ks.test(rstandard(fit.minutes),'pnorm')
```

```
## Warning in ks.test(rstandard(fit.minutes), "pnorm"): ties should not be present
## for the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.minutes)
## D = 0.10693, p-value = 4.18e-09
## alternative hypothesis: two-sided
```

```
require(nortest)
lillie.test(rstandard(fit.minutes))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.minutes)
## D = 0.10704, p-value < 2.2e-16
```

```
cor(playersSalary$PTS,playersSalary$Salary)
```

```
## [1] 0.622702
```

```
cor.test(playersSalary$PTS,playersSalary$Salary)
```

```
##
```

```
##  Pearson's product-moment correlation
##
## data:  playersSalary$PTS and playersSalary$Salary
## t = 23.5, df = 872, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5803534 0.6616918
## sample estimates:
##      cor
## 0.622702
```

```
summary(fit.PTS)
```

```
##
## Call:
## lm(formula = playersSalary$Salary ~ playersSalary$PTS, data = playersSalary)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0621 -0.9229 -0.2654  0.6465  9.4313
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.86544    0.08697   9.951   <2e-16 ***
## playersSalary$PTS   0.07122    0.00303  23.500   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.799 on 872 degrees of freedom
## Multiple R-squared:  0.3878, Adjusted R-squared:  0.3871
## F-statistic: 552.3 on 1 and 872 DF,  p-value: < 2.2e-16
```

```
cor(playersSalary$SCA,playersSalary$Salary)
```

```
## [1] NA
```

```
cor.test(playersSalary$SCA,playersSalary$Salary)
```

```
##
##  Pearson's product-moment correlation
##
## data:  playersSalary$SCA and playersSalary$Salary
## t = 19.913, df = 871, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5119692 0.6032831
## sample estimates:
##       cor
## 0.5593206
```

```
summary(fit.SCA)
```

```
##
## Call:
## lm(formula = playersSalary$Salary ~ playersSalary$SCA, data = playersSalary)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -4.2587 -1.1300 -0.1482  0.5020 10.4213
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.5876621  0.1086068   5.411 8.11e-08 ***
## playersSalary$SCA 0.0145907  0.0007327  19.913  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.907 on 871 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.3128, Adjusted R-squared:  0.3121
## F-statistic: 396.5 on 1 and 871 DF,  p-value: < 2.2e-16
```

```
cor(playersSalary$Chosen,playersSalary$Salary)
```

```
## [1] 0.5633451
```

```
cor.test(playersSalary$Chosen,playersSalary$Salary)
```

```
##
##  Pearson's product-moment correlation
##
## data:  playersSalary$Chosen and playersSalary$Salary
## t = 20.134, df = 872, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5163201 0.6069833
## sample estimates:
##       cor
## 0.5633451
```

```
summary(fit.chosen)
```

```
##
## Call:
## lm(formula = playersSalary$Salary ~ playersSalary$Chosen, data = playersSalary)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.1630 -0.8890 -0.5861  0.8460  9.0482
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.386142   0.079410   17.46   <2e-16 ***
## playersSalary$Chosen 0.142626   0.007084   20.13   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.9 on 872 degrees of freedom
## Multiple R-squared:  0.3174, Adjusted R-squared:  0.3166
## F-statistic: 405.4 on 1 and 872 DF,  p-value: < 2.2e-16
```

```
cor(playersSalary$TOI.GP,playersSalary$Salary)
```

```
## [1] 0.5952072
```

```r
cor.test(playersSalary$TOI.GP,playersSalary$Salary)
```

```
##
##  Pearson's product-moment correlation
##
## data:  playersSalary$TOI.GP and playersSalary$Salary
## t = 21.873, df = 872, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5506272 0.6364017
## sample estimates:
##       cor
## 0.5952072
```

```r
summary(fit.minutes)
```

```
##
## Call:
## lm(formula = playersSalary$Salary ~ playersSalary$TOI.GP, data = playersSalary)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.2725 -1.0685 -0.2238  0.7512  9.8383
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          -2.50087    0.22933  -10.90   <2e-16 ***
## playersSalary$TOI.GP  0.32078    0.01467   21.87   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.848 on 872 degrees of freedom
## Multiple R-squared:  0.3543, Adjusted R-squared:  0.3535
## F-statistic: 478.4 on 1 and 872 DF,  p-value: < 2.2e-16
```

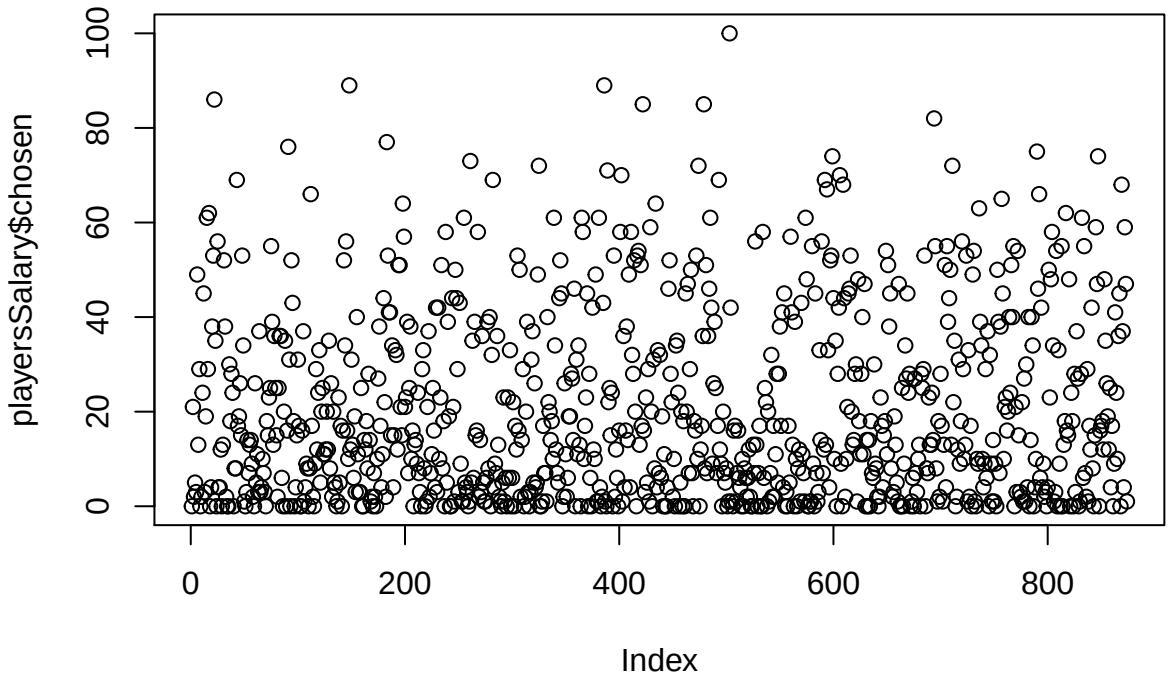
###ZAKLJUČAK

Došli smo do zaključka da varijable koje smo odabrali pozitivno utječu na visinu plaće, ali i da postoji određeni stupanj korelacije među njima, što bi nam dosta otežalo izvedbu višestruke regresije iz ovih varijabli.

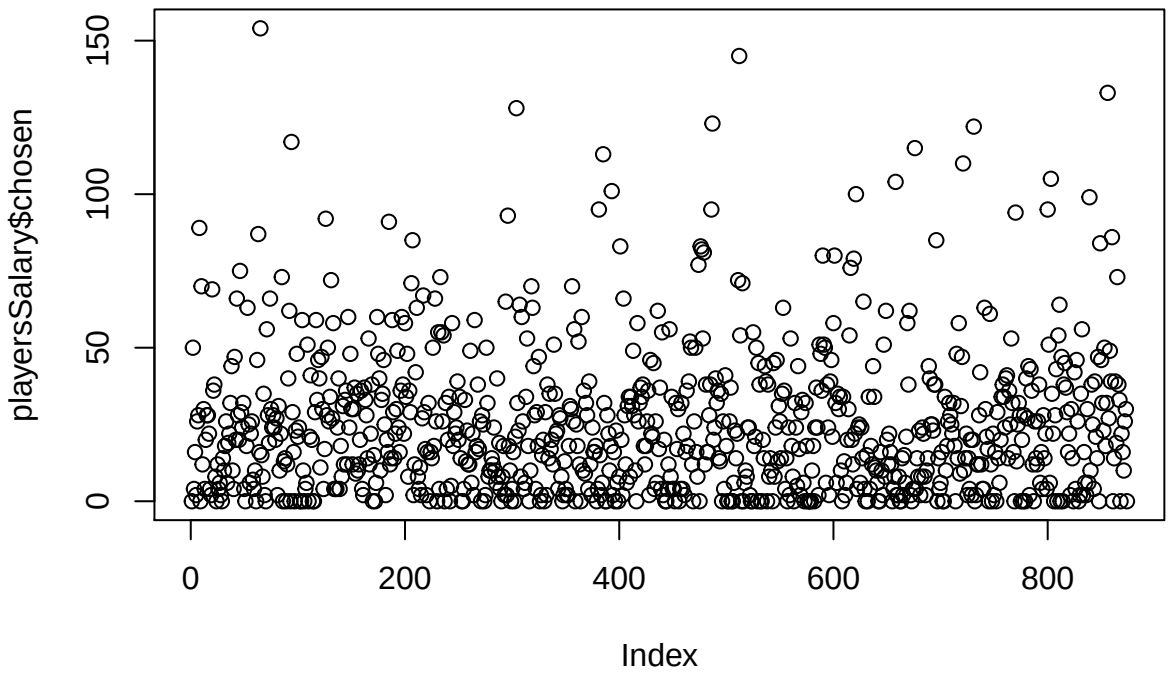
U gotovo svakom sportu biraju se najbolji od njaboljih.Za NHL vrijedi da svake godine se biraju 3 najbolje postave.Stoga ćemo probati uzeti u obzir neke varijable koje bi mogle ovisiti o izboru u najbolju postavu.

###ZADATAK 11 Opet koristimo model linearne regresije

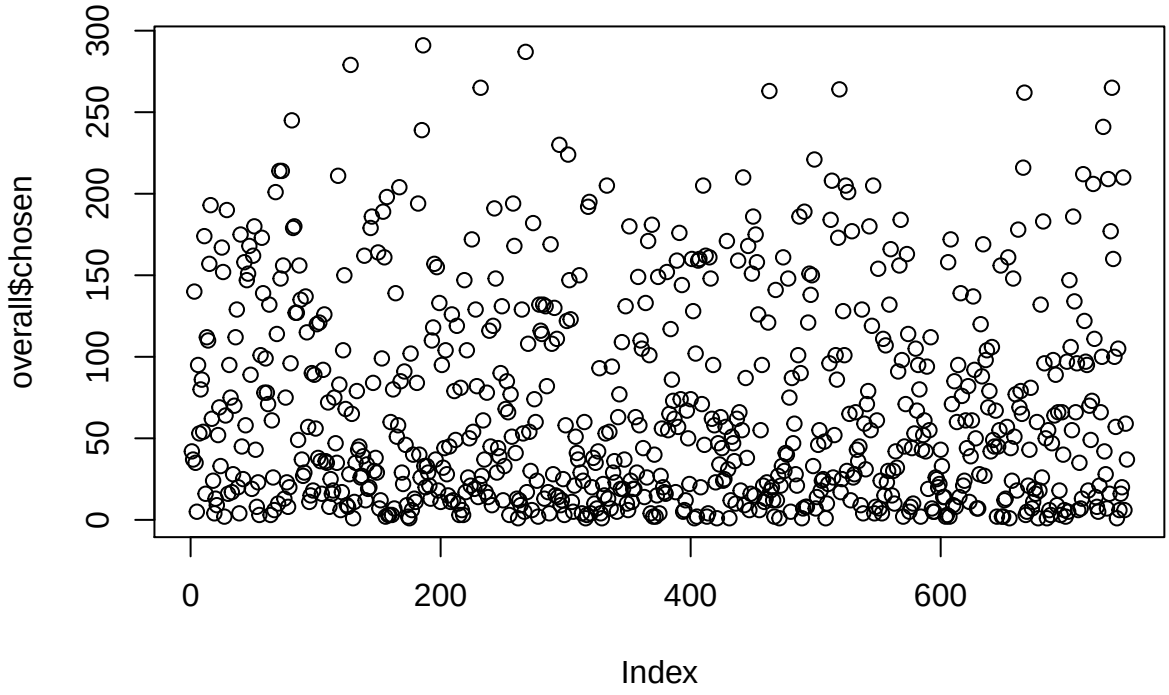
```r
overall <- AllSits_copy[, c("First.Name", "Last.Name", "Ovrl")]
overall <- overall[!is.na(overall$Ovrl), ]

overall <-  merge(overall, playersSalary, by = c("Last.Name", "First.Name"), all.x=TRUE)

overall <- overall[, (names(overall) %in% c("Last.Name", "First.Name", "Chosen", "Ovrl"))]
overall <- overall[!is.na(overall$Chosen), ]


plot(playersSalary$PTS, playersSalary$chosen)
```



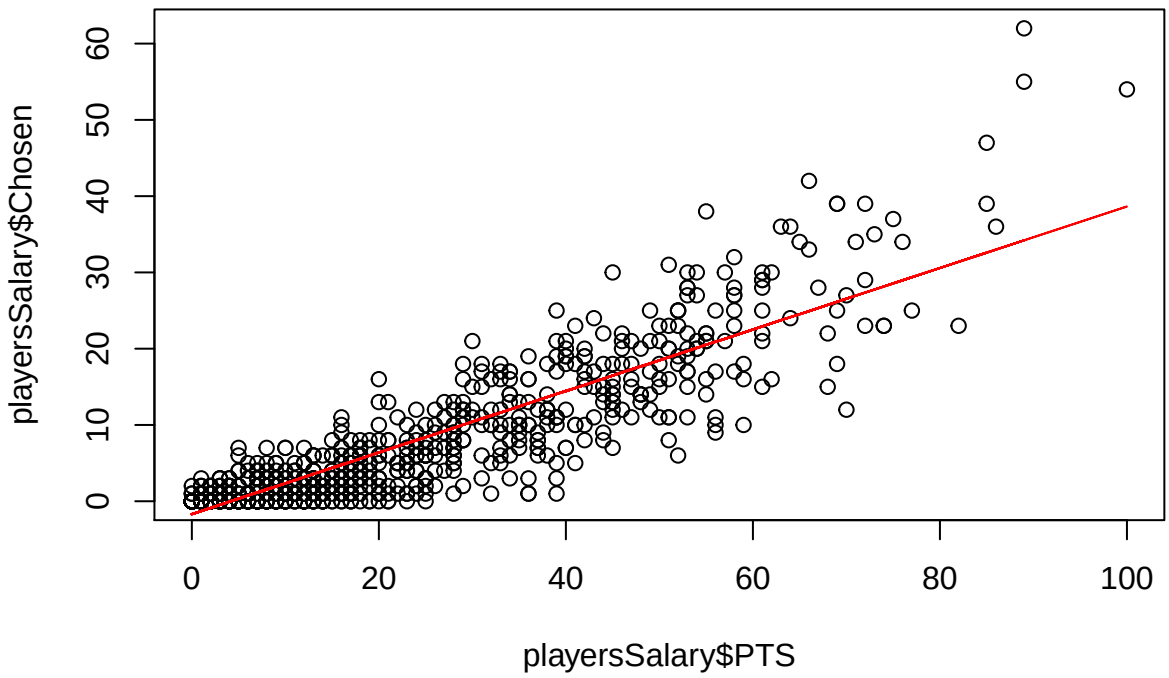

```
plot(playersSalary$PIM, playersSalary$chosen)
```



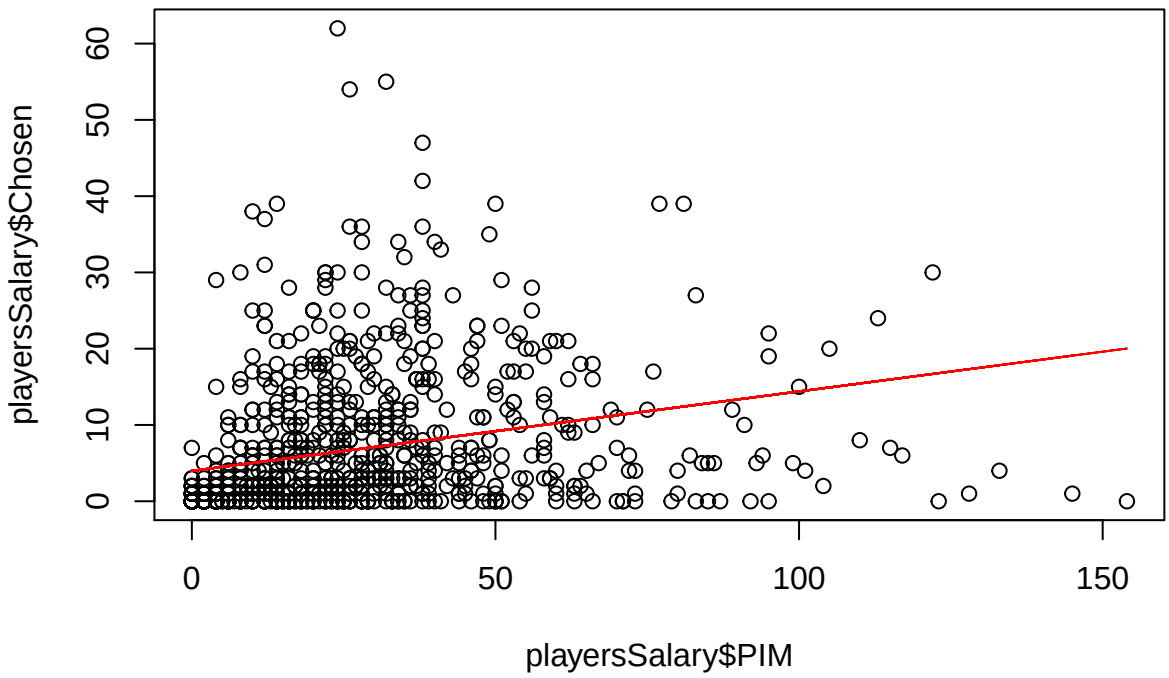

```
plot(overall$Ovrl, overall$chosen)
```



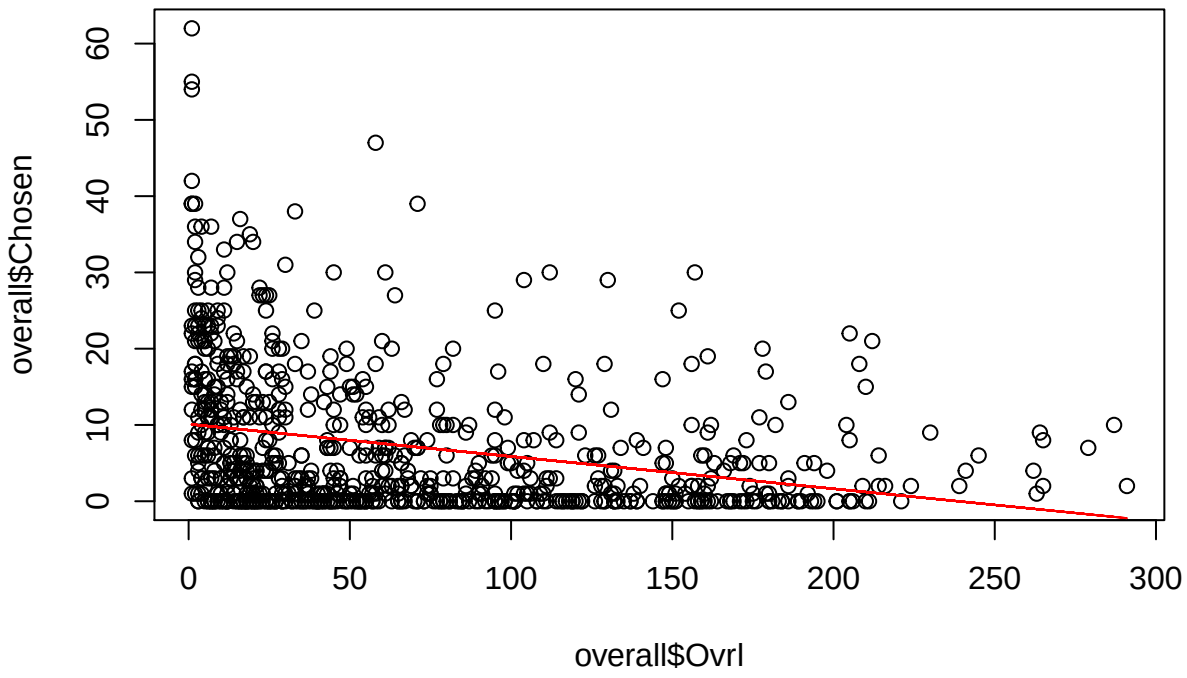

```
fit.pts = lm(playersSalary$Chosen~playersSalary$PTS, data=playersSalary)
fit.pim = lm(playersSalary$Chosen~playersSalary$PIM, data=playersSalary)
fit.ovrl = lm(overall$Chosen~overall$Ovrl, data=overall)


plot(playersSalary$PTS,  playersSalary$Chosen)
lines(playersSalary$PTS, fit.pts$fitted.values, col='red')
```



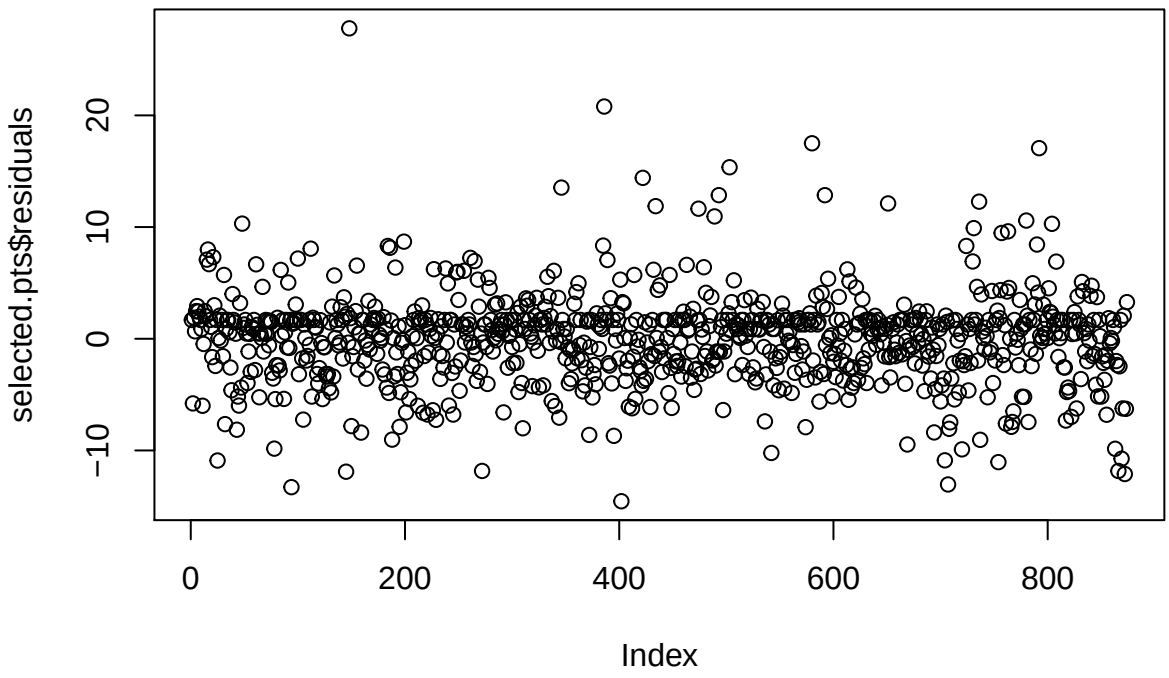

```
plot(playersSalary$PIM, playersSalary$Chosen)
lines(playersSalary$PIM, fit.pim$fitted.values, col='red')
```



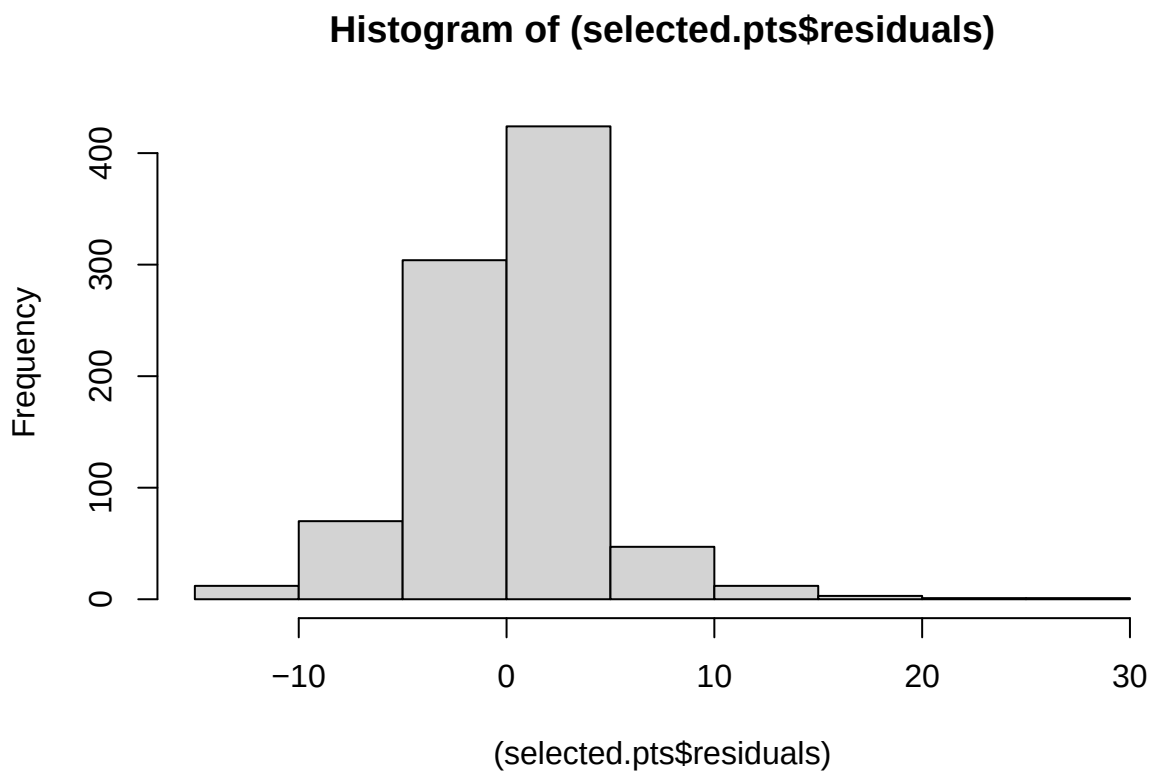

```
plot(overall$Ovrl, overall$Chosen)
lines(overall$Ovrl, fit.ovrl$fitted.values, col='red')
```



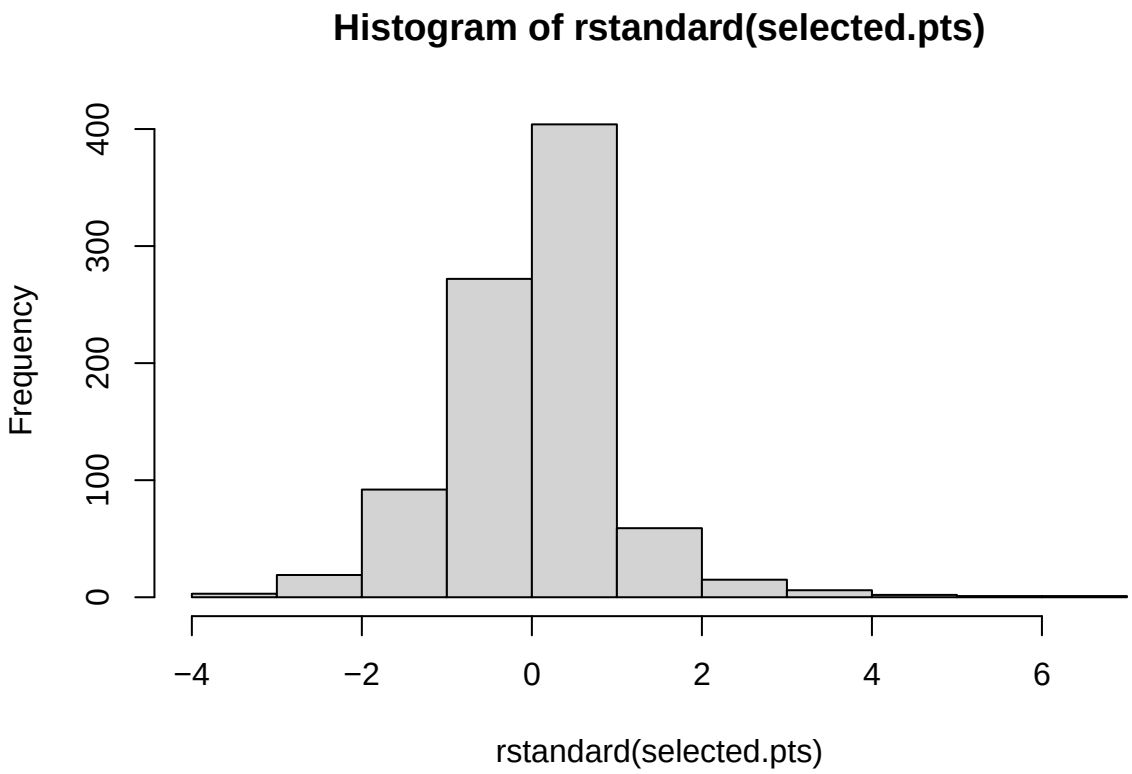

```
####  Bodovi
selected.pts = fit.pts

plot(selected.pts$residuals)
```



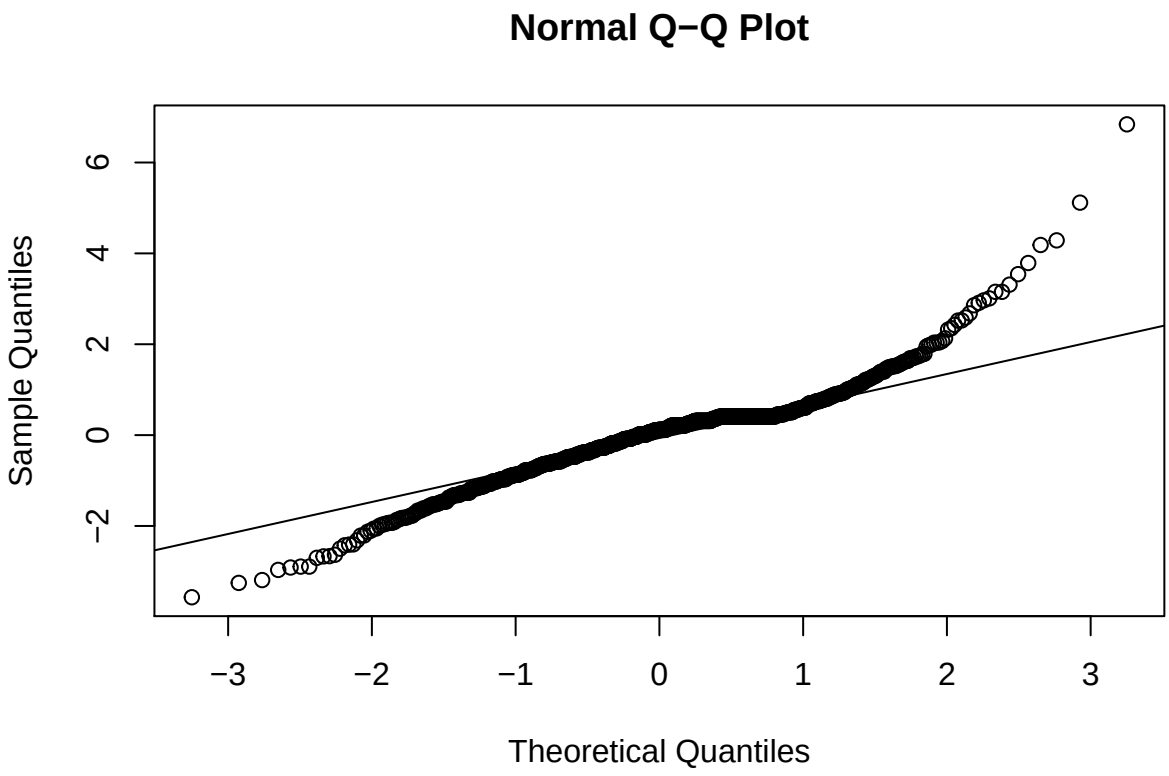

```
#histogram je vrlo interpretativan
hist((selected.pts$residuals))
```



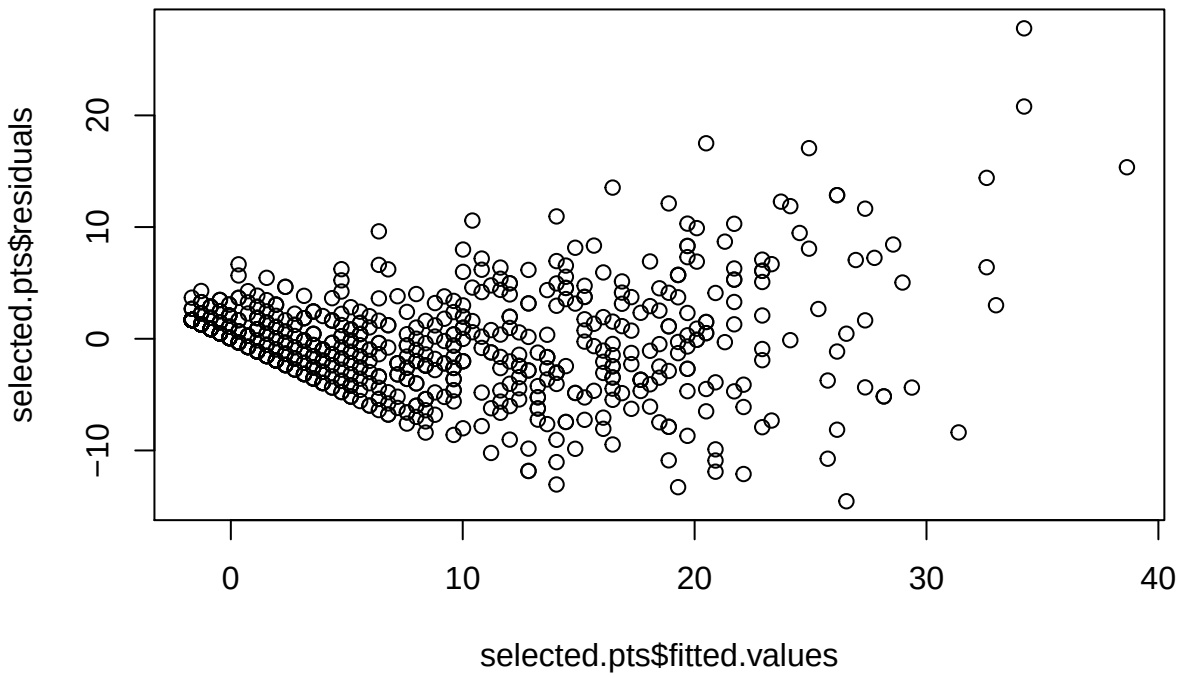

```
hist(rstandard(selected.pts))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.pts))
qqline(rstandard(selected.pts))
```




```
plot(selected.pts$fitted.values,selected.pts$residuals)
```




```
#KS test na normalnost
ks.test(rstandard(fit.pts),'pnorm')
```

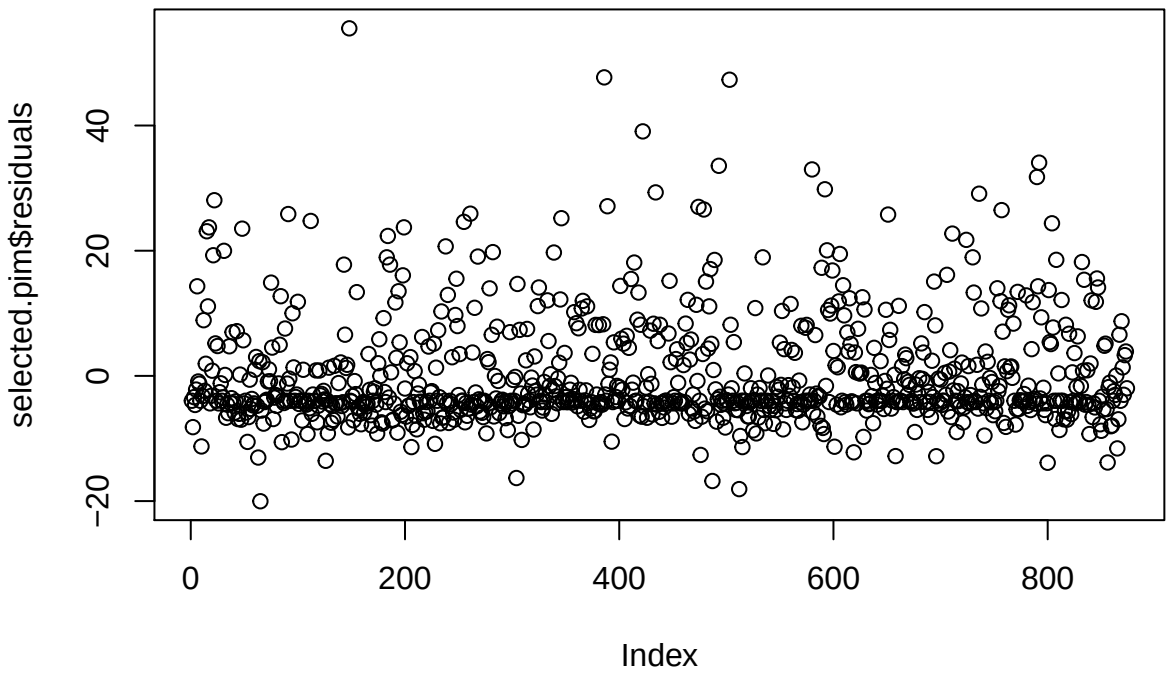
```
## Warning in ks.test(rstandard(fit.pts), "pnorm"): ties should not be present for
## the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.pts)
## D = 0.13102, p-value = 1.856e-13
## alternative hypothesis: two-sided
```

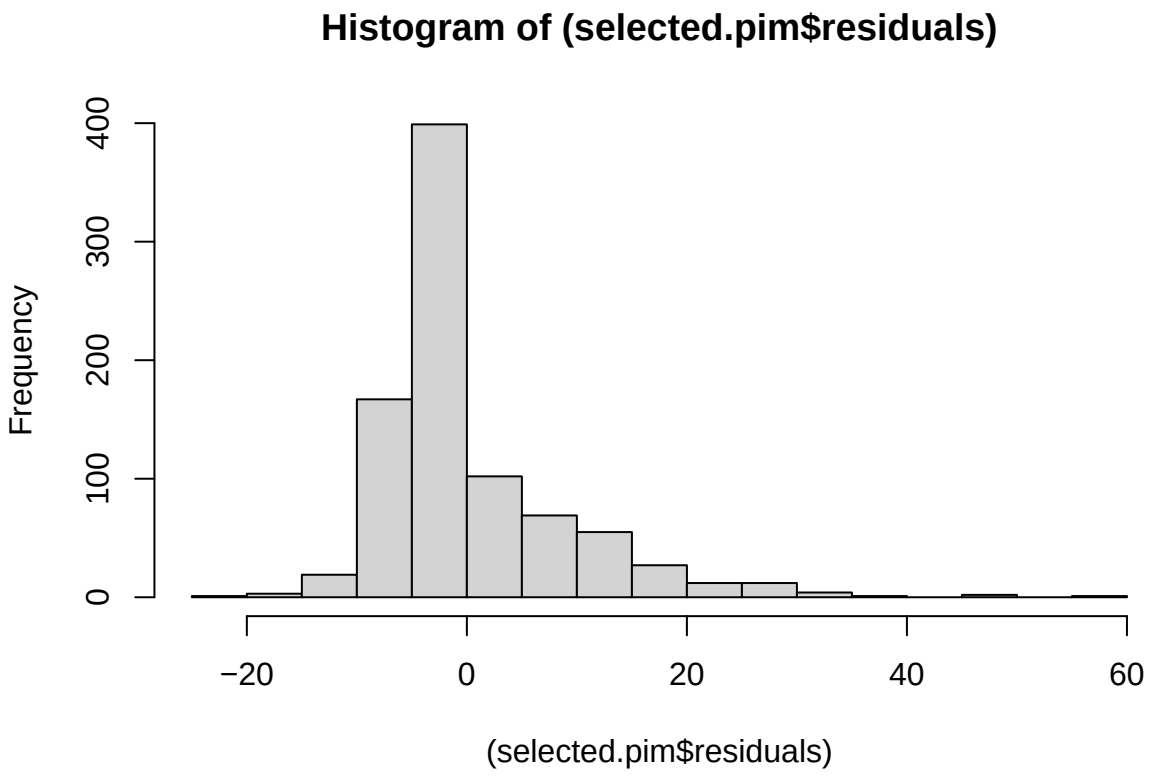
```
require(nortest)
lillie.test(rstandard(fit.pts))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.pts)
## D = 0.13139, p-value < 2.2e-16
```

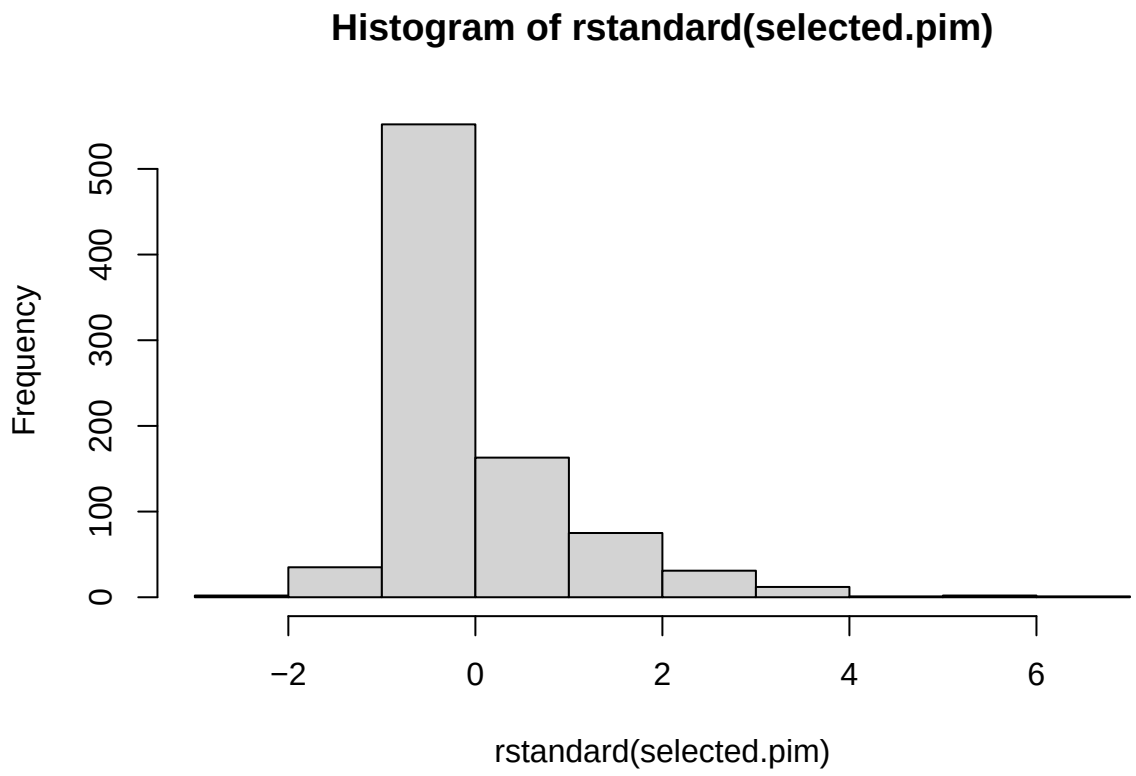
```
####  PIM
selected.pim = fit.pim

plot(selected.pim$residuals)
```



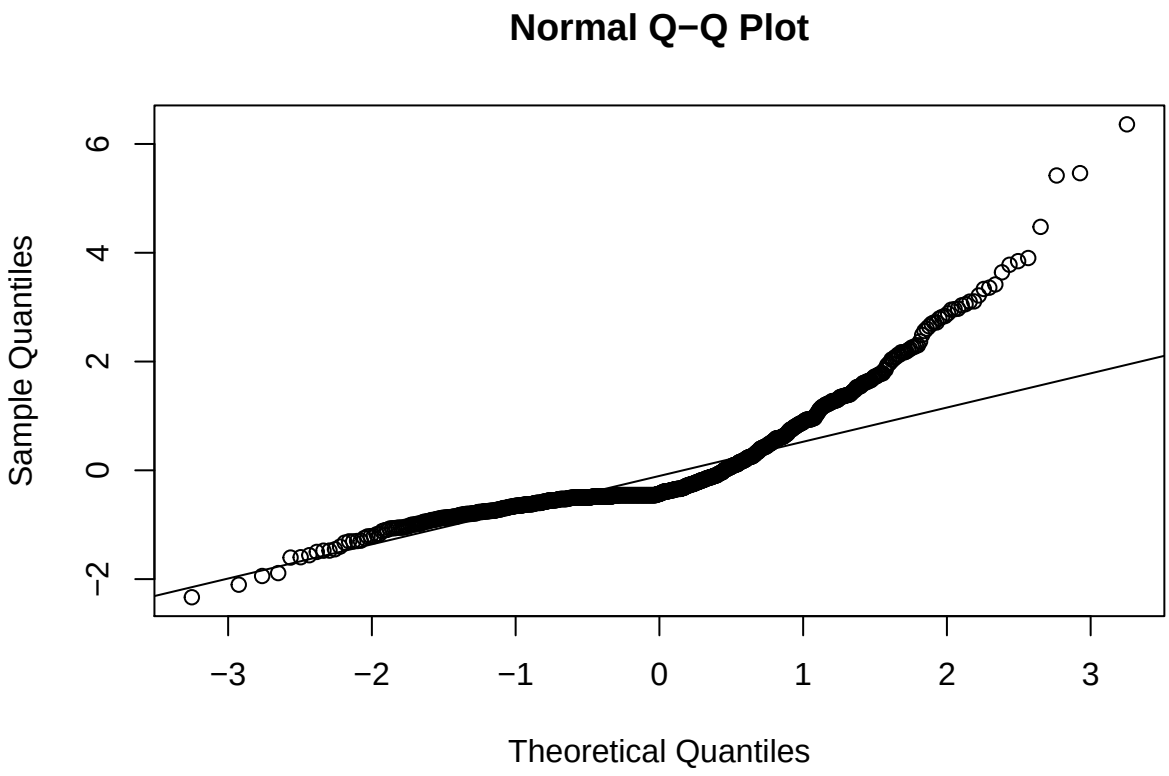

```
#histogram je vrlo interpretativan
hist((selected.pim$residuals))
```



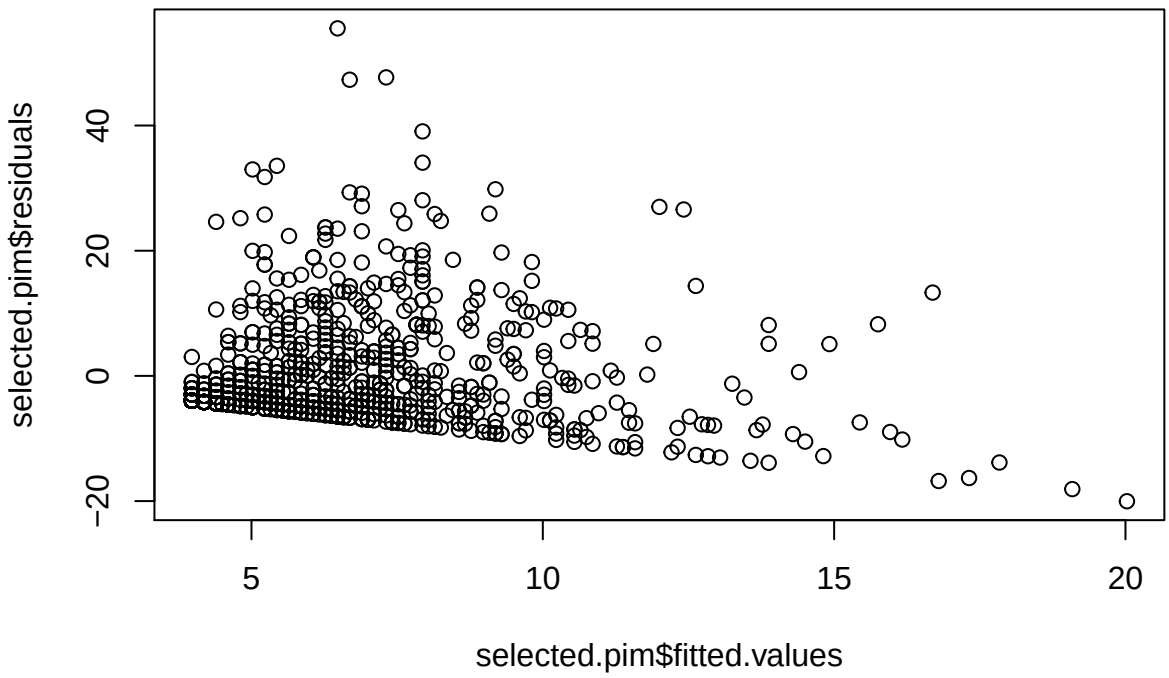

```
hist(rstandard(selected.pim))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.pim))
qqline(rstandard(selected.pim))
```




```
plot(selected.pim$fitted.values,selected.pim$residuals) #reziduale je dobro prikazati u ovisnosti o pro
```




```
#KS test na normalnost
ks.test(rstandard(fit.pim),'pnorm')
```

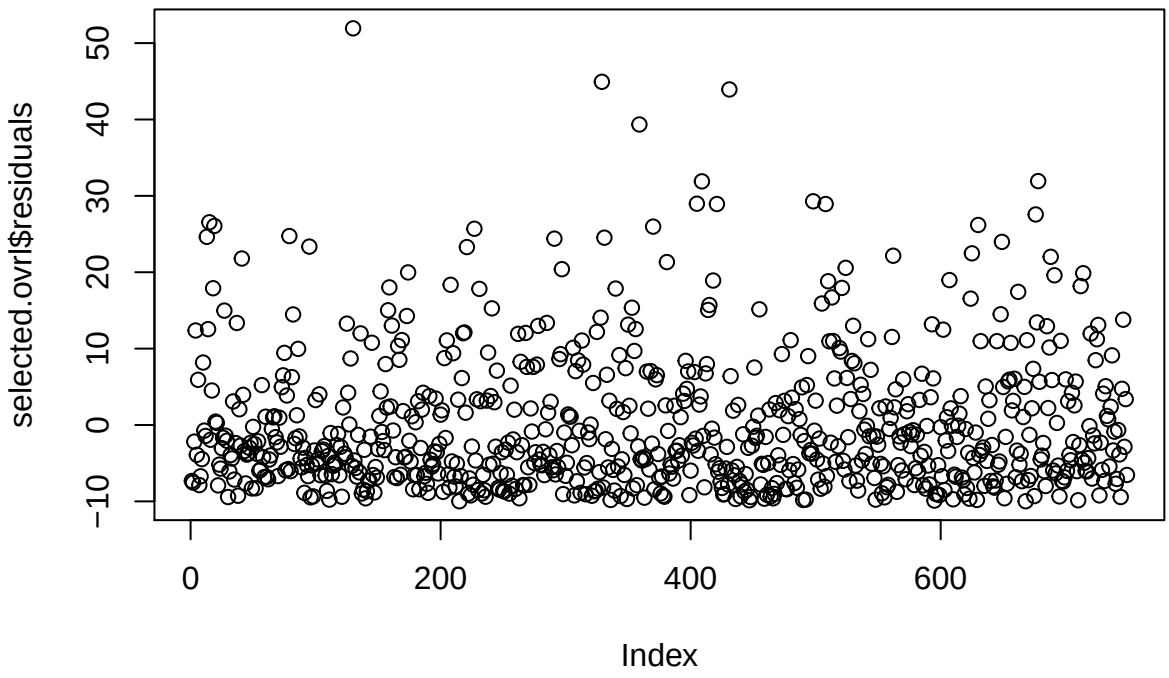
```
## Warning in ks.test(rstandard(fit.pim), "pnorm"): ties should not be present for
## the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.pim)
## D = 0.19493, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

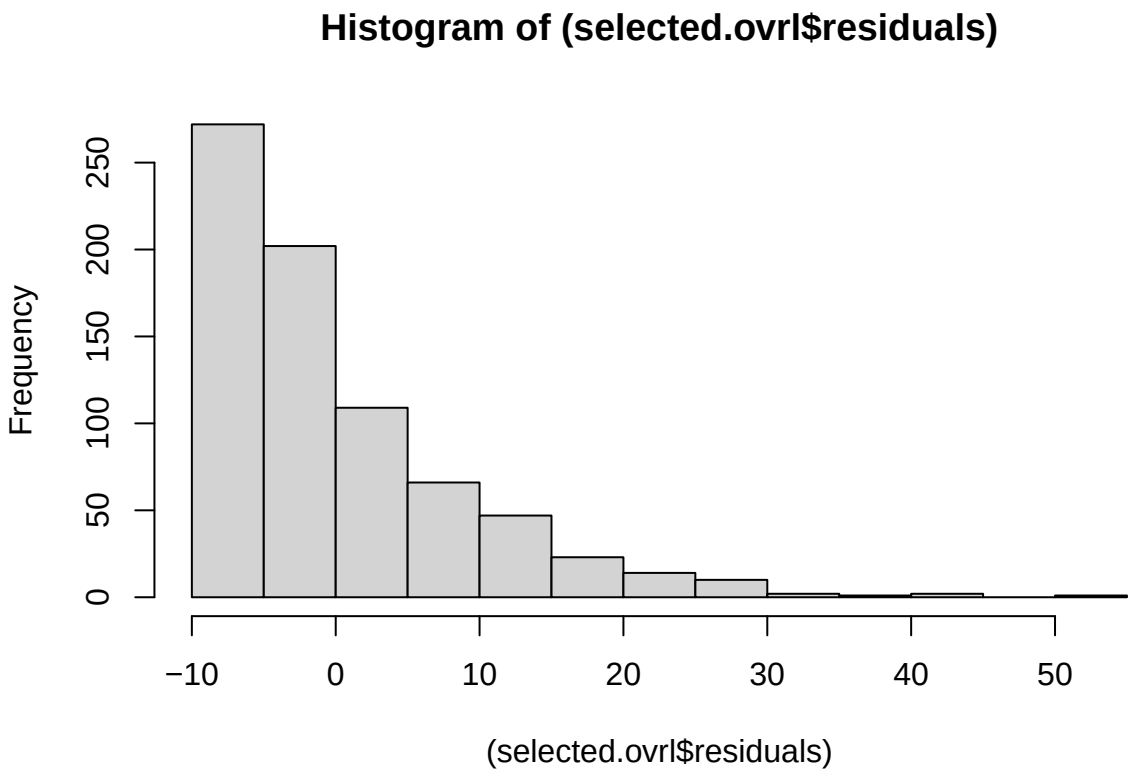
```
require(nortest)
lillie.test(rstandard(fit.pim))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.pim)
## D = 0.19473, p-value < 2.2e-16
```

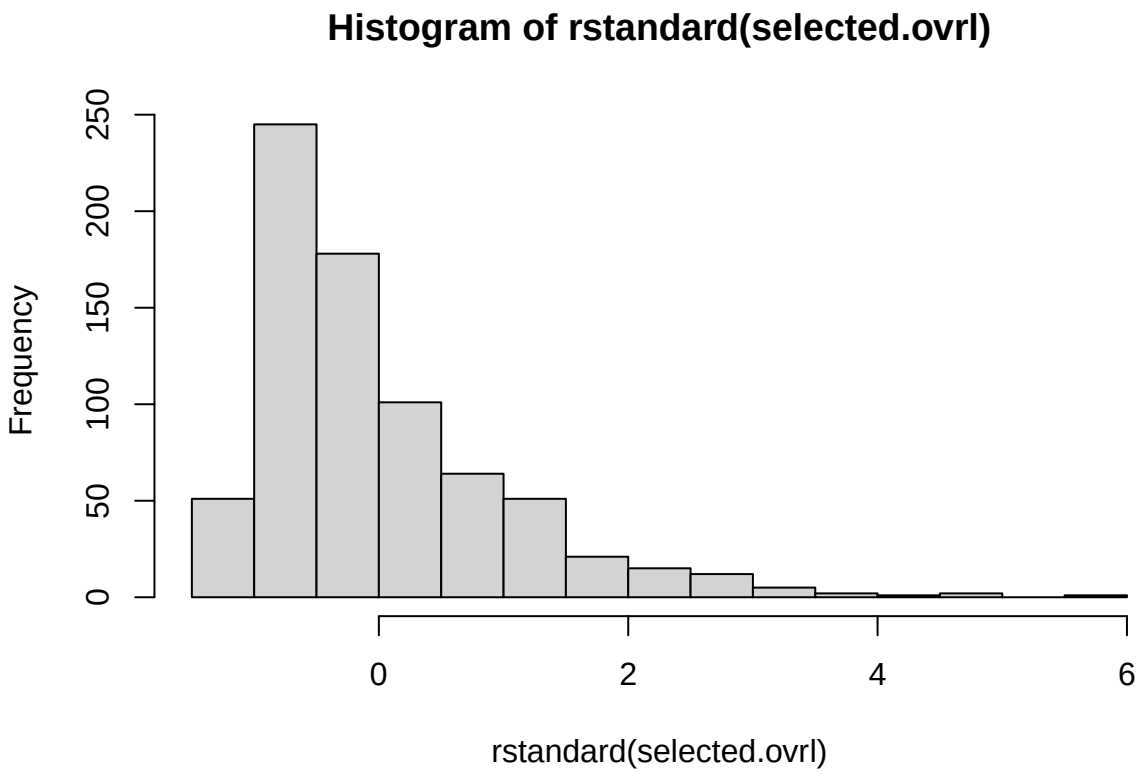
```
####  Overall
selected.ovrl = fit.ovrl

plot(selected.ovrl$residuals)
```



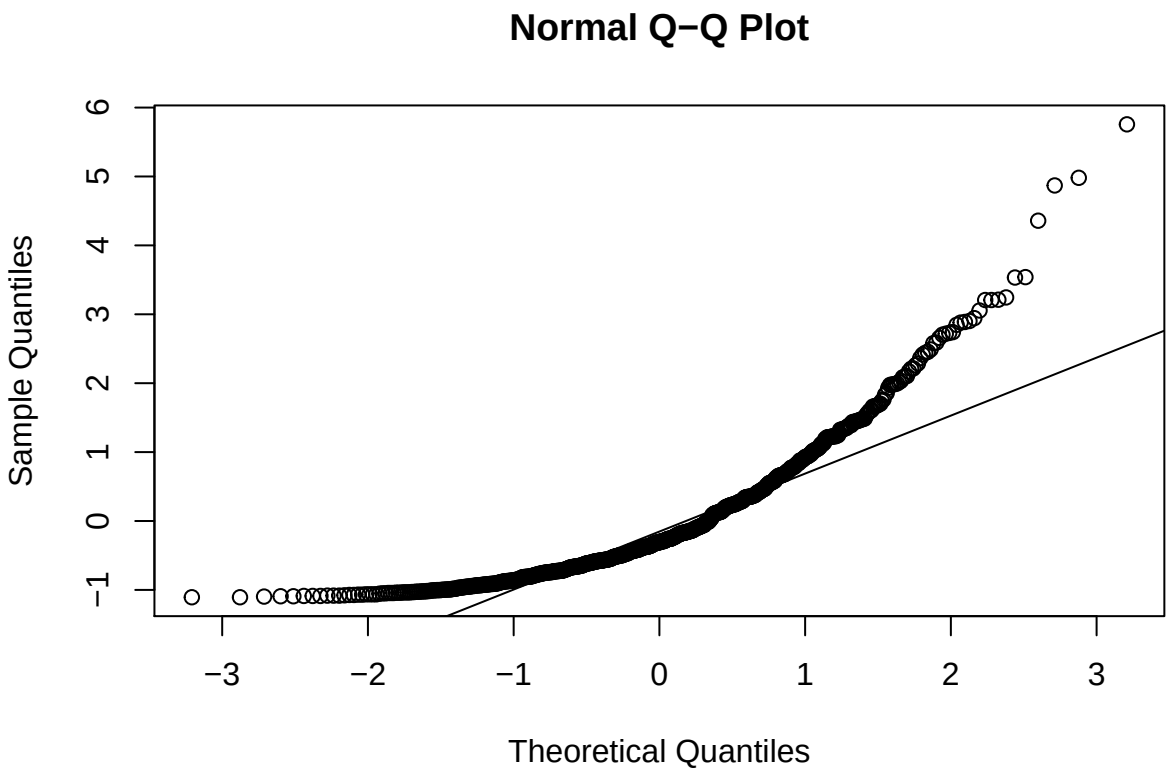

```
#histogram je vrlo interpretativan
hist((selected.ovrl$residuals))
```



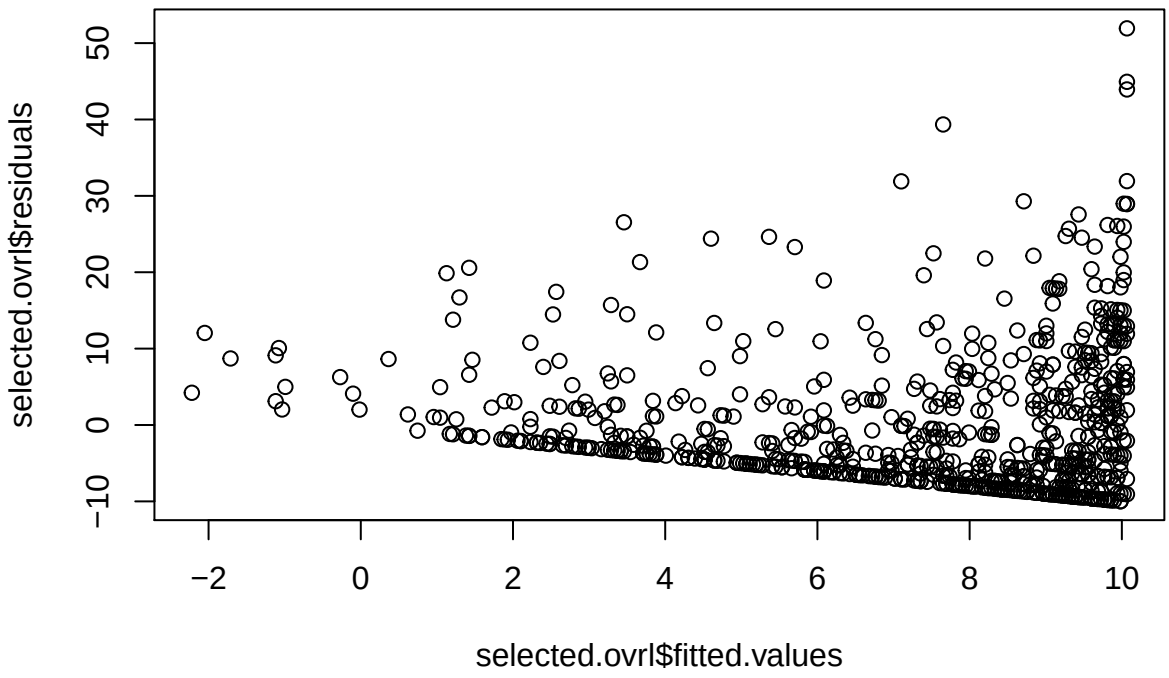

```
hist(rstandard(selected.ovrl))
```




```
#q-q plot reziduala s linijom normalne distribucije
qqnorm(rstandard(selected.ovrl))
qqline(rstandard(selected.ovrl))
```




```
plot(selected.ovrl$fitted.values,selected.ovrl$residuals)
```




```
#KS test na normalnost
ks.test(rstandard(fit.ovrl),'pnorm')
```

```
## Warning in ks.test(rstandard(fit.ovrl), "pnorm"): ties should not be present for
## the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.ovrl)
## D = 0.14451, p-value = 5.185e-14
## alternative hypothesis: two-sided
```

```
require(nortest)
lillie.test(rstandard(fit.ovrl))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  rstandard(fit.ovrl)
## D = 0.14455, p-value < 2.2e-16
```

###ZAKLJUČAK

Kao i očekivano, broj postignutih golova i asistencija te pozicija na draftu (niža pozicija je bolja pozicija) pozitivno utječu na broj koliko je puta igrač izabran u najbolji tim, te tu uočavamo uzročno posljedične veze. Kod PIM (minute isključenja) smo došli do zaključka da ne utječu previše na osvojene nagrade, što se može najbolje vidjeti iz scatter plota.

——————-ZAKLJUČAK RADA——————- Tijekom čitave ove analize pokušali smo pomoću situacija iz svakodnevice motivirati zanimljive probleme koje smo preslikali na naš skup podataka iz NHL,te smo prilikom rješavanja pojedinih problema izračunali neke korisne mjere,primjenili neke korisne modele i testove te isto tako napravili vizualizaciju odgovora na naše probleme.Da bi smo uspješno savladali ove probleme poslužili smo se znanjem Statističke Analize Podataka kojeg smo osim upotrebe još dodatno utvrdili.