model.rmd

Daniel Karstad Maren Sognefest

Modeller

Leser inn data

```
pm2 <- read_csv("data/pm2.csv", show_col_types = FALSE)
```

Vi er ute etter fylkesnummeret, for å lage en ny fylke faktorvariabel. (De to første siffrene i kommunenummeret). Vi bruker mutate() som funksjon for å hente ut deler av en tekststreng. Lager også en faktorvariabel fra årsvariabelen. Og skalerer variabelen Trade_pc.

```
pm2 <- pm2 %>%
  mutate(
    fnr = (str_sub(knr, 1,2))
  )
```

```
pm2 <- pm2 %>%
  mutate(
    fnr = parse_factor(fnr, levels = fnr),
    aar_f = parse_factor(as.character(aar))
  )
```

```
pm2 <- pm2 %>%mutate(Trade_pc_100K = trade_pc/100000)
```

```
head(pm2, n = 4)
```

```
## # A tibble: 4 x 19
##   knr   knavn    aar   pm2 ya_menn ya_kvinner ya_total inc_k1 inc_k5 uni_k_mf
##   <chr> <chr>  <dbl> <dbl>   <dbl>      <dbl>    <dbl>  <dbl>  <dbl>    <dbl>
## 1 0101  Halden  2008 13427    59.7       56.8     58.3   24.5   13.6     17.8
## 2 0101  Halden  2009 13095    59.8       57.0     58.4   24.4   14.1     18.2
## 3 0101  Halden  2010 13832    59.6       57.1     58.3   23.9   13.7     18.6
## 4 0101  Halden  2011 14915    59.8       57.2     58.5   24     14       19
## # ... with 9 more variables: uni_k_m <dbl>, uni_k_f <dbl>, uni_l_mf <dbl>,
## #   uni_l_m <dbl>, uni_l_f <dbl>, trade_pc <dbl>, fnr <fct>, aar_f <fct>,
## #   Trade_pc_100K <dbl>
```

##Modell

```r
mod1 <- 'pm2 ~ aar_f + ya_total + inc_k1 + inc_k5 + uni_k_mf + uni_l_mf + Trade_pc_100K'
```

i.Generer et lm objekt (lm1) utfra mod1 og datasettet pm2.

```r
lm1 <- lm(mod1, data = pm2)
```

```r
lm1 %>%
  summary()
```

```
##
## Call:
## lm(formula = mod1, data = pm2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9089.8 -1525.7    -0.9  1526.2 16322.2
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -23574.01    2293.29 -10.280  < 2e-16 ***
## aar_f2009        235.79     237.53   0.993 0.320967
## aar_f2010        856.73     236.16   3.628 0.000291 ***
## aar_f2011       1573.77     233.50   6.740 1.94e-11 ***
## aar_f2012       2147.08     232.51   9.234  < 2e-16 ***
## aar_f2013       2696.89     233.74  11.538  < 2e-16 ***
## aar_f2014       2690.92     236.01  11.402  < 2e-16 ***
## aar_f2015       3273.99     236.74  13.830  < 2e-16 ***
## aar_f2016       3935.87     240.54  16.363  < 2e-16 ***
## aar_f2017       4750.67     241.87  19.641  < 2e-16 ***
## ya_total         640.19      33.87  18.902  < 2e-16 ***
## inc_k1          -380.90      25.34 -15.030  < 2e-16 ***
## inc_k5           198.36      19.24  10.309  < 2e-16 ***
## uni_k_mf        -122.68      27.26  -4.501 7.06e-06 ***
## uni_l_mf        1262.10      40.27  31.340  < 2e-16 ***
## Trade_pc_100K   1184.06     205.44   5.764 9.19e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2611 on 2649 degrees of freedom
## Multiple R-squared:  0.8196, Adjusted R-squared:  0.8186
## F-statistic: 802.5 on 15 and 2649 DF,  p-value: < 2.2e-16
```

ii. Legg residualene fra den lineære modellen til datasettet pm2.

```
pm2 <- pm2 %>%
  add_residuals(lm1)
```

Forklaring:

i.

Vi observerer at koeffisientene er signifikante på et 0.5%-nivå, og de fleste har solide t-verdier. Års-koeffisientene viser hvor mye prisene øker pr kvadratmeter fra år til år.

ii.

Fortegnene er som ventet. De illustrerer en økning fra år til år, som nevnt.

Test for heteroskedastisitet

i.

```
bptest(lm1)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  lm1
## BP = 315.98, df = 15, p-value < 2.2e-16
```

ii.

Vi ser at p-verdien er under 0,5%. Dermed er det ikke grunnlag for heteroskedastisitet. Nullhypotesen forkastes ettersom p-verdien er under 0.5%.

```
library(gvlma)
gvlma(lm1)
```

```
##
## Call:
## lm(formula = mod1, data = pm2)
##
## Coefficients:
##   (Intercept)      aar_f2009      aar_f2010      aar_f2011      aar_f2012
##      -23574.0          235.8          856.7         1573.8         2147.1
##     aar_f2013      aar_f2014      aar_f2015      aar_f2016      aar_f2017
##        2696.9         2690.9         3274.0         3935.9         4750.7
##      ya_total         inc_k1         inc_k5       uni_k_mf       uni_l_mf
##         640.2         -380.9          198.4         -122.7         1262.1
```

```
## Trade_pc_100K
##       1184.1
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance =  0.05
##
## Call:
##  gvlma(x = lm1)
##
##                     Value   p-value                   Decision
## Global Stat        871.80 0.000e+00 Assumptions NOT satisfied!
## Skewness            65.91 4.441e-16 Assumptions NOT satisfied!
## Kurtosis           691.06 0.000e+00 Assumptions NOT satisfied!
## Link Function       84.17 0.000e+00 Assumptions NOT satisfied!
## Heteroscedasticity  30.66 3.077e-08 Assumptions NOT satisfied!
```

iii.

```
coeftest(lm1, vcov = vcovHC(lm1, type = "HC3"))
```

```
##
## t test of coefficients:
##
##                   Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)   -23574.015   2526.378  -9.3312 < 2.2e-16 ***
## aar_f2009        235.786    203.381   1.1593 0.2464254
## aar_f2010        856.728    194.239   4.4107 1.072e-05 ***
## aar_f2011       1573.772    198.548   7.9264 3.294e-15 ***
## aar_f2012       2147.080    212.046  10.1255 < 2.2e-16 ***
## aar_f2013       2696.888    207.517  12.9960 < 2.2e-16 ***
## aar_f2014       2690.915    217.566  12.3683 < 2.2e-16 ***
## aar_f2015       3273.988    231.201  14.1608 < 2.2e-16 ***
## aar_f2016       3935.865    249.553  15.7717 < 2.2e-16 ***
## aar_f2017       4750.672    265.244  17.9106 < 2.2e-16 ***
## ya_total         640.195     37.917  16.8841 < 2.2e-16 ***
## inc_k1          -380.905     25.433 -14.9768 < 2.2e-16 ***
## inc_k5           198.359     22.263   8.9099 < 2.2e-16 ***
## uni_k_mf        -122.684     34.785  -3.5270 0.0004275 ***
## uni_l_mf        1262.100     69.257  18.2233 < 2.2e-16 ***
## Trade_pc_100K   1184.061    218.171   5.4272 6.242e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

iv.

```
pm2 <- pm2 %>%
  add_residuals(lm1)
```

v.

```
pm2 <- pm2 %>%
  mutate(aar_d =make_date(aar))
```

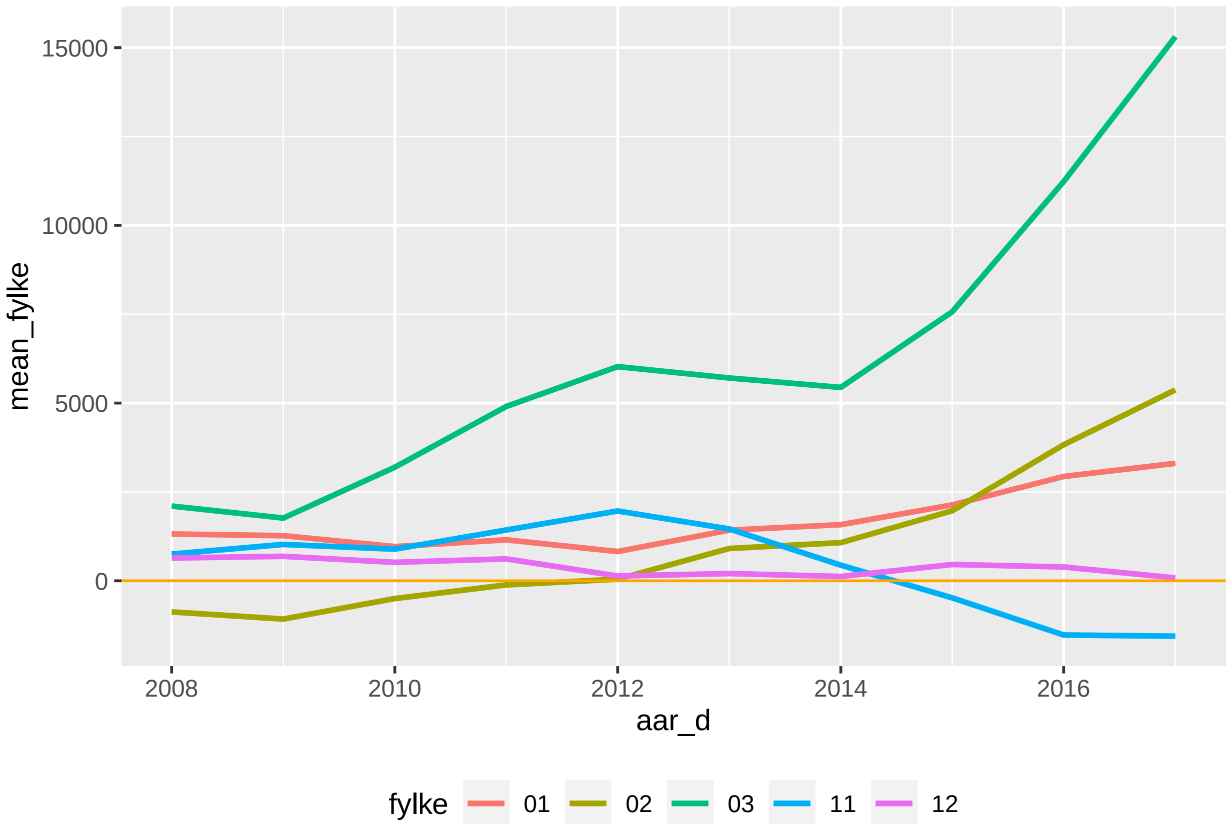
vi.

```
pm2 <- pm2 %>%
  mutate(fylke = substr(knr, start = 1, stop = 2))
```

vii + viii + ix + x

```
pm2 %>%
  filter(fylke %in% c("01", "02", "03", "11", "12")) %>%
  unnest(c(fylke)) %>%
  group_by(fylke, aar_d) %>%
  summarize(mean_fylke = mean(resid)
            ) %>%
  ggplot(aes(x = aar_d, y = mean_fylke, colour = fylke)) +
  geom_line(lwd=1) +
  theme(legend.position = "bottom")+
  geom_hline(yintercept = 0, colour = "orange")
```

```
## `summarise()` has grouped output by 'fylke'. You can override using the `.groups` arg
```



Dummy fylke og år

i + ii.

Her innfører vi en dummy for hvert fylke hvert år. + Genererer lm2 fra modell 2 og datasettet pm2.

```
mod2 <- 'pm2 ~ aar_f*fnr + ya_total + inc_k1 + inc_k5 + uni_k_mf + uni_l_mf + Trade_pc_1
lm2 <- lm(mod2, data = pm2)
summary(lm2)
```

```
##
## Call:
## lm(formula = mod2, data = pm2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7923.8 -1315.5    16.7  1307.7 13524.6
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -21008.194   2229.110  -9.424  < 2e-16 ***
## aar_f2009           117.254    779.646   0.150 0.880466
```

```
## aar_f2010           427.365    791.428   0.540 0.589252
## aar_f2011          1290.937    791.696   1.631 0.103103
## aar_f2012          1475.509    792.404   1.862 0.062712 .
## aar_f2013          2518.738    792.973   3.176 0.001510 **
## aar_f2014          2755.249    782.555   3.521 0.000438 ***
## aar_f2015          3823.769    782.737   4.885 1.10e-06 ***
## aar_f2016          5286.236    795.176   6.648 3.65e-11 ***
## aar_f2017          6448.890    783.851   8.227 3.06e-16 ***
## fnr02             -1576.497    745.081  -2.116 0.034455 *
## fnr03              2858.301   2325.006   1.229 0.219049
## fnr04              -993.568    822.150  -1.208 0.226971
## fnr05             -1836.135    792.893  -2.316 0.020654 *
## fnr06             -2065.423    770.483  -2.681 0.007396 **
## fnr07              -647.551   1148.302  -0.564 0.572859
## fnr08             -3331.237    878.035  -3.794 0.000152 ***
## fnr09             -1665.045    903.551  -1.843 0.065481 .
## fnr10              -336.884    879.838  -0.383 0.701832
## fnr11              -750.142    772.929  -0.971 0.331883
## fnr12             -1106.505    768.237  -1.440 0.149905
## fnr14             -3038.500   1019.569  -2.980 0.002909 **
## fnr15             -3457.883    758.964  -4.556 5.47e-06 ***
## fnr16             -1190.346    793.732  -1.500 0.133825
## fnr17             -2715.149    858.902  -3.161 0.001590 **
## fnr18             -1485.901    822.635  -1.806 0.070998 .
## fnr19             -3255.562   1253.077  -2.598 0.009431 **
## fnr20             -2725.899   1074.474  -2.537 0.011243 *
## ya_total            527.285     32.099  16.427  < 2e-16 ***
## inc_k1             -271.324     23.167 -11.711  < 2e-16 ***
## inc_k5              236.241     20.114  11.745  < 2e-16 ***
## uni_k_mf            131.552     26.599   4.946 8.09e-07 ***
## uni_l_mf            815.724     40.991  19.900  < 2e-16 ***
## Trade_pc_100K      1376.572    181.391   7.589 4.54e-14 ***
## aar_f2009:fnr02     -71.931   1031.037  -0.070 0.944386
## aar_f2010:fnr02     774.604   1039.933   0.745 0.456427
## aar_f2011:fnr02     976.443   1039.975   0.939 0.347869
## aar_f2012:fnr02    1529.124   1040.006   1.470 0.141607
## aar_f2013:fnr02    1893.317   1040.167   1.820 0.068849 .
## aar_f2014:fnr02    1996.251   1031.509   1.935 0.053071 .
## aar_f2015:fnr02    2482.620   1032.027   2.406 0.016220 *
## aar_f2016:fnr02    3546.831   1040.857   3.408 0.000666 ***
## aar_f2017:fnr02    4777.595   1032.312   4.628 3.88e-06 ***
## aar_f2009:fnr03      14.345   3259.610   0.004 0.996489
## aar_f2010:fnr03    1913.904   3262.569   0.587 0.557509
## aar_f2011:fnr03    3747.122   3262.937   1.148 0.250919
## aar_f2012:fnr03    5474.015   3263.386   1.677 0.093590 .
```

```
## aar_f2013:fnr03   4844.502   3264.116   1.484 0.137891
## aar_f2014:fnr03   4686.445   3261.996   1.437 0.150936
## aar_f2015:fnr03   6763.321   3264.005   2.072 0.038360 *
## aar_f2016:fnr03   9816.586   3267.606   3.004 0.002689 **
## aar_f2017:fnr03  13692.078   3265.735   4.193 2.85e-05 ***
## aar_f2009:fnr04   -444.565   1130.748  -0.393 0.694235
## aar_f2010:fnr04   -136.985   1138.855  -0.120 0.904268
## aar_f2011:fnr04   -814.258   1138.925  -0.715 0.474716
## aar_f2012:fnr04   -883.900   1138.982  -0.776 0.437797
## aar_f2013:fnr04  -1448.025   1139.100  -1.271 0.203777
## aar_f2014:fnr04  -1290.240   1140.388  -1.131 0.257995
## aar_f2015:fnr04  -2018.714   1108.852  -1.821 0.068797 .
## aar_f2016:fnr04  -2342.956   1148.279  -2.040 0.041416 *
## aar_f2017:fnr04  -3652.695   1122.818  -3.253 0.001157 **
## aar_f2009:fnr05    308.104   1102.501   0.279 0.779916
## aar_f2010:fnr05    570.141   1128.329   0.505 0.613395
## aar_f2011:fnr05    397.338   1103.369   0.360 0.718793
## aar_f2012:fnr05    682.055   1096.902   0.622 0.534130
## aar_f2013:fnr05   -515.624   1103.659  -0.467 0.640402
## aar_f2014:fnr05   -248.069   1102.980  -0.225 0.822069
## aar_f2015:fnr05  -1446.840   1088.628  -1.329 0.183955
## aar_f2016:fnr05  -2295.593   1111.090  -2.066 0.038926 *
## aar_f2017:fnr05  -2587.929   1095.274  -2.363 0.018214 *
## aar_f2009:fnr06   -284.365   1115.258  -0.255 0.798762
## aar_f2010:fnr06    199.040   1095.054   0.182 0.855784
## aar_f2011:fnr06   -273.079   1087.600  -0.251 0.801770
## aar_f2012:fnr06    338.417   1087.688   0.311 0.755725
## aar_f2013:fnr06    -30.035   1095.053  -0.027 0.978121
## aar_f2014:fnr06    141.659   1086.822   0.130 0.896306
## aar_f2015:fnr06     50.755   1095.323   0.046 0.963044
## aar_f2016:fnr06  -1265.800   1095.330  -1.156 0.247943
## aar_f2017:fnr06   -618.578   1086.918  -0.569 0.569332
## aar_f2009:fnr07    123.862   1615.949   0.077 0.938908
## aar_f2010:fnr07    742.421   1621.634   0.458 0.647120
## aar_f2011:fnr07    290.257   1621.737   0.179 0.857969
## aar_f2012:fnr07   1045.407   1621.717   0.645 0.519227
## aar_f2013:fnr07    880.167   1621.718   0.543 0.587359
## aar_f2014:fnr07    620.068   1616.312   0.384 0.701285
## aar_f2015:fnr07   1134.524   1616.177   0.702 0.482758
## aar_f2016:fnr07    392.252   1621.759   0.242 0.808902
## aar_f2017:fnr07    970.177   1563.951   0.620 0.535093
## aar_f2009:fnr08    459.896   1232.238   0.373 0.709017
## aar_f2010:fnr08   1394.612   1239.740   1.125 0.260731
## aar_f2011:fnr08    653.046   1239.741   0.527 0.598408
## aar_f2012:fnr08    613.058   1224.304   0.501 0.616599
```

```
## aar_f2013:fnr08    435.360   1224.295   0.356 0.722170
## aar_f2014:fnr08    723.664   1232.312   0.587 0.557096
## aar_f2015:fnr08   -966.484   1216.771  -0.794 0.427096
## aar_f2016:fnr08  -2335.970   1211.230  -1.929 0.053896 .
## aar_f2017:fnr08  -2518.929   1203.695  -2.093 0.036481 *
## aar_f2009:fnr09    496.821   1250.596   0.397 0.691205
## aar_f2010:fnr09    845.272   1242.706   0.680 0.496450
## aar_f2011:fnr09    586.783   1275.878   0.460 0.645625
## aar_f2012:fnr09   1078.728   1275.855   0.845 0.397917
## aar_f2013:fnr09   -161.737   1276.067  -0.127 0.899151
## aar_f2014:fnr09   -572.407   1269.304  -0.451 0.652057
## aar_f2015:fnr09   -775.032   1268.702  -0.611 0.541332
## aar_f2016:fnr09  -1827.137   1297.677  -1.408 0.159255
## aar_f2017:fnr09  -2917.203   1290.742  -2.260 0.023902 *
## aar_f2009:fnr10   -381.991   1216.757  -0.314 0.753591
## aar_f2010:fnr10    430.096   1224.244   0.351 0.725382
## aar_f2011:fnr10   -451.023   1224.302  -0.368 0.712612
## aar_f2012:fnr10   -554.393   1224.482  -0.453 0.650764
## aar_f2013:fnr10   -793.190   1240.041  -0.640 0.522461
## aar_f2014:fnr10   -848.570   1233.045  -0.688 0.491397
## aar_f2015:fnr10  -1677.160   1232.946  -1.360 0.173864
## aar_f2016:fnr10  -3029.946   1258.763  -2.407 0.016154 *
## aar_f2017:fnr10  -4488.173   1217.432  -3.687 0.000232 ***
## aar_f2009:fnr11    356.831   1059.170   0.337 0.736223
## aar_f2010:fnr11    542.002   1074.015   0.505 0.613850
## aar_f2011:fnr11    952.291   1062.307   0.896 0.370107
## aar_f2012:fnr11   1895.127   1057.622   1.792 0.073275 .
## aar_f2013:fnr11    891.229   1057.851   0.842 0.399595
## aar_f2014:fnr11   -184.236   1054.843  -0.175 0.861363
## aar_f2015:fnr11  -1379.012   1061.217  -1.299 0.193907
## aar_f2016:fnr11  -3293.368   1064.259  -3.095 0.001993 **
## aar_f2017:fnr11  -3711.962   1050.680  -3.533 0.000419 ***
## aar_f2009:fnr12    132.179   1065.544   0.124 0.901287
## aar_f2010:fnr12    250.201   1062.208   0.236 0.813803
## aar_f2011:fnr12    183.677   1062.212   0.173 0.862729
## aar_f2012:fnr12     33.874   1052.540   0.032 0.974329
## aar_f2013:fnr12   -490.927   1048.180  -0.468 0.639568
## aar_f2014:fnr12   -603.292   1044.182  -0.578 0.563475
## aar_f2015:fnr12   -725.510   1044.000  -0.695 0.487163
## aar_f2016:fnr12  -1715.402   1044.383  -1.643 0.100613
## aar_f2017:fnr12  -2337.025   1044.352  -2.238 0.025325 *
## aar_f2009:fnr14   -426.236   1525.949  -0.279 0.780019
## aar_f2010:fnr14   -644.540   1434.484  -0.449 0.653242
## aar_f2011:fnr14    440.606   1434.376   0.307 0.758735
## aar_f2012:fnr14    457.978   1357.528   0.337 0.735873
```

```
## aar_f2013:fnr14   -384.308   1378.066  -0.279 0.780364
## aar_f2014:fnr14   -936.722   1371.465  -0.683 0.494666
## aar_f2015:fnr14  -1554.032   1333.875  -1.165 0.244111
## aar_f2016:fnr14  -3120.327   1403.357  -2.223 0.026274 *
## aar_f2017:fnr14  -2571.348   1319.536  -1.949 0.051447 .
## aar_f2009:fnr15    261.649   1042.149   0.251 0.801783
## aar_f2010:fnr15    471.393   1046.216   0.451 0.652339
## aar_f2011:fnr15    282.472   1031.058   0.274 0.784136
## aar_f2012:fnr15    560.923   1024.943   0.547 0.584241
## aar_f2013:fnr15   -264.161   1027.923  -0.257 0.797212
## aar_f2014:fnr15   -813.723   1019.126  -0.798 0.424685
## aar_f2015:fnr15   -861.554   1029.604  -0.837 0.402796
## aar_f2016:fnr15  -2262.057   1027.905  -2.201 0.027853 *
## aar_f2017:fnr15  -2834.650   1022.436  -2.772 0.005605 **
## aar_f2009:fnr16   -246.995   1120.133  -0.221 0.825496
## aar_f2010:fnr16   -166.038   1119.104  -0.148 0.882066
## aar_f2011:fnr16   -340.432   1103.490  -0.309 0.757724
## aar_f2012:fnr16    308.793   1097.092   0.281 0.778377
## aar_f2013:fnr16    -71.436   1097.365  -0.065 0.948102
## aar_f2014:fnr16     -5.831   1096.103  -0.005 0.995756
## aar_f2015:fnr16     96.743   1095.855   0.088 0.929661
## aar_f2016:fnr16  -1427.167   1091.278  -1.308 0.191065
## aar_f2017:fnr16  -2135.764   1089.054  -1.961 0.049977 *
## aar_f2009:fnr17    728.760   1217.308   0.599 0.549451
## aar_f2010:fnr17     53.380   1208.813   0.044 0.964781
## aar_f2011:fnr17    271.488   1195.638   0.227 0.820392
## aar_f2012:fnr17    558.253   1195.861   0.467 0.640669
## aar_f2013:fnr17    163.080   1184.640   0.138 0.890519
## aar_f2014:fnr17    -21.005   1177.212  -0.018 0.985765
## aar_f2015:fnr17   -966.679   1166.258  -0.829 0.407257
## aar_f2016:fnr17  -1820.576   1184.636  -1.537 0.124465
## aar_f2017:fnr17  -2008.188   1188.256  -1.690 0.091148 .
## aar_f2009:fnr18   -264.823   1140.066  -0.232 0.816334
## aar_f2010:fnr18    240.262   1138.884   0.211 0.832934
## aar_f2011:fnr18   -204.436   1123.906  -0.182 0.855678
## aar_f2012:fnr18    539.456   1124.472   0.480 0.631453
## aar_f2013:fnr18    128.551   1123.848   0.114 0.908942
## aar_f2014:fnr18   -555.694   1109.804  -0.501 0.616617
## aar_f2015:fnr18   -948.114   1076.882  -0.880 0.378715
## aar_f2016:fnr18  -1751.338   1117.818  -1.567 0.117302
## aar_f2017:fnr18  -2579.542   1109.505  -2.325 0.020155 *
## aar_f2009:fnr19   1114.445   1692.031   0.659 0.510186
## aar_f2010:fnr19    136.839   1769.674   0.077 0.938372
## aar_f2011:fnr19    153.142   1647.754   0.093 0.925959
## aar_f2012:fnr19    807.952   1611.002   0.502 0.616049
```

```
## aar_f2013:fnr19   -619.025   1560.988  -0.397 0.691726
## aar_f2014:fnr19    827.923   1537.296   0.539 0.590240
## aar_f2015:fnr19   -813.901   1537.453  -0.529 0.596588
## aar_f2016:fnr19   -459.857   1543.410  -0.298 0.765767
## aar_f2017:fnr19  -1726.045   1555.795  -1.109 0.267353
## aar_f2009:fnr20  -1033.459   1564.091  -0.661 0.508840
## aar_f2010:fnr20   -429.231   1569.873  -0.273 0.784555
## aar_f2011:fnr20   -916.024   1570.019  -0.583 0.559645
## aar_f2012:fnr20   -300.249   1648.151  -0.182 0.855462
## aar_f2013:fnr20  -1072.631   1648.536  -0.651 0.515328
## aar_f2014:fnr20  -1324.069   1564.804  -0.846 0.397548
## aar_f2015:fnr20  -2197.055   1564.890  -1.404 0.160454
## aar_f2016:fnr20  -1780.031   1515.941  -1.174 0.240425
## aar_f2017:fnr20  -3151.542   1564.756  -2.014 0.044109 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2238 on 2469 degrees of freedom
## Multiple R-squared:  0.8765, Adjusted R-squared:  0.8667
## F-statistic: 89.86 on 195 and 2469 DF,  p-value: < 2.2e-16
```

iii.

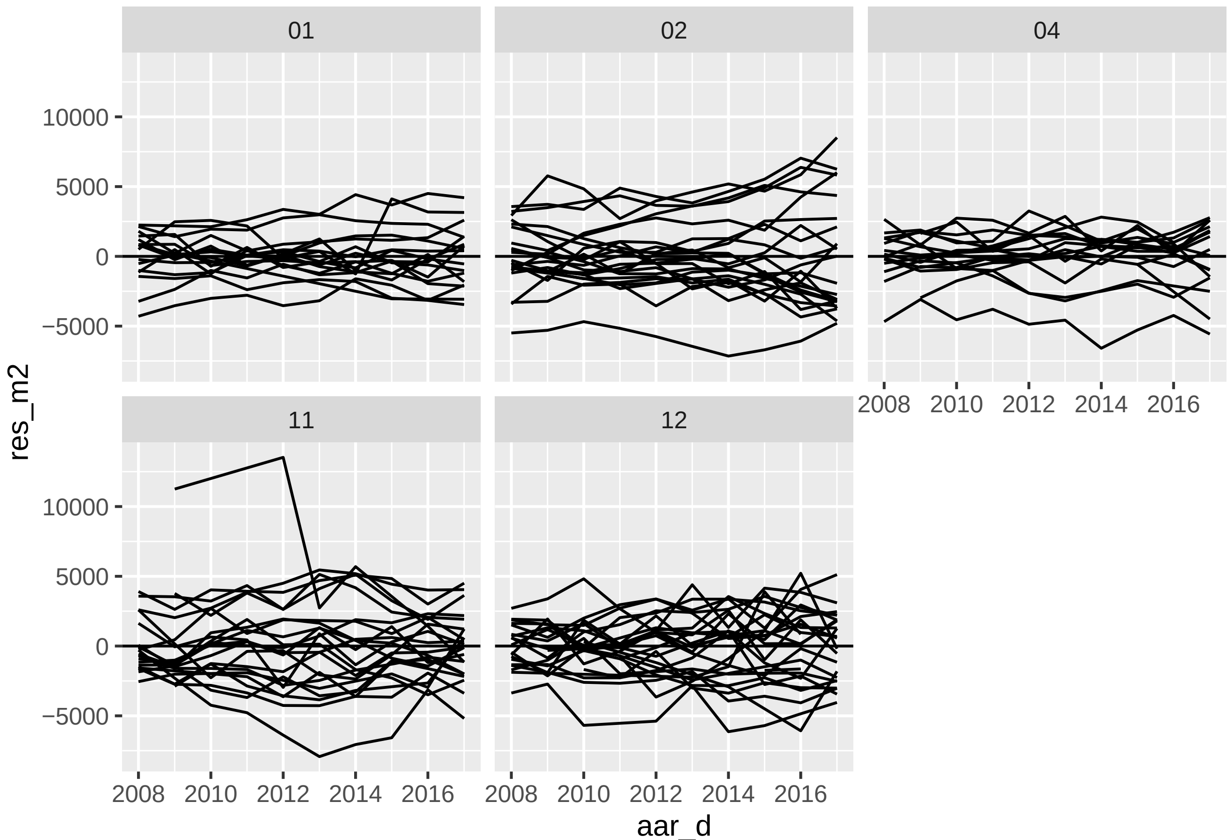
Legger residualene fra lm2 til pm2 og kaller dem res_m2.

```r
pm2 <- pm2 %>%
  mutate(res_m2 = resid(lm2))
```

iv.

Lager del-plott for hvert fylke.

```r
pm2 %>% filter(fnr %in% c("01", "02", "04", "11", "12")) %>%
ggplot(mapping = aes(x = aar_d, y = res_m2)) +
geom_line(aes(group = knavn)) +
scale_size_manual(values = c(seq(2.0, 0.5, by = -0.1))) +
geom_hline(yintercept = 0) +
theme(legend.position = 'bottom') +
  facet_wrap(~fylke)
```



i.

Kvaliteten på modellen er ikke optimal, ettersom den har noen manglende variabler. Det er stor variasjon, noe som kan være forårsaket av heteroskedastisitet i modellen.

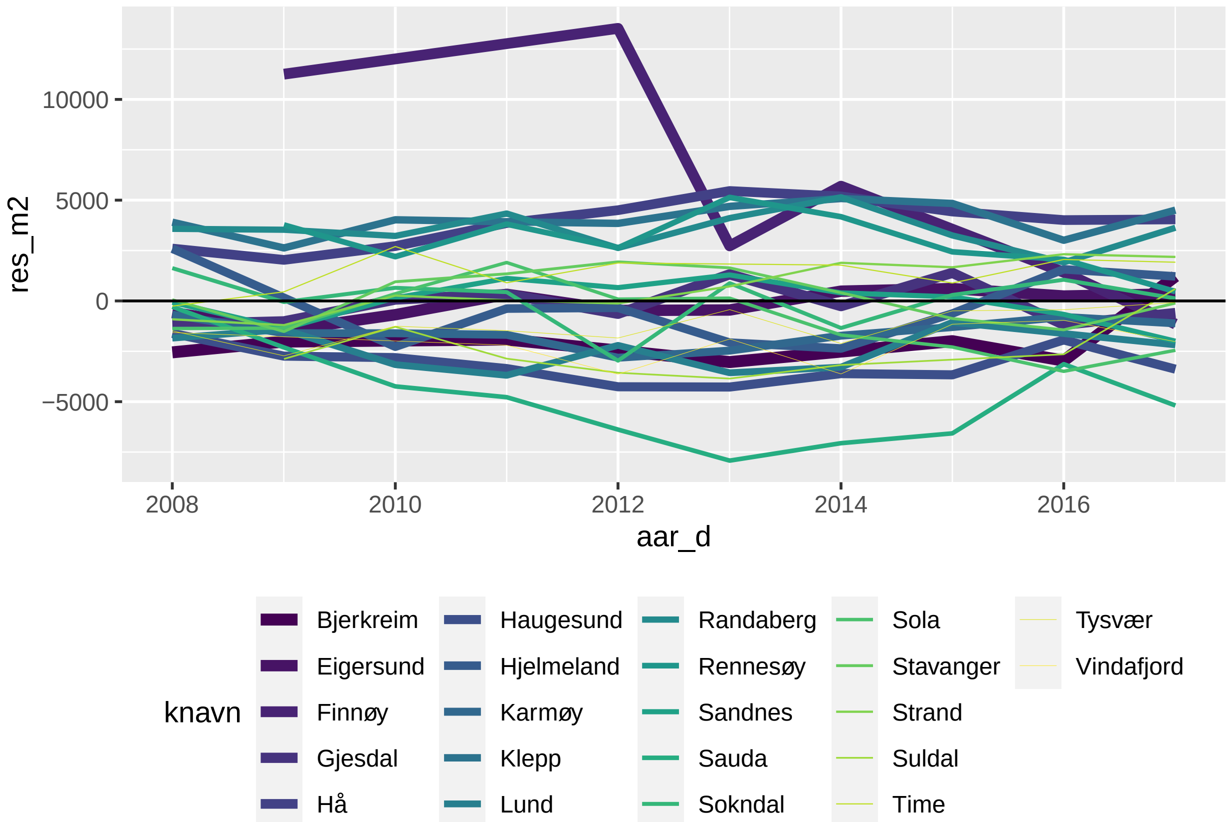
ii.

Ja, det er grunn til å mistenke at vi mangler viktige variabler i modell 2.

iii.

Filtrert med hensyn på fylke "11".

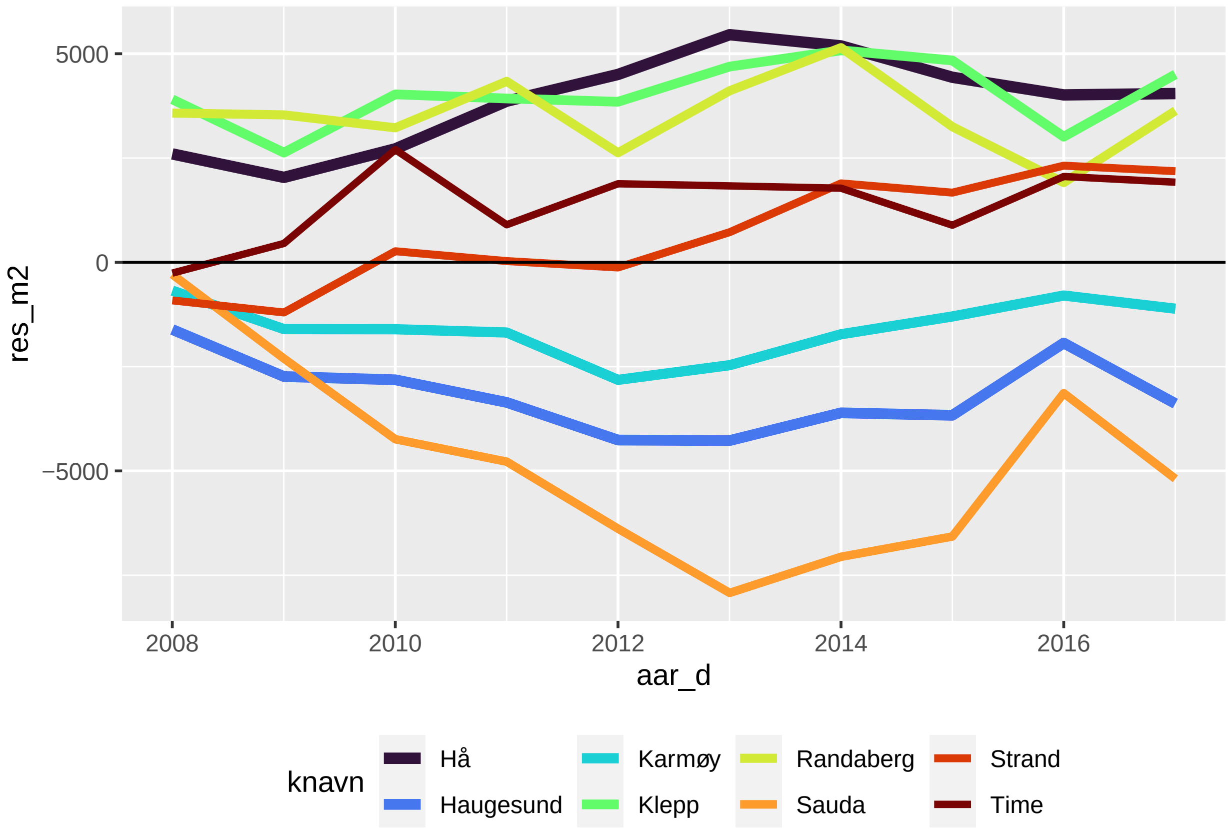
```
pm2 %>% filter(fnr %in% c("11")) %>%
ggplot(mapping = aes(x = aar_d, y = res_m2)) +
scale_color_viridis(discrete = TRUE, option = "D") +
geom_line(aes(group = knavn, colour = knavn, size =knavn)) +
scale_size_manual(values = c(seq(2.1, 0, by = -0.1))) +
geom_hline(yintercept = 0) +
theme(legend.position = 'bottom')
```



i.

Gjentar plottet ovenfor men med sortering for enkelte kommuner.

```
pm2 %>% filter(knr %in% c("1119", "1120", "1127", "1121", "1130", "1135", "1106", "1149"
ggplot(mapping = aes(x = aar_d, y = res_m2)) +
scale_color_viridis(discrete = TRUE, option = "H") +
geom_line(aes(group = knavn, colour = knavn, size =knavn)) +
scale_size_manual(values = c(seq(2.0, 0.5, by = -0.1))) +
geom_hline(yintercept = 0) +
theme(legend.position = 'bottom')
```



ii.

De kommunene som ligger nærme Stavanger overvuderes hva gjelder pris per kvadratmeter.

Feil! Modelen undervurderer prisen derfor positiv residual.

Modell for hvert år

i.

```
pm2 <- pm2 %>%
  mutate(
    aar_d = date(paste0(aar, "-01-01"))
  )
```

```
pm2_n <- pm2 %>%
  select(pm2, fnr, knr, aar_d, aar, aar_f, ya_menn, ya_kvinner, ya_total, inc_k1, inc_k5
  group_by(aar_d) %>%
  nest()
```

```
pm2_n
```

```
## # A tibble: 10 x 2
```

```
## # Groups:   aar_d [10]
##    aar_d      data
##    <date>     <list>
##  1 2008-01-01 <tibble [241 x 13]>
##  2 2009-01-01 <tibble [243 x 13]>
##  3 2010-01-01 <tibble [250 x 13]>
##  4 2011-01-01 <tibble [265 x 13]>
##  5 2012-01-01 <tibble [276 x 13]>
##  6 2013-01-01 <tibble [275 x 13]>
##  7 2014-01-01 <tibble [273 x 13]>
##  8 2015-01-01 <tibble [285 x 13]>
##  9 2016-01-01 <tibble [276 x 13]>
## 10 2017-01-01 <tibble [281 x 13]>
```

```
pm2_n$data[[1]] %>%
head(n = 5)
```

```
## # A tibble: 5 x 13
##     pm2 fnr   knr     aar aar_f ya_menn ya_kvinner ya_total inc_k1 inc_k5
##   <dbl> <fct> <chr> <dbl> <fct>   <dbl>      <dbl>    <dbl>  <dbl>  <dbl>
## 1 13427 01    0101   2008 2008     59.7       56.8     58.3   24.5   13.6
## 2 18299 01    0104   2008 2008     60.7       58.7     59.7   22.8   16.2
## 3 14981 01    0105   2008 2008     60.9       58.1     59.5   22.2   13.6
## 4 15671 01    0106   2008 2008     59.8       57.8     58.8   21.8   16.2
## 5 18844 01    0111   2008 2008     61.7       61.3     61.5   17.8   19
## # ... with 3 more variables: uni_k_mf <dbl>, uni_l_mf <dbl>,
## #   Trade_pc_100K <dbl>
```

i.

Funksjonen *kom_model* for å kjøre hvert enkelt år:

```
kom_model <- function(a_df) {
  lm(pm2 ~ fnr + ya_total + inc_k1 + inc_k5 + uni_k_mf + uni_l_mf + Trade_pc_100K, data
}
```

i.

```
pm2_n <- pm2_n %>%
  mutate(
    model = map(data, .f = kom_model)
    )
```

```
# summary 2008
pm2_n$model[[1]] %>%
  summary()
```

```
##
## Call:
## lm(formula = pm2 ~ fnr + ya_total + inc_k1 + inc_k5 + uni_k_mf +
##     uni_l_mf + Trade_pc_100K, data = a_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4654.7 -1159.6   108.5  1074.2  4907.5
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -19995.26    5886.89  -3.397 0.000812 ***
## fnr02             94.71     661.30   0.143 0.886251
## fnr03           4409.33    2011.32   2.192 0.029429 *
## fnr04          -1846.50     670.37  -2.754 0.006380 **
## fnr05          -2384.11     638.14  -3.736 0.000239 ***
## fnr06          -1791.29     619.59  -2.891 0.004231 **
## fnr07           -452.89     921.19  -0.492 0.623476
## fnr08          -3705.14     705.30  -5.253 3.57e-07 ***
## fnr09          -1867.00     726.18  -2.571 0.010812 *
## fnr10           -422.80     723.49  -0.584 0.559570
## fnr11            201.24     697.48   0.289 0.773227
## fnr12           -562.95     668.02  -0.843 0.400325
## fnr14          -2970.92     849.89  -3.496 0.000574 ***
## fnr15          -3270.48     665.51  -4.914 1.76e-06 ***
## fnr16          -1398.42     646.10  -2.164 0.031531 *
## fnr17          -3260.93     713.61  -4.570 8.21e-06 ***
## fnr18          -2036.75     681.74  -2.988 0.003137 **
## fnr19          -3524.96    1004.86  -3.508 0.000549 ***
## fnr20          -3077.48     873.75  -3.522 0.000522 ***
## ya_total         467.07      85.56   5.459 1.31e-07 ***
## inc_k1          -116.06      65.47  -1.773 0.077714 .
## inc_k5           191.86      53.09   3.614 0.000375 ***
## uni_k_mf         191.90      74.65   2.571 0.010822 *
## uni_l_mf         606.16     126.96   4.774 3.32e-06 ***
## Trade_pc_100K   1338.87     513.05   2.610 0.009699 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1768 on 216 degrees of freedom
## Multiple R-squared:  0.8652, Adjusted R-squared:  0.8502
## F-statistic: 57.76 on 24 and 216 DF,  p-value: < 2.2e-16
```

i.

Funksjonen *glance* + *mod_summary* og *unnest ()*

```
pm2_n %>%
  filter(aar_d == "2008-01-01") %>%
  .$model %>%
  map_df(glance)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl> <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.865         0.850 1768.      57.8 2.25e-80    24 -2131. 4314. 4404.
## # ... with 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

```
mod_sum <- pm2_n %>%
  mutate(mod_summary = map(.x = model, .f = glance)) %>%
  unnest(mod_summary) %>%
  print()
```

```
## # A tibble: 10 x 15
## # Groups:   aar_d [10]
##    aar_d      data       model r.squared adj.r.squared sigma statistic   p.value
##    <date>     <list>     <lis>     <dbl>         <dbl> <dbl>     <dbl>     <dbl>
##  1 2008-01-01 <tibble [~ <lm>      0.865         0.850 1768.      57.8 2.25e- 80
##  2 2009-01-01 <tibble [~ <lm>      0.857         0.841 1867.      54.2 2.66e- 78
##  3 2010-01-01 <tibble [~ <lm>      0.876         0.862 1842.      66.0 5.80e- 88
##  4 2011-01-01 <tibble [~ <lm>      0.873         0.861 2023.      69.0 1.45e- 93
##  5 2012-01-01 <tibble [~ <lm>      0.848         0.834 2302.      58.4 1.70e- 88
##  6 2013-01-01 <tibble [~ <lm>      0.884         0.873 2095.      79.7 9.97e-103
##  7 2014-01-01 <tibble [~ <lm>      0.862         0.849 2308.      64.5 2.02e- 92
##  8 2015-01-01 <tibble [~ <lm>      0.868         0.855 2438.      70.9 1.20e- 99
##  9 2016-01-01 <tibble [~ <lm>      0.871         0.859 2535.      70.5 3.01e- 97
## 10 2017-01-01 <tibble [~ <lm>      0.877         0.865 2703.      75.7 7.97e-102
## # ... with 7 more variables: df <dbl>, logLik <dbl>, AIC <dbl>, BIC <dbl>,
## #   deviance <dbl>, df.residual <int>, nobs <int>
```

i.

Ny variabel av type date i *coef_df* som angir år.

```
coef_df <- mod_sum$model %>%
  map_df(1) %>%
  tibble()
```

```
coef_df <- coef_df %>%
  mutate(
    aar = ymd(paste(2008:2017, "-01-01", sep = ""))
```

```
  ) %>%
  select(aar, everything())
```

ii.

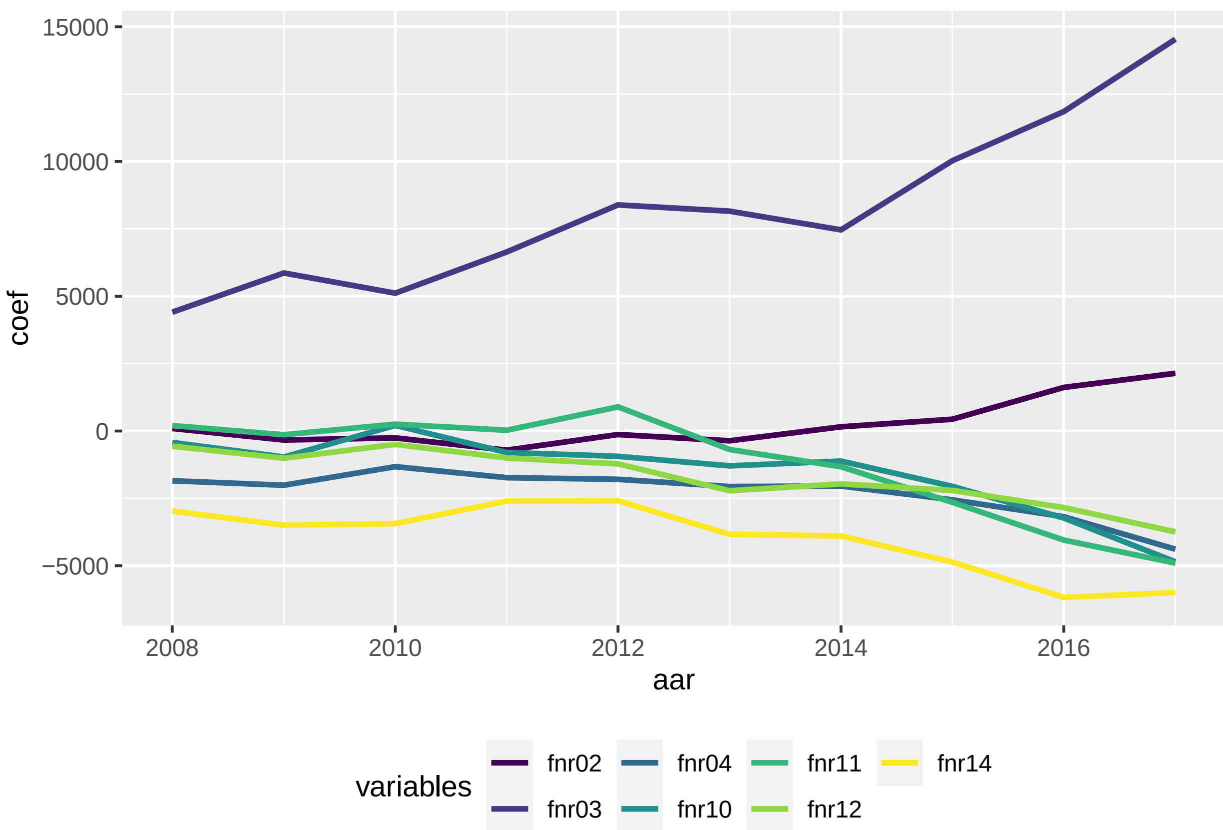
Bruk pivot_longer på *coef_df* for å gjøre om til *coef_df_long*

```
coef_df_long <- coef_df %>%
  pivot_longer(
    cols = `(Intercept)`:`Trade_pc_100K`,
    names_to = "variables",
    values_to = "coef")
```

iii.

Lager så et plot av fylke-faktorvariablenes koeffisienter ved bruk av *coef_df_long*.

```
coef_df_long %>%
  select(aar, variables, coef) %>%
  filter(
    variables %in% c("fnr02", "fnr03", "fnr04", "fnr10", "fnr11", "fnr12", "fnr14")
  ) %>%
  ggplot(mapping = aes(x = aar, y = coef, colour = variables)) +
  scale_color_viridis(discrete = TRUE, option = "D") +
  geom_line(aes(group = variables), lwd = 1) +
  theme(legend.position = 'bottom')
```



iv.

Plot-et sier oss at prisuktviklingen i fnr03 har økt kontinuerlig fra år 2008 til 2017, bortsett fra en liten nedgang fra 2012 til 2014.

Fnr02 skiller seg ut ved at den har så og si flat utvikling, har steget noe fra 2015 til 2017.

Resten av variablene har sunket over tidsrommet 2008 til 2017. Dette vil si at det har vært negativ prisutvikling i dette tidsrommet.

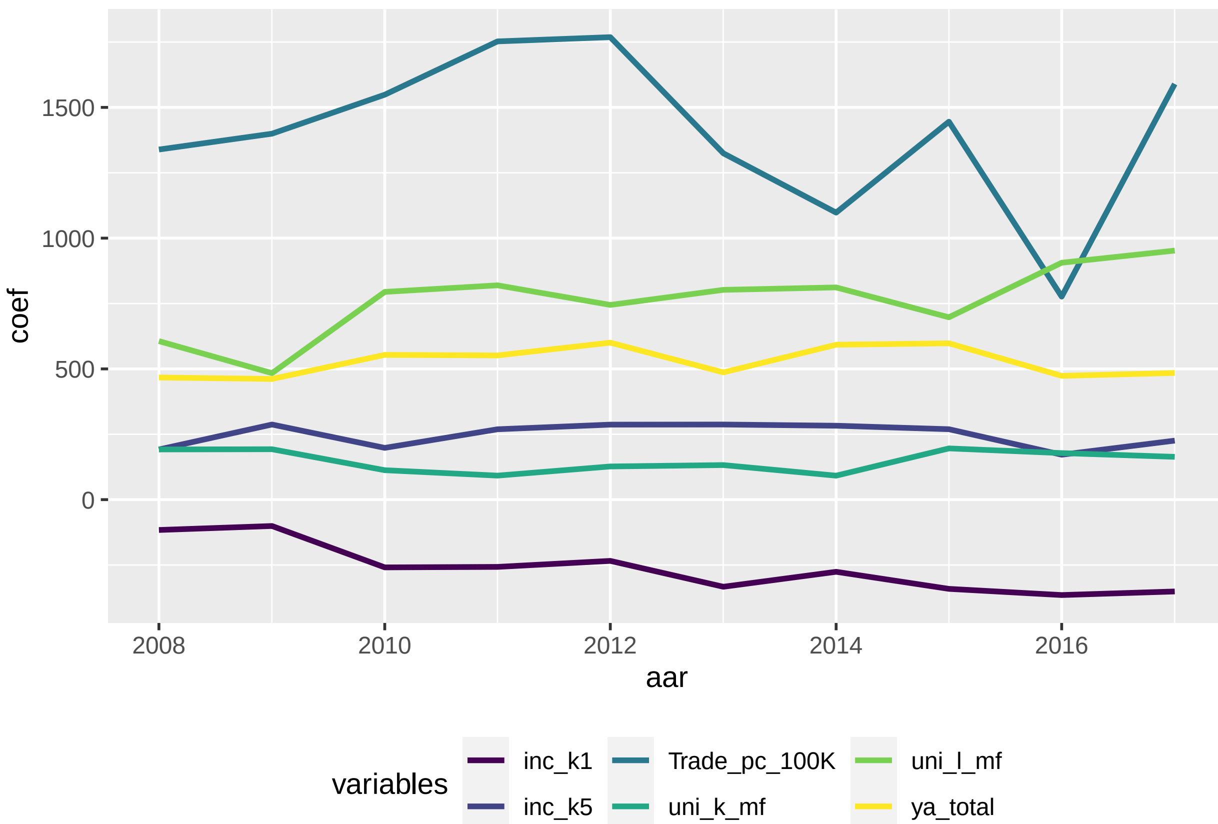
v.

I 2014 gikk vi gjennom en oljekrise. Oljeprisen gikk fra 114 dollar fatet til under 30 usd på et halvt år. Norsk økonomi som er veldig avhengig av olje og norske kronen som er veldig avhengig av oljeprisen gikk dermed inn i en nedgangskonjuktur. Derav fikk vi mange oppsigelser direkte i oljerelaterte næringer men også økende arbeidsledighet generelt. Dette rammet særlig områder som har høy tetthet av oljerelatert næring, for eksempel Stavanger-regionen. Noe som forårsaket en negativ boligpris utvikling og generelt lavere attraktivitet for hele området.

i.

Lage et tilsvarende plot, med nye variabler.

```r
coef_df_long %>%
select(aar, variables, coef) %>%
filter(
variables %in% c("ya_total", "inc_k1", "inc_k5", "uni_k_mf", "uni_l_mf", "Trade_pc_100K"
) %>%
ggplot(mapping = aes(x = aar, y = coef, colour = variables)) +
scale_color_viridis(discrete = TRUE, option = "D") +
geom_line(aes(group = variables), lwd = 1) +
theme(legend.position = 'bottom')
```



ii.

3/4 av variablene ser ut til å være stabile over tid. Unntaket er *trade_pc_100k* og *inc_k1*. *uni_I_mf* svinger godt et par ganger også, så den er ikke helt stabil på kort sikt, men relativt stabil over tid. *trade_pc_100k* er desidert mest ustabil over tidsrommet 2008-2017. *inc_k1* har et par kraftige fall, og over tid er den ikke stabil, med tanke på at den har en nedadgående trend over tid.

```r
#siste
```