European Survey - What is the political sentiment in German speaking Countries Before Covid?

Milica Pajkic and Marco Rieder

2022-11-24

Contents

Getting the Data and installing libraries

1. Introduction

The European Social Survey (ESS) is a large scale survey conducted in over 38 countries within Europe. Focusing on public attitudes and values and changes within time. This paper is evaluating the ninth version of the survey (ESS9) form 2018.

1.1 Selection of parameters

The survey is really comprehensive and consists out of 572 variables. Some of them are related to others and only answered by a subset of participants.

1.2 Goal

2. Descriptive Statistics

First

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0      30      60     187      90    9999
```

```
## [1] 0
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   1.000   1.000   1.486   2.000   2.000
```

```
## [1] 0
```

```
## [1] "double"
```

```
## tibble [2,358 x 572] (S3: tbl_df/tbl/data.frame)
##  $ dweight : num [1:2358] 0.999 0.999 0.999 0.999 0.999 ...
##  $ pspwght : num [1:2358] 1.275 0.854 0.76 1.079 1.27 ...
##  $ pweight : num [1:2358] 3.04 3.04 3.04 3.04 3.04 ...
##  $ name    : chr [1:2358] "ESS9e03_1" "ESS9e03_1" "ESS9e03_1" "ESS9e03_1" ...
##  $ essround: num [1:2358] 9 9 9 9 9 9 9 9 9 9 ...
##  $ edition : num [1:2358] 3.1 3.1 3.1 3.1 3.1 3.1 3.1 3.1 3.1 3.1 ...
##  $ proddate: chr [1:2358] "17.02.2021" "17.02.2021" "17.02.2021" "17.02.2021" ...
##  $ idno    : num [1:2358] 9 10 64 65 91 119 150 212 255 270 ...
```

```
##  $ cntry   : chr [1:2358] "DE" "DE" "DE" "DE" ...
##  $ anweight: num [1:2358] 3.87 2.59 2.31 3.28 3.86 ...
##  $ prob    : num [1:2358] 0.000122 0.000122 0.000122 0.000122 0.000122 ...
##  $ stratum : num [1:2358] 336 284 307 338 297 323 320 295 294 278 ...
##  $ psu     : num [1:2358] 5856 5755 5798 5861 5779 ...
##  $ nwspol  : num [1:2358] 8 60 120 300 0 30 60 30 60 30 ...
##  $ netusoft: num [1:2358] 5 1 3 2 5 5 4 5 4 1 ...
##  $ netustm : num [1:2358] 480 6666 6666 6666 60 ...
##  $ ppltrst : num [1:2358] 5 7 7 7 5 3 3 6 7 8 ...
##  $ pplfair : num [1:2358] 5 8 7 6 6 6 4 5 7 9 ...
##  $ pplhlp  : num [1:2358] 5 5 5 3 6 5 4 5 8 9 ...
##  $ polintr : num [1:2358] 3 1 1 1 3 2 1 3 2 2 ...
##  $ psppsgva: num [1:2358] 4 4 2 3 2 3 3 3 2 3 ...
##  $ actrolga: num [1:2358] 3 4 2 2 3 1 3 1 3 2 ...
##  $ psppipla: num [1:2358] 3 4 3 2 3 8 3 2 2 3 ...
##  $ cptppola: num [1:2358] 3 4 3 3 4 4 3 2 4 3 ...
##  $ trstprl : num [1:2358] 2 7 3 3 4 9 10 5 5 7 ...
##  $ trstlgl : num [1:2358] 4 8 5 4 5 7 10 7 8 8 ...
##  $ trstplc : num [1:2358] 5 8 6 4 7 7 10 7 9 8 ...
##  $ trstplt : num [1:2358] 0 6 3 3 5 8 10 5 5 6 ...
##  $ trstprt : num [1:2358] 2 6 5 3 5 8 4 5 5 6 ...
##  $ trstep  : num [1:2358] 3 4 5 2 4 10 10 6 3 7 ...
##  $ trstun  : num [1:2358] 0 5 6 2 5 10 10 6 7 7 ...
##  $ vote    : num [1:2358] 2 1 1 1 1 1 1 1 1 1 ...
##  $ prtvtcat: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtdbe: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtdbg: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtgch: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtbcy: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtecz: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvede1: num [1:2358] 66 1 2 2 9 2 1 2 3 1 ...
##  $ prtvede2: num [1:2358] 66 1 2 2 88 1 1 2 3 1 ...
##  $ prtvtddk: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtgee: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtees: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtdfi: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtdfr: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtcgb: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtahr: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtfhu: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtcie: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtcis: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtcit: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvblt1: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvblt2: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvblt3: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtalv: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtme : num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtgnl: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtbno: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtdpl: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtcpt: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtrs : num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtcse: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
```

```
##  $ prtvtfsi: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtvtdsk: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ contplt : num [1:2358] 2 2 2 1 2 2 1 2 2 1 ...
##  $ wrkprty : num [1:2358] 2 2 2 2 2 2 2 2 2 1 ...
##  $ wrkorg  : num [1:2358] 1 2 1 1 1 2 1 2 2 2 ...
##  $ badge   : num [1:2358] 2 2 2 2 2 2 2 2 2 2 ...
##  $ sgnptit : num [1:2358] 2 2 2 2 1 2 1 2 1 2 ...
##  $ pbldmn  : num [1:2358] 2 1 2 2 2 2 2 2 1 2 ...
##  $ bctprd  : num [1:2358] 2 1 2 1 1 2 1 2 1 2 ...
##  $ pstplonl: num [1:2358] 2 2 2 2 1 2 2 2 2 2 ...
##  $ clsprty : num [1:2358] 2 1 1 1 2 2 1 1 1 1 ...
##  $ prtcldat: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtcldbe: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtcldbg: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclgch: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclbcy: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclecz: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclede: num [1:2358] 66 1 2 2 66 66 1 2 3 1 ...
##  $ prtclddk: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclgee: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclfes: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclefi: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclffr: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclcgb: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclahr: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclghu: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtcleie: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclcis: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtcldit: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclblt: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclalv: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclme : num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclfnl: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclbno: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclhpl: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclept: num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##  $ prtclrs : num [1:2358] NA NA NA NA NA NA NA NA NA NA ...
##   [list output truncated]
```

2.1 Cleaning of the data

```
##    1    2    3 NA's
##   22 1902  136  298
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##       0    1700    2800    9931    4500  200000    1235
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##       0   19494   33600   38700   50000  288000    1235
```

```
##    0    1    2    3    4    5    6    7    8    9   10 NA's
##   56   45  119  206  214  285  297  422  424  177   76   37
```

```
## not at all   a little      quite       very completely       NA's
##        224        755        802        408        151         18
```

```
##          Yes           No Not eligible         NA's
```

```
##         1860          259          238            1
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##  0.0000  0.0000  0.0000  0.1938  0.0000  1.0000 
```

```
## 
##    0    1 
## 1901  457 
```

```
## 
##    1    2    8 
##  457 1899    2 
```

```
##    1    2    3    4    5 NA's 
##  346  839  834  297   23   19 
```

```
##    0    1    2    3    4    5    6    7    8    9   10 NA's 
##  134   73  161  254  258  396  301  320  281   88   61   31 
```

3. Models

3.1 Linear Model

Question: What is an influence on the

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##    0.00   30.00   45.00   62.98   75.00 1200.00       1 
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##    0.00   12.00   14.00   14.29   16.00   30.00       7 
```

```
##    1    2 
## 1212 1146 
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##   15.00   34.00   51.00   49.65   64.00   90.00       4 
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##   0.000   4.000   6.000   5.887   8.000  10.000      37 
```

```
##    0    1    2    3    4    5    6    7    8    9   10 NA's 
##   56   45  119  206  214  285  297  422  424  177   76   37 
```

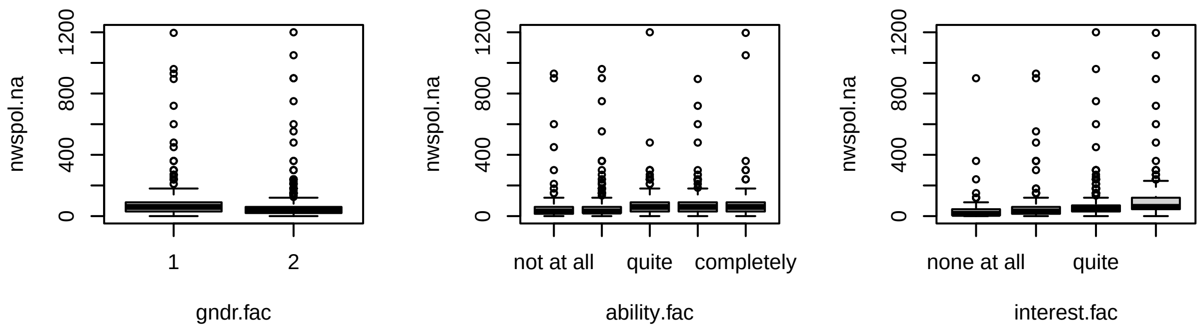
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's 
##   1.000   2.000   3.000   2.789   3.000   5.000      18 
```

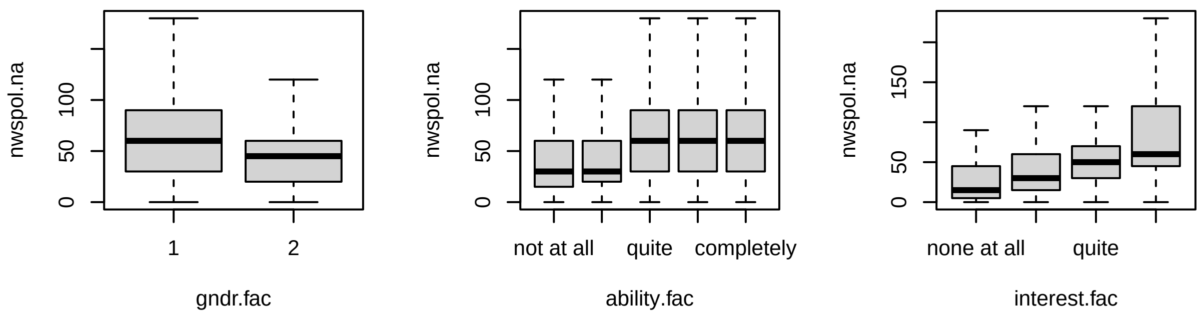
```
## not at all   a little      quite       very completely       NA's 
##        224        755        802        408        151         18 
```

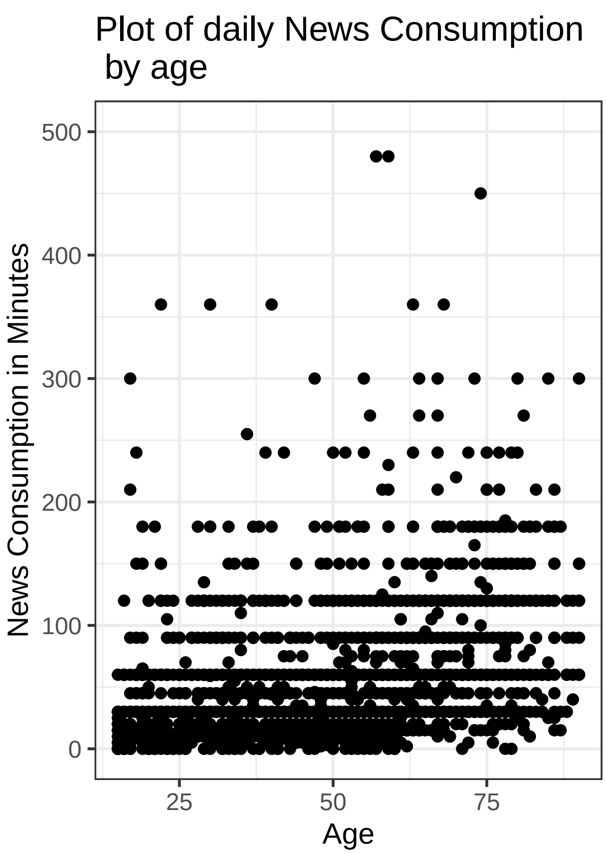
```
## none at all      hardly       quite        very        NA's 
##         100         691         997         569           1 
```

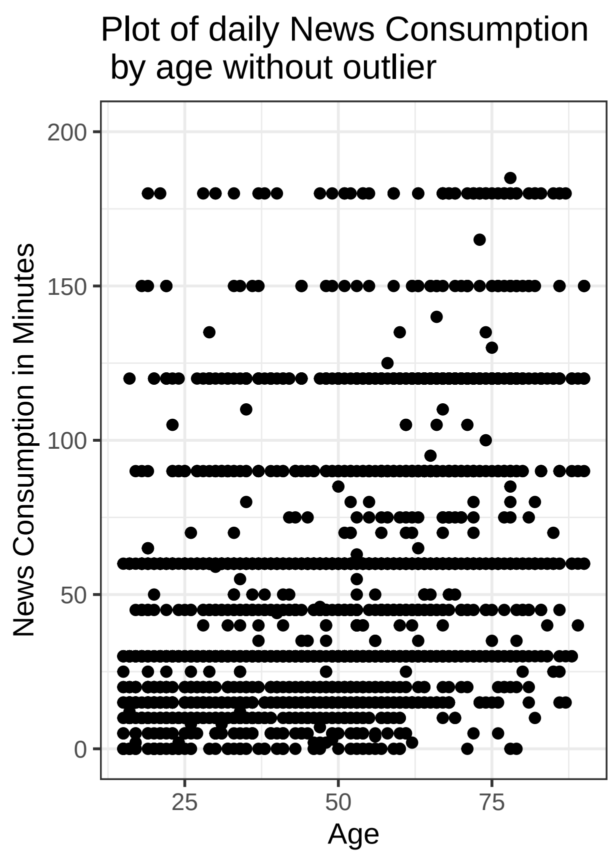
```
##    1    2    3    4 NA's 
##  569  997  691  100    1 
```

Graphical analysis

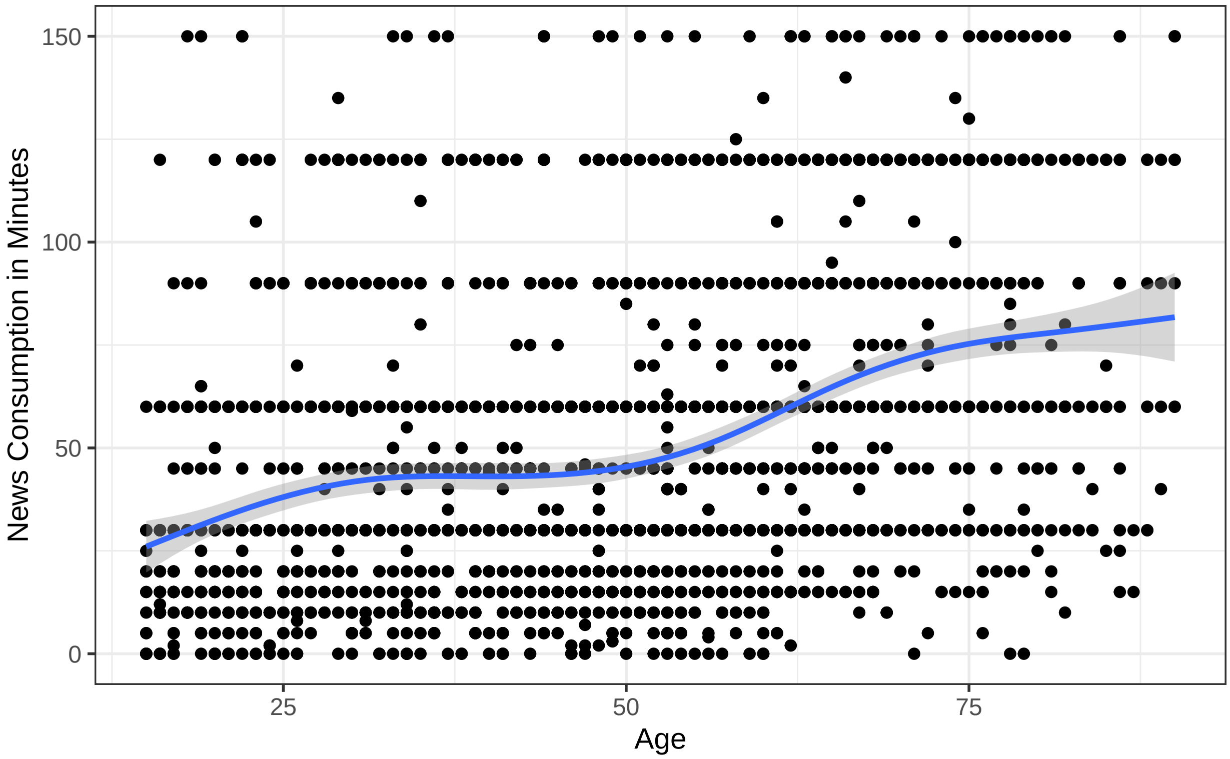




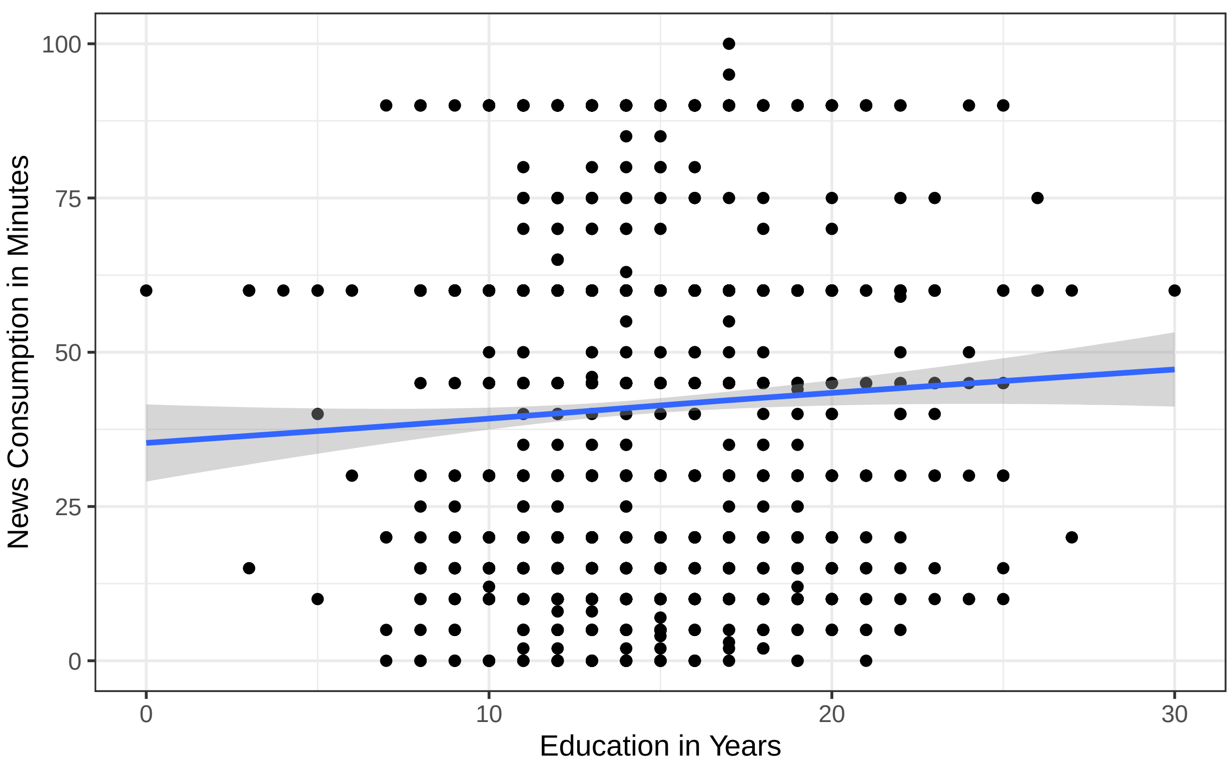




Plot of daily News Consumption
by age without outlier



Plot of daily News Consumption
by age without outlier

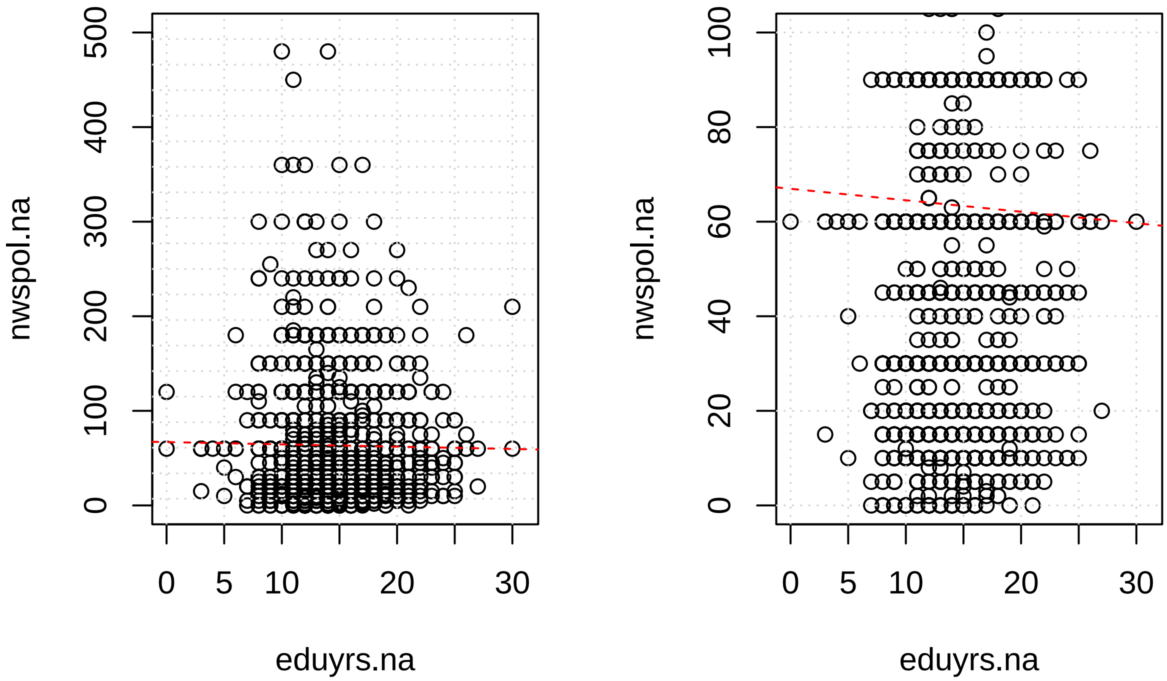


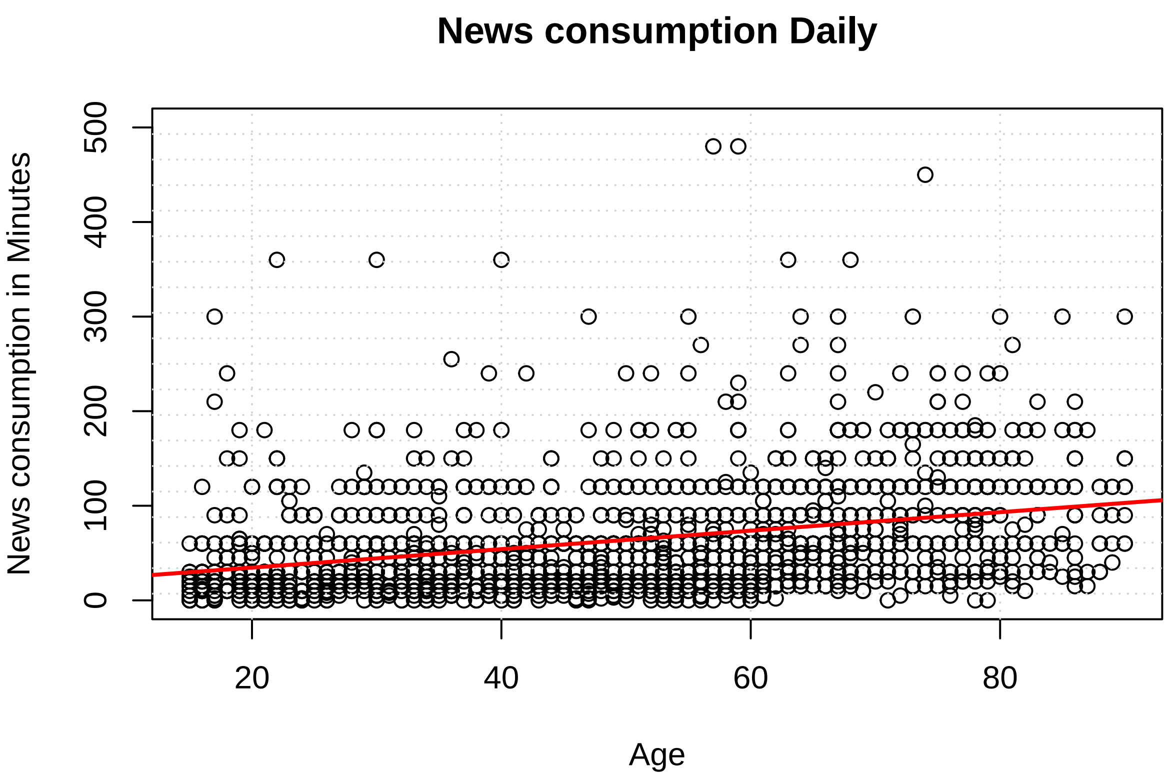
Fitting Model

```
##
## Call:
## lm(formula = nwspol.na ~ eduyrs.na, data = ess9_linear)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -65.26  -34.53  -16.98   11.20 1136.20
##
## Coefficients:
##             Estimate Std. Error t value            Pr(>|t|)
## (Intercept)  66.9534     7.1991   9.300 <0.0000000000000002 ***
## eduyrs.na    -0.2424     0.4876  -0.497               0.619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 81.35 on 2294 degrees of freedom
## Multiple R-squared:  0.0001077,  Adjusted R-squared:  -0.0003281
## F-statistic: 0.2472 on 1 and 2294 DF,  p-value: 0.6191
```

```
##               Estimate Std. Error    t value                    Pr(>|t|)
## (Intercept) 66.9534190  7.1990896  9.3002620 0.0000000000000000000003161288
## eduyrs.na   -0.2424129  0.4875839 -0.4971717 0.6191156521844023430389825l
```

```
##
## Call:
## lm(formula = nwspol.na ~ age.na, data = ess9_linear)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -92.23  -34.89  -15.36   13.66 1119.49
##
## Coefficients:
##             Estimate Std. Error t value             Pr(>|t|)
## (Intercept) 15.09729    4.62268   3.266              0.00111 **
## age.na       0.97636    0.08713  11.206 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 79.21 on 2294 degrees of freedom
## Multiple R-squared:  0.0519, Adjusted R-squared:  0.05149
## F-statistic: 125.6 on 1 and 2294 DF,  p-value: < 0.00000000000000022
```





Interpretation: we can see that the line is nearly flat and the p-value confirms this observation. The p-Value is nearly 1 and the parameter is nearly 0.

```
##
## Call:
## lm(formula = nwspol.outlier ~ eduyrs.na, data = ess9_de, na.omit = TRUE)
##
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
## -59.30 -29.10 -12.57  11.57 302.81
##
## Coefficients:
##             Estimate Std. Error t value            Pr(>|t|)
## (Intercept)  55.2679     4.3602  12.675 <0.0000000000000002 ***
## eduyrs.na     0.1918     0.2962   0.647               0.517
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 50.05 on 2333 degrees of freedom
##   (23 Beobachtungen als fehlend gelöscht)
## Multiple R-squared:  0.0001797,  Adjusted R-squared:  -0.0002489
## F-statistic: 0.4192 on 1 and 2333 DF,  p-value: 0.5174
```

Adding additional predictors

```
##
## Call:
## lm(formula = nwspol.na ~ age.na + eduyrs.na, data = ess9_linear)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
##  -91.77  -34.82  -15.19   13.78 1119.62
##
## Coefficients:
##             Estimate Std. Error t value            Pr(>|t|)
## (Intercept) 13.35443    8.49027   1.573               0.116
## age.na       0.97780    0.08734  11.195 <0.0000000000000002 ***
## eduyrs.na    0.11649    0.47597   0.245               0.807
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 79.23 on 2293 degrees of freedom
## Multiple R-squared:  0.05193,    Adjusted R-squared:  0.0511
## F-statistic:  62.8 on 2 and 2293 DF,  p-value: < 0.00000000000000022
```

```
## (Intercept)      age.na   eduyrs.na
##  13.3544308   0.9778000   0.1164948
```

```
##               Estimate Std. Error    t value
## (Intercept) 13.3544308 8.49026594  1.5729108
## age.na       0.9778000 0.08734175 11.1951044
## eduyrs.na    0.1164948 0.47596568  0.2447547
##                                    Pr(>|t|)
## (Intercept) 0.1158774401912398832603656728679198
## age.na      0.0000000000000000000000000002309964
## eduyrs.na   0.8066682762931414174545352580025792
```

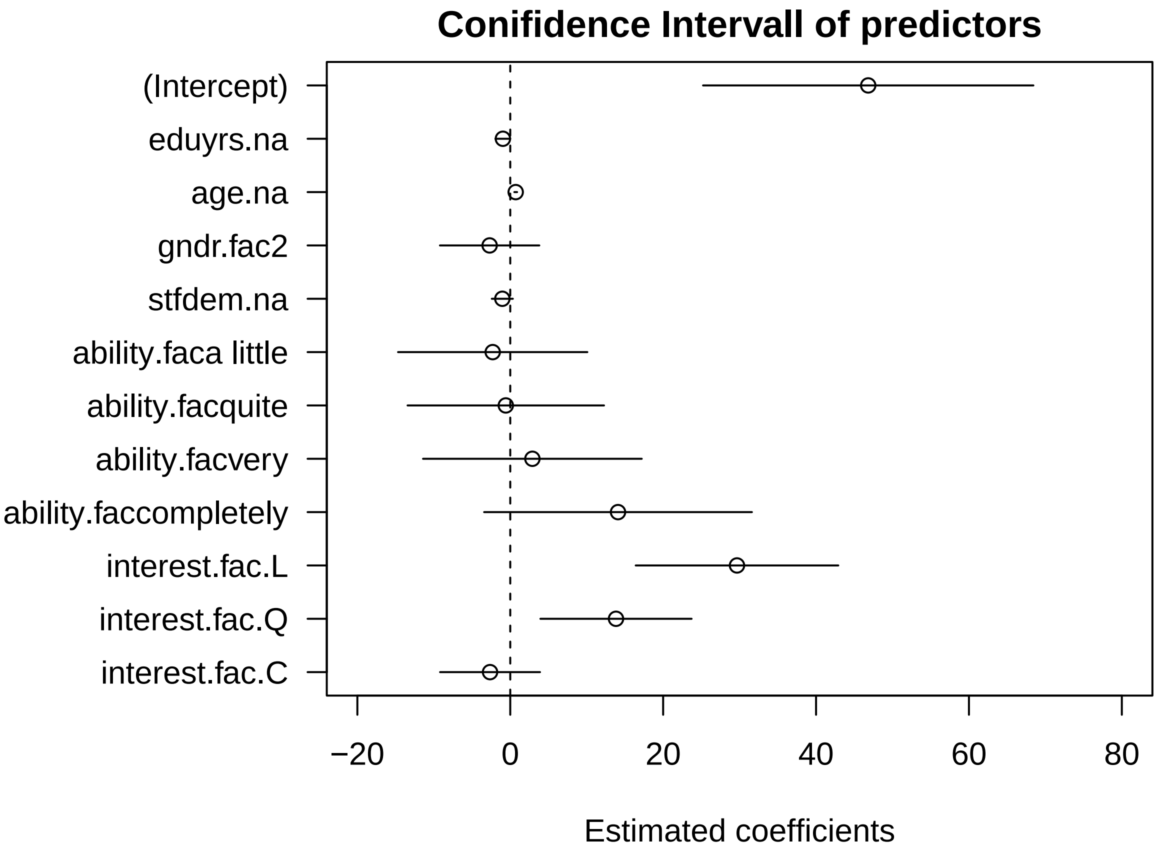
```
##                  2.5 %    97.5 %
## (Intercept) -3.2949730 30.003835
## age.na       0.8065229  1.149077
## eduyrs.na   -0.8168734  1.049863
```

We can see that the predictor age has a low p-value and it is significant.

```
##
```

```
## Call:
## lm(formula = nwspol.na ~ eduyrs.na + age.na + gndr.fac + stfdem.na +
##     ability.fac + interest.fac, data = ess9_linear)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -100.69  -31.90  -12.93   13.30 1127.04
##
## Coefficients:
##                        Estimate Std. Error t value           Pr(>|t|)
## (Intercept)            46.82360   11.01789   4.250 0.0000222574534076 ***
## eduyrs.na              -0.96445    0.49499  -1.948            0.05149 .
## age.na                  0.71236    0.09329   7.636 0.0000000000000328 ***
## gndr.fac2              -2.69782    3.31452  -0.814            0.41576
## stfdem.na              -1.03706    0.70387  -1.473            0.14079
## ability.faca little    -2.29699    6.31408  -0.364            0.71605
## ability.facquite       -0.58392    6.55242  -0.089            0.92900
## ability.facvery         2.89396    7.29519   0.397            0.69163
## ability.faccompletely 14.09631    8.93151   1.578            0.11464
## interest.fac.L         29.65914    6.75701   4.389 0.0000118846240809 ***
## interest.fac.Q         13.83210    5.04290   2.743            0.00614 **
## interest.fac.C         -2.64438    3.33380  -0.793            0.42774
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 77.76 on 2284 degrees of freedom
## Multiple R-squared:  0.09036,    Adjusted R-squared:  0.08598
## F-statistic: 20.62 on 11 and 2284 DF,  p-value: < 0.00000000000000022

##                         Estimate  Std. Error     t value              Pr(>|t|)
## (Intercept)           46.8235974 11.01788591  4.24978056 0.00002225745340756224
## eduyrs.na             -0.9644454  0.49499357 -1.94839985 0.05148945434962375750
## age.na                 0.7123618  0.09329339  7.63571519 0.00000000000003278539
## gndr.fac2             -2.6978177  3.31451767 -0.81393977 0.41576430595330315931
## stfdem.na             -1.0370582  0.70386803 -1.47337019 0.14078903977644036116
## ability.faca little   -2.2969899  6.31407706 -0.36378870 0.71604947094609150415
## ability.facquite      -0.5839200  6.55241874 -0.08911518 0.92899818723781546481
## ability.facvery        2.8939586  7.29518851  0.39669415 0.69163007455917546729
## ability.faccompletely 14.0963060  8.93151086  1.57826668 0.11464283451138868042
## interest.fac.L        29.6591412  6.75701149  4.38938741 0.00001188462408086408
## interest.fac.Q        13.8320989  5.04290452  2.74288336 0.00613777558879450603
## interest.fac.C        -2.6443798  3.33379622 -0.79320378 0.42774149293002872163

##                            2.5 %       97.5 %
## (Intercept)            25.2174882 68.429706625
## eduyrs.na              -1.9351294  0.006238564
## age.na                  0.5294131  0.895310385
## gndr.fac2              -9.1975974  3.801961940
## stfdem.na              -2.4173456  0.343229267
## ability.faca little   -14.6789151 10.084935267
## ability.facquite      -13.4332339 12.265393970
## ability.facvery       -11.4119292 17.199846419
## ability.faccompletely  -3.4184151 31.611027139
## interest.fac.L         16.4086202 42.909662132
## interest.fac.Q          3.9429471 23.721250676
```

```
## interest.fac.C         -9.1819647  3.893205185
```



Interaction: Age * Interest

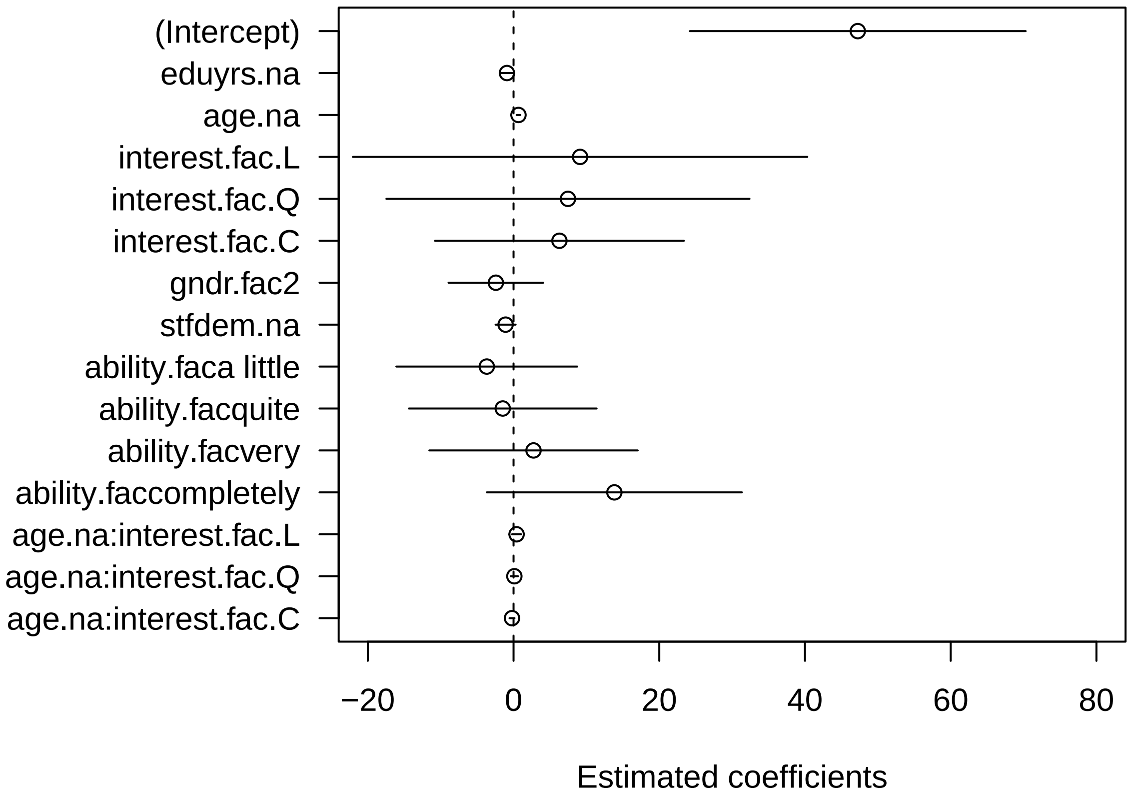
```
##
## Call:
## lm(formula = nwspol.na ~ eduyrs.na + age.na * interest.fac +
##     gndr.fac + stfdem.na + ability.fac, data = ess9_linear)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -105.53  -31.57  -13.78   13.58 1124.69
##
## Coefficients:
##                        Estimate Std. Error t value    Pr(>|t|)
## (Intercept)             47.2447    11.7517   4.020 0.000060026 ***
## eduyrs.na               -0.9064     0.4954  -1.830      0.0674 .
## age.na                   0.6758     0.1302   5.191 0.000000228 ***
## interest.fac.L           9.1358    15.9051   0.574      0.5658
## interest.fac.Q           7.4665    12.7089   0.587      0.5569
## interest.fac.C           6.2965     8.7081   0.723      0.4697
## gndr.fac2               -2.4274     3.3147  -0.732      0.4641
## stfdem.na               -1.0906     0.7037  -1.550      0.1213
## ability.faca little     -3.6725     6.3316  -0.580      0.5620
## ability.facquite        -1.4807     6.5698  -0.225      0.8217
## ability.facvery          2.7415     7.2937   0.376      0.7070
## ability.faccompletely   13.8350     8.9299   1.549      0.1215
## age.na:interest.fac.L    0.4165     0.3182   1.309      0.1906
```

```
## age.na:interest.fac.Q   0.1103     0.2566   0.430      0.6673
## age.na:interest.fac.C  -0.2110     0.1758  -1.201      0.2301
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 77.69 on 2281 degrees of freedom
## Multiple R-squared:  0.09316,    Adjusted R-squared:  0.08759
## F-statistic: 16.74 on 14 and 2281 DF,  p-value: < 0.00000000000000022
```

```
##           (Intercept)             eduyrs.na                age.na
##            47.2447097            -0.9064125             0.6758316
##       interest.fac.L        interest.fac.Q        interest.fac.C
##             9.1357745             7.4664644             6.2965187
##             gndr.fac2             stfdem.na   ability.faca little
##            -2.4274402            -1.0905985            -3.6724999
##      ability.facquite       ability.facvery ability.faccompletely
##            -1.4807135             2.7415278            13.8349518
## age.na:interest.fac.L age.na:interest.fac.Q age.na:interest.fac.C
##             0.4165108             0.1103203            -0.2110107
```

```
##                         Estimate Std. Error    t value        Pr(>|t|)
## (Intercept)           47.2447097 11.7517104  4.0202411 0.0000600263264
## eduyrs.na             -0.9064125  0.4954314 -1.8295417 0.0674489633340
## age.na                 0.6758316  0.1301902  5.1911105 0.0000002275338
## interest.fac.L         9.1357745 15.9050505  0.5743946 0.5657574699023
## interest.fac.Q         7.4664644 12.7089193  0.5874980 0.5569274601306
## interest.fac.C         6.2965187  8.7080929  0.7230652 0.4697139785251
## gndr.fac2             -2.4274402  3.3147418 -0.7323165 0.4640506195895
## stfdem.na             -1.0905985  0.7036944 -1.5498183 0.1213238491175
## ability.faca little   -3.6724999  6.3316232 -0.5800250 0.5619550348906
## ability.facquite      -1.4807135  6.5697715 -0.2253828 0.8217016811062
## ability.facvery        2.7415278  7.2937490  0.3758736 0.7070458104483
## ability.faccompletely 13.8349518  8.9299356  1.5492779 0.1214536545791
## age.na:interest.fac.L  0.4165108  0.3181606  1.3091212 0.1906251586660
## age.na:interest.fac.Q  0.1103203  0.2566211  0.4298956 0.6673121709087
## age.na:interest.fac.C -0.2110107  0.1757617 -1.2005499 0.2300505141922
```

```
##                             2.5 %      97.5 %
## (Intercept)            24.1995523 70.28986719
## eduyrs.na              -1.8779558  0.06513082
## age.na                  0.4205281  0.93113517
## interest.fac.L        -22.0541017 40.32565078
## interest.fac.Q        -17.4557841 32.38871283
## interest.fac.C        -10.7800910 23.37312831
## gndr.fac2              -8.9276639  4.07278341
## stfdem.na              -2.4705464  0.28934943
## ability.faca little   -16.0888417  8.74384199
## ability.facquite      -14.3640653 11.40263833
## ability.facvery       -11.5615471 17.04460271
## ability.faccompletely  -3.6766924 31.34659597
## age.na:interest.fac.L  -0.2074036  1.04042509
## age.na:interest.fac.Q  -0.3929149  0.61355545
## age.na:interest.fac.C  -0.5556802  0.13365880
```



The interaction is not significant

Testing several Variables

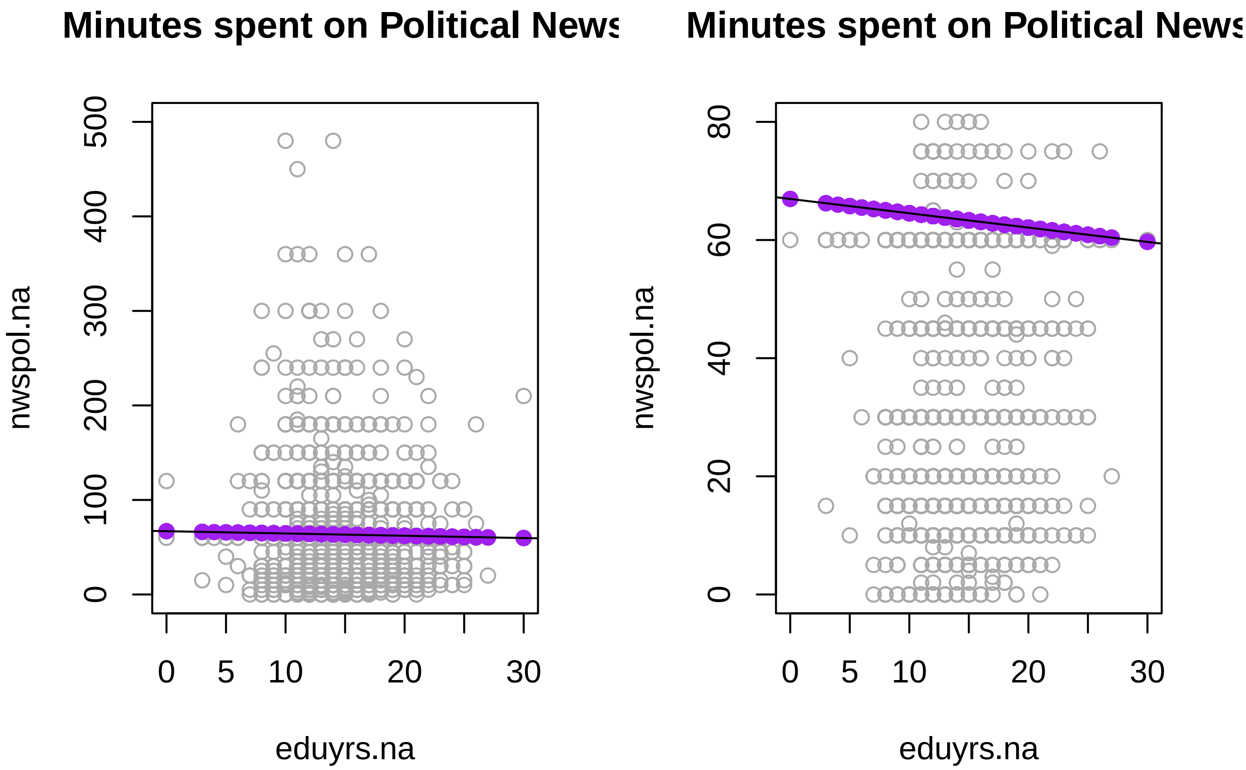
```
## nwspol.na ~ age.na + eduyrs.na

## nwspol.na ~ eduyrs.na + age.na + gndr.fac + stfdem.na + ability.fac +
##     interest.fac

## nwspol.na ~ eduyrs.na + age.na * interest.fac + gndr.fac + stfdem.na +
##     ability.fac

## nwspol.na ~ eduyrs.na + age.na + gndr.fac + stfdem.na + ability.fac +
##     interest.fac

## Single term deletions
##
## Model:
## nwspol.na ~ eduyrs.na + age.na + gndr.fac + stfdem.na + ability.fac +
##     interest.fac
##              Df Sum of Sq      RSS   AIC F value              Pr(>F)
## <none>                    13809380 20004
## eduyrs.na     1     22953 13832333 20006  3.7963             0.05149 .
## age.na        1    352515 14161895 20060 58.3041 0.0000000000003279 ***
## gndr.fac      1      4006 13813386 20002  0.6625             0.41576
## stfdem.na     1     13125 13822505 20004  2.1708             0.14079
## ability.fac   4     33085 13842466 20001  1.3680             0.24256
## interest.fac  3    404127 14213507 20064 22.2802 0.0000000000003227 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Fitted Values and Residuals

Fitted Values

```
##  Named num [1:2296] 63.8 64.3 62.1 64 63.6 ...
##  - attr(*, "names")= chr [1:2296] "1" "2" "3" "4" ...
```

```
##        1        2        3        4        5        6
## 63.80205 64.28688 62.10516 64.04446 63.55964 63.80205
```



There are 67 unique values, therefore, the other observations are overlapping. Here we zoomed in to see the negative relationship between news consumption and education years.

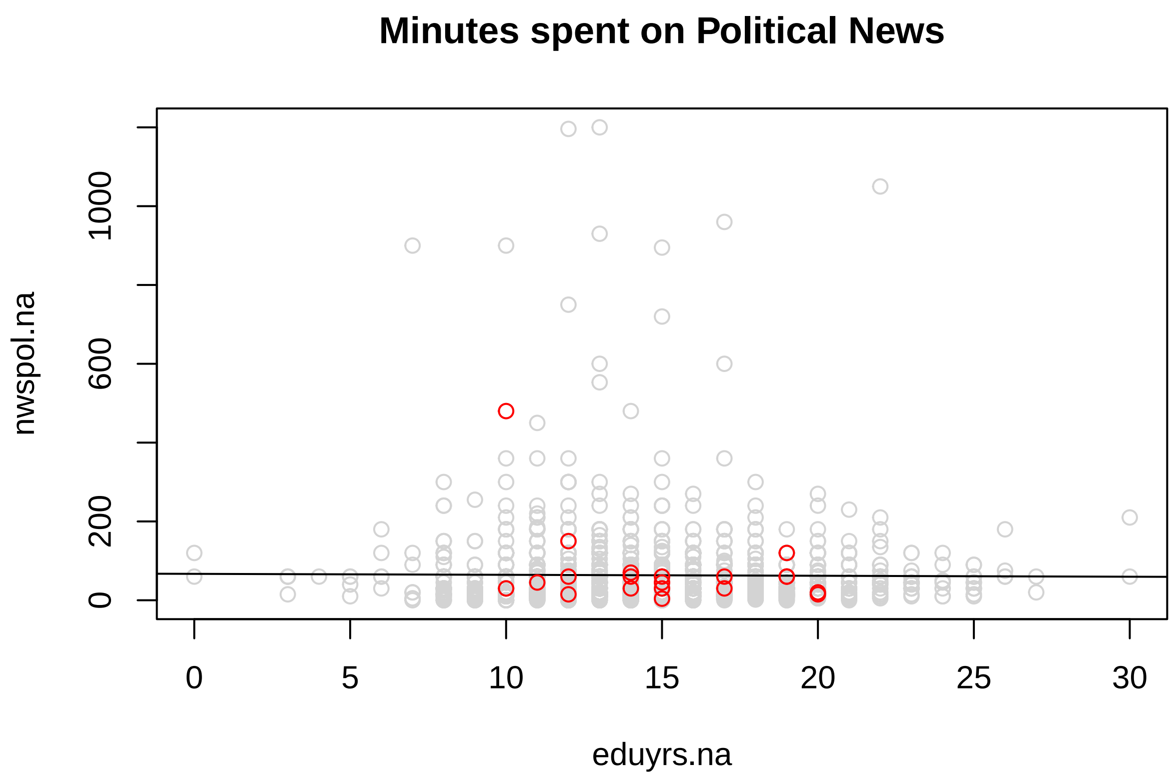
Residuals

```
## [1] 2296
```

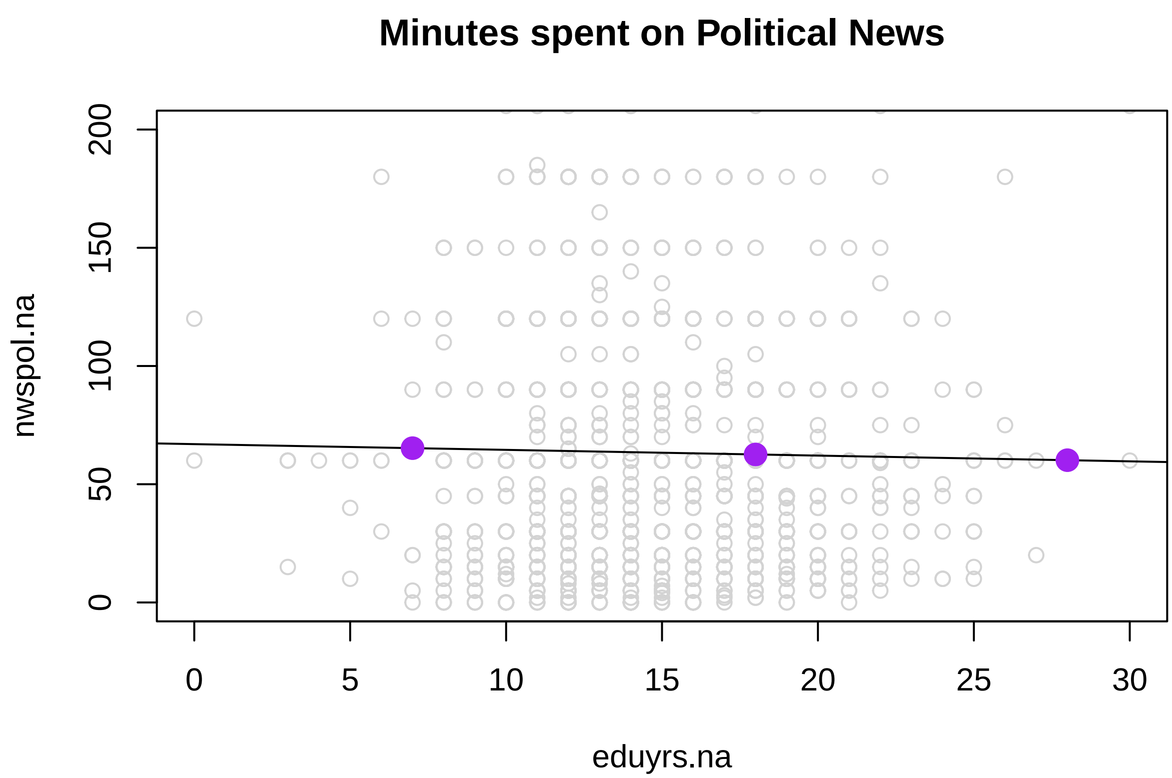
```
##          1          2          3          4          5          6
## -55.802052  -4.286877  57.894839 235.955536 -63.559639 -33.802052
```

```
##        166       1215       1899       1912       1666       1890        541
##  -3.559639 -42.105161 -19.286877 -33.317226  -4.044464   6.440361 -49.044464
##        990       1837        511        127        323        957        633
## 415.470710  -2.832400 -59.317226 -32.832400  -2.347574 -33.559639  -3.317226
##        456       1983        180        969        981        465
##  57.652426 -34.529290 -18.317226  -2.347574 -47.105161  85.955536
```

```
##      166     1215     1899     1912     1666     1890      541      990
## 63.55964 62.10516 64.28688 63.31723 64.04446 63.55964 64.04446 64.52929
##     1837      511      127      323      957      633      456     1983
## 62.83240 63.31723 62.83240 62.34757 63.55964 63.31723 62.34757 64.52929
##      180      969      981      465
## 63.31723 62.34757 62.10516 64.04446
```

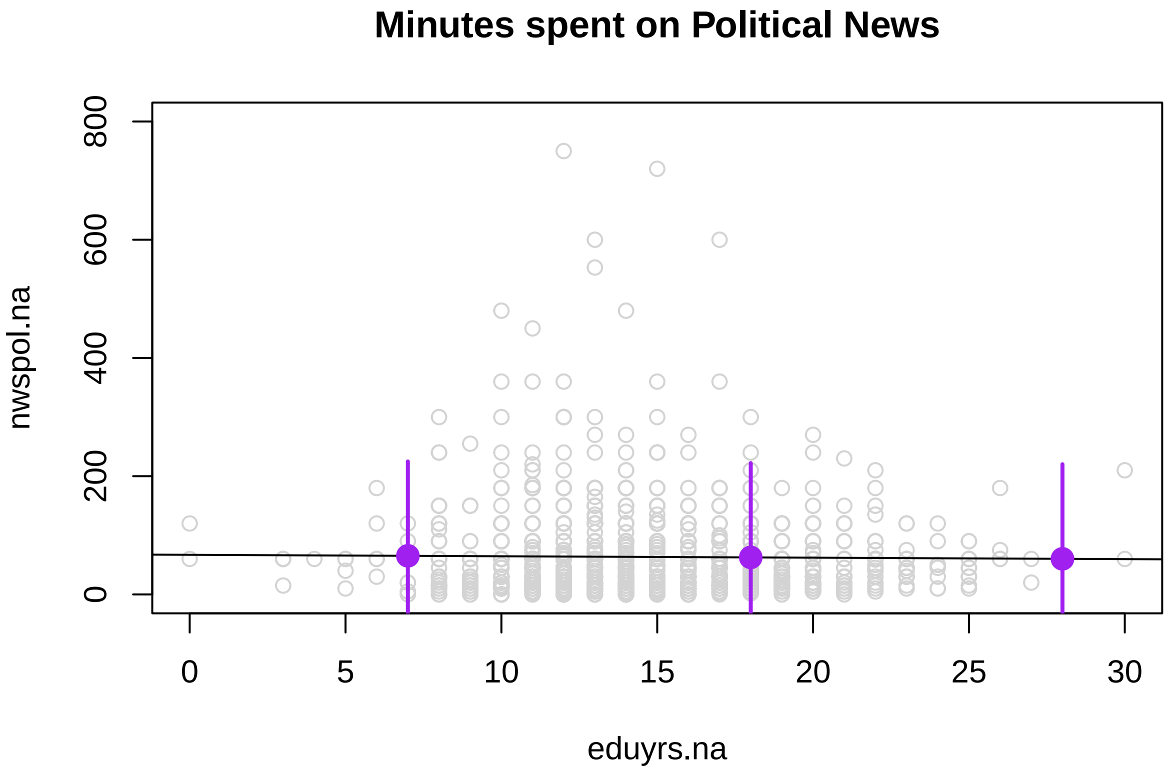


Predicted values




```
##        fit       lwr      upr
## 1 62.58999 -97.00055 222.1805
## 2 65.25653 -94.45044 224.9635
```

```
## 3 60.16586 -99.91952 220.2512
```



3.2 GLM Poisson

The goal of this chapter is to apply a generalized linear model of the poisson type on the ESS9 data set. The poisson model is applied to count data, in this case the information of how many minutes the participants spend per day for consuming media, such as newspaper, TV new shows or online resources. In this chapter the model is used to simulate the data, based on the provided data from the survey. First the data is cleaned and prepared for using it. A subset of the data is used, containing variables which should help to model the news consumption of a person. These parameters are used: Depended variable:

- How much politics are you watching ("nwspol") Independent variables:

The goal of this chapter is to apply a generalized linear model of the poisson type on the ESS9 data set. The poisson model is applied to count data, in this case the information of how many minutes the participants spend per day for consuming media, such as newspaper, TV new shows or online resources. In this chapter the model is used to simulate the data, based on the provided data from the survey. First the data is cleaned and prepared for using it. A subset of the data is used, containing variables which should help to model the news consumption of a person. These parameters are used: Depended variable:

- How much politics are you watching ("nwspol") Independent variables:
- Interest in politics ("polintr)
- Trust into the current parlament ("trstprl")
- Highest level of education ("eisced")
- Years of education ("eduyrs")
- Satisfaction with the general economical situation ("stfeco")
- Satisfaction with the current government ("stfgov")
- Gender ("gndr")

- Age (“agea”)
- Level of person religion believe (“rlgdgr”)
- Time spent online (“netusoft”)
- Posibility for political participation (“psppsgva”)
- Yearly gross income (combination of two factors into a new one - “yrpy”)
- Interest in politics (“polintr)
- Trust into the current parlament (“trstprl”)
- Highest level of education (“eisced”)
- Years of education (“eduyrs”)
- Satisfaction with the general economical situation (“stfeco”)
- Satisfaction with the current government (“stfgov”)
- Gender (“gndr”)
- Age (“agea”)
- Level of person religion believe (“rlgdgr”)
- Time spent online (“netusoft”)
- Posibility for political participation (“psppsgva”)
- Yearly gross income (combination of two factors into a new one - “yrpy”)

3.2.1 Data preparation

The data have to be prepared and transformed. For poisson count data are needed for the predictive variable. The time in minutes is modeled therefore.

```
## [1] 572
```

```
## [1] 708
```

```
## [1] 1752
```

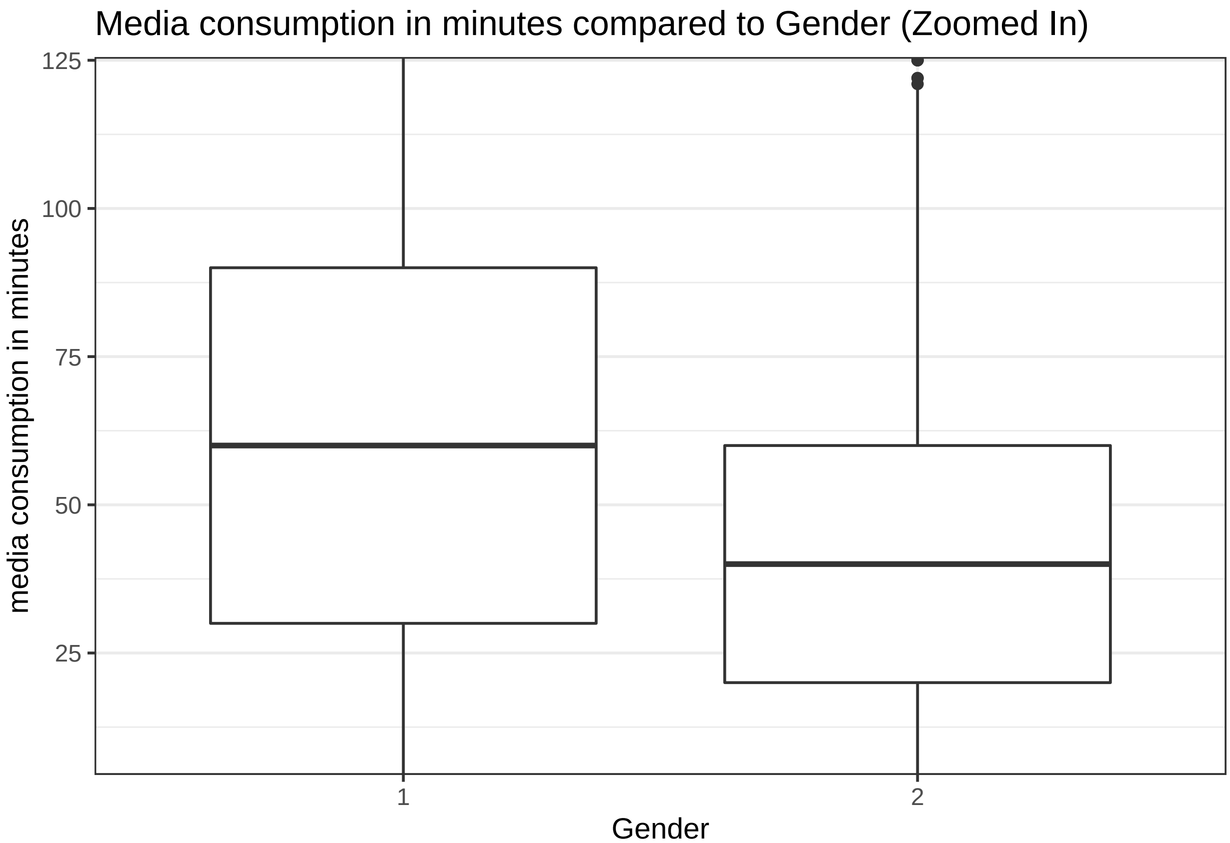
```
## [1] 47
```

Subset of data for this task, as 14 parameters in total. The rows with NA are dropped, as the data set is large enough for droping theses values.

```
## tibble [17,029 x 14] (S3: tbl_df/tbl/data.frame)
##  $ nwspol  : num [1:17029] 60 45 60 120 15 60 30 30 10 10 ...
##  $ polintr : Factor w/ 4 levels "1","2","3","4": 4 3 4 2 4 1 2 3 2 2 ...
##  $ trstprl : Factor w/ 11 levels "0","1","2","3",..: 7 1 1 7 5 4 4 8 8 6 ...
##  $ eisced  : Factor w/ 8 levels "1","2","3","4",..: 2 3 3 3 3 4 3 6 2 7 ...
##  $ eduyrs  : num [1:17029] 12 11 12 12 13 21 18 17 9 17 ...
##  $ stfeco  : Factor w/ 11 levels "0","1","2","3",..: 6 7 2 11 10 8 7 8 7 9 ...
##  $ stfgov  : Factor w/ 11 levels "0","1","2","3",..: 7 9 4 11 9 3 8 3 8 7 ...
##  $ stfdem  : Factor w/ 11 levels "0","1","2","3",..: 7 7 4 11 8 4 11 7 9 8 ...
##  $ gndr    : Factor w/ 2 levels "1","2": 2 1 1 1 2 1 1 1 1 1 ...
##  $ agea    : num [1:17029] 40 63 56 48 41 27 49 42 50 35 ...
##  $ rlgdgr  : Factor w/ 11 levels "0","1","2","3",..: 5 2 9 1 4 4 3 1 4 3 ...
##  $ netusoft: Factor w/ 5 levels "1","2","3","4",..: 4 5 1 1 4 5 5 5 5 5 ...
##  $ psppsgva: Factor w/ 5 levels "1","2","3","4",..: 2 2 2 5 1 3 1 1 2 3 ...
##  $ yrpy    : num [1:17029] 31200 30600 18000 31200 37200 18000 20400 17400 45600 70000 ...
```

The plot shows the media consumption for male and female. Males have a higher mean and the variance of the first and thrid quantile are larger.

The plot shows the media consumption for male and female. Males have a higher mean and the variance of the first and third quantile are larger.



3.2.2 Fitting the poisson model

Using the parameter specified above to model the consumption.

```
##
## Call:
## glm(formula = nwspol ~ polintr + eisced + trstprl + eduyrs +
##     netusoft + stfeco + stfgov + stfdem + gndr + agea + rlgdgr +
##     yrpy, family = "poisson", data = ess9_poisson)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -17.160   -6.476   -3.322    0.474   74.725
##
## Coefficients:
##                     Estimate       Std. Error  z value             Pr(>|z|)
## (Intercept)  4.9551478455834  0.0105343674796  470.379 < 0.0000000000000002 ***
## polintr2    -0.3492665558265  0.0026678349780 -130.918 < 0.0000000000000002 ***
## polintr3    -0.5376726514740  0.0028816344085 -186.586 < 0.0000000000000002 ***
## polintr4    -0.6563093286586  0.0037354537637 -175.697 < 0.0000000000000002 ***
## eisced2     -0.1458105762706  0.0064736859414  -22.524 < 0.0000000000000002 ***
## eisced3     -0.2098897454780  0.0063900096516  -32.847 < 0.0000000000000002 ***
```

```
## eisced4     -0.0801180098272  0.0062959701053  -12.725 < 0.0000000000000002 ***
## eisced5     -0.1793751705951  0.0066035938535  -27.163 < 0.0000000000000002 ***
## eisced6     -0.0136582683597  0.0068047744981   -2.007             0.0447 *
## eisced7     -0.0747622892396  0.0070310740541  -10.633 < 0.0000000000000002 ***
## eisced55     0.1991819464428  0.0184339911655   10.805 < 0.0000000000000002 ***
## trstprl1     0.0438116718767  0.0056421385757    7.765 0.000000000000008159 ***
## trstprl2     0.0878787242680  0.0050348881462   17.454 < 0.0000000000000002 ***
## trstprl3     0.0575992821977  0.0049226015149   11.701 < 0.0000000000000002 ***
## trstprl4     0.1649699717288  0.0049617090096   33.249 < 0.0000000000000002 ***
## trstprl5     0.1785393616823  0.0046962221481   38.018 < 0.0000000000000002 ***
## trstprl6     0.0926411783091  0.0050274726710   18.427 < 0.0000000000000002 ***
## trstprl7     0.1654323040088  0.0050166399191   32.977 < 0.0000000000000002 ***
## trstprl8     0.2017014330997  0.0053305193535   37.839 < 0.0000000000000002 ***
## trstprl9     0.1425488919881  0.0068020761893   20.957 < 0.0000000000000002 ***
## trstprl10    0.1034979576788  0.0080156227526   12.912 < 0.0000000000000002 ***
## eduyrs      -0.0047838263191  0.0003248431772  -14.727 < 0.0000000000000002 ***
## netusoft2   -0.1307787280215  0.0062823662011  -20.817 < 0.0000000000000002 ***
## netusoft3   -0.1124713541777  0.0061051654140  -18.422 < 0.0000000000000002 ***
## netusoft4   -0.1650624943097  0.0053911538839  -30.617 < 0.0000000000000002 ***
## netusoft5   -0.2565499520473  0.0047119730917  -54.446 < 0.0000000000000002 ***
## stfeco1      0.2042630443847  0.0078684572864   25.960 < 0.0000000000000002 ***
## stfeco2      0.1454844795976  0.0065882376682   22.082 < 0.0000000000000002 ***
## stfeco3      0.1097594146381  0.0063537352177   17.275 < 0.0000000000000002 ***
## stfeco4      0.1173993792901  0.0063956760982   18.356 < 0.0000000000000002 ***
## stfeco5      0.1060677186837  0.0062503682707   16.970 < 0.0000000000000002 ***
## stfeco6      0.0454950862456  0.0063907869403    7.119 0.000000000001088272 ***
## stfeco7     -0.0695855130540  0.0063960463742  -10.879 < 0.0000000000000002 ***
## stfeco8     -0.0844850487183  0.0065231751269  -12.952 < 0.0000000000000002 ***
## stfeco9     -0.1059951308680  0.0073197601880  -14.481 < 0.0000000000000002 ***
## stfeco10    -0.2077250234061  0.0087105452425  -23.848 < 0.0000000000000002 ***
## stfgov1     -0.0371282966740  0.0053909882570   -6.887 0.000000000005693961 ***
## stfgov2     -0.1628748230687  0.0050296948069  -32.383 < 0.0000000000000002 ***
## stfgov3     -0.1492579265587  0.0049582011543  -30.103 < 0.0000000000000002 ***
## stfgov4     -0.0842706452037  0.0050657063724  -16.636 < 0.0000000000000002 ***
## stfgov5     -0.0532296925917  0.0049390418045  -10.777 < 0.0000000000000002 ***
## stfgov6     -0.1646096319626  0.0052200210811  -31.534 < 0.0000000000000002 ***
## stfgov7     -0.0601493306535  0.0052887394475  -11.373 < 0.0000000000000002 ***
## stfgov8     -0.1257885537272  0.0059854927793  -21.016 < 0.0000000000000002 ***
## stfgov9      0.1018727906473  0.0080519498858   12.652 < 0.0000000000000002 ***
## stfgov10     0.0782407518655  0.0098315985238    7.958 0.000000000000001747 ***
## stfdem1     -0.0563098592349  0.0069113969117   -8.147 0.000000000000000372 ***
## stfdem2     -0.0025138003015  0.0061045169098   -0.412             0.6805
## stfdem3     -0.1749548063588  0.0060699082915  -28.823 < 0.0000000000000002 ***
## stfdem4     -0.0834553121230  0.0061336706557  -13.606 < 0.0000000000000002 ***
## stfdem5     -0.0745510981878  0.0059070855389  -12.621 < 0.0000000000000002 ***
## stfdem6     -0.1142834927352  0.0061219691225  -18.668 < 0.0000000000000002 ***
## stfdem7     -0.0757921344766  0.0060426970153  -12.543 < 0.0000000000000002 ***
## stfdem8     -0.1281854636406  0.0062024480020  -20.667 < 0.0000000000000002 ***
## stfdem9     -0.1819244111818  0.0069479462167  -26.184 < 0.0000000000000002 ***
## stfdem10    -0.1081689324763  0.0079802880965  -13.555 < 0.0000000000000002 ***
## gndr2       -0.0955221839468  0.0018928147422  -50.466 < 0.0000000000000002 ***
## agea         0.0039442838404  0.0000752630756   52.407 < 0.0000000000000002 ***
## rlgdgr1      0.0524594264550  0.0039129662333   13.407 < 0.0000000000000002 ***
## rlgdgr2      0.0229590182679  0.0037164081451    6.178 0.000000000650240197 ***
```

```
## rlgdgr3      0.1252369296839  0.0035904157083   34.881 < 0.0000000000000002 ***
## rlgdgr4      0.0911051329206  0.0040471525705   22.511 < 0.0000000000000002 ***
## rlgdgr5      0.0008947887099  0.0032698289371    0.274             0.7844
## rlgdgr6      0.0599344542248  0.0035893248803   16.698 < 0.0000000000000002 ***
## rlgdgr7     -0.0077316837152  0.0035782505209   -2.161             0.0307 *
## rlgdgr8      0.0079295142834  0.0038527509859    2.058             0.0396 *
## rlgdgr9     -0.0750978361239  0.0059677036213  -12.584 < 0.0000000000000002 ***
## rlgdgr10     0.1758154007801  0.0047371153192   37.114 < 0.0000000000000002 ***
## yrpy        -0.0000000085380  0.0000000005367  -15.907 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1622523  on 17028  degrees of freedom
## Residual deviance: 1532064  on 16960  degrees of freedom
## AIC: 1622455
##
## Number of Fisher Scoring iterations: 6
```

```
##        (Intercept)           polintr2           polintr3           polintr4
##  4.955147845583398 -0.349266555826538 -0.537672651474009 -0.656309328658579
##           eisced2            eisced3            eisced4            eisced5
## -0.145810576270611 -0.209889745477968 -0.080118009827152 -0.179375170595105
##           eisced6            eisced7           eisced55           trstprl1
## -0.013658268359742 -0.074762289239633  0.199181946442816  0.043811671876679
##          trstprl2           trstprl3           trstprl4           trstprl5
##  0.087878724267998  0.057599282197725  0.164969971728847  0.178539361682290
##          trstprl6           trstprl7           trstprl8           trstprl9
##  0.092641178309068  0.165432304008800  0.201701433099665  0.142548891988131
##         trstprl10             eduyrs          netusoft2          netusoft3
##  0.103497957678787 -0.004783826319139 -0.130778728021462 -0.112471354177693
##         netusoft4          netusoft5            stfeco1            stfeco2
## -0.165062494309659 -0.256549952047336  0.204263044384658  0.145484479597629
##           stfeco3            stfeco4            stfeco5            stfeco6
##  0.109759414638132  0.117399379290054  0.106067718683748  0.045495086245550
##           stfeco7            stfeco8            stfeco9           stfeco10
## -0.069585513054008 -0.084485048718322 -0.105995130868024 -0.207725023406075
##           stfgov1            stfgov2            stfgov3            stfgov4
## -0.037128296673999 -0.162874823068651 -0.149257926558705 -0.084270645203717
##           stfgov5            stfgov6            stfgov7            stfgov8
## -0.053229692591687 -0.164609631962625 -0.060149330653520 -0.125788553727176
##           stfgov9           stfgov10            stfdem1            stfdem2
##  0.101872790647289  0.078240751865529 -0.056309859234943 -0.002513800301504
##           stfdem3            stfdem4            stfdem5            stfdem6
## -0.174954806358804 -0.083455312122970 -0.074551098187797 -0.114283492735243
##           stfdem7            stfdem8            stfdem9           stfdem10
## -0.075792134476610 -0.128185463640629 -0.181924411181762 -0.108168932476321
##             gndr2               agea            rlgdgr1            rlgdgr2
## -0.095522183946788  0.003944283840447  0.052459426455015  0.022959018267916
##           rlgdgr3            rlgdgr4            rlgdgr5            rlgdgr6
##  0.125236929683865  0.091105132920631  0.000894788709914  0.059934454224848
##           rlgdgr7            rlgdgr8            rlgdgr9           rlgdgr10
## -0.007731683715226  0.007929514283363 -0.075097836123885  0.175815400780054
```

```
##               yrpy
## -0.000000008538016
```

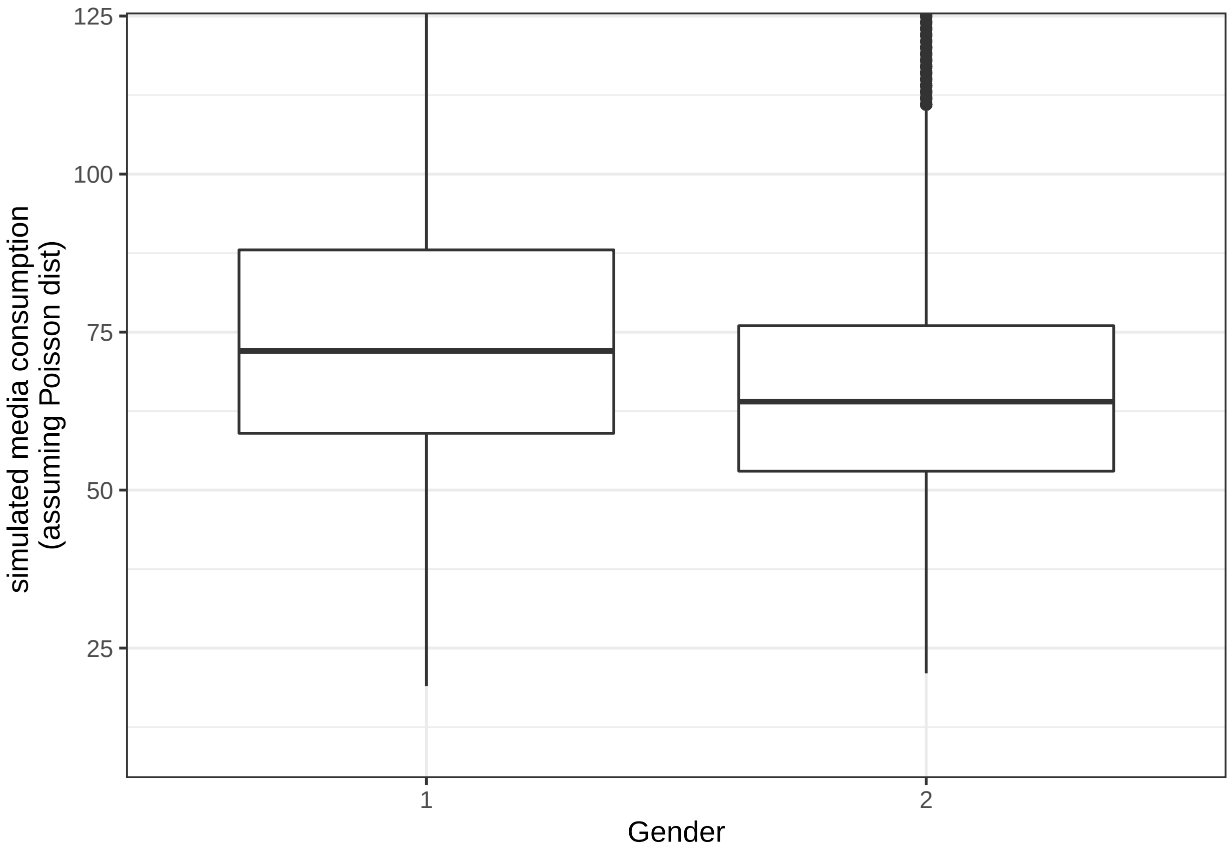
```
## (Intercept)
##    141.9036
```

3.2.3 Simulation

Now we simulate with our model the average media consumption for Females and Males.

```
## [1] 17029
```

```
##   sim_1
## 1    48
## 2    47
## 3    53
## 4    81
## 5    44
## 6    79
```



The simulated data look similar as the real data based of the survey. The variance is smaller for the simulated one. The higher mean for males are in both plots visible.

The simulated data look similar as the real data based of the survey. The variance is smaller for the simulated one. The higher mean and the higher variance for the male group is indicated in the real data as well as in the simulated data.

3.2.4 GLM Quasi-Poisson

```
##
```

```
## Call:
## glm(formula = nwspol ~ polintr + eisced + trstprl + eduyrs +
##     netusoft + stfeco + stfgov + stfdem + gndr + agea + rlgdgr +
##     yrpy, family = "quasipoisson", data = ess9_poisson)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -17.160   -6.476   -3.322    0.474   74.725
##
## Coefficients:
##                   Estimate      Std. Error t value             Pr(>|t|)
## (Intercept)  4.955147845583  0.151318671203  32.746 < 0.0000000000000002 ***
## polintr2    -0.349266555827  0.038321545611  -9.114 < 0.0000000000000002 ***
## polintr3    -0.537672651474  0.041392621858 -12.990 < 0.0000000000000002 ***
## polintr4    -0.656309328659  0.053657127583 -12.232 < 0.0000000000000002 ***
## eisced2     -0.145810576271  0.092989878732  -1.568             0.116894
## eisced3     -0.209889745478  0.091787928543  -2.287             0.022227 *
## eisced4     -0.080118009827  0.090437117569  -0.886             0.375685
## eisced5     -0.179375170595  0.094855913182  -1.891             0.058638 .
## eisced6     -0.013658268360  0.097745729573  -0.140             0.888873
## eisced7     -0.074762289240  0.100996361201  -0.740             0.459160
## eisced55     0.199181946443  0.264791128042   0.752             0.451927
## trstprl1     0.043811671877  0.081045294240   0.541             0.588802
## trstprl2     0.087878724268  0.072322575173   1.215             0.224347
## trstprl3     0.057599282198  0.070709657845   0.815             0.415319
## trstprl4     0.164969971729  0.071271409098   2.315             0.020643 *
## trstprl5     0.178539361682  0.067457879792   2.647             0.008136 **
## trstprl6     0.092641178309  0.072216057163   1.283             0.199568
## trstprl7     0.165432304009  0.072060452413   2.296             0.021703 *
## trstprl8     0.201701433100  0.076569106494   2.634             0.008440 **
## trstprl9     0.142548891988  0.097706970293   1.459             0.144599
## trstprl10    0.103497957679  0.115138700652   0.899             0.368720
## eduyrs      -0.004783826319  0.004666140422  -1.025             0.305273
## netusoft2   -0.130778728021  0.090241707193  -1.449             0.147299
## netusoft3   -0.112471354178  0.087696344343  -1.283             0.199682
## netusoft4   -0.165062494310  0.077440078253  -2.131             0.033063 *
## netusoft5   -0.256549952047  0.067684130856  -3.790             0.000151 ***
## stfeco1      0.204263044385  0.113024773751   1.807             0.070742 .
## stfeco2      0.145484479598  0.094635332540   1.537             0.124234
## stfeco3      0.109759414638  0.091266872185   1.203             0.229140
## stfeco4      0.117399379290  0.091869323004   1.278             0.201304
## stfeco5      0.106067718684  0.089782079758   1.181             0.237464
## stfeco6      0.045495086246  0.091799093740   0.496             0.620187
## stfeco7     -0.069585513054  0.091874641754  -0.757             0.448823
## stfeco8     -0.084485048718  0.093700755564  -0.902             0.367257
## stfeco9     -0.105995130868  0.105143131500  -1.008             0.313419
## stfeco10    -0.207725023406  0.125120766301  -1.660             0.096893 .
## stfgov1     -0.037128296674  0.077437699141  -0.479             0.631617
## stfgov2     -0.162874823069  0.072247976560  -2.254             0.024185 *
## stfgov3     -0.149257926559  0.071221021260  -2.096             0.036124 *
## stfgov4     -0.084270645204  0.072765256999  -1.158             0.246833
## stfgov5     -0.053229692592  0.070945810873  -0.750             0.453093
## stfgov6     -0.164609631963  0.074981877665  -2.195             0.028154 *
## stfgov7     -0.060149330654  0.075968967960  -0.792             0.428511
```

```
## stfgov8     -0.125788553727  0.085977332347  -1.463             0.143474
## stfgov9      0.101872790647  0.115660513996   0.881             0.378443
## stfgov10     0.078240751866  0.141223896670   0.554             0.579573
## stfdem1     -0.056309859235  0.099277284456  -0.567             0.570587
## stfdem2     -0.002513800302  0.087687029042  -0.029             0.977130
## stfdem3     -0.174954806359  0.087189900939  -2.007             0.044809 *
## stfdem4     -0.083455312123  0.088105801797  -0.947             0.343542
## stfdem5     -0.074551098188  0.084851068293  -0.879             0.379625
## stfdem6     -0.114283492735  0.087937717623  -1.300             0.193757
## stfdem7     -0.075792134477  0.086799030374  -0.873             0.382571
## stfdem8     -0.128185463641  0.089093739295  -1.439             0.150234
## stfdem9     -0.181924411182  0.099802289140  -1.823             0.068344 .
## stfdem10    -0.108168932476  0.114631143533  -0.944             0.345374
## gndr2       -0.095522183947  0.027188933002  -3.513             0.000444 ***
## agea         0.003944283840  0.001081100371   3.648             0.000265 ***
## rlgdgr1      0.052459426455  0.056206967531   0.933             0.350665
## rlgdgr2      0.022959018268  0.053383550864   0.430             0.667145
## rlgdgr3      0.125236929684  0.051573759421   2.428             0.015180 *
## rlgdgr4      0.091105132921  0.058134458505   1.567             0.117099
## rlgdgr5      0.000894788710  0.046968759233   0.019             0.984801
## rlgdgr6      0.059934454225  0.051558090456   1.162             0.245063
## rlgdgr7     -0.007731683715  0.051399015186  -0.150             0.880431
## rlgdgr8      0.007929514283  0.055342018474   0.143             0.886069
## rlgdgr9     -0.075097836124  0.085721803787  -0.876             0.381007
## rlgdgr10     0.175815400780  0.068045281012   2.584             0.009780 **
## yrpy        -0.000000008538  0.000000007710  -1.107             0.268132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 206.3327)
##
##     Null deviance: 1622523  on 17028  degrees of freedom
## Residual deviance: 1532064  on 16960  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 6
```

```
## Analysis of Deviance Table
##
## Model 1: nwspol ~ polintr + eisced + trstprl + eduyrs + netusoft + stfeco +
##     stfgov + stfdem + gndr + agea + rlgdgr + yrpy
## Model 2: nwspol ~ polintr + eisced + trstprl + eduyrs + netusoft + stfeco +
##     stfgov + stfdem + agea + rlgdgr + yrpy
##   Resid. Df Resid. Dev Df Deviance     F    Pr(>F)
## 1     16960    1532064
## 2     16961    1534617 -1  -2552.4 12.37 0.0004373 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yes, Gender plays a role for the quasi poisson model.

```
## Analysis of Deviance Table
##
## Model 1: nwspol ~ polintr + eisced + trstprl + eduyrs + netusoft + stfeco +
##     stfgov + stfdem + gndr + agea + rlgdgr + yrpy
```

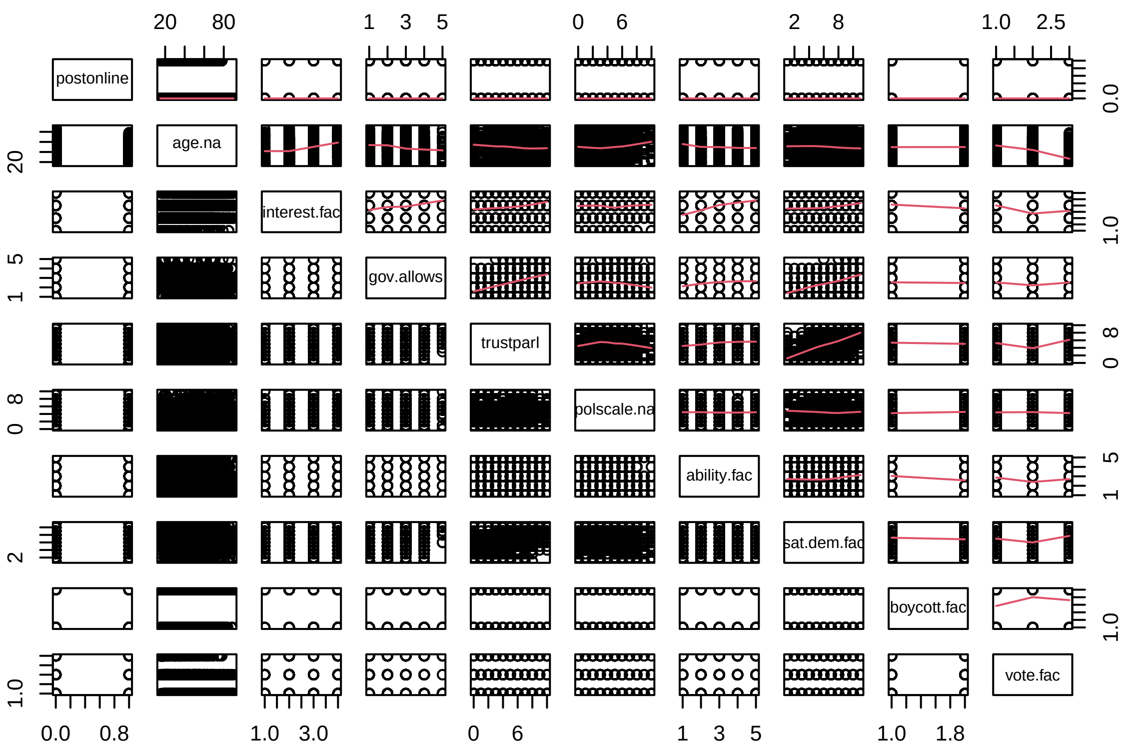
```
## Model 2: nwspol ~ polintr + eisced + trstprl + eduyrs + netusoft + stfeco +
##     stfgov + gndr + agea + rlgdgr + yrpy
##   Resid. Df Resid. Dev  Df Deviance      F Pr(>F)
## 1     16960    1532064
## 2     16970    1534358 -10  -2293.3 1.1114 0.3488
```

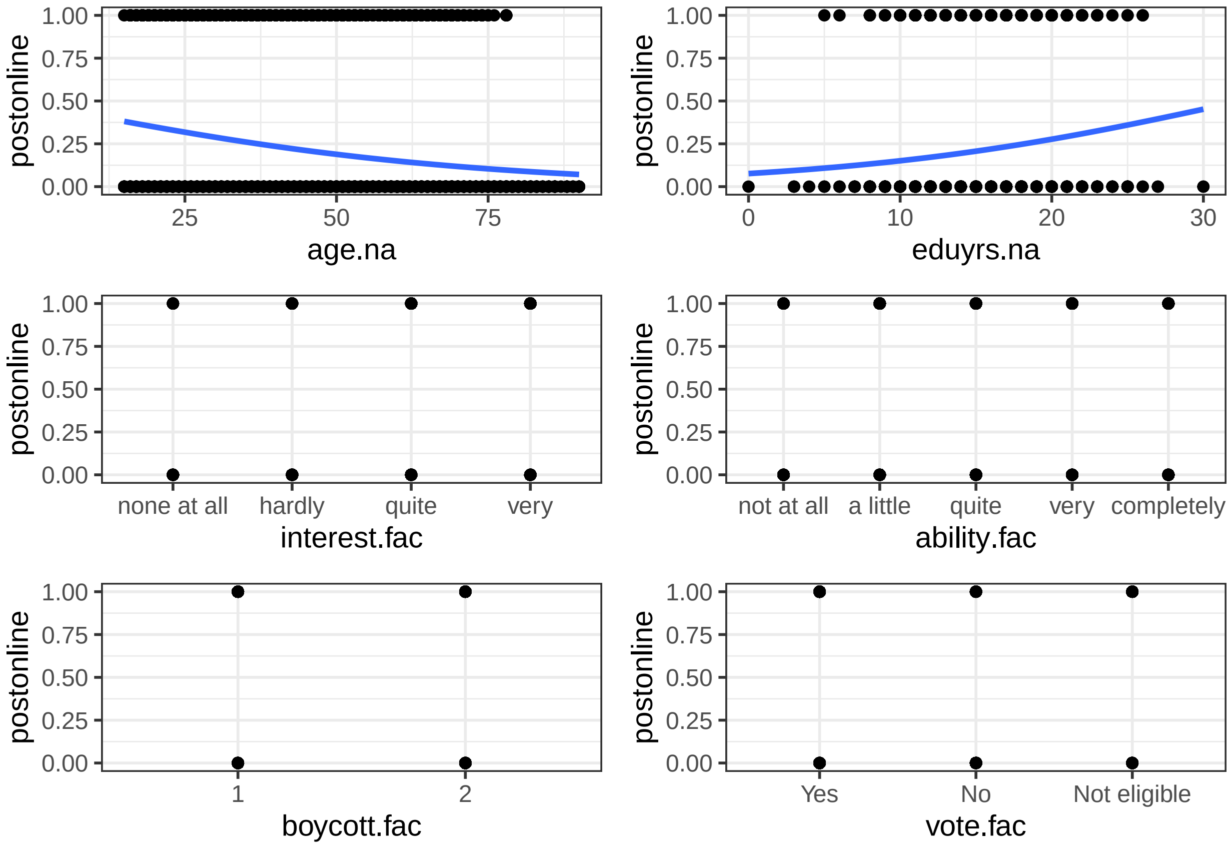
To compare it with the variable of "stfdem" about how satisfied the participants of the survey are about how the democracy in their country works. This variable doensn't play a role to predict the media consumption according to the quasipoisson model.

3.3 GLM Binomial

```
## tibble [2,212 x 11] (S3: tbl_df/tbl/data.frame)
##  $ postonline  : num [1:2212] 0 0 0 0 1 0 0 0 0 0 ...
##  $ age.na      : num [1:2212] 26 65 74 64 54 20 71 41 62 65 ...
##  $ interest.fac: Ord.factor w/ 4 levels "none at all"<..: 2 4 4 4 2 3 4 2 3 3 ...
##  $ gov.allows  : Factor w/ 5 levels "1","2","3","4",..: 4 4 2 3 2 3 3 3 2 3 ...
##  $ trustparl   : num [1:2212] 2 7 3 3 4 9 10 5 5 7 ...
##  $ polscale.na : num [1:2212] 5 5 3 2 1 5 5 2 3 5 ...
##  $ eduyrs.na   : num [1:2212] 13 11 20 12 14 13 12 14 16 14 ...
##  $ ability.fac : Factor w/ 5 levels "not at all","a little",..: 3 4 3 3 4 4 3 2 4 3 ...
##  $ sat.dem.fac : Factor w/ 11 levels "0","1","2","3",..: 8 10 7 5 6 11 11 8 7 10 ...
##  $ boycott.fac : Factor w/ 2 levels "1","2": 2 1 2 1 1 2 1 2 1 2 ...
##  $ vote.fac    : Factor w/ 3 levels "Yes","No","Not eligible": 2 1 1 1 1 1 1 1 1 1 ...
```

Graphical Analysis





Fitting the binary model (logistic regression

```
##
## Call:
## glm(formula = postonline ~ age.na, family = "binomial", data = ess9_binary.na)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9792  -0.7141  -0.5696  -0.4445   2.1626
##
## Coefficients:
##              Estimate Std. Error z value            Pr(>|z|)
## (Intercept) -0.069127   0.140640  -0.492               0.623
## age.na      -0.027793   0.002954  -9.408 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2232.0  on 2211  degrees of freedom
## Residual deviance: 2138.1  on 2210  degrees of freedom
## AIC: 2142.1
##
## Number of Fisher Scoring iterations: 4
```

```
##
## Call:
```

```
## glm(formula = postonline ~ age.na + eduyrs.na + interest.fac,
##     family = "binomial", data = ess9_binary.na)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.3493  -0.7204  -0.5271  -0.3221   2.7383
##
## Coefficients:
##                 Estimate Std. Error z value             Pr(>|z|)
## (Intercept)    -0.623148   0.292110  -2.133              0.03290 *
## age.na         -0.038233   0.003369 -11.349 < 0.0000000000000002 ***
## eduyrs.na       0.052708   0.016123   3.269              0.00108 **
## interest.fac.L  1.344213   0.288990   4.651            0.0000033 ***
## interest.fac.Q  0.100668   0.221055   0.455              0.64882
## interest.fac.C  0.005344   0.132833   0.040              0.96791
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2232.0  on 2211  degrees of freedom
## Residual deviance: 2041.4  on 2206  degrees of freedom
## AIC: 2053.4
##
## Number of Fisher Scoring iterations: 5
```

Estimating the performance of a binary model

```
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16
## 0.31 0.13 0.11 0.14 0.17 0.35 0.11 0.23 0.14 0.13 0.20 0.13 0.20 0.12 0.15 0.14
##   17   18   19   20   21   22   23   24   25   26   27   28   29   30   31   32
## 0.18 0.11 0.13 0.22 0.10 0.31 0.09 0.31 0.12 0.28 0.18 0.14 0.10 0.12 0.23 0.08
##   33   34   35   36   37   38   39   40   41   42   43   44   45   46   47   48
## 0.34 0.19 0.13 0.18 0.34 0.11 0.12 0.19 0.10 0.29 0.33 0.12 0.17 0.09 0.11 0.09
##   49   50   51   52   53   54   55   56   57   58   59   60   61   62   63   64
## 0.31 0.35 0.24 0.34 0.12 0.16 0.19 0.16 0.19 0.23 0.13 0.29 0.35 0.13 0.15 0.12
##   65   66   67   68   69   70   71   72   73   74   75   76   77   78   79   80
## 0.36 0.34 0.17 0.07 0.07 0.12 0.16 0.13 0.34 0.16 0.34 0.14 0.09 0.22 0.13 0.13
##   81   82   83   84   85   86   87   88   89   90   91   92   93   94   95   96
## 0.27 0.19 0.09 0.25 0.12 0.31 0.22 0.16 0.37 0.36 0.14 0.26 0.34 0.09 0.14 0.17
##   97   98   99  100  101  102  103  104  105  106  107  108  109  110  111  112
## 0.37 0.16 0.10 0.26 0.22 0.18 0.16 0.38 0.17 0.13 0.14 0.22 0.14 0.16 0.13 0.35
##  113  114  115  116  117  118  119  120  121  122  123  124  125  126  127  128
## 0.18 0.18 0.21 0.28 0.13 0.19 0.35 0.19 0.20 0.26 0.16 0.26 0.37 0.21 0.09 0.14
##  129  130  131  132  133  134  135  136  137  138  139  140  141  142  143  144
## 0.14 0.35 0.15 0.14 0.38 0.16 0.12 0.18 0.14 0.15 0.19 0.28 0.19 0.37 0.10 0.16
##  145  146  147  148  149  150  151  152  153  154  155  156  157  158  159  160
## 0.09 0.13 0.28 0.09 0.31 0.20 0.28 0.21 0.17 0.26 0.24 0.14 0.12 0.21 0.21 0.18
##  161  162  163  164  165  166  167  168  169  170  171  172  173  174  175  176
## 0.15 0.18 0.33 0.37 0.09 0.32 0.15 0.35 0.37 0.27 0.18 0.32 0.26 0.07 0.14 0.16
##  177  178  179  180  181  182  183  184  185  186  187  188  189  190  191  192
## 0.16 0.23 0.16 0.20 0.23 0.19 0.15 0.31 0.23 0.16 0.16 0.25 0.21 0.22 0.14 0.29
##  193  194  195  196  197  198  199  200  201  202  203  204  205  206  207  208
## 0.20 0.11 0.27 0.16 0.20 0.13 0.20 0.15 0.10 0.08 0.18 0.14 0.29 0.12 0.17 0.10
```

```
##  209  210  211  212  213  214  215  216  217  218  219  220  221  222  223  224
## 0.37 0.13 0.08 0.35 0.14 0.21 0.16 0.22 0.25 0.24 0.14 0.22 0.10 0.14 0.21 0.12
##  225  226  227  228  229  230  231  232  233  234  235  236  237  238  239  240
## 0.33 0.28 0.08 0.16 0.10 0.33 0.37 0.18 0.26 0.20 0.12 0.24 0.34 0.37 0.19 0.32
##  241  242  243  244  245  246  247  248  249  250  251  252  253  254  255  256
## 0.13 0.26 0.07 0.35 0.12 0.20 0.18 0.15 0.11 0.09 0.22 0.34 0.26 0.26 0.10 0.15
##  257  258  259  260  261  262  263  264  265  266  267  268  269  270  271  272
## 0.34 0.11 0.25 0.25 0.17 0.13 0.25 0.18 0.25 0.27 0.31 0.20 0.29 0.27 0.09 0.11
##  273  274  275  276  277  278  279  280  281  282  283  284  285  286  287  288
## 0.22 0.25 0.31 0.14 0.18 0.32 0.11 0.09 0.19 0.32 0.17 0.31 0.14 0.34 0.09 0.21
##  289  290  291  292  293  294  295  296  297  298  299  300  301  302  303  304
## 0.16 0.29 0.10 0.27 0.15 0.13 0.14 0.25 0.08 0.22 0.32 0.15 0.21 0.18 0.21 0.37
##  305  306  307  308  309  310  311  312  313  314  315  316  317  318  319  320
## 0.21 0.12 0.29 0.26 0.15 0.24 0.18 0.20 0.17 0.11 0.15 0.12 0.10 0.23 0.16 0.28
##  321  322  323  324  325  326  327  328  329  330  331  332  333  334  335  336
## 0.18 0.15 0.16 0.09 0.17 0.14 0.17 0.16 0.21 0.37 0.09 0.25 0.09 0.23 0.14 0.20
##  337  338  339  340  341  342  343  344  345  346  347  348  349  350  351  352
## 0.09 0.25 0.11 0.28 0.11 0.14 0.29 0.16 0.12 0.14 0.19 0.27 0.17 0.23 0.19 0.18
##  353  354  355  356  357  358  359  360  361  362  363  364  365  366  367  368
## 0.14 0.12 0.34 0.12 0.37 0.15 0.09 0.14 0.17 0.14 0.32 0.08 0.20 0.18 0.36 0.30
##  369  370  371  372  373  374  375  376  377  378  379  380  381  382  383  384
## 0.28 0.19 0.17 0.27 0.25 0.15 0.23 0.19 0.37 0.10 0.14 0.13 0.35 0.11 0.22 0.33
##  385  386  387  388  389  390  391  392  393  394  395  396  397  398  399  400
## 0.31 0.25 0.18 0.09 0.18 0.34 0.16 0.28 0.29 0.15 0.17 0.20 0.13 0.34 0.12 0.15
##  401  402  403  404  405  406  407  408  409  410  411  412  413  414  415  416
## 0.27 0.11 0.23 0.22 0.14 0.19 0.14 0.18 0.10 0.32 0.13 0.15 0.10 0.26 0.08 0.13
##  417  418  419  420  421  422  423  424  425  426  427  428  429  430  431  432
## 0.08 0.33 0.21 0.11 0.23 0.16 0.35 0.27 0.20 0.12 0.14 0.20 0.31 0.13 0.12 0.31
##  433  434  435  436  437  438  439  440  441  442  443  444  445  446  447  448
## 0.37 0.17 0.09 0.14 0.18 0.18 0.14 0.12 0.19 0.10 0.34 0.10 0.15 0.16 0.10 0.21
##  449  450  451  452  453  454  455  456  457  458  459  460  461  462  463  464
## 0.27 0.20 0.17 0.17 0.34 0.27 0.28 0.18 0.11 0.09 0.13 0.11 0.14 0.19 0.29 0.33
##  465  466  467  468  469  470  471  472  473  474  475  476  477  478  479  480
## 0.23 0.15 0.10 0.16 0.09 0.16 0.16 0.24 0.18 0.15 0.18 0.23 0.13 0.12 0.15 0.23
##  481  482  483  484  485  486  487  488  489  490  491  492  493  494  495  496
## 0.15 0.16 0.14 0.32 0.38 0.29 0.21 0.31 0.18 0.15 0.29 0.16 0.26 0.21 0.29 0.22
##  497  498  499  500  501  502  503  504  505  506  507  508  509  510  511  512
## 0.25 0.12 0.20 0.07 0.16 0.34 0.10 0.23 0.23 0.18 0.28 0.09 0.13 0.10 0.22 0.10
##  513  514  515  516  517  518  519  520  521  522  523  524  525  526  527  528
## 0.11 0.12 0.26 0.09 0.16 0.36 0.21 0.11 0.35 0.12 0.31 0.12 0.18 0.19 0.15 0.09
##  529  530  531  532  533  534  535  536  537  538  539  540  541  542  543  544
## 0.33 0.29 0.10 0.24 0.31 0.38 0.18 0.25 0.34 0.19 0.35 0.24 0.09 0.19 0.10 0.31
##  545  546  547  548  549  550  551  552  553  554  555  556  557  558  559  560
## 0.35 0.07 0.19 0.23 0.11 0.18 0.23 0.29 0.35 0.11 0.36 0.18 0.27 0.34 0.17 0.33
##  561  562  563  564  565  566  567  568  569  570  571  572  573  574  575  576
## 0.16 0.37 0.23 0.18 0.12 0.14 0.11 0.15 0.10 0.20 0.12 0.30 0.20 0.34 0.37 0.10
##  577  578  579  580  581  582  583  584  585  586  587  588  589  590  591  592
## 0.22 0.18 0.15 0.30 0.16 0.16 0.18 0.17 0.11 0.36 0.26 0.15 0.34 0.08 0.11 0.29
##  593  594  595  596  597  598  599  600  601  602  603  604  605  606  607  608
## 0.10 0.12 0.27 0.13 0.11 0.13 0.20 0.13 0.21 0.10 0.25 0.34 0.14 0.17 0.10 0.18
##  609  610  611  612  613  614  615  616  617  618  619  620  621  622  623  624
## 0.31 0.13 0.30 0.07 0.23 0.14 0.15 0.23 0.10 0.15 0.35 0.13 0.07 0.18 0.12 0.16
##  625  626  627  628  629  630  631  632  633  634  635  636  637  638  639  640
## 0.17 0.29 0.35 0.15 0.18 0.18 0.29 0.25 0.23 0.18 0.15 0.31 0.30 0.34 0.16 0.21
```

```
##  641  642  643  644  645  646  647  648  649  650  651  652  653  654  655  656
## 0.09 0.14 0.18 0.20 0.13 0.12 0.28 0.11 0.18 0.34 0.35 0.20 0.14 0.37 0.34 0.19
##  657  658  659  660  661  662  663  664  665  666  667  668  669  670  671  672
## 0.13 0.13 0.17 0.37 0.10 0.14 0.25 0.14 0.16 0.26 0.28 0.23 0.37 0.36 0.26 0.20
##  673  674  675  676  677  678  679  680  681  682  683  684  685  686  687  688
## 0.20 0.23 0.29 0.17 0.11 0.34 0.15 0.12 0.30 0.14 0.13 0.10 0.14 0.30 0.32 0.32
##  689  690  691  692  693  694  695  696  697  698  699  700  701  702  703  704
## 0.37 0.25 0.14 0.24 0.28 0.16 0.28 0.18 0.20 0.11 0.34 0.20 0.14 0.19 0.34 0.28
##  705  706  707  708  709  710  711  712  713  714  715  716  717  718  719  720
## 0.14 0.13 0.20 0.15 0.25 0.22 0.18 0.22 0.26 0.16 0.14 0.17 0.27 0.17 0.30 0.10
##  721  722  723  724  725  726  727  728  729  730  731  732  733  734  735  736
## 0.10 0.13 0.37 0.15 0.31 0.12 0.13 0.23 0.12 0.33 0.30 0.18 0.16 0.16 0.22 0.13
##  737  738  739  740  741  742  743  744  745  746  747  748  749  750  751  752
## 0.23 0.09 0.35 0.35 0.12 0.24 0.34 0.11 0.15 0.14 0.21 0.17 0.19 0.35 0.18 0.28
##  753  754  755  756  757  758  759  760  761  762  763  764  765  766  767  768
## 0.09 0.21 0.20 0.35 0.34 0.24 0.31 0.15 0.10 0.14 0.11 0.12 0.18 0.12 0.35 0.26
##  769  770  771  772  773  774  775  776  777  778  779  780  781  782  783  784
## 0.32 0.25 0.32 0.21 0.28 0.13 0.22 0.18 0.37 0.09 0.30 0.14 0.10 0.10 0.34 0.17
##  785  786  787  788  789  790  791  792  793  794  795  796  797  798  799  800
## 0.21 0.27 0.08 0.38 0.31 0.23 0.14 0.26 0.12 0.27 0.24 0.09 0.24 0.19 0.18 0.16
##  801  802  803  804  805  806  807  808  809  810  811  812  813  814  815  816
## 0.12 0.15 0.12 0.23 0.20 0.13 0.23 0.18 0.20 0.18 0.32 0.36 0.12 0.18 0.29 0.18
##  817  818  819  820  821  822  823  824  825  826  827  828  829  830  831  832
## 0.14 0.28 0.28 0.28 0.18 0.11 0.16 0.16 0.26 0.14 0.35 0.21 0.17 0.35 0.16 0.23
##  833  834  835  836  837  838  839  840  841  842  843  844  845  846  847  848
## 0.21 0.25 0.21 0.21 0.23 0.32 0.08 0.29 0.11 0.30 0.22 0.32 0.31 0.20 0.19 0.11
##  849  850  851  852  853  854  855  856  857  858  859  860  861  862  863  864
## 0.34 0.17 0.33 0.36 0.18 0.18 0.18 0.35 0.27 0.15 0.12 0.38 0.22 0.17 0.22 0.09
##  865  866  867  868  869  870  871  872  873  874  875  876  877  878  879  880
## 0.09 0.37 0.17 0.27 0.14 0.10 0.32 0.37 0.18 0.16 0.20 0.23 0.12 0.11 0.20 0.08
##  881  882  883  884  885  886  887  888  889  890  891  892  893  894  895  896
## 0.08 0.13 0.09 0.22 0.14 0.35 0.11 0.23 0.10 0.19 0.18 0.22 0.12 0.31 0.29 0.33
##  897  898  899  900  901  902  903  904  905  906  907  908  909  910  911  912
## 0.36 0.26 0.16 0.32 0.29 0.20 0.32 0.14 0.30 0.12 0.31 0.12 0.11 0.24 0.16 0.12
##  913  914  915  916  917  918  919  920  921  922  923  924  925  926  927  928
## 0.16 0.13 0.18 0.19 0.31 0.35 0.23 0.14 0.12 0.28 0.28 0.11 0.21 0.08 0.21 0.25
##  929  930  931  932  933  934  935  936  937  938  939  940  941  942  943  944
## 0.13 0.18 0.19 0.19 0.09 0.12 0.33 0.25 0.19 0.14 0.09 0.28 0.30 0.14 0.21 0.07
##  945  946  947  948  949  950  951  952  953  954  955  956  957  958  959  960
## 0.18 0.37 0.37 0.22 0.35 0.15 0.14 0.15 0.11 0.15 0.15 0.15 0.18 0.14 0.27 0.16
##  961  962  963  964  965  966  967  968  969  970  971  972  973  974  975  976
## 0.10 0.14 0.14 0.09 0.14 0.17 0.15 0.13 0.19 0.13 0.16 0.26 0.25 0.26 0.35 0.15
##  977  978  979  980  981  982  983  984  985  986  987  988  989  990  991  992
## 0.17 0.18 0.35 0.14 0.20 0.11 0.15 0.08 0.26 0.11 0.26 0.15 0.33 0.25 0.37 0.14
##  993  994  995  996  997  998  999 1000 1001 1002 1003 1004 1005 1006 1007 1008
## 0.36 0.30 0.12 0.16 0.12 0.09 0.10 0.35 0.20 0.35 0.13 0.31 0.20 0.23 0.23 0.26
## 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024
## 0.13 0.23 0.15 0.20 0.37 0.14 0.13 0.14 0.18 0.21 0.12 0.14 0.16 0.25 0.31 0.18
## 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040
## 0.13 0.10 0.37 0.26 0.22 0.15 0.23 0.14 0.17 0.33 0.13 0.29 0.10 0.18 0.33 0.11
## 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056
## 0.28 0.32 0.24 0.12 0.15 0.09 0.13 0.37 0.13 0.30 0.20 0.16 0.24 0.15 0.17 0.07
## 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072
## 0.28 0.13 0.22 0.14 0.26 0.26 0.18 0.35 0.26 0.10 0.20 0.13 0.07 0.32 0.32 0.17
```

```
## 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088
## 0.34 0.28 0.37 0.18 0.19 0.25 0.14 0.19 0.12 0.16 0.30 0.26 0.10 0.20 0.18 0.10
## 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104
## 0.34 0.18 0.13 0.12 0.37 0.22 0.26 0.34 0.16 0.13 0.18 0.15 0.32 0.13 0.18 0.10
## 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120
## 0.13 0.37 0.08 0.15 0.14 0.32 0.35 0.08 0.27 0.23 0.16 0.20 0.20 0.10 0.19 0.31
## 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136
## 0.33 0.34 0.16 0.22 0.10 0.22 0.14 0.29 0.22 0.15 0.36 0.35 0.22 0.20 0.18 0.18
## 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152
## 0.17 0.14 0.31 0.13 0.20 0.19 0.23 0.28 0.18 0.11 0.26 0.16 0.33 0.09 0.15 0.09
## 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168
## 0.15 0.14 0.27 0.33 0.15 0.14 0.11 0.14 0.25 0.10 0.11 0.18 0.14 0.14 0.23 0.35
## 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184
## 0.18 0.18 0.10 0.29 0.31 0.37 0.11 0.26 0.37 0.18 0.13 0.26 0.17 0.17 0.38 0.10
## 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
## 0.18 0.16 0.22 0.34 0.07 0.10 0.08 0.37 0.12 0.16 0.21 0.17 0.09 0.37 0.37 0.33
## 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216
## 0.35 0.16 0.18 0.12 0.29 0.23 0.36 0.11 0.36 0.10 0.14 0.14 0.29 0.12 0.32 0.31
## 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232
## 0.13 0.35 0.18 0.21 0.15 0.29 0.31 0.13 0.15 0.11 0.11 0.19 0.16 0.14 0.18 0.34
## 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248
## 0.17 0.11 0.25 0.29 0.15 0.20 0.18 0.22 0.25 0.18 0.37 0.33 0.10 0.15 0.29 0.30
## 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264
## 0.31 0.09 0.10 0.23 0.20 0.15 0.13 0.09 0.13 0.10 0.23 0.10 0.37 0.30 0.23 0.12
## 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280
## 0.09 0.29 0.28 0.10 0.13 0.26 0.11 0.11 0.29 0.13 0.18 0.15 0.18 0.35 0.22 0.20
## 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296
## 0.29 0.38 0.15 0.14 0.11 0.15 0.19 0.15 0.36 0.26 0.12 0.19 0.23 0.30 0.10 0.09
## 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312
## 0.27 0.29 0.25 0.31 0.37 0.14 0.12 0.09 0.13 0.26 0.12 0.18 0.22 0.37 0.19 0.38
## 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328
## 0.19 0.11 0.34 0.21 0.28 0.12 0.28 0.18 0.29 0.30 0.34 0.29 0.29 0.12 0.16 0.27
## 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344
## 0.12 0.17 0.13 0.15 0.16 0.19 0.18 0.27 0.29 0.28 0.13 0.13 0.29 0.15 0.21 0.11
## 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360
## 0.25 0.15 0.08 0.24 0.21 0.13 0.27 0.12 0.23 0.18 0.09 0.19 0.26 0.09 0.29 0.17
## 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376
## 0.15 0.37 0.34 0.23 0.23 0.21 0.17 0.09 0.16 0.26 0.27 0.38 0.28 0.19 0.23 0.18
## 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392
## 0.15 0.13 0.21 0.20 0.23 0.15 0.29 0.31 0.12 0.21 0.10 0.25 0.18 0.14 0.15 0.22
## 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408
## 0.19 0.09 0.32 0.10 0.37 0.17 0.10 0.37 0.30 0.36 0.27 0.18 0.17 0.15 0.13 0.13
## 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424
## 0.29 0.17 0.32 0.11 0.25 0.12 0.15 0.19 0.20 0.16 0.12 0.22 0.21 0.12 0.15 0.13
## 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440
## 0.11 0.14 0.34 0.13 0.14 0.08 0.10 0.28 0.13 0.24 0.13 0.21 0.17 0.15 0.33 0.36
## 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456
## 0.37 0.20 0.16 0.26 0.13 0.29 0.29 0.17 0.21 0.14 0.11 0.31 0.11 0.37 0.26 0.16
## 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472
## 0.34 0.23 0.10 0.31 0.17 0.37 0.29 0.13 0.15 0.19 0.12 0.13 0.22 0.10 0.18 0.16
## 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488
## 0.11 0.21 0.10 0.25 0.20 0.17 0.14 0.15 0.37 0.11 0.16 0.34 0.25 0.14 0.34 0.15
## 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504
## 0.22 0.34 0.15 0.22 0.16 0.13 0.28 0.13 0.36 0.15 0.23 0.35 0.28 0.17 0.32 0.12
```

```
## 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520
## 0.14 0.33 0.23 0.26 0.27 0.25 0.09 0.33 0.33 0.23 0.23 0.12 0.35 0.15 0.24 0.20
## 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536
## 0.35 0.09 0.19 0.11 0.28 0.13 0.19 0.29 0.23 0.35 0.18 0.37 0.18 0.16 0.19 0.23
## 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552
## 0.23 0.14 0.18 0.12 0.13 0.10 0.29 0.23 0.14 0.31 0.25 0.17 0.14 0.10 0.18 0.13
## 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568
## 0.16 0.31 0.28 0.30 0.16 0.19 0.10 0.16 0.14 0.35 0.12 0.27 0.20 0.18 0.27 0.27
## 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584
## 0.07 0.16 0.10 0.35 0.25 0.14 0.35 0.34 0.21 0.37 0.14 0.32 0.18 0.11 0.29 0.10
## 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600
## 0.32 0.33 0.32 0.07 0.32 0.08 0.27 0.21 0.17 0.27 0.13 0.32 0.16 0.18 0.29 0.14
## 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616
## 0.08 0.16 0.21 0.29 0.28 0.27 0.27 0.22 0.28 0.10 0.21 0.11 0.13 0.18 0.15 0.13
## 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632
## 0.14 0.22 0.32 0.11 0.18 0.17 0.10 0.29 0.28 0.11 0.31 0.20 0.18 0.14 0.12 0.14
## 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648
## 0.31 0.11 0.17 0.11 0.23 0.10 0.17 0.35 0.16 0.24 0.13 0.24 0.29 0.23 0.14 0.30
## 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664
## 0.26 0.11 0.24 0.09 0.16 0.30 0.14 0.22 0.28 0.13 0.14 0.18 0.08 0.31 0.17 0.37
## 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680
## 0.09 0.09 0.22 0.27 0.15 0.09 0.18 0.35 0.15 0.16 0.30 0.10 0.19 0.20 0.32 0.10
## 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696
## 0.16 0.20 0.18 0.19 0.21 0.32 0.25 0.14 0.32 0.15 0.29 0.38 0.11 0.32 0.21 0.17
## 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712
## 0.20 0.13 0.14 0.10 0.20 0.32 0.34 0.26 0.15 0.11 0.18 0.10 0.19 0.30 0.27 0.22
## 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728
## 0.21 0.18 0.15 0.23 0.29 0.21 0.18 0.10 0.17 0.09 0.34 0.10 0.32 0.14 0.14 0.15
## 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744
## 0.20 0.24 0.09 0.20 0.38 0.10 0.13 0.37 0.23 0.14 0.13 0.15 0.14 0.17 0.11 0.34
## 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760
## 0.18 0.15 0.09 0.15 0.10 0.08 0.13 0.16 0.14 0.31 0.29 0.17 0.26 0.25 0.07 0.15
## 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776
## 0.27 0.20 0.37 0.29 0.27 0.09 0.15 0.19 0.11 0.13 0.24 0.10 0.17 0.18 0.18 0.19
## 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792
## 0.31 0.33 0.12 0.11 0.15 0.16 0.23 0.33 0.13 0.31 0.26 0.15 0.19 0.12 0.35 0.19
## 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808
## 0.24 0.14 0.11 0.15 0.14 0.16 0.17 0.17 0.29 0.09 0.35 0.16 0.19 0.20 0.21 0.10
## 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824
## 0.14 0.20 0.13 0.27 0.29 0.12 0.16 0.14 0.08 0.25 0.31 0.36 0.15 0.15 0.21 0.15
## 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840
## 0.09 0.14 0.16 0.22 0.23 0.35 0.18 0.18 0.09 0.25 0.09 0.15 0.28 0.20 0.27 0.18
## 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856
## 0.34 0.21 0.25 0.09 0.26 0.28 0.38 0.28 0.22 0.26 0.09 0.27 0.36 0.27 0.17 0.28
## 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872
## 0.15 0.23 0.16 0.14 0.16 0.17 0.21 0.26 0.18 0.15 0.10 0.20 0.15 0.12 0.36 0.14
## 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888
## 0.24 0.34 0.16 0.11 0.14 0.14 0.17 0.34 0.17 0.14 0.17 0.31 0.31 0.34 0.12 0.11
## 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904
## 0.16 0.33 0.21 0.31 0.29 0.11 0.12 0.13 0.15 0.33 0.10 0.16 0.38 0.38 0.19 0.34
## 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920
## 0.16 0.37 0.26 0.10 0.17 0.21 0.35 0.13 0.21 0.11 0.11 0.10 0.24 0.14 0.28 0.22
## 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936
## 0.28 0.32 0.32 0.18 0.28 0.08 0.17 0.21 0.07 0.32 0.16 0.34 0.15 0.26 0.13 0.09
```

```
## 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952
## 0.32 0.14 0.12 0.23 0.31 0.27 0.21 0.11 0.36 0.36 0.16 0.13 0.14 0.14 0.17 0.08
## 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968
## 0.29 0.10 0.12 0.12 0.15 0.18 0.27 0.11 0.22 0.23 0.13 0.16 0.13 0.35 0.25 0.25
## 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984
## 0.17 0.13 0.16 0.09 0.36 0.12 0.27 0.07 0.16 0.14 0.26 0.10 0.22 0.27 0.36 0.17
## 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000
## 0.23 0.25 0.17 0.08 0.35 0.28 0.11 0.09 0.12 0.31 0.32 0.32 0.23 0.20 0.34 0.26
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
## 0.16 0.09 0.27 0.12 0.14 0.11 0.14 0.23 0.10 0.15 0.34 0.36 0.30 0.14 0.20 0.19
## 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032
## 0.15 0.25 0.17 0.09 0.35 0.26 0.27 0.20 0.15 0.35 0.21 0.16 0.09 0.14 0.13 0.10
## 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048
## 0.12 0.27 0.13 0.35 0.24 0.14 0.21 0.29 0.10 0.14 0.20 0.17 0.36 0.23 0.25 0.14
## 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064
## 0.29 0.26 0.17 0.15 0.25 0.18 0.15 0.33 0.11 0.28 0.14 0.12 0.27 0.11 0.23 0.14
## 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080
## 0.13 0.27 0.32 0.20 0.21 0.19 0.18 0.13 0.37 0.18 0.28 0.16 0.35 0.08 0.34 0.16
## 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096
## 0.19 0.12 0.22 0.36 0.36 0.26 0.18 0.19 0.34 0.10 0.28 0.22 0.10 0.11 0.17 0.15
## 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112
## 0.34 0.12 0.19 0.18 0.12 0.15 0.10 0.36 0.14 0.34 0.34 0.29 0.14 0.28 0.11 0.16
## 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128
## 0.37 0.19 0.16 0.29 0.19 0.37 0.11 0.11 0.14 0.11 0.08 0.12 0.30 0.27 0.34 0.23
## 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144
## 0.18 0.17 0.22 0.27 0.29 0.23 0.26 0.12 0.32 0.10 0.18 0.34 0.19 0.22 0.22 0.14
## 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160
## 0.14 0.27 0.18 0.27 0.18 0.13 0.12 0.27 0.15 0.19 0.14 0.29 0.24 0.32 0.20 0.19
## 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176
## 0.19 0.34 0.17 0.25 0.17 0.23 0.09 0.17 0.17 0.12 0.30 0.28 0.29 0.16 0.34 0.16
## 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192
## 0.27 0.14 0.23 0.27 0.22 0.12 0.35 0.07 0.11 0.11 0.19 0.18 0.09 0.13 0.11 0.13
## 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208
## 0.22 0.22 0.13 0.37 0.22 0.16 0.19 0.18 0.10 0.27 0.19 0.28 0.07 0.11 0.19 0.11
## 2209 2210 2211 2212
## 0.18 0.26 0.22 0.26
```

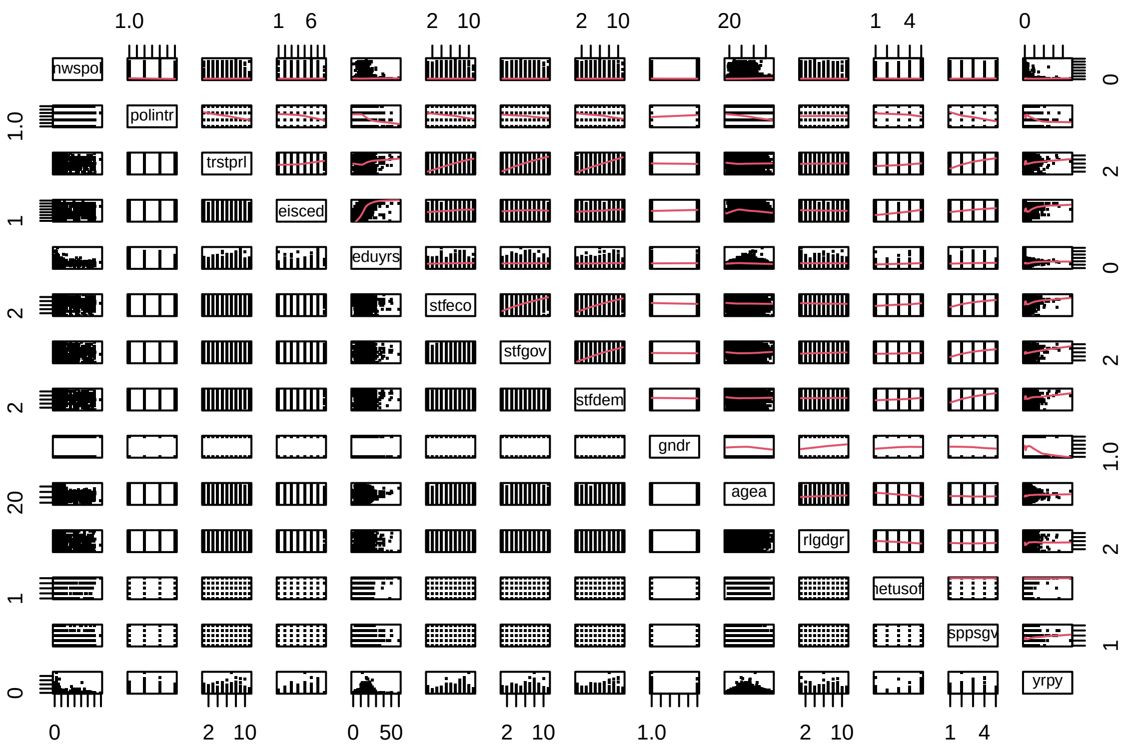
```
## 1 2 3 4 5 6
## 0 1 1 1 1 0
```

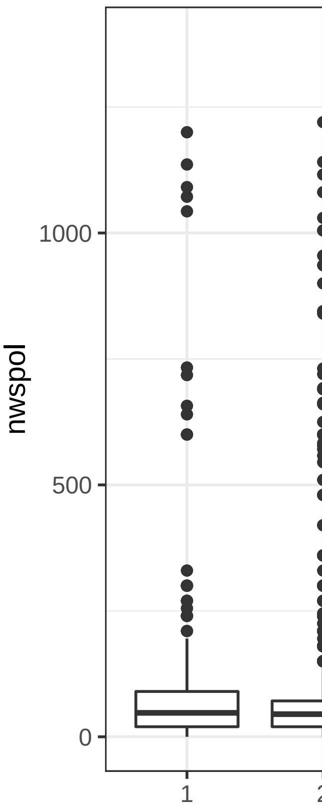
```
##    0    1
##  363 1849
```

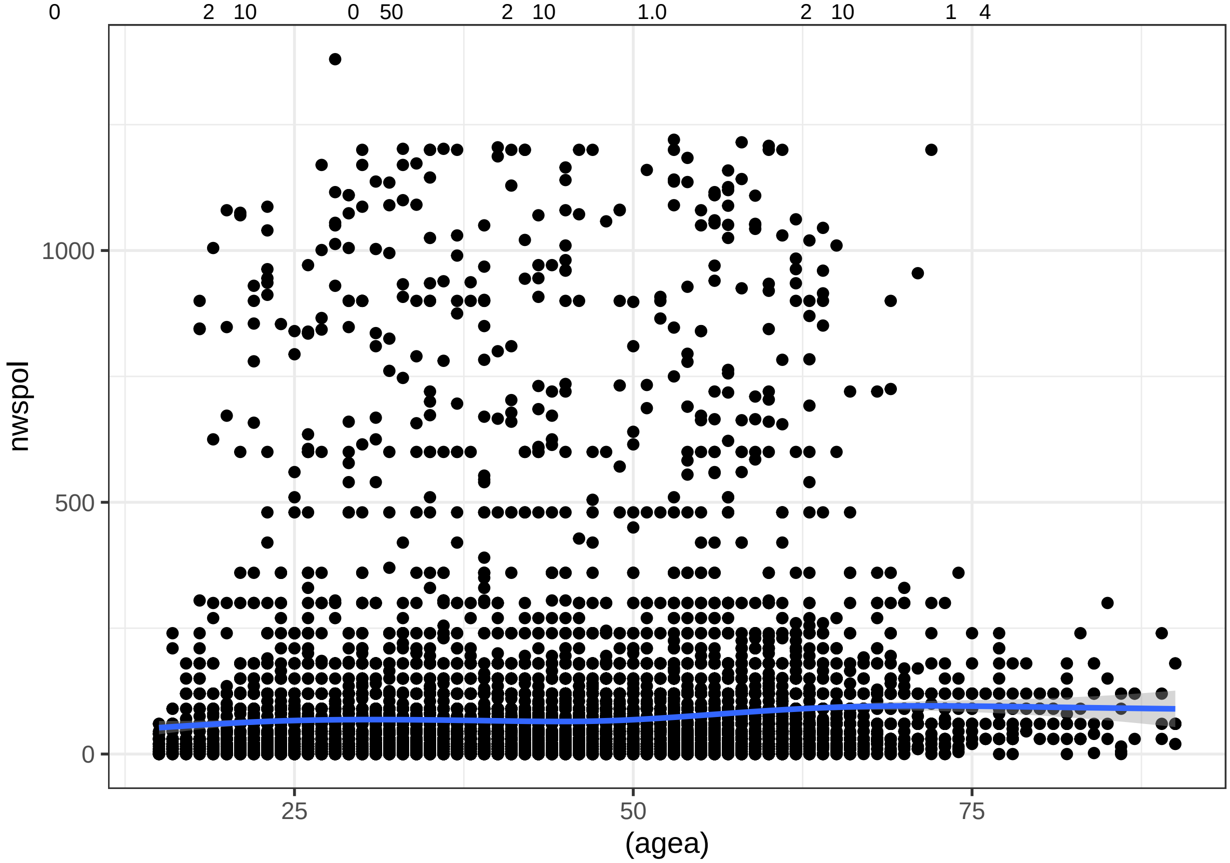
```
##    fitted
## obs    0    1
##   0  255 1508
##   1  108  341
```

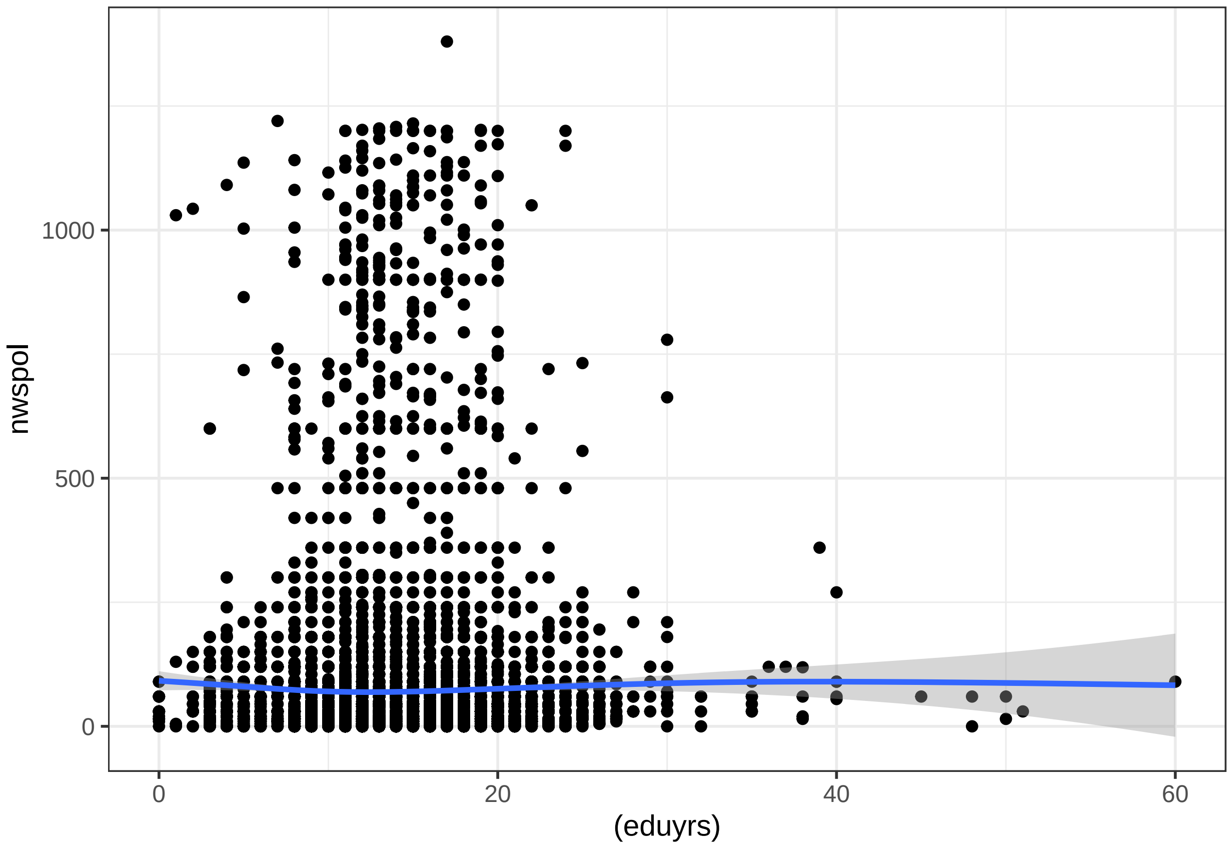
3.4 GAM

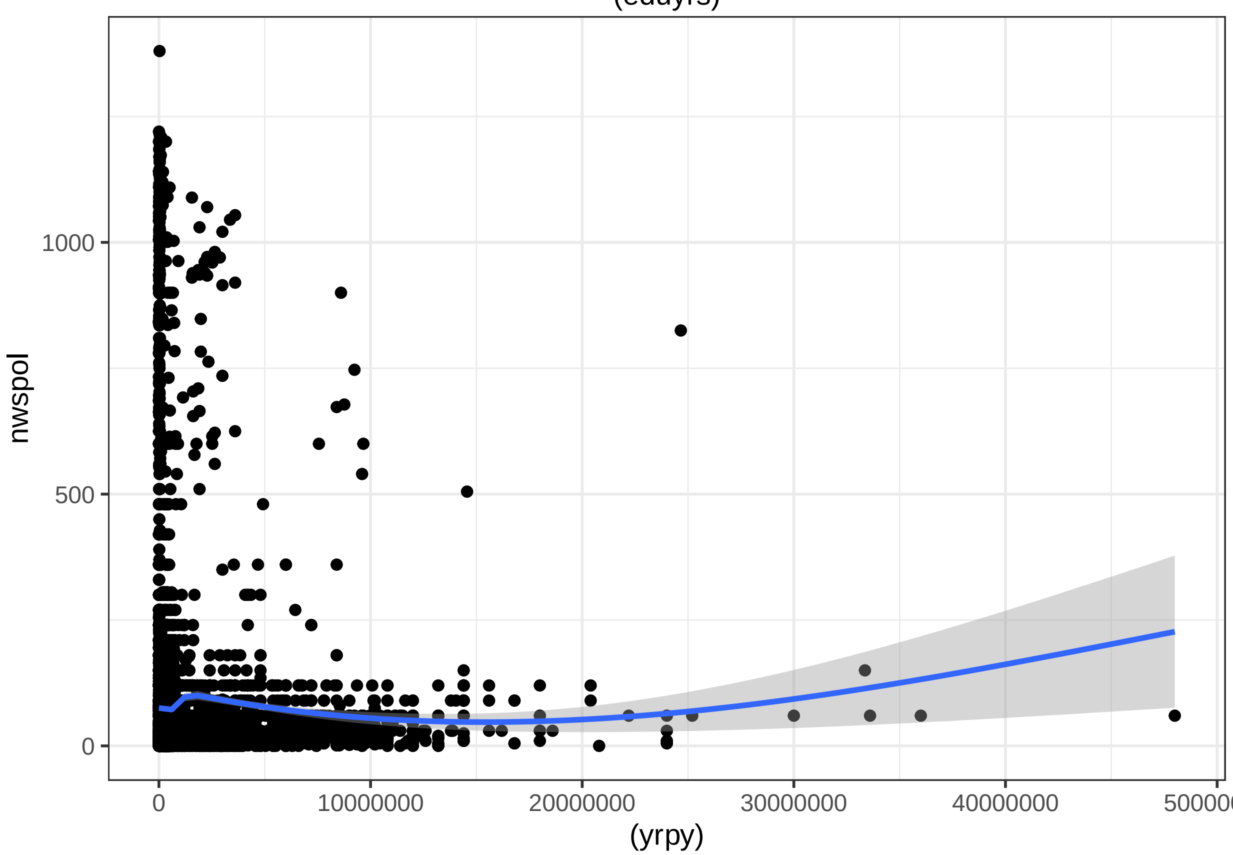
In this chapter a generalized additive model (GAM) is applied to the data set. First the data is visually analyzed and checked. The variables don't show a strong relation ship and the line is mostly horizontal.

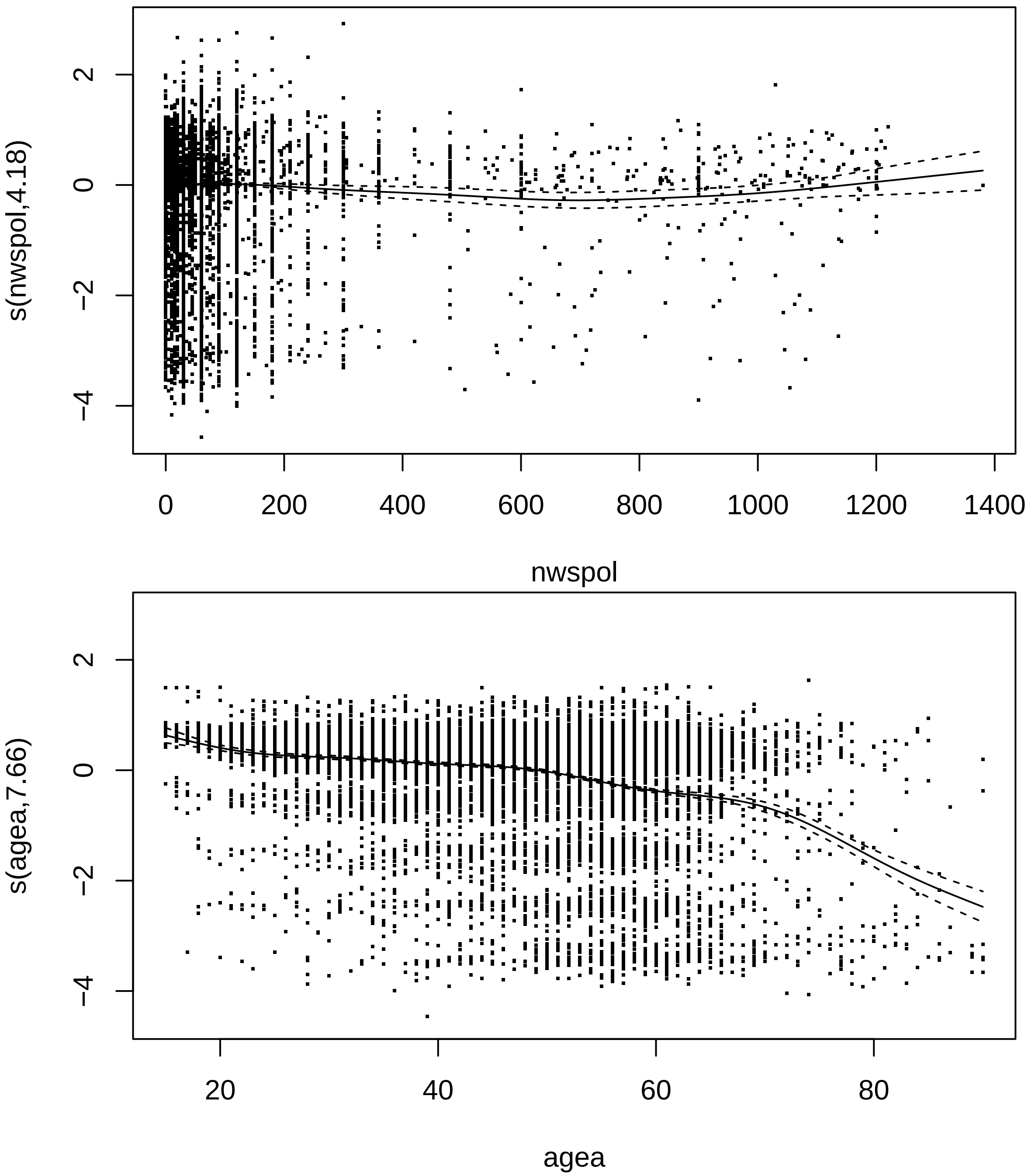

nwspol
polintr
trstprl
eisced
eduyrs
stfeco
stfgov
stfdem
gndr
agea
rlgdgr
etusof
sppsgv
yrpy
1.0
1 6
2 10
2 10
20
1 4
0
0
2
0
2
1.0
2
1
1.0
1
2
2
20
1
0
0
2 10
0 50
2 10
1.0
2 10
1 4

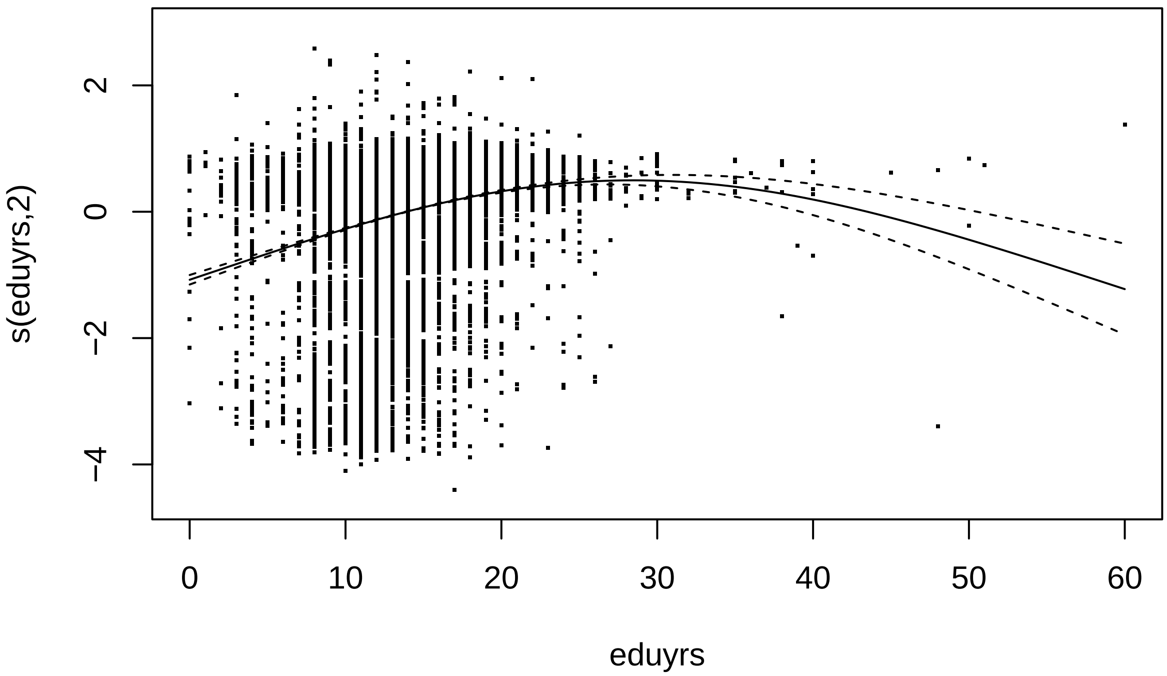

nwspol
1000
500
0
1

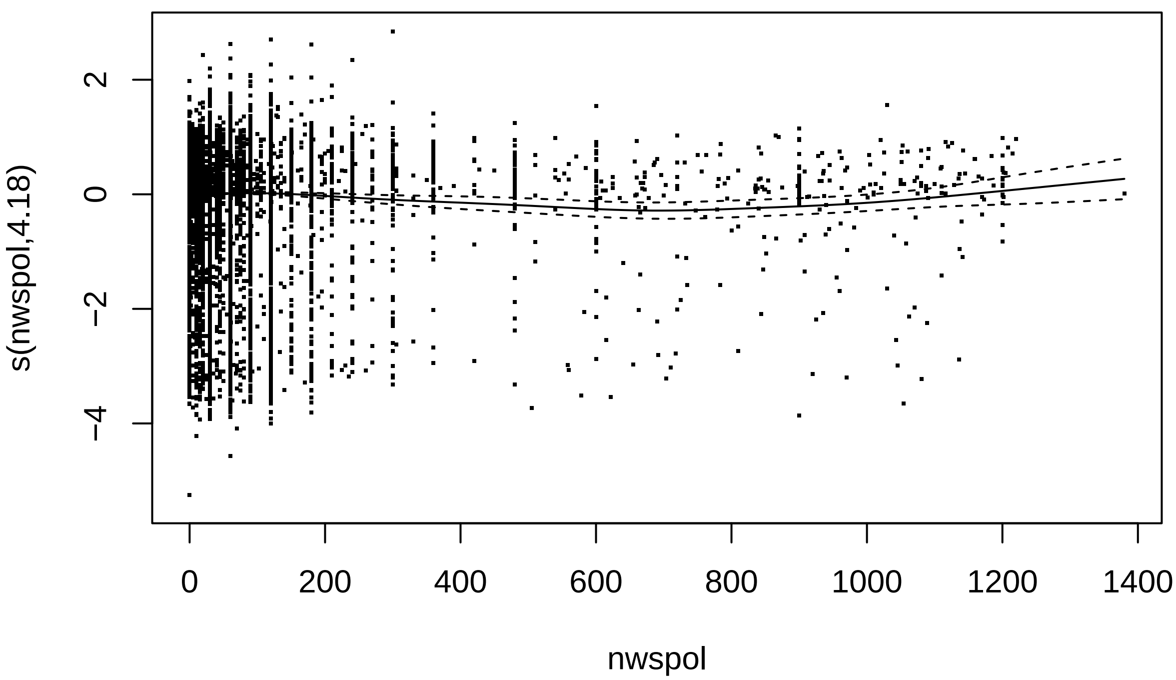

nwspol
1000
500
0
25
50
75
(agea)

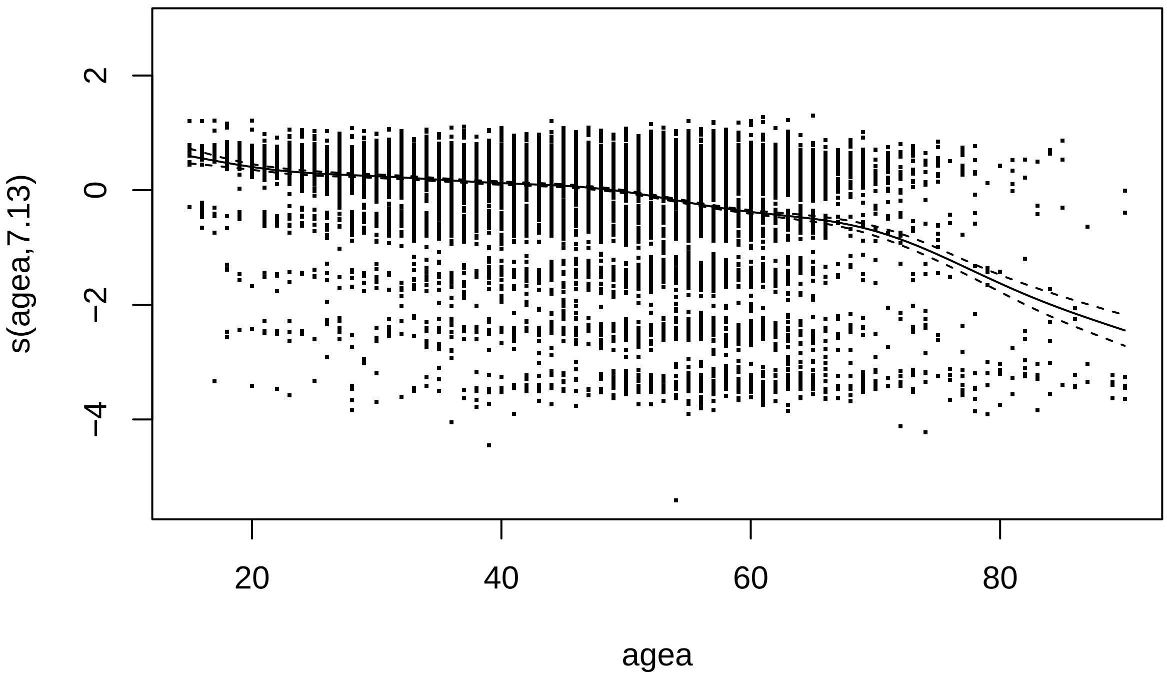

nwspol
1000
500
0
0
20
40
60
(eduyrs)


nwspol
1000
500
0
0
10000000
20000000
30000000
40000000
500000
(yrpy)


s(nwspol,4.18)
2
0
−2
−4
0
200
400
600
800
1000
1200
1400
nwspol
s(agea,7.66)
2
0
−2
−4
20
40
60
80
agea


s(eduyrs,2)
2
0
−2
−4
0
10
20
30
40
50
60
eduyrs


s(nwspol,4.18)
2
0
−2
−4
0
200
400
600
800
1000
1200
1400
nwspol



3.5 Neural Network

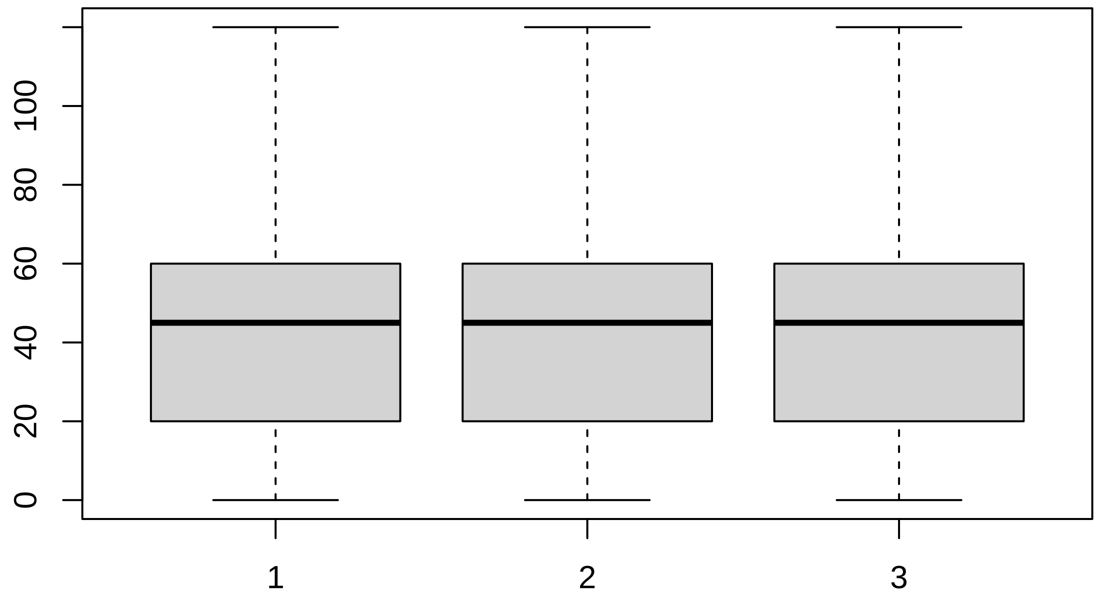
Additional to “classic” linear models and GLM options, there are neural networks for analyzing and predicting data.

3.5.1 Neural Network

Within this chapter two types of neural network model are applied to two different questions.

1. Predicting the values for the continuous variable
2. Predicting the vallues for the categorical variable inidicating the 5 levels of ### 3.5.1 continious variable

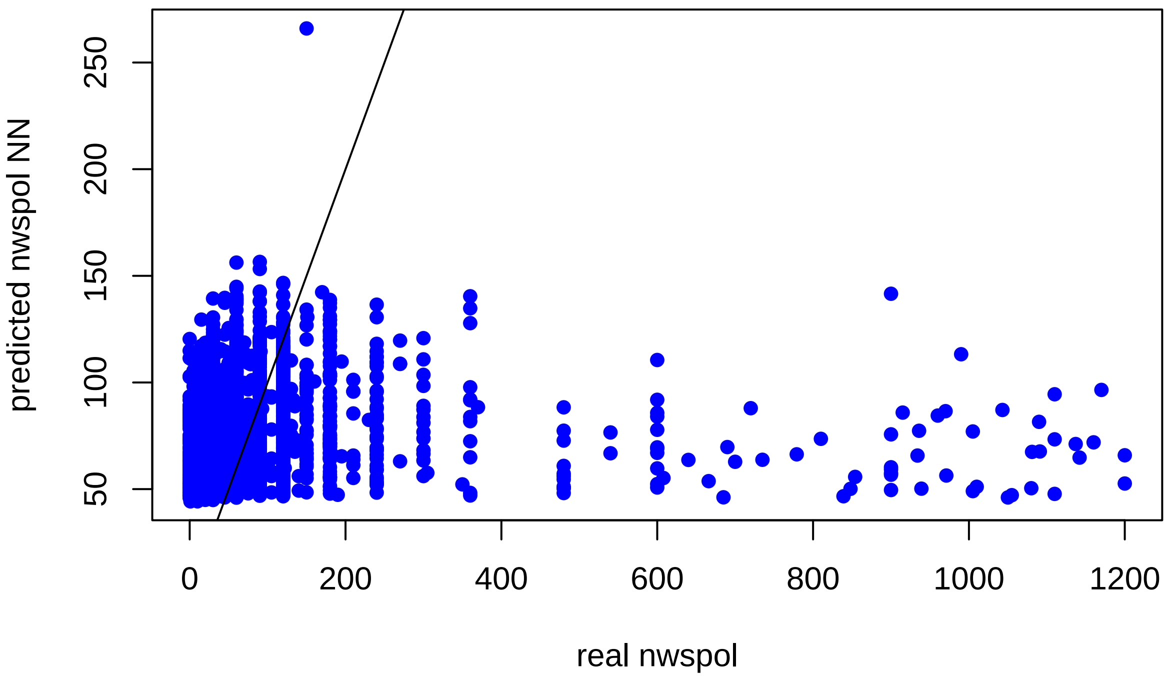
```
## tibble [17,029 x 14] (S3: tbl_df/tbl/data.frame)
##  $ nwspol  : num [1:17029] 60 45 60 120 15 60 30 30 10 10 ...
##  $ polintr : num [1:17029] 4 3 4 2 4 1 2 3 2 2 ...
##  $ trstprl : num [1:17029] 6 0 0 6 4 3 3 7 7 5 ...
##  $ eisced  : num [1:17029] 2 3 3 3 3 4 3 6 2 7 ...
##  $ eduyrs  : num [1:17029] 12 11 12 12 13 21 18 17 9 17 ...
##  $ stfeco  : num [1:17029] 5 6 1 10 9 7 6 7 6 8 ...
##  $ stfgov  : num [1:17029] 6 8 3 10 8 2 7 2 7 6 ...
##  $ stfdem  : num [1:17029] 6 6 3 10 7 3 10 6 8 7 ...
##  $ gndr    : num [1:17029] 2 1 1 1 2 1 1 1 1 1 ...
##  $ agea    : num [1:17029] 40 63 56 48 41 27 49 42 50 35 ...
##  $ rlgdgr  : num [1:17029] 4 1 8 0 3 3 2 0 3 2 ...
##  $ netusoft: num [1:17029] 4 5 1 1 4 5 5 5 5 5 ...
##  $ psppsgva: num [1:17029] 2 2 2 5 1 3 1 1 2 3 ...
##  $ yrpy    : num [1:17029] 31200 30600 18000 31200 37200 18000 20400 17400 45600 70000 ...
```



Prepare for Training

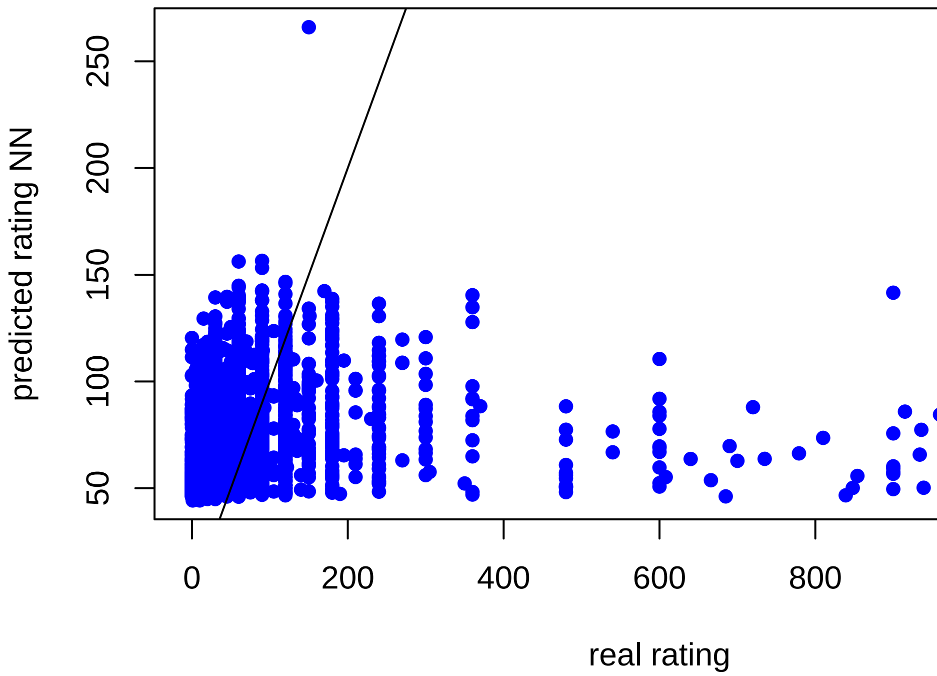
Fit the Network

Predict the test set So what went wrong here? Obviously we also need to use the scaled data to make predictions and then scale back to real nwspol values

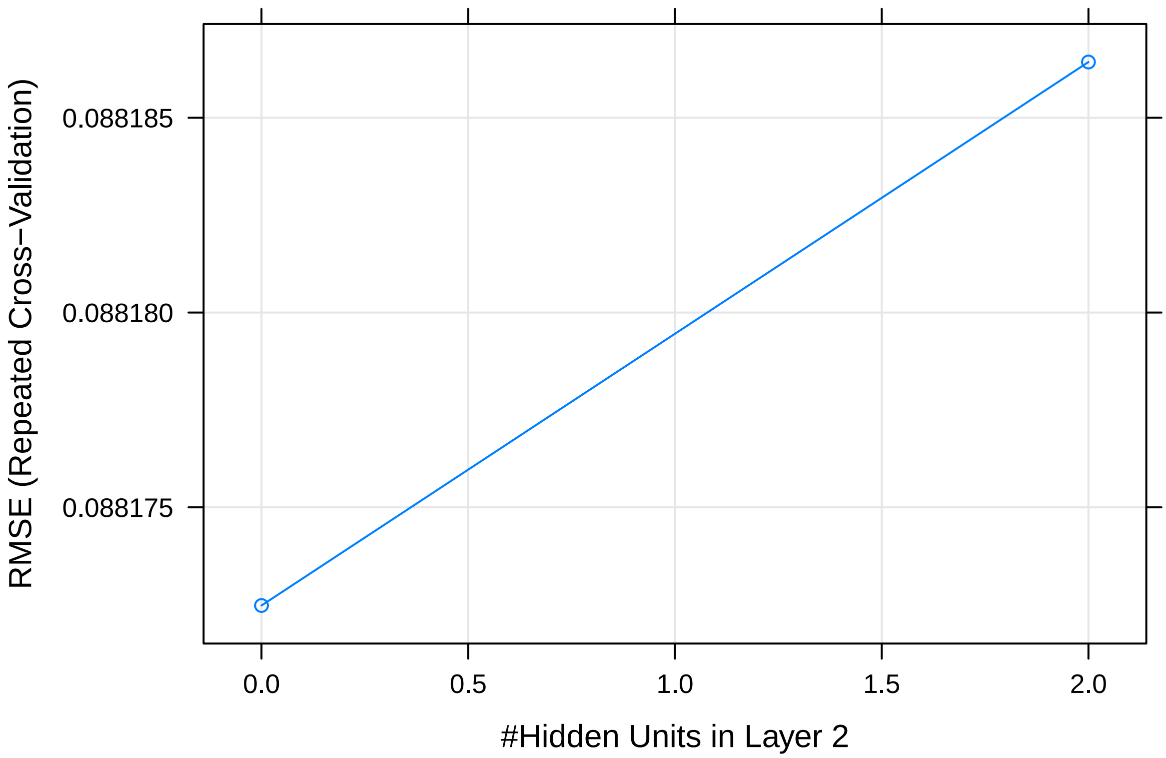


And calculate the RMSE

```
## [1] 123.4398
```



Redo with caret and Cross Validation

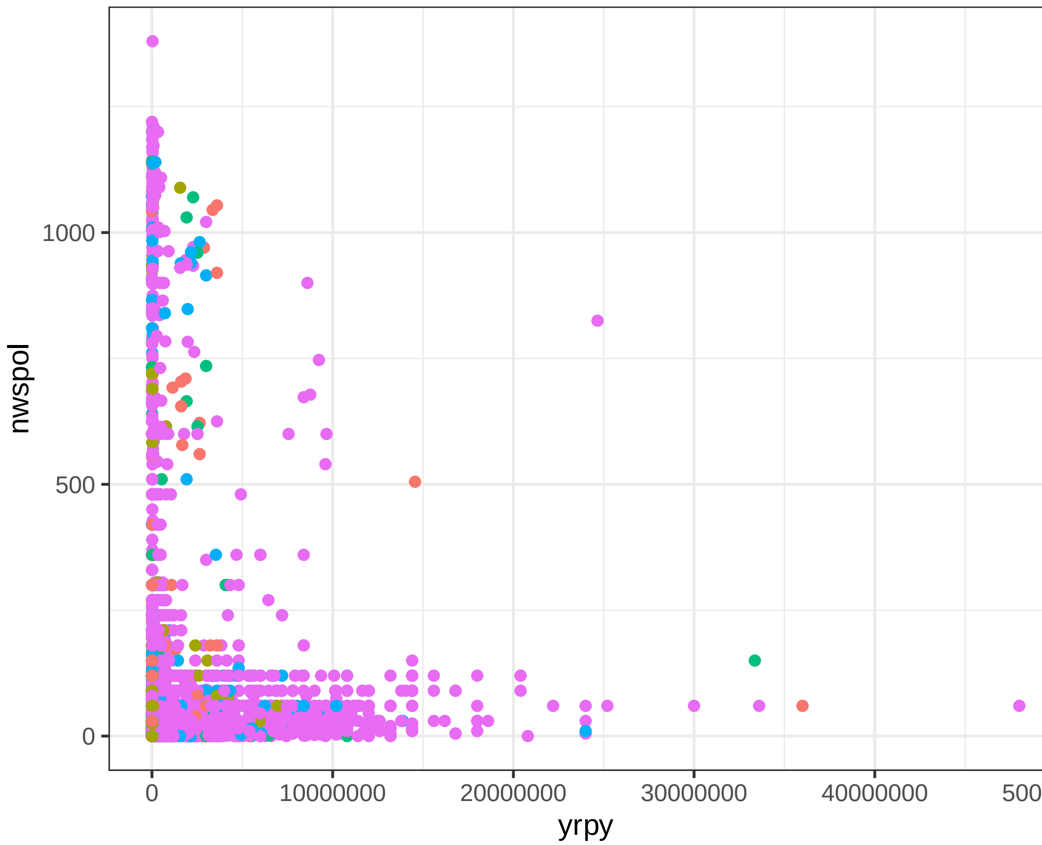


and extract the best model

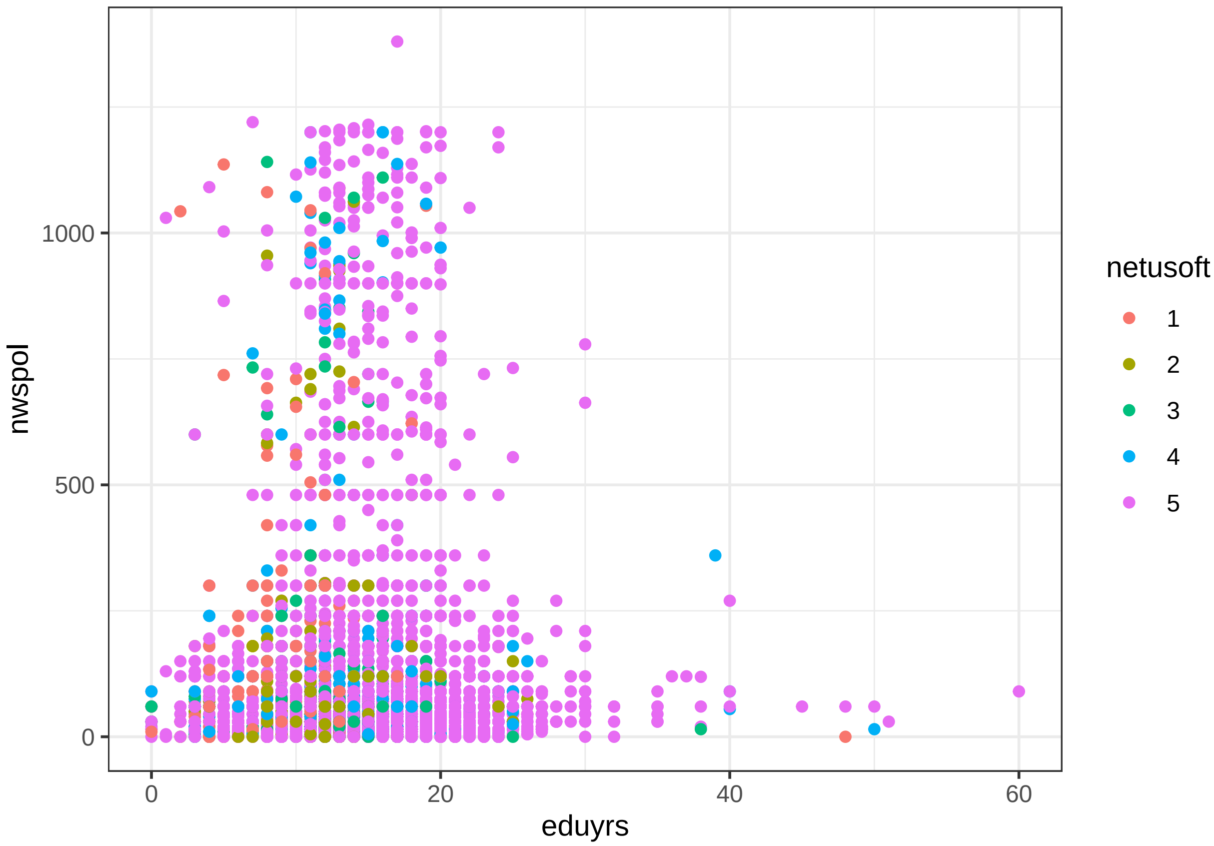
3.5.2 Categorical variable

```
## tibble [17,029 x 14] (S3: tbl_df/tbl/data.frame)
##  $ nwspol  : num [1:17029] 60 45 60 120 15 60 30 30 10 10 ...
##  $ polintr : num [1:17029] 4 3 4 2 4 1 2 3 2 2 ...
##  $ trstprl : num [1:17029] 6 0 0 6 4 3 3 7 7 5 ...
##  $ eisced  : num [1:17029] 2 3 3 3 3 4 3 6 2 7 ...
```

```
##  $ eduyrs  : num [1:17029] 12 11 12 12 13 21 18 17 9 17 ...
##  $ stfeco  : num [1:17029] 5 6 1 10 9 7 6 7 6 8 ...
##  $ stfgov  : num [1:17029] 6 8 3 10 8 2 7 2 7 6 ...
##  $ stfdem  : num [1:17029] 6 6 3 10 7 3 10 6 8 7 ...
##  $ gndr    : num [1:17029] 2 1 1 1 2 1 1 1 1 1 ...
##  $ agea    : num [1:17029] 40 63 56 48 41 27 49 42 50 35 ...
##  $ rlgdgr  : num [1:17029] 4 1 8 0 3 3 2 0 3 2 ...
##  $ netusoft: Factor w/ 5 levels "1","2","3","4",..: 4 5 1 1 4 5 5 5 5 5 ...
##  $ psppsgva: num [1:17029] 2 2 2 5 1 3 1 1 2 3 ...
##  $ yrpy    : num [1:17029] 31200 30600 18000 31200 37200 18000 20400 17400 45600 70000 ...
```



Have a quick look at the Data



Prepare the Data for Training

Why do we use the caret function Look at the distribution of our different prediction classes in the train and test datasets:

```
##          train
## netusoft FALSE  TRUE
##        1    90   515
##        2    91   516
##        3   106   604
##        4   210  1193
##        5  2055 11649
```

Now compare this to the "random" approach:

```
##          train
## netusoft FALSE  TRUE
##        1   104   501
##        2    89   518
##        3   106   604
##        4   197  1206
##        5  2033 11671
```

So using the `createDataPartition` function makes sure that no class is over- or underrepresented relative to the total occurance in the two sets

Create some easy Variables to access Data

Train the Neural Network Call the `neuralnet` function creating a network with two hidden layers containing 4 and 3 neurons (probably way too complex for our problem here).

Plot the resulting network including the weights

Make Predictions Find class (i.e. output neuron) with the highest probability and convert this back into a factor

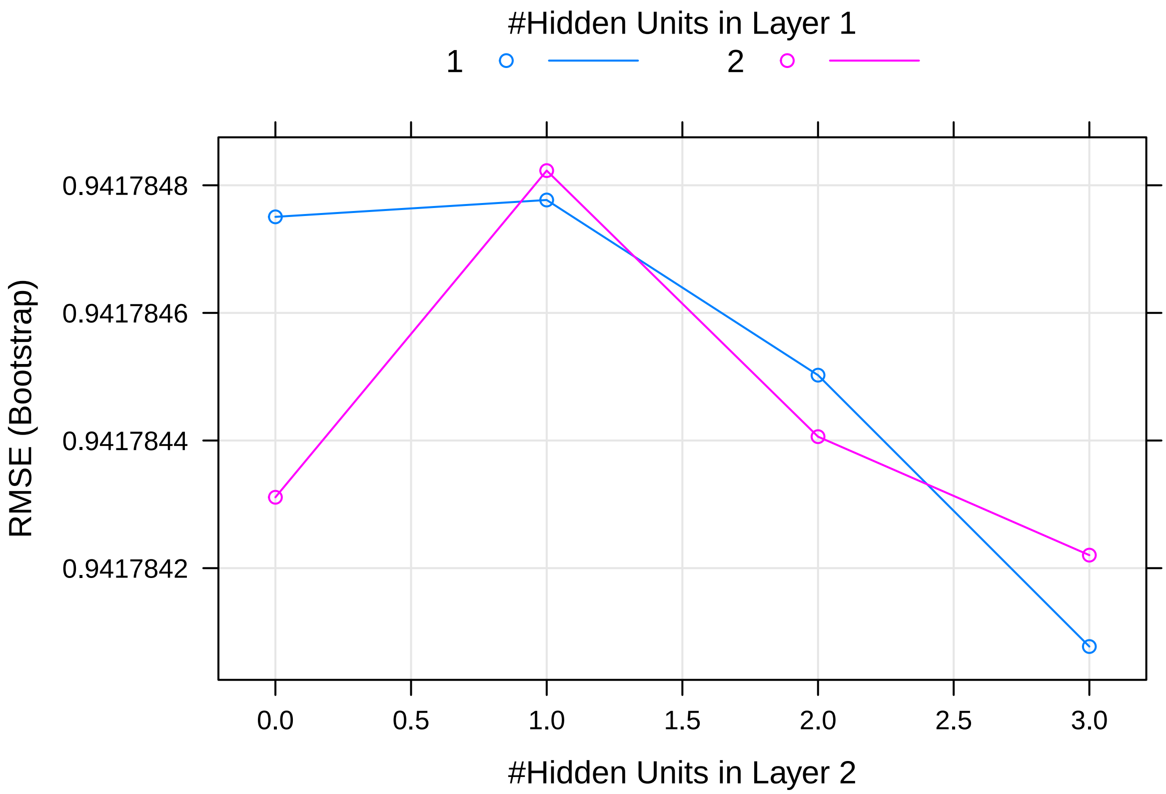
Evaluate the Results

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    1    2    3    4    5
##          1    1    0    0    0   89
##          2    0    0    0    0   91
##          3    0    0    0    0  106
##          4    0    0    0    0  210
##          5    0    0    0    0 2055
##
## Overall Statistics
##
##                Accuracy : 0.8056
##                  95% CI : (0.7897, 0.8208)
##     No Information Rate : 0.9996
##     P-Value [Acc > NIR] : 1
##
##                   Kappa : 0.0036
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                       Class: 1 Class: 2 Class: 3 Class: 4 Class: 5
## Sensitivity          1.0000000       NA       NA       NA 0.805566
## Specificity          0.9651117  0.96434  0.95846  0.91771 1.000000
## Pos Pred Value       0.0111111       NA       NA       NA 1.000000
## Neg Pred Value       1.0000000       NA       NA       NA 0.002012
## Prevalence           0.0003918  0.00000  0.00000  0.00000 0.999608
## Detection Rate       0.0003918  0.00000  0.00000  0.00000 0.805251
## Detection Prevalence 0.0352665  0.03566  0.04154  0.08229 0.805251
## Balanced Accuracy    0.9825559       NA       NA       NA 0.902783
```

Optimize Network Structure First we need to remodel the data due to some limitations of `caret`

```
##   netusoft netusoft2 netusoft3 netusoft4 netusoft5
## 1        1         0         0         1         0
## 2        1         0         0         0         1
## 3        1         0         0         0         0
## 4        1         0         0         0         0
## 5        1         0         0         1         0
## 6        1         0         0         0         1
```

And have a look at the different models



And try out the best model

3.6 Support Vector Machine

Loading the Packages

We are going to use the `e1071` package for this example and also some helper functions from `caret`

```
##   left center  right   NA's 
##    987   1149    125     97
```

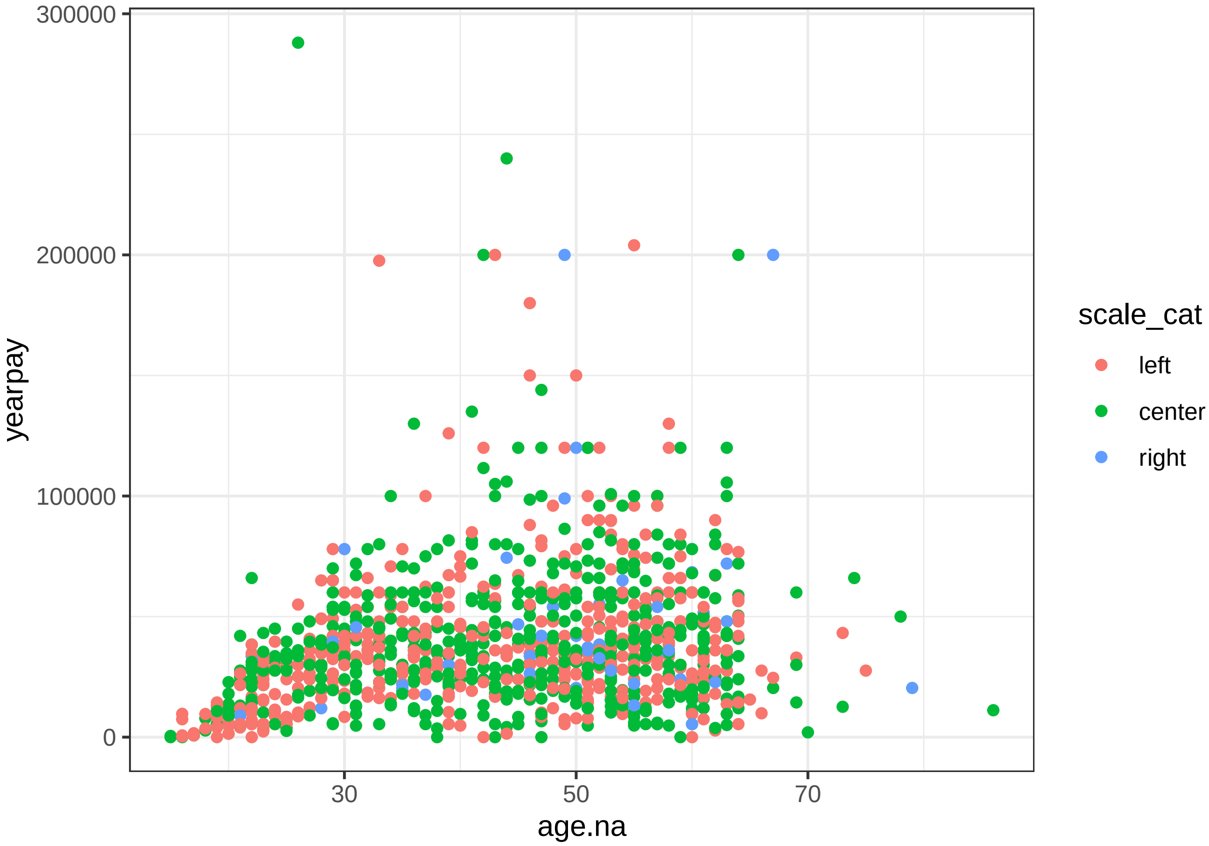
```
##    0    1    2    3    4    5    6    7    8    9   10 NA's 
##   95   64  186  340  302  826  198  125   86   14   25   97
```

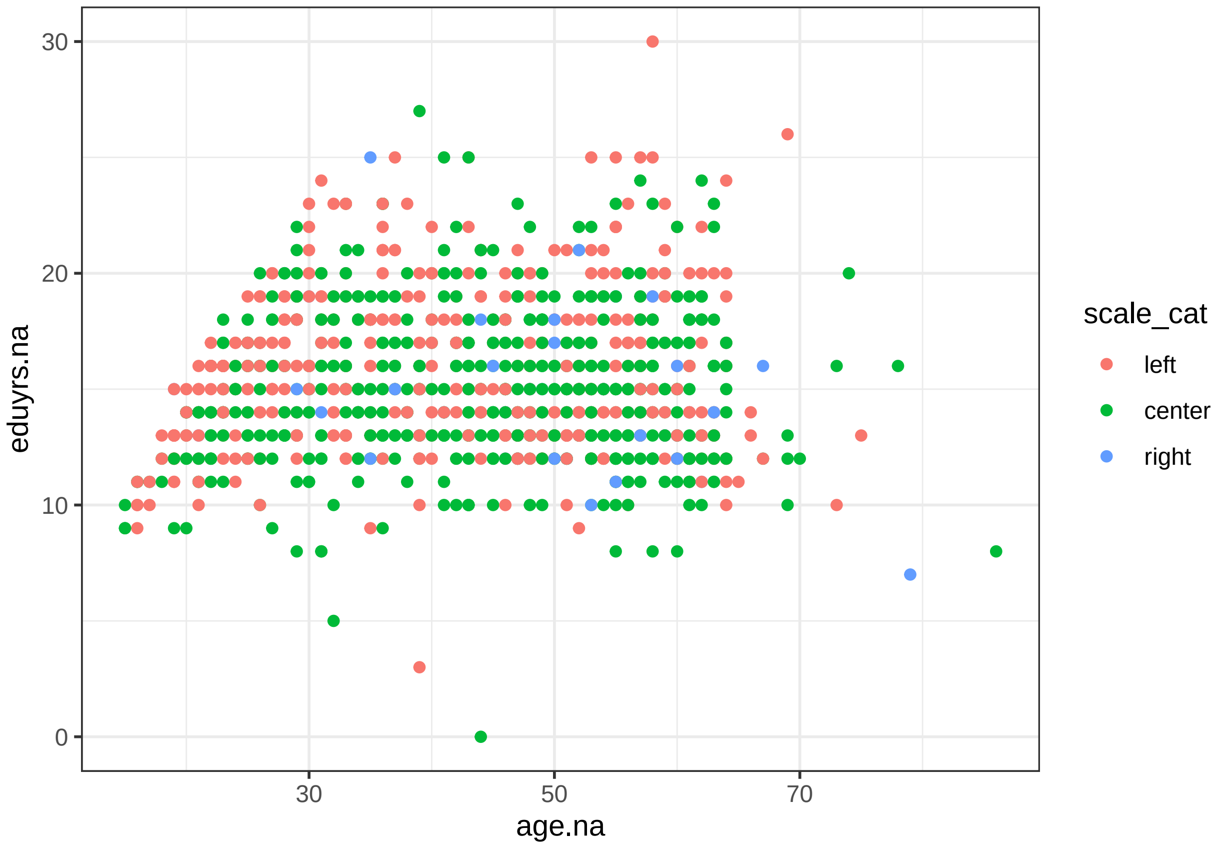
```
## [1] "integer"
```

```
##   left center  right   NA's 
##    987   1149    125     97
```

```
## tibble [1,089 x 7] (S3: tbl_df/tbl/data.frame)
##  $ age.na      : num [1:1089] 26 54 41 47 48 60 44 51 26 79 ...
##  $ interest.fac: Ord.factor w/ 4 levels "none at all"<..: 2 2 2 3 4 4 3 3 2 3 ...
##  $ scale_cat   : Factor w/ 3 levels "left","center",..: 2 1 1 1 2 1 2 2 2 3 ...
##  $ eduyrs.na   : num [1:1089] 13 14 14 13 18 16 0 14 10 7 ...
##  $ sat.dem.fac : Factor w/ 11 levels "0","1","2","3",..: 8 6 8 6 10 9 6 6 5 5 ...
##  $ yearpay     : num [1:1089] 288000 72000 19200 26400 72000 ...
##  $ vote.fac    : Factor w/ 3 levels "Yes","No","Not eligible": 2 1 1 1 1 1 2 1 1 1 ...
```

Have a quick look at the Data



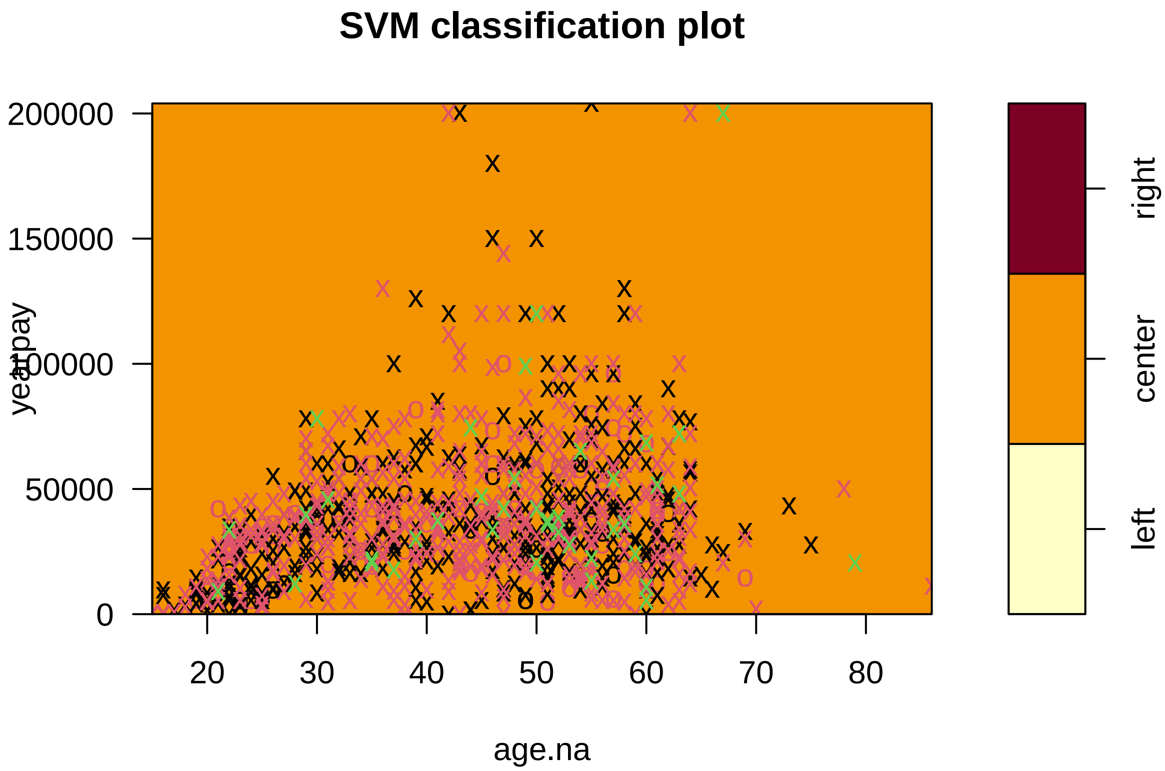


Prepare the Data for Training

Create some easy Variables to access Data

Train the Neural Network

```
##
## Call:
## svm(formula = scale_cat ~ ., data = train, kernel = "linear", cost = 10,
##     scale = TRUE)
##
##
## Parameters:
##    SVM-Type:  C-classification
##  SVM-Kernel:  linear
##        cost:  10
##
## Number of Support Vectors:  833
##
##  ( 385 408 40 )
##
##
## Number of Classes:  3
##
## Levels:
##  left center right
```




```
##
## Parameter tuning of 'svm':
##
## - sampling method: 10-fold cross validation
##
## - best parameters:
##  cost
##    10
##
## - best performance: 0.4334777
##
## - Detailed performance results:
##       cost     error dispersion
## 1   0.001 0.4838430 0.04493581
## 2   0.010 0.4784064 0.04000725
## 3   0.100 0.4380309 0.03559308
## 4   1.000 0.4408087 0.03350829
## 5   5.000 0.4343952 0.04236453
## 6  10.000 0.4334777 0.04530840
## 7 100.000 0.4362300 0.04347618
```

```
##
## Parameter tuning of 'svm':
##
## - sampling method: 10-fold cross validation
##
## - best parameters:
##  cost
##    10
##
```

```
## - best performance: 0.4334777
##
## - Detailed performance results:
##      cost     error dispersion
## 1   0.001 0.4838430 0.04493581
## 2   0.010 0.4784064 0.04000725
## 3   0.100 0.4380309 0.03559308
## 4   1.000 0.4408087 0.03350829
## 5   5.000 0.4343952 0.04236453
## 6  10.000 0.4334777 0.04530840
## 7 100.000 0.4362300 0.04347618
```

```
##
## Call:
## best.tune(METHOD = svm, train.x = scale_cat ~ ., data = d.ess9_supportvec,
##     ranges = list(cost = c(0.001, 0.01, 0.1, 1, 5, 10, 100)), kernel = "linear")
##
##
## Parameters:
##    SVM-Type:  C-classification
##  SVM-Kernel:  linear
##        cost:  10
##
## Number of Support Vectors:  984
##
##  ( 485 452 47 )
##
##
## Number of Classes:  3
##
## Levels:
##  left center right
```

Make Predictions

```
## test_pred
##   left center  right
##     68     95      0
```

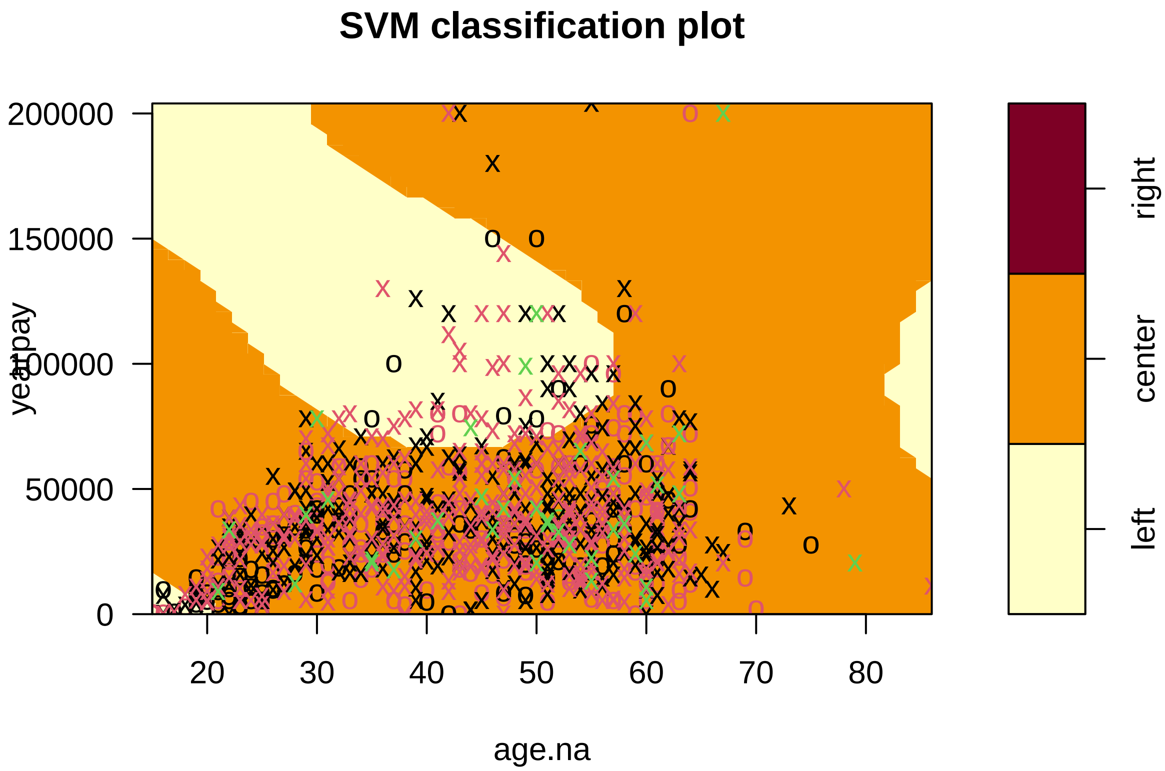
Evaluate the Results

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction left center right
##     left     35     32     1
##     center   37     52     6
##     right     0      0     0
##
## Overall Statistics
##
##                Accuracy : 0.5337
##                  95% CI : (0.4541, 0.6121)
##     No Information Rate : 0.5153
##     P-Value [Acc > NIR] : 0.3479
```

```
##
##                   Kappa : 0.0953
##
##  Mcnemar's Test P-Value : 0.0612
##
## Statistics by Class:
##
##                      Class: left Class: center Class: right
## Sensitivity               0.4861        0.6190      0.00000
## Specificity               0.6374        0.4557      1.00000
## Pos Pred Value            0.5147        0.5474          NaN
## Neg Pred Value            0.6105        0.5294      0.95706
## Prevalence                0.4417        0.5153      0.04294
## Detection Rate            0.2147        0.3190      0.00000
## Detection Prevalence      0.4172        0.5828      0.00000
## Balanced Accuracy         0.5617        0.5374      0.50000
```

Try again with other parameters

```
##
## Call:
## svm(formula = scale_cat ~ ., data = train, kernel = "radial", cost = 100000,
##     scale = TRUE)
##
##
## Parameters:
##    SVM-Type:  C-classification
##  SVM-Kernel:  radial
##        cost:  100000
##
## Number of Support Vectors:  661
##
##  ( 300 322 39 )
##
##
## Number of Classes:  3
##
## Levels:
##  left center right
```




```
## test_pred1
##   left center  right
##     85     67     11
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction left center right
##     left     40     41     4
##     center   29     36     2
##     right     3      7     1
##
## Overall Statistics
##
##                Accuracy : 0.4724
##                  95% CI : (0.3938, 0.552)
##     No Information Rate : 0.5153
##     P-Value [Acc > NIR] : 0.8801
##
##                   Kappa : 0.0492
##
##  Mcnemar's Test P-Value : 0.1734
##
## Statistics by Class:
##
##                      Class: left Class: center Class: right
## Sensitivity               0.5556        0.4286     0.142857
## Specificity               0.5055        0.6076     0.935897
## Pos Pred Value            0.4706        0.5373     0.090909
## Neg Pred Value            0.5897        0.5000     0.960526
```

```
## Prevalence                0.4417         0.5153     0.042945
## Detection Rate            0.2454         0.2209     0.006135
## Detection Prevalence      0.5215         0.4110     0.067485
## Balanced Accuracy         0.5305         0.5181     0.539377
```