Assignment 4

Ingrid-Liv Morkken Thibiga Kuddyar

```
suppressPackageStartupMessages({
library(tidyverse)
library(lubridate)
library(modelr)
library(broom)
library(lmtest)
library(sandwich)
library(viridis)
})
knitr::opts_chunk$set(echo = TRUE, include = TRUE)
```

Modeller

Leser inn data

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

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

```
head(pm2)
```

```
## # A tibble: 6 x 18
##   knr     aar knavn    pm2 Menn_ya_p Kvinner_ya_p Total_ya_p inc_k1 inc_k5
##   <chr> <dbl> <chr>  <dbl>     <dbl>        <dbl>      <dbl>  <dbl>  <dbl>
## 1 0101   2008 Halden 13427      59.7         56.8       58.3   24.5   13.6
## 2 0101   2009 Halden 13095      59.8         57.0       58.4   24.4   14.1
## 3 0101   2010 Halden 13832      59.6         57.1       58.3   23.9   13.7
## 4 0101   2011 Halden 14915      59.8         57.2       58.5   24     14
## 5 0101   2012 Halden 15473      59.5         57.0       58.2   23.9   14
## 6 0101   2013 Halden 15461      59.0         56.7       57.9   24.1   13.4
## # ... with 9 more variables: uni_k_mf <dbl>, uni_k_m <dbl>, uni_k_f <dbl>,
## #   uni_l_mf <dbl>, uni_l_m <dbl>, uni_l_f <dbl>, Trade_p <dbl>, fnr <chr>,
## #   aar_f <chr>
```

```
pm2 %>%
  mutate(
    fnr = parse_factor(fnr, levels =fnr),
    aar_f = parse_factor(aar_f, levels = aar_f)
  )
```

```
## # A tibble: 2,140 x 18
```

```
##    knr     aar knavn    pm2 Menn_ya_p Kvinner_ya_p Total_ya_p inc_k1 inc_k5
##    <chr> <dbl> <chr>  <dbl>     <dbl>        <dbl>      <dbl>  <dbl>  <dbl>
##  1 0101   2008 Halden 13427      59.7         56.8       58.3   24.5   13.6
##  2 0101   2009 Halden 13095      59.8         57.0       58.4   24.4   14.1
##  3 0101   2010 Halden 13832      59.6         57.1       58.3   23.9   13.7
##  4 0101   2011 Halden 14915      59.8         57.2       58.5   24     14
##  5 0101   2012 Halden 15473      59.5         57.0       58.2   23.9   14
##  6 0101   2013 Halden 15461      59.0         56.7       57.9   24.1   13.4
##  7 0101   2014 Halden 17164      58.8         56.7       57.7   23.9   13.5
##  8 0101   2015 Halden 17427      58.7         56.8       57.8   24     13.7
##  9 0101   2016 Halden 18941      58.7         56.6       57.7   24     13.8
## 10 0101   2017 Halden 20143      58.9         56.9       57.9   23.7   14
## # ... with 2,130 more rows, and 9 more variables: uni_k_mf <dbl>,
## #   uni_k_m <dbl>, uni_k_f <dbl>, uni_l_mf <dbl>, uni_l_m <dbl>, uni_l_f <dbl>,
## #   Trade_p <dbl>, fnr <fct>, aar_f <fct>
```

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

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

```
## # A tibble: 4 x 19
##   knr     aar knavn    pm2 Menn_ya_p Kvinner_ya_p Total_ya_p inc_k1 inc_k5
##   <chr> <dbl> <chr>  <dbl>     <dbl>        <dbl>      <dbl>  <dbl>  <dbl>
## 1 0101   2008 Halden 13427      59.7         56.8       58.3   24.5   13.6
## 2 0101   2009 Halden 13095      59.8         57.0       58.4   24.4   14.1
## 3 0101   2010 Halden 13832      59.6         57.1       58.3   23.9   13.7
## 4 0101   2011 Halden 14915      59.8         57.2       58.5   24     14
## # ... with 10 more variables: uni_k_mf <dbl>, uni_k_m <dbl>, uni_k_f <dbl>,
## #   uni_l_mf <dbl>, uni_l_m <dbl>, uni_l_f <dbl>, Trade_p <dbl>, fnr <chr>,
## #   aar_f <chr>, Trade_pc_100K <dbl>
```

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

Modell

i.

```r
lm1 <- lm(mod1, data = pm2, subset = complete.cases(pm2))
```

```r
summary(lm1)
```

```
##
## Call:
## lm(formula = mod1, data = pm2, subset = complete.cases(pm2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8516.6 -1472.1   -29.9  1467.3 15736.3
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -20400.74    2663.02  -7.661 2.79e-14 ***
## aar_f2009        104.15     244.77   0.426 0.670512
```

```
## aar_f2010        908.13     245.16   3.704 0.000217 ***
## aar_f2011       1663.93     245.86   6.768 1.68e-11 ***
## aar_f2012       2240.48     247.10   9.067  < 2e-16 ***
## aar_f2013       2869.30     248.31  11.555  < 2e-16 ***
## aar_f2014       2863.22     250.54  11.428  < 2e-16 ***
## aar_f2015       3525.22     253.08  13.929  < 2e-16 ***
## aar_f2016       4274.99     255.81  16.711  < 2e-16 ***
## aar_f2017       5146.33     258.50  19.909  < 2e-16 ***
## Total_ya_p       582.44      38.94  14.957  < 2e-16 ***
## inc_k1          -376.99      30.29 -12.445  < 2e-16 ***
## inc_k5           194.35      22.87   8.498  < 2e-16 ***
## uni_k_mf         -82.02      29.42  -2.788 0.005357 **
## uni_l_mf        1206.86      42.22  28.585  < 2e-16 ***
## Trade_pc_100K    871.99     218.42   3.992 6.77e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2531 on 2124 degrees of freedom
## Multiple R-squared:  0.8346, Adjusted R-squared:  0.8334
## F-statistic: 714.3 on 15 and 2124 DF,  p-value: < 2.2e-16
```

ii.

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

```
## # A tibble: 2,140 x 20
##    knr     aar knavn    pm2 Menn_ya_p Kvinner_ya_p Total_ya_p inc_k1 inc_k5
##    <chr> <dbl> <chr>  <dbl>     <dbl>        <dbl>      <dbl>  <dbl>  <dbl>
##  1 0101   2008 Halden 13427      59.7         56.8       58.3   24.5   13.6
##  2 0101   2009 Halden 13095      59.8         57.0       58.4   24.4   14.1
##  3 0101   2010 Halden 13832      59.6         57.1       58.3   23.9   13.7
##  4 0101   2011 Halden 14915      59.8         57.2       58.5   24     14
##  5 0101   2012 Halden 15473      59.5         57.0       58.2   23.9   14
##  6 0101   2013 Halden 15461      59.0         56.7       57.9   24.1   13.4
##  7 0101   2014 Halden 17164      58.8         56.7       57.7   23.9   13.5
##  8 0101   2015 Halden 17427      58.7         56.8       57.8   24     13.7
##  9 0101   2016 Halden 18941      58.7         56.6       57.7   24     13.8
## 10 0101   2017 Halden 20143      58.9         56.9       57.9   23.7   14
## # ... with 2,130 more rows, and 11 more variables: uni_k_mf <dbl>,
## #   uni_k_m <dbl>, uni_k_f <dbl>, uni_l_mf <dbl>, uni_l_m <dbl>, uni_l_f <dbl>,
## #   Trade_p <dbl>, fnr <chr>, aar_f <chr>, Trade_pc_100K <dbl>, resid <dbl>
```

Forklaring av modell og diskusjon av fortegnet.

i.

Års-koeffisientene viser hvor mye y (pm2) øker i kvadratmeter fra år til år. I 2009 vil økningen være 104.15, i 2010 908, i 2011 1663.93, osv. opp til år 2017 hvor modellen viser en økning på 5146.33 per kvadratmeter.

ii.

Vi ser at alle fortegnene er positive i modellen, utenom skjæringspunktet, inc__k1 og uni__k__mf. De positive fortegnene illustrerer en økning i y (pm2) fra år til år, som nevnt over. Uni__k__mf med negativt fortegn viser at personer som er bosatt i et område med lavere kvadratmeterpris, også har lavere utdanning. Videre er det to kvintiler i tabellen. De to kvintilene illustrerer ulikt nivå hvor kvintil 1, som har et negtivt fortegn, både er mindre i tall (etter navnet "inc_k**1**") og i utfallet av tabellen (-376.99). Den andre kvintilen, inc_k5, er større både i tall og i utfallet illustrert i tabellen (194.35).

Heteroskedastisitet

i.

```
bptest(lm1)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  lm1
## BP = 352.89, df = 15, p-value < 2.2e-16
```

ii.

Dersom p-verdien er større enn 0.05 kan vi forkaste H0 og dermed kan vi ha heteroskedastisitet. I denne testen har vi ikke heteroskedastisitet siden p-verdien er veldig lav.

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

```
##
## Call:
## lm(formula = mod1, data = pm2, subset = complete.cases(pm2))
##
## Coefficients:
##   (Intercept)      aar_f2009      aar_f2010      aar_f2011      aar_f2012
##     -20400.74         104.15         908.13        1663.93        2240.48
##     aar_f2013      aar_f2014      aar_f2015      aar_f2016      aar_f2017
##       2869.30        2863.22        3525.22        4274.99        5146.33
##    Total_ya_p         inc_k1         inc_k5       uni_k_mf       uni_l_mf
##        582.44        -376.99         194.35         -82.02        1206.86
## Trade_pc_100K
##        871.99
##
##
## 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        733.35 0.000e+00 Assumptions NOT satisfied!
## Skewness            48.82 2.804e-12 Assumptions NOT satisfied!
## Kurtosis           538.05 0.000e+00 Assumptions NOT satisfied!
## Link Function       96.62 0.000e+00 Assumptions NOT satisfied!
## Heteroscedasticity  49.86 1.652e-12 Assumptions NOT satisfied!
```

iii.

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

```
##
## t test of coefficients:
##
##                  Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)    -20400.742   3049.260  -6.6904 2.838e-11 ***
## aar_f2009         104.150    206.348   0.5047 0.6138025
## aar_f2010         908.129    204.590   4.4388 9.511e-06 ***
## aar_f2011        1663.926    212.628   7.8255 7.907e-15 ***
## aar_f2012        2240.475    216.923  10.3285 < 2.2e-16 ***
## aar_f2013        2869.297    221.858  12.9330 < 2.2e-16 ***
## aar_f2014        2863.224    231.248  12.3816 < 2.2e-16 ***
## aar_f2015        3525.223    251.782  14.0011 < 2.2e-16 ***
## aar_f2016        4274.990    274.021  15.6010 < 2.2e-16 ***
## aar_f2017        5146.326    299.039  17.2095 < 2.2e-16 ***
## Total_ya_p        582.436     46.559  12.5096 < 2.2e-16 ***
## inc_k1           -376.989     28.318 -13.3128 < 2.2e-16 ***
## inc_k5            194.354     23.395   8.3074 < 2.2e-16 ***
## uni_k_mf          -82.023     38.935  -2.1067 0.0352645 *
## uni_l_mf         1206.857     73.917  16.3272 < 2.2e-16 ***
## Trade_pc_100K     871.993    246.774   3.5336 0.0004187 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

iv.

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

v.

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

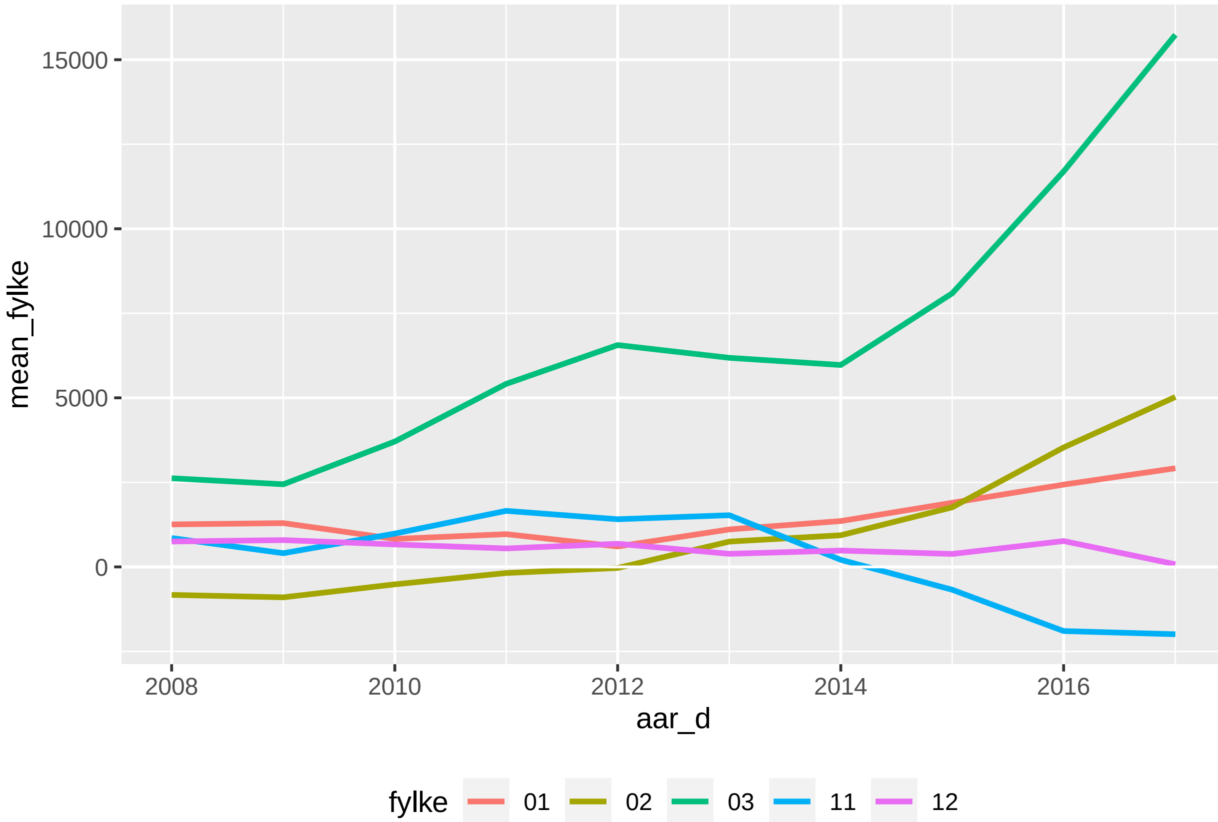
vi.

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

vii-x

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

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



Dummy: Fylke og år

i & ii.

```
mod2 <- 'pm2 ~ aar_f*fnr + Total_ya_p + inc_k1 + inc_k5 + uni_k_mf + uni_l_mf + Trade_pc_100K'
lm2 <- lm(mod2, data = pm2)
summary(lm2)
```

```
##
## Call:
## lm(formula = mod2, data = pm2)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -8546  -1191     32   1198   8328
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -21200.688   2521.645  -8.407  < 2e-16 ***
## aar_f2009            94.009    744.240   0.126 0.899496
## aar_f2010           417.129    744.379   0.560 0.575290
## aar_f2011          1280.914    744.731   1.720 0.085597 .
## aar_f2012          1455.525    745.679   1.952 0.051088 .
## aar_f2013          2479.533    746.367   3.322 0.000910 ***
## aar_f2014          2795.831    747.254   3.741 0.000188 ***
## aar_f2015          3987.973    748.109   5.331 1.09e-07 ***
```

```
## aar_f2016          5264.965    749.169   7.028 2.89e-12 ***
## aar_f2017          6618.572    749.430   8.831  < 2e-16 ***
## fnr02             -1482.789    702.970  -2.109 0.035045 *
## fnr03              3248.234   2190.443   1.483 0.138260
## fnr04             -1049.219    774.264  -1.355 0.175537
## fnr05             -1937.388    758.293  -2.555 0.010696 *
## fnr06             -2172.731    772.094  -2.814 0.004941 **
## fnr07              -737.995   1080.348  -0.683 0.494620
## fnr08             -3213.279    878.620  -3.657 0.000262 ***
## fnr09             -1219.813    913.691  -1.335 0.182020
## fnr10              -281.375    852.265  -0.330 0.741323
## fnr11              -565.360    771.927  -0.732 0.464012
## fnr12              -903.071    742.464  -1.216 0.224012
## fnr14             -3339.829   1182.013  -2.826 0.004768 **
## fnr15             -3619.198    715.832  -5.056 4.69e-07 ***
## fnr16             -1093.217    759.677  -1.439 0.150296
## fnr17             -2005.965    917.216  -2.187 0.028860 *
## fnr18             -1567.503    774.530  -2.024 0.043126 *
## fnr19             -2856.881   1326.142  -2.154 0.031341 *
## fnr20             -2656.315   1180.088  -2.251 0.024500 *
## Total_ya_p          511.787     36.100  14.177  < 2e-16 ***
## inc_k1             -243.050     27.007  -9.000  < 2e-16 ***
## inc_k5              251.645     22.916  10.981  < 2e-16 ***
## uni_k_mf            178.253     28.157   6.331 3.02e-10 ***
## uni_l_mf            732.442     42.235  17.342  < 2e-16 ***
## Trade_pc_100K      1067.760    190.885   5.594 2.54e-08 ***
## aar_f2009:fnr02     -40.505    978.026  -0.041 0.966969
## aar_f2010:fnr02     792.694    978.020   0.811 0.417747
## aar_f2011:fnr02     992.480    978.070   1.015 0.310359
## aar_f2012:fnr02    1565.161    978.102   1.600 0.109716
## aar_f2013:fnr02    1953.373    978.298   1.997 0.045996 *
## aar_f2014:fnr02    2019.269    978.649   2.063 0.039214 *
## aar_f2015:fnr02    2401.120    979.036   2.453 0.014273 *
## aar_f2016:fnr02    3656.344    979.067   3.735 0.000193 ***
## aar_f2017:fnr02    4707.776    979.374   4.807 1.65e-06 ***
## aar_f2009:fnr03      84.133   3068.211   0.027 0.978127
## aar_f2010:fnr03    2004.378   3068.354   0.653 0.513677
## aar_f2011:fnr03    3891.025   3068.768   1.268 0.204970
## aar_f2012:fnr03    5674.403   3069.281   1.849 0.064642 .
## aar_f2013:fnr03    5108.375   3070.149   1.664 0.096297 .
## aar_f2014:fnr03    4938.603   3071.105   1.608 0.107979
## aar_f2015:fnr03    6985.367   3073.112   2.273 0.023131 *
## aar_f2016:fnr03   10264.572   3074.072   3.339 0.000856 ***
## aar_f2017:fnr03   13986.613   3075.071   4.548 5.74e-06 ***
## aar_f2009:fnr04    -330.219   1089.318  -0.303 0.761813
## aar_f2010:fnr04    -191.813   1089.355  -0.176 0.860250
## aar_f2011:fnr04    -775.700   1089.399  -0.712 0.476523
## aar_f2012:fnr04    -808.528   1089.510  -0.742 0.458115
## aar_f2013:fnr04   -1206.685   1089.615  -1.107 0.268240
## aar_f2014:fnr04   -1456.367   1089.708  -1.336 0.181550
## aar_f2015:fnr04   -1912.336   1089.754  -1.755 0.079446 .
## aar_f2016:fnr04   -2459.017   1089.893  -2.256 0.024169 *
## aar_f2017:fnr04   -3549.658   1089.920  -3.257 0.001146 **
## aar_f2009:fnr05     416.862   1069.758   0.390 0.696816
```

```
## aar_f2010:fnr05    655.342   1069.794   0.613 0.540221
## aar_f2011:fnr05    183.865   1069.834   0.172 0.863563
## aar_f2012:fnr05    820.104   1070.017   0.766 0.443507
## aar_f2013:fnr05   -198.536   1070.094  -0.186 0.852832
## aar_f2014:fnr05   -254.055   1070.253  -0.237 0.812388
## aar_f2015:fnr05  -1326.089   1070.254  -1.239 0.215480
## aar_f2016:fnr05  -2117.228   1070.338  -1.978 0.048059 *
## aar_f2017:fnr05  -2397.820   1070.176  -2.241 0.025165 *
## aar_f2009:fnr06   -163.759   1089.292  -0.150 0.880516
## aar_f2010:fnr06    189.332   1089.409   0.174 0.862046
## aar_f2011:fnr06     33.963   1089.394   0.031 0.975132
## aar_f2012:fnr06    800.976   1089.455   0.735 0.462302
## aar_f2013:fnr06    410.281   1089.375   0.377 0.706497
## aar_f2014:fnr06    571.152   1089.474   0.524 0.600167
## aar_f2015:fnr06     22.631   1089.626   0.021 0.983431
## aar_f2016:fnr06   -598.671   1089.701  -0.549 0.582801
## aar_f2017:fnr06     60.036   1089.704   0.055 0.956069
## aar_f2009:fnr07    134.353   1525.051   0.088 0.929808
## aar_f2010:fnr07    728.914   1525.112   0.478 0.632745
## aar_f2011:fnr07    275.017   1525.266   0.180 0.856930
## aar_f2012:fnr07   1047.940   1525.235   0.687 0.492122
## aar_f2013:fnr07    890.998   1525.236   0.584 0.559173
## aar_f2014:fnr07    582.123   1525.332   0.382 0.702772
## aar_f2015:fnr07    990.944   1525.354   0.650 0.515996
## aar_f2016:fnr07    447.813   1525.278   0.294 0.769099
## aar_f2017:fnr07    960.018   1525.236   0.629 0.529146
## aar_f2009:fnr08    329.317   1240.237   0.266 0.790631
## aar_f2010:fnr08   1281.636   1240.345   1.033 0.301597
## aar_f2011:fnr08    646.495   1240.336   0.521 0.602269
## aar_f2012:fnr08   1090.416   1240.413   0.879 0.379470
## aar_f2013:fnr08    575.599   1240.249   0.464 0.642628
## aar_f2014:fnr08    689.084   1240.251   0.556 0.578548
## aar_f2015:fnr08   -776.910   1240.290  -0.626 0.531130
## aar_f2016:fnr08  -1716.491   1240.468  -1.384 0.166595
## aar_f2017:fnr08  -2045.538   1240.415  -1.649 0.099294 .
## aar_f2009:fnr09    686.715   1288.922   0.533 0.594245
## aar_f2010:fnr09    986.486   1288.914   0.765 0.444149
## aar_f2011:fnr09    599.582   1288.944   0.465 0.641860
## aar_f2012:fnr09   1071.846   1289.011   0.832 0.405779
## aar_f2013:fnr09     64.585   1289.204   0.050 0.960050
## aar_f2014:fnr09   -186.541   1289.179  -0.145 0.884965
## aar_f2015:fnr09  -1242.730   1289.232  -0.964 0.335201
## aar_f2016:fnr09  -1987.219   1289.181  -1.541 0.123368
## aar_f2017:fnr09  -3223.036   1289.344  -2.500 0.012510 *
## aar_f2009:fnr10    231.288   1199.909   0.193 0.847172
## aar_f2010:fnr10    924.121   1199.916   0.770 0.441302
## aar_f2011:fnr10    168.648   1199.944   0.141 0.888243
## aar_f2012:fnr10    321.458   1200.216   0.268 0.788856
## aar_f2013:fnr10   -515.180   1200.200  -0.429 0.667793
## aar_f2014:fnr10   -674.319   1200.339  -0.562 0.574335
## aar_f2015:fnr10  -1492.749   1200.502  -1.243 0.213856
## aar_f2016:fnr10  -3090.918   1200.777  -2.574 0.010124 *
## aar_f2017:fnr10  -3807.142   1200.767  -3.171 0.001545 **
## aar_f2009:fnr11   -414.412   1069.772  -0.387 0.698515
```

```
## aar_f2010:fnr11    642.468   1069.866   0.601 0.548235
## aar_f2011:fnr11   1243.418   1070.024   1.162 0.245359
## aar_f2012:fnr11   1467.212   1070.665   1.370 0.170728
## aar_f2013:fnr11   1179.371   1071.062   1.101 0.270979
## aar_f2014:fnr11   -183.391   1071.523  -0.171 0.864124
## aar_f2015:fnr11  -1489.385   1072.451  -1.389 0.165063
## aar_f2016:fnr11  -3274.743   1072.946  -3.052 0.002303 **
## aar_f2017:fnr11  -3863.610   1073.185  -3.600 0.000326 ***
## aar_f2009:fnr12     21.853   1036.805   0.021 0.983186
## aar_f2010:fnr12    381.898   1036.801   0.368 0.712658
## aar_f2011:fnr12    165.379   1036.901   0.159 0.873297
## aar_f2012:fnr12    669.171   1037.128   0.645 0.518864
## aar_f2013:fnr12    -69.430   1037.183  -0.067 0.946636
## aar_f2014:fnr12   -147.825   1037.277  -0.143 0.886690
## aar_f2015:fnr12   -711.755   1037.476  -0.686 0.492767
## aar_f2016:fnr12   -901.775   1037.688  -0.869 0.384941
## aar_f2017:fnr12  -2046.447   1038.104  -1.971 0.048828 *
## aar_f2009:fnr14   -220.698   1663.985  -0.133 0.894498
## aar_f2010:fnr14    536.844   1663.957   0.323 0.747009
## aar_f2011:fnr14   1984.847   1664.012   1.193 0.233090
## aar_f2012:fnr14   1739.551   1664.177   1.045 0.296018
## aar_f2013:fnr14    208.353   1664.208   0.125 0.900381
## aar_f2014:fnr14    253.302   1664.812   0.152 0.879084
## aar_f2015:fnr14  -1695.187   1665.139  -1.018 0.308783
## aar_f2016:fnr14  -1552.417   1665.259  -0.932 0.351330
## aar_f2017:fnr14  -2074.192   1665.271  -1.246 0.213077
## aar_f2009:fnr15    205.720    998.429   0.206 0.836779
## aar_f2010:fnr15    548.008    998.671   0.549 0.583249
## aar_f2011:fnr15    463.880    998.884   0.464 0.642414
## aar_f2012:fnr15    463.860    999.265   0.464 0.642556
## aar_f2013:fnr15      7.994    999.213   0.008 0.993617
## aar_f2014:fnr15   -481.056    999.093  -0.481 0.630220
## aar_f2015:fnr15   -587.449    999.385  -0.588 0.556727
## aar_f2016:fnr15  -1872.887    999.582  -1.874 0.061126 .
## aar_f2017:fnr15  -2799.827    999.681  -2.801 0.005149 **
## aar_f2009:fnr16   -346.631   1069.772  -0.324 0.745955
## aar_f2010:fnr16   -237.962   1069.934  -0.222 0.824020
## aar_f2011:fnr16   -497.945   1069.952  -0.465 0.641705
## aar_f2012:fnr16    380.682   1070.437   0.356 0.722154
## aar_f2013:fnr16   -347.235   1070.757  -0.324 0.745754
## aar_f2014:fnr16   -229.362   1070.812  -0.214 0.830418
## aar_f2015:fnr16   -139.973   1070.880  -0.131 0.896019
## aar_f2016:fnr16  -1074.143   1070.970  -1.003 0.316004
## aar_f2017:fnr16  -2278.453   1070.923  -2.128 0.033499 *
## aar_f2009:fnr17   -288.412   1288.940  -0.224 0.822969
## aar_f2010:fnr17   -422.338   1289.001  -0.328 0.743214
## aar_f2011:fnr17    257.671   1289.086   0.200 0.841590
## aar_f2012:fnr17    637.493   1289.624   0.494 0.621133
## aar_f2013:fnr17    203.405   1289.762   0.158 0.874704
## aar_f2014:fnr17    -61.073   1289.824  -0.047 0.962239
## aar_f2015:fnr17   -867.834   1289.740  -0.673 0.501107
## aar_f2016:fnr17  -1612.215   1290.487  -1.249 0.211703
## aar_f2017:fnr17  -2761.733   1290.527  -2.140 0.032479 *
## aar_f2009:fnr18   -148.285   1089.412  -0.136 0.891744
```

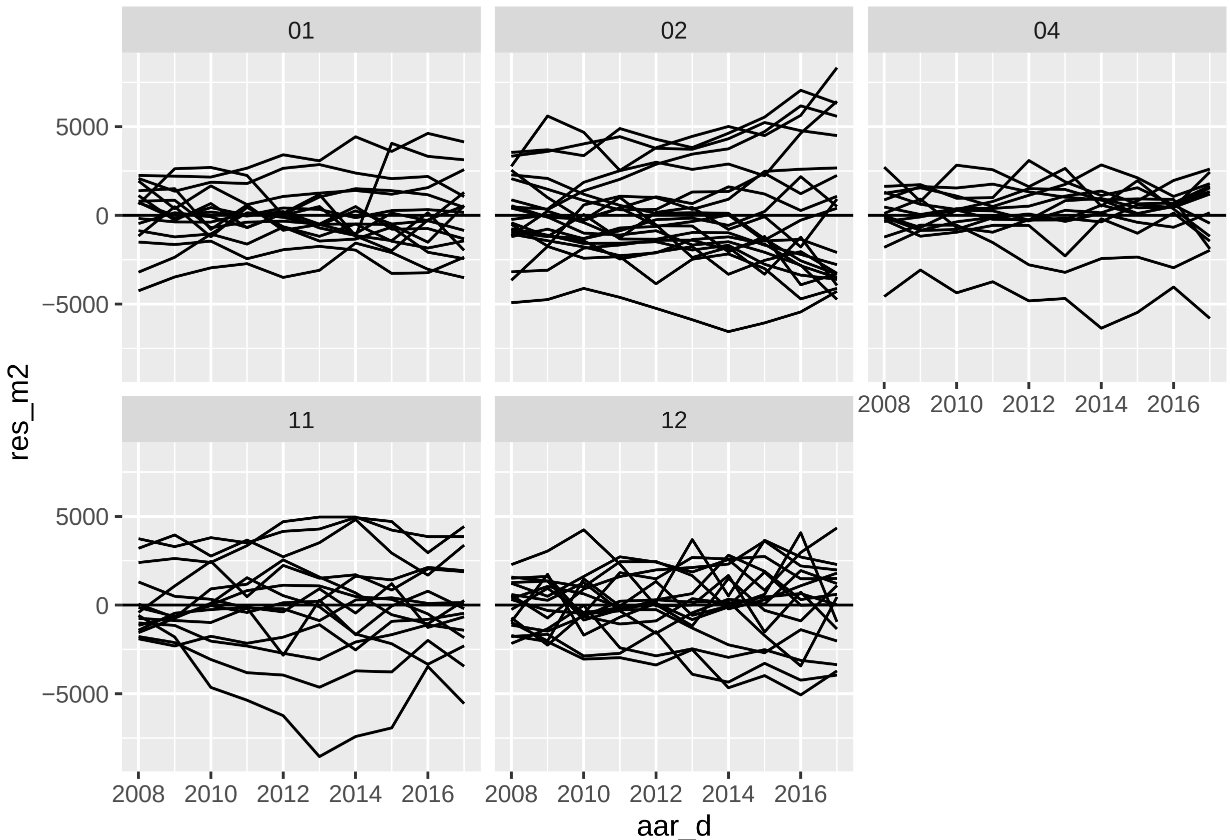
```
## aar_f2010:fnr18    402.939   1089.510   0.370 0.711545
## aar_f2011:fnr18    252.454   1089.674   0.232 0.816812
## aar_f2012:fnr18    482.679   1089.761   0.443 0.657871
## aar_f2013:fnr18    201.272   1090.026   0.185 0.853524
## aar_f2014:fnr18   -393.115   1090.258  -0.361 0.718459
## aar_f2015:fnr18   -439.127   1090.372  -0.403 0.687190
## aar_f2016:fnr18  -1361.291   1090.771  -1.248 0.212178
## aar_f2017:fnr18  -2661.041   1090.689  -2.440 0.014785 *
## aar_f2009:fnr19    453.061   1872.733   0.242 0.808864
## aar_f2010:fnr19    982.125   1872.779   0.524 0.600045
## aar_f2011:fnr19   -669.729   1872.850  -0.358 0.720682
## aar_f2012:fnr19    727.671   1872.902   0.389 0.697670
## aar_f2013:fnr19    278.261   1873.128   0.149 0.881921
## aar_f2014:fnr19   1688.165   1873.121   0.901 0.367563
## aar_f2015:fnr19    369.085   1873.412   0.197 0.843839
## aar_f2016:fnr19    906.286   1873.612   0.484 0.628646
## aar_f2017:fnr19   -716.410   1873.886  -0.382 0.702272
## aar_f2009:fnr20   -927.061   1664.164  -0.557 0.577542
## aar_f2010:fnr20   -547.207   1664.063  -0.329 0.742313
## aar_f2011:fnr20   -542.321   1664.293  -0.326 0.744568
## aar_f2012:fnr20   -378.342   1664.741  -0.227 0.820240
## aar_f2013:fnr20  -1110.163   1664.836  -0.667 0.504960
## aar_f2014:fnr20  -1563.827   1665.176  -0.939 0.347778
## aar_f2015:fnr20  -3266.760   1665.444  -1.961 0.049964 *
## aar_f2016:fnr20  -3169.910   1665.821  -1.903 0.057200 .
## aar_f2017:fnr20  -3922.387   1665.464  -2.355 0.018615 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2105 on 1944 degrees of freedom
## Multiple R-squared:  0.8953, Adjusted R-squared:  0.8848
## F-statistic: 85.21 on 195 and 1944 DF,  p-value: < 2.2e-16
```

iii.

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

iv.

```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)
```



0-linjen: Diskusjon

i.

Kvaliteten til modellen er preget av store variasjoner ettersom relevante variabler ikke er inkludert etter filtrering. Konklusjonen er dermed at modellen har forbedringspotensiale.

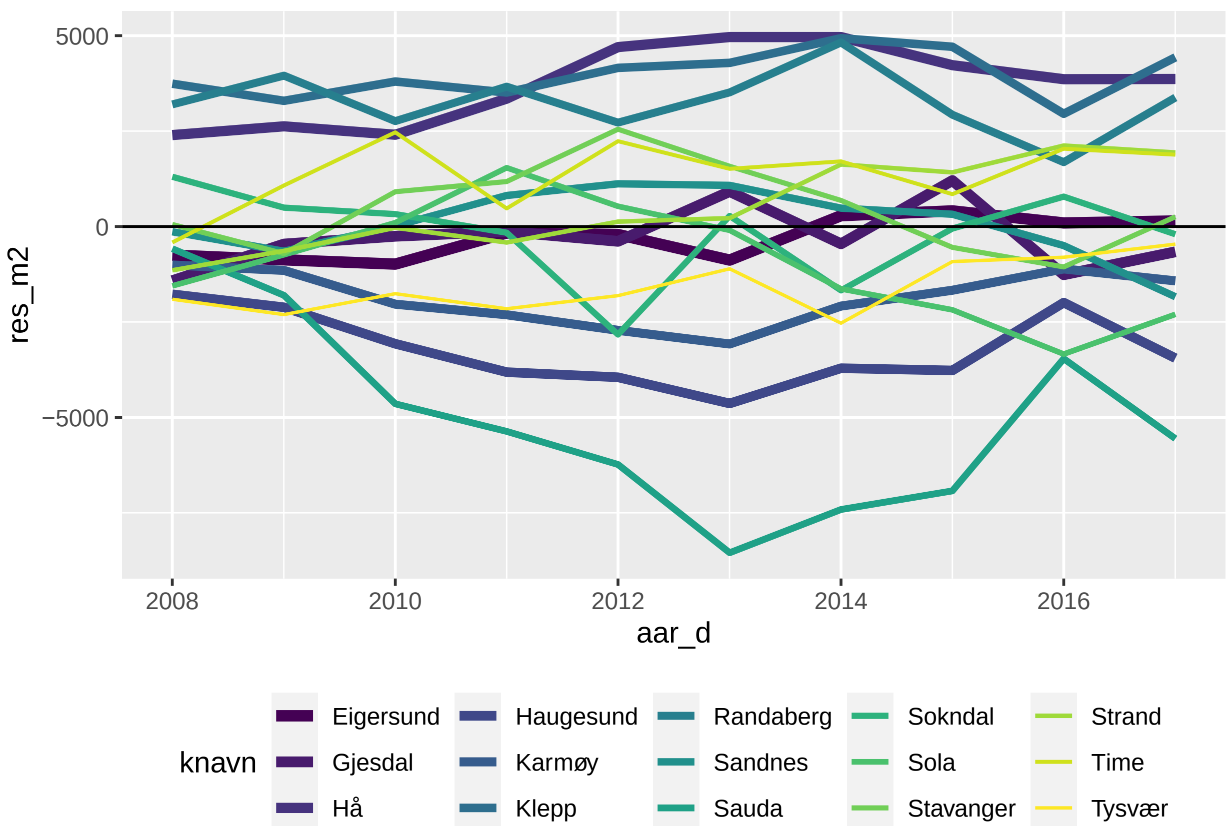
ii.

Som presentert i svaret over, mangler modellen sannelig viktige variabler.

iii.

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

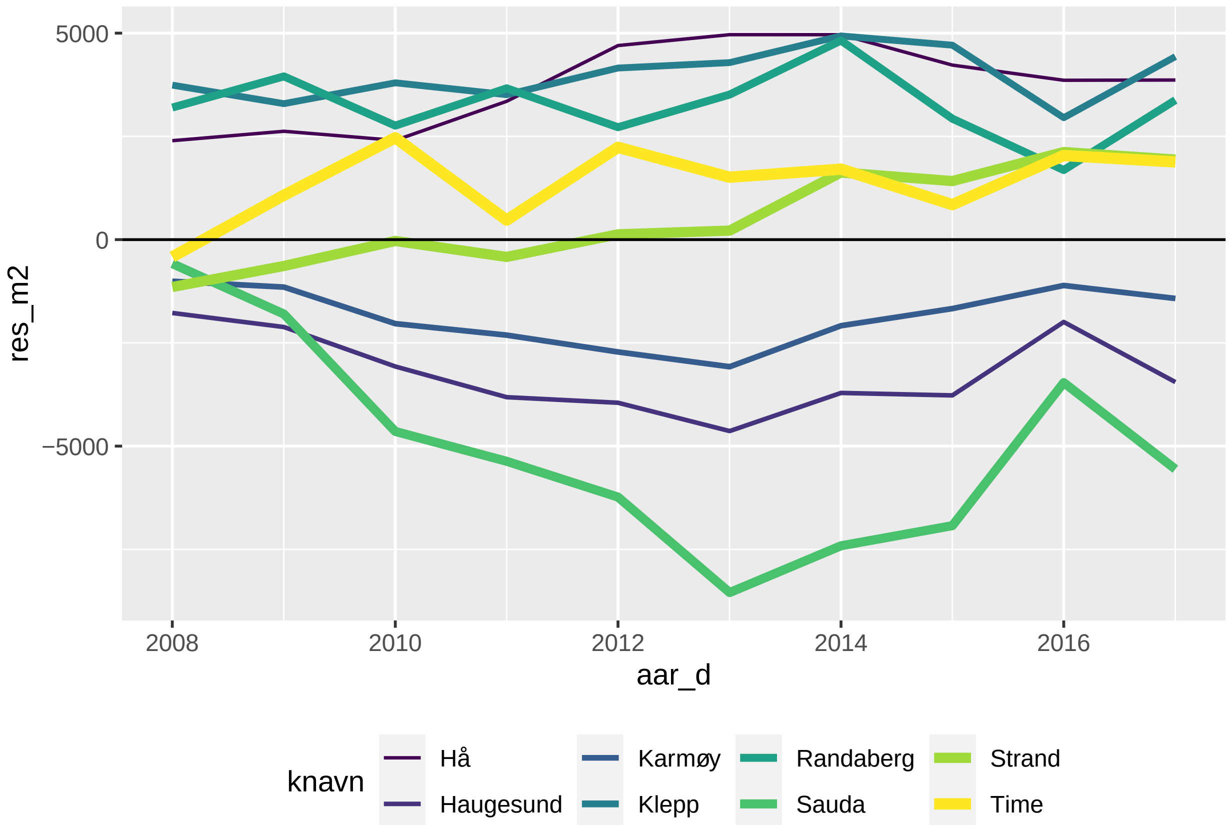
```
pm2 %>% filter(fylke %in% c("11")) %>%
  ggplot(mapping=aes(x = aar_d, y = res_m2)) +
  scale_color_viridis(discrete=TRUE, option = "D") +
  geom_line(aes(group = knavn, colour = knavn, size = knavn)) +
  scale_size_manual(values = c(seq(2.0, 0.5, by = -0.1))) +
  geom_hline(yintercept = 0) +
  theme(legend.position='bottom')
```



Figur

i.

```
pm2 %>%
  filter(knr %in% c("1106", "1119", "1120", "1121", "1127", "1130", "1135",  "1149")) %>%
  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(0.6, 2.0, by = 0.2))) +
  geom_hline(yintercept = 0) +
  theme(legend.position="bottom")
```



ii.

Modellen viser at Stavanger kommune er overvurdert, og ligger over null-linjen. Forflytter vi oss nærmere Haugesund, ser vi at flere kommuner er undervurdert inkludert Haugesund kommune.

> Feil! Det blir omvendt. Husk at det er residualene vi her betrakter. Positive residualer innebærer at modellen har undervurdert prisen, mens negative residualer innebærer at modellen overvurderer prisen. Vår modell ser altså ut til systematisk å undervurdere pm2 i kommunene rundt Stavanger, mens den overvurderer pm2 i kommunene på Haugalandet. Det er et nokså klart tegn på at modellen vår mangler en viktig variabel.

Modell for hvert år

i.

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

```
pm2_n <- pm2 %>%
  select(pm2, fnr, knr, aar, aar_f, aar_d, Menn_ya_p, Kvinner_ya_p,
         Total_ya_p, inc_k1, inc_k5, uni_k_mf, uni_l_mf, Trade_pc_100K) %>%
    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 [214 x 13]>
##  2 2009-01-01 <tibble [214 x 13]>
##  3 2010-01-01 <tibble [214 x 13]>
##  4 2011-01-01 <tibble [214 x 13]>
##  5 2012-01-01 <tibble [214 x 13]>
##  6 2013-01-01 <tibble [214 x 13]>
##  7 2014-01-01 <tibble [214 x 13]>
##  8 2015-01-01 <tibble [214 x 13]>
##  9 2016-01-01 <tibble [214 x 13]>
## 10 2017-01-01 <tibble [214 x 13]>
```

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

```
## # A tibble: 5 x 13
##     pm2 fnr   knr     aar aar_f Menn_ya_p Kvinner_ya_p Total_ya_p inc_k1 inc_k5
##   <dbl> <chr> <chr> <dbl> <chr>     <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>
```

pm2$n_data

i

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

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

i.

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

```
## # A tibble: 0 x 0
```

```
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    df
##    <date>     <lis> <lis>     <dbl>         <dbl> <dbl>     <dbl>    <dbl> <dbl>
##  1 2008-01-01 <tib~ <lm>      0.873         0.857 1701.      54.2 1.19e-71    24
##  2 2009-01-01 <tib~ <lm>      0.886         0.871 1614.      61.2 5.63e-76    24
##  3 2010-01-01 <tib~ <lm>      0.888         0.874 1743.      62.4 1.13e-76    24
##  4 2011-01-01 <tib~ <lm>      0.883         0.868 1925.      59.4 6.50e-75    24
##  5 2012-01-01 <tib~ <lm>      0.891         0.877 1953.      64.2 1.06e-77    24
##  6 2013-01-01 <tib~ <lm>      0.895         0.881 2026.      67.0 3.03e-79    24
##  7 2014-01-01 <tib~ <lm>      0.884         0.869 2149.      60.1 2.30e-75    24
##  8 2015-01-01 <tib~ <lm>      0.879         0.863 2361.      57.1 1.57e-73    24
##  9 2016-01-01 <tib~ <lm>      0.883         0.869 2467.      59.7 4.19e-75    24
## 10 2017-01-01 <tib~ <lm>      0.895         0.882 2614.      67.0 2.84e-79    24
## # ... with 6 more variables: logLik <dbl>, AIC <dbl>, BIC <dbl>,
## #   deviance <dbl>, df.residual <int>, nobs <int>
```

Coef_df

i.

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

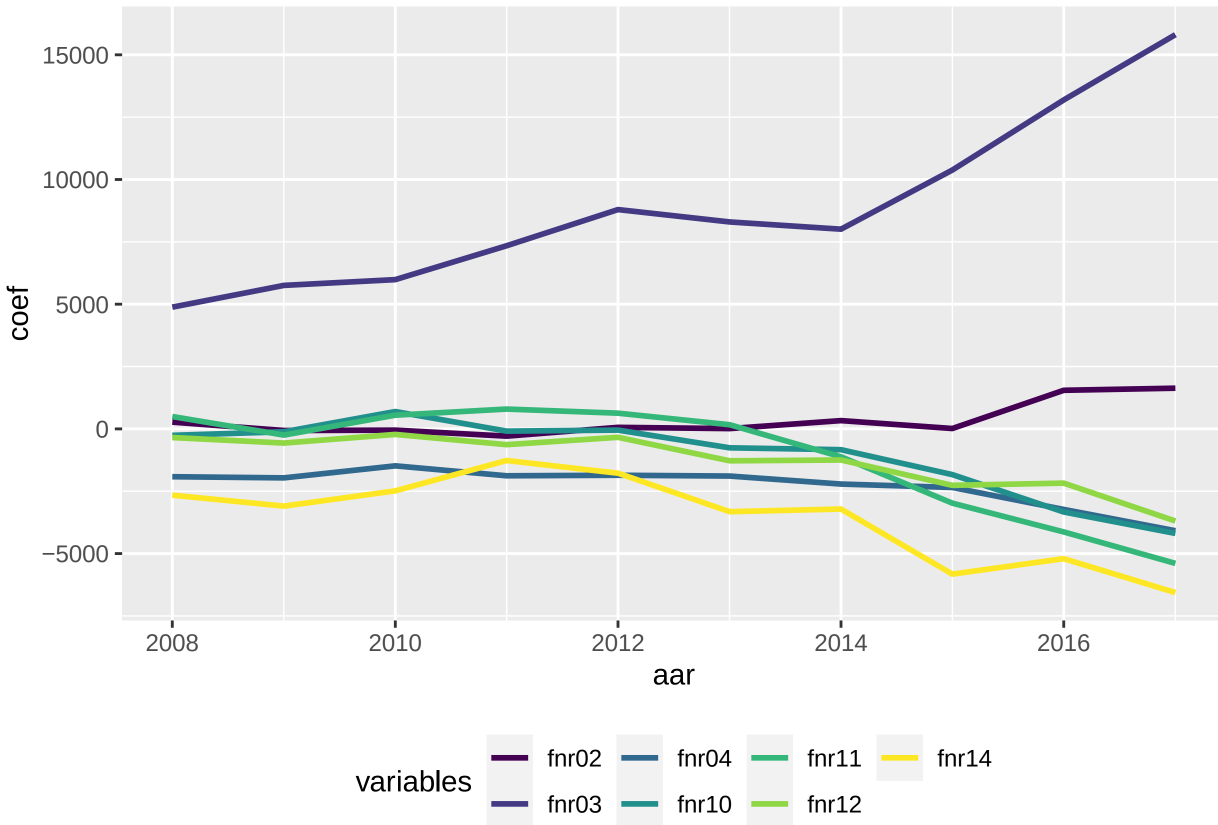
```
coef_df <- coef_df %>%
  mutate(
    aar = ymd(paste(2008:2017, "-01-01", sep = ""))
  ) %>%
  select(aar, everything())
```

ii.

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

iii.

```
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.

Grafen viser en prisstigning for fnr03, hvor fnr02 har hatt en stabil utvikling og opplevde prisstigning fra 2015 til 2017. De resterende fylkene viser en nedadgående prisutvikling i perioden rundt 2012 til 2017. I perioden før dette var samtlige fylker relativt stabile i sin utvikling.

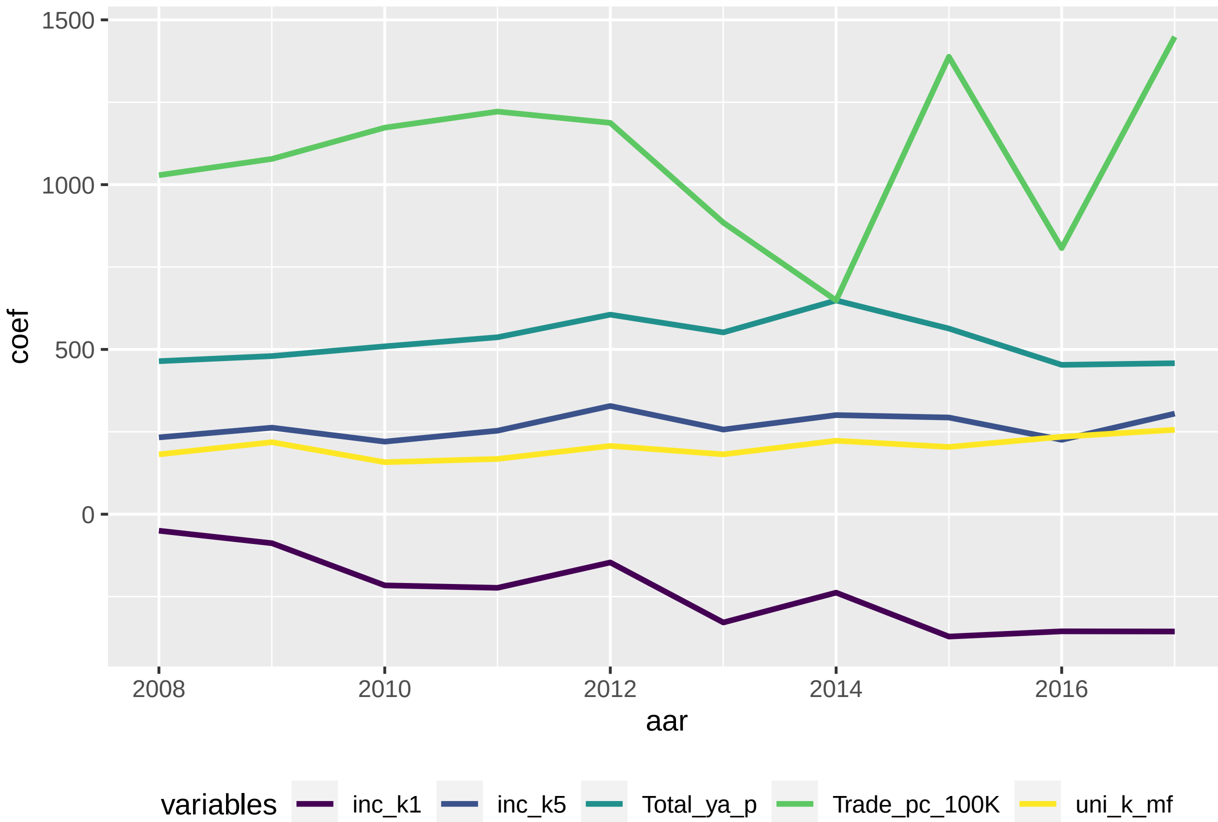
v.

2014 var preget av oljekrise, som igjen hadde en påvirking på prisutvikling. Prisen på olje gikk ned, og flere i oljebransjen mistet jobbene sine. Stavanger by ble dermed også preget ettersom det er en populær by kjent for olje og oljeindustri.

Modell; coef_df_long

i.

```
coef_df_long %>%
  select(aar, variables, coef) %>%
  filter(
    variables %in% c("Total_ya_p", "inc_k1", "inc_k5", "uni_k_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.

Grafen viser at **inc__k5** og **uni__k__mf** har vært mest stabile over tid. **Inc__k1** viser en relativt stabil variasjon, hvor årene 2012 til 2015 viser størst variasjon. **Total__ya__p** viser stabilitet frem til 2012 hvor grafen viser et lite oppsving med samme scenario i 2014, før grafen igjen stabiliseres i 2016. Avlutningsvis, viser **Trade__pc__100K** størst variasjon preget av både oppsving og nedsving fra 2012 til 2017.