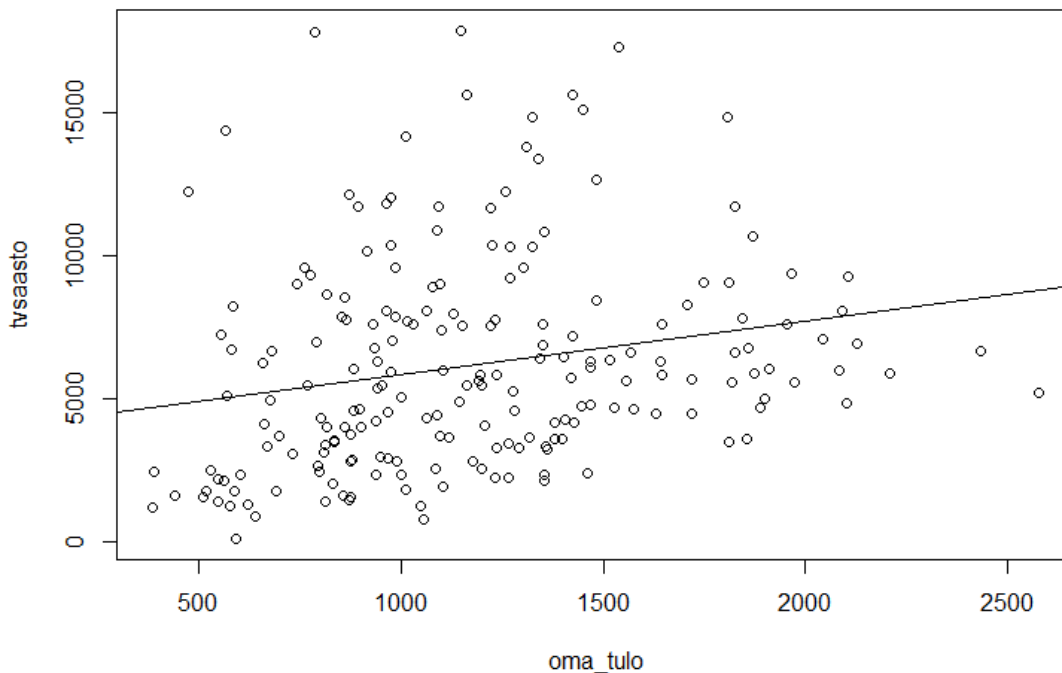


# Tilastollisten mallien peruskurssi – H4

2.

Sirontakuvio omien nettotulojen ja tutkimusvuoden säästöjen korrelaatiosta



## Pearson's product-moment correlation

```
data: oma_tulo and tvsaasto
t = 3.2139, df = 199, p-value = 0.001528
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.08639321 0.34977404
sample estimates:
      cor
0.222132
```

Pearsonin suoraviivainen  
korrelaatiokerroin = 0.222

## Spearman's rank correlation rho

```
data: oma_tulo and tvsaasto
S = 937576, p-value = 9.141e-06
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.3072442
```

Spearmanin monotoninen  
korrelaatiokerroin = 0.307

Pearsonin korrelaatio osoittaa heikkoa suoraviivaista riippuvuutta ja Spearmanin korrelaatio osoittaa kohtalaista positiivista riippuvuutta, joten muuttujien välinen riippuvuus ei ole kovin suoraviivaista.

3.

```
> lm.saasto <- lm(tvsaasto~oma_tulo)
> summary(lm.saasto)

Call:
lm(formula = tvsaasto ~ oma_tulo)

Residuals:
    Min       1Q   Median       3Q      Max
-5165  -2764  -1026   1812  12381

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3969.8557   731.4533   5.427 1.65e-07 ***
oma_tulo      1.8758     0.5837   3.214 0.00153 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

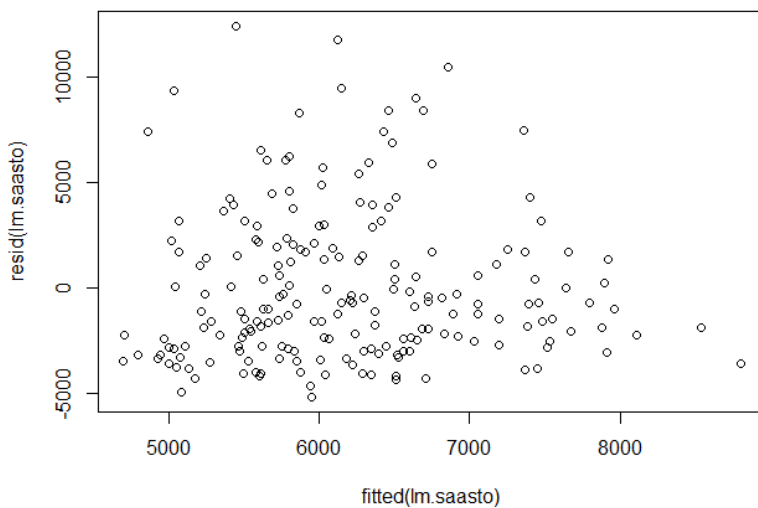
Residual standard error: 3601 on 199 degrees of freedom
Multiple R-squared:  0.04934,    Adjusted R-squared:  0.04457
F-statistic: 10.33 on 1 and 199 DF,  p-value: 0.001528
```

Mallin selitysaste = 0.049

Malli on tilastollisesti  
merkitsevä (p=0.002)

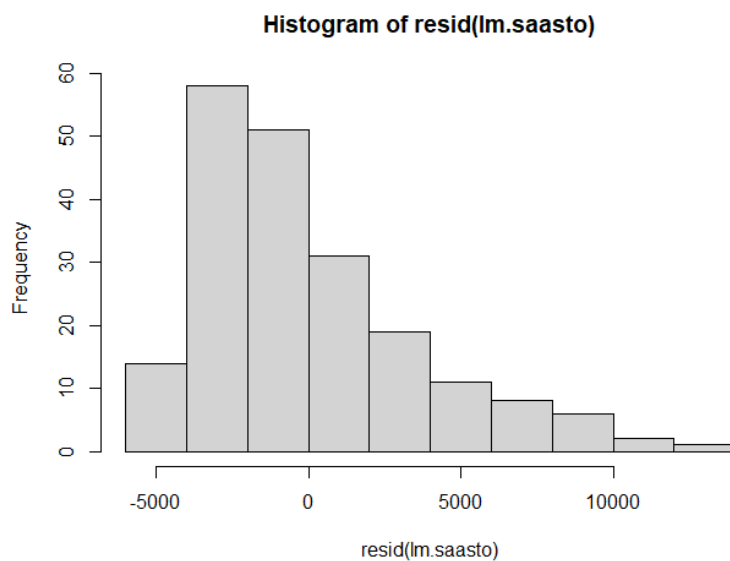
Selittäjä on tilastollisesti  
merkitsevä (p=0.002)

Regressiosuoran  
kulmakerroin = 1.88



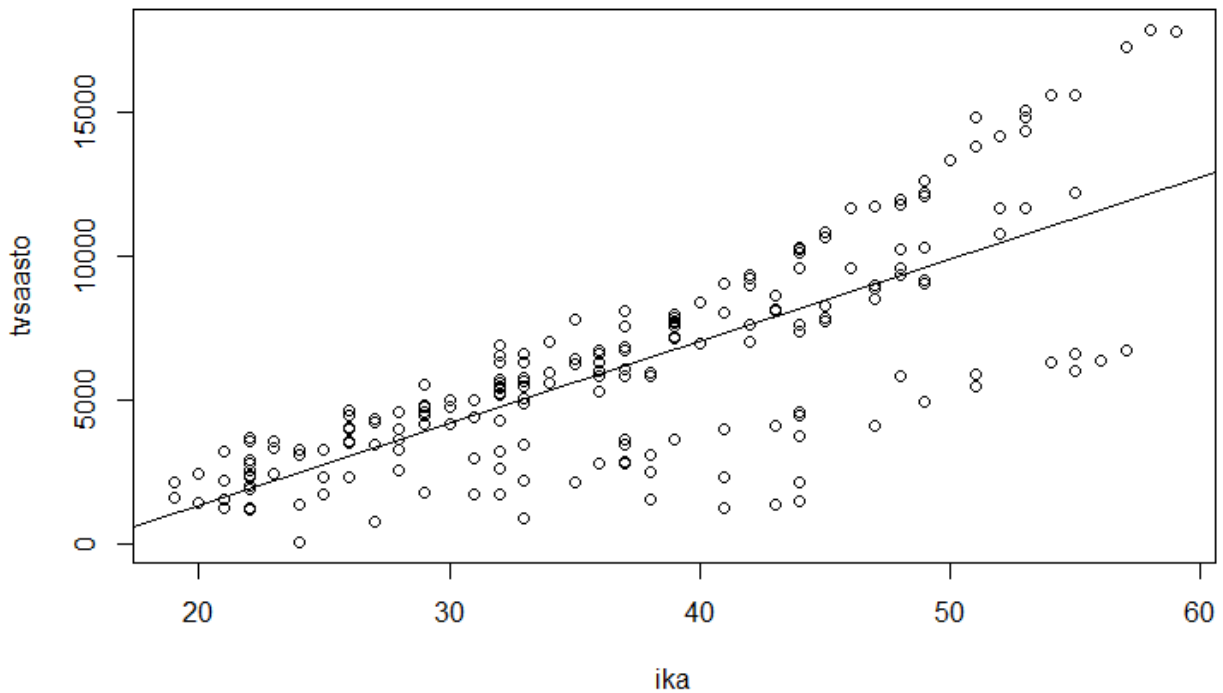
Jäännösten jakauma on  
symmetrinen

Jäännösten hajonta ei ole vakio



4.

län ja tutkimusvuoden säästöjen sirontakuvio



```
Pearson's product-moment correlation
data:  ika and tvsaasto
t = 17.13, df = 199, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.709328 0.822471
sample estimates:
      cor 
0.7719432
```

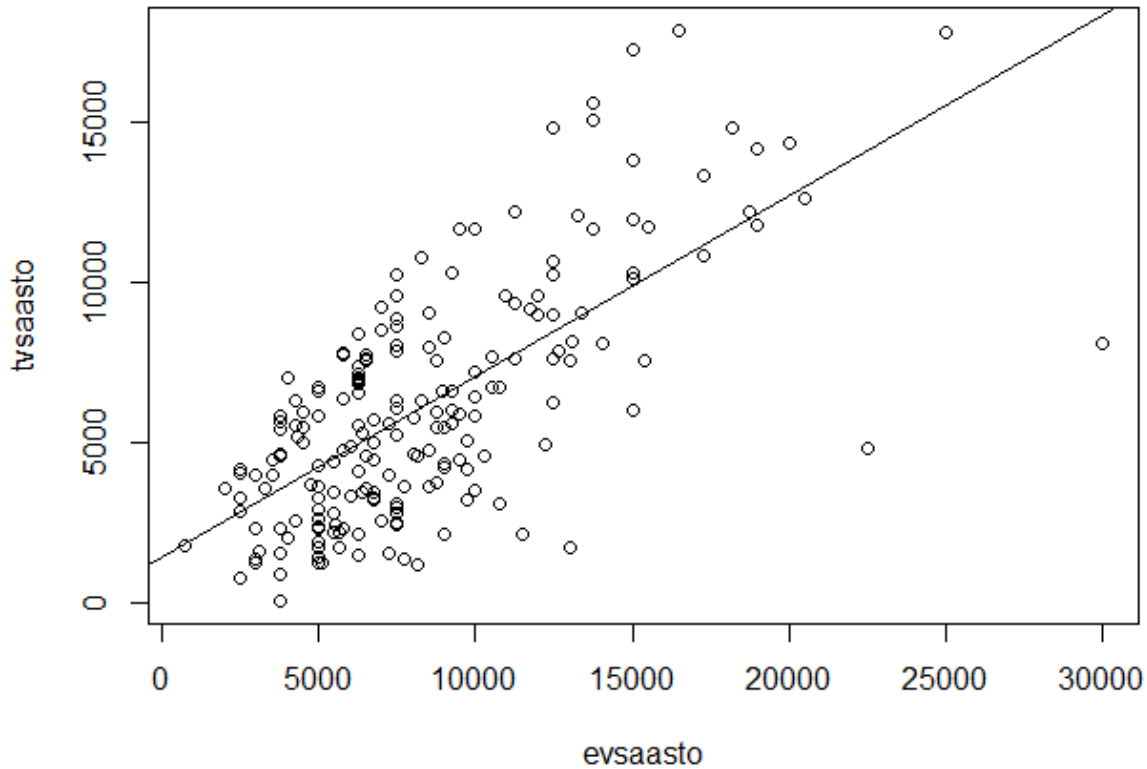
Pearsonin suoraviivainen  
korrelaatiokerroin = 0.772

```
Spearman's rank correlation rho
data:  ika and tvsaasto
S = 322338, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho 
0.7618311
```

Spearmanin monotoninen  
korrelaatiokerroin = 0.762

Suoraviivainen riippuvuus vallitsee

Edellisen vuoden säästöjen ja tutkimusvuoden säästöjen sirontakuvio



#### Pearson's product-moment correlation

```
data: evsaasto and tvsaasto
t = 13.314, df = 199, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6054975 0.7532212
sample estimates:
      cor
0.6863756
```

Pearsonin suoraviivainen  
korrelaatiokerroin = 0.686

#### Spearman's rank correlation rho

```
data: evsaasto and tvsaasto
S = 481791, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.6440141
```

Spearmanin monotoninen  
korrelaatiokerroin = 0.644

Suoraviivainen riippuvuus vallitsee

5.

Regressiomalli: Iän ja edellisen vuoden säästöjen vaikutus tutkimusvuoden säästöihin

```
Call:
lm(formula = tvsaasto ~ ika + evsaasto)

Residuals:
    Min       1Q   Median       3Q      Max
-5825.7  -880.4   377.9  1338.1  4960.2

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -4039.1324    562.4006  -7.182 1.36e-11 ***
ika          208.0552     17.8097   11.682 < 2e-16 ***
evsaasto      0.3004      0.0396    7.587 1.25e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2072 on 198 degrees of freedom
Multiple R-squared:  0.6869,    Adjusted R-squared:  0.6837
F-statistic: 217.2 on 2 and 198 DF,  p-value: < 2.2e-16
```

Mallin selitysaste = 0.687

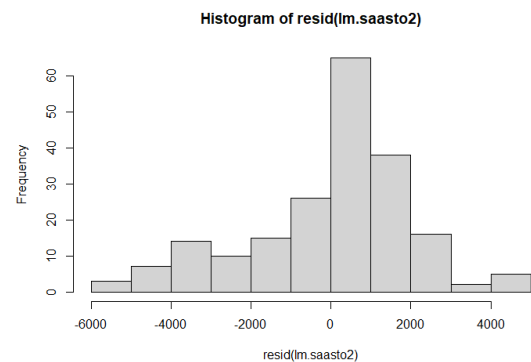
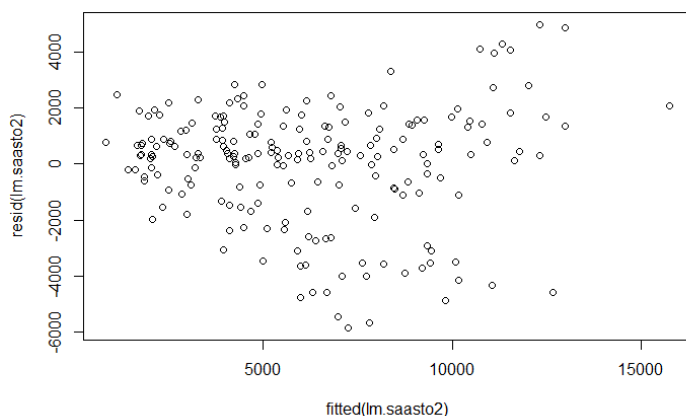
Mallin vertailukelpoinen selitysaste = 0.684

Malli on tilastollisesti merkitsevä,  $p < 0.001$

Ikä on tilastollisesti merkitsevä selittäjä,  $p < 0.001$

Edellisen vuoden säästöt on tilastollisesti merkitsevä selittäjä,  $p < 0.001$

Regressioyhtälö:  $\text{tvsaasto(€)} = -4039\text{€} + 208\text{€} \cdot \text{ika} + 0.3 \cdot \text{evsaasto(€)}$



Jäännösten jakauma ei ole symmetrinen

Jäännösten hajonta ei ole vakio

6.

Asunnonomistajien t-testit tutkimusvuoden ja edellisen vuoden säästöille

```
welch Two Sample t-test

data: tvsaasto by omistusa
t = -1.1821, df = 148.72, p-value = 0.239
alternative hypothesis: true difference in means between group ei and group on is not equal to 0
95 percent confidence interval:
 -1733.5455  435.7904
sample estimates:
mean in group ei mean in group on
      5899.922      6548.800
```

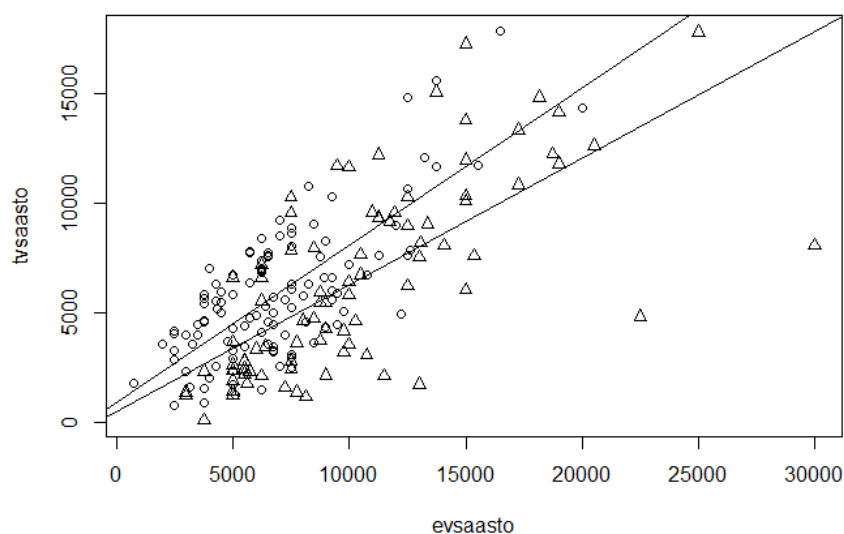
Havaittu ero ei ole tilastollisesti merkitsevä,  $p=0.239$

```
welch Two Sample t-test

data: evsaasto by omistusa
t = -5.533, df = 133.74, p-value = 1.598e-07
alternative hypothesis: true difference in means between group ei and group on is not equal to 0
95 percent confidence interval:
 -4751.636 -2249.122
sample estimates:
mean in group ei mean in group on
      6940.103      10440.482
```

Havaittu ero on tilastollisesti merkitsevä,  $p<0.001$

7.



Jyrkempi regressiosuora on ei-omistusasuntoa omaavien ja loivempi omistusasunnon omaavien.

8.

Ei omistusasuntoa:

```
Pearson's product-moment correlation

data:  evsaasto and tvsaasto
t = 11.825, df = 114, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6475339 0.8143454
sample estimates:
      cor 
0.7422261
```

Pearsonin suoraviivainen  
korrelaatiokerroin =  
0.742

```
Spearman's rank correlation rho

data:  evsaasto and tvsaasto
S = 102964, p-value = 6.947e-13
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho 
0.6041823
```

Spearmanin  
monotoninen  
korrelaatiokerroin =  
0.604

On suoraviivainen

Omistusasunto:

```
Pearson's product-moment correlation

data:  evsaasto and tvsaasto
t = 8.7407, df = 83, p-value = 2.158e-13
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5621274 0.7890168
sample estimates:
      cor 
0.6923106
```

Pearsonin suoraviivainen  
korrelaatiokerroin =  
0.692

```
Spearman's rank correlation rho

data:  evsaasto and tvsaasto
S = 26462, p-value = 4.857e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho 
0.7414326
```

Spearmanin  
monotoninen  
korrelaatiokerroin =  
0.741

Kelpaa myös suoraviivaiseksi

9.

```
Call:
lm(formula = tvsaasto ~ evsaasto * omistusa)

Residuals:
    Min       1Q   Median       3Q      Max
-9754.5 -1796.0  -243.8  1661.0  8104.6

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    906.36741   563.16289     1.609   0.109
evsaasto         0.71952    0.07339     9.804 <2e-16 ***
omistusaon     -388.05759   853.78691    -0.455   0.650
evsaasto:omistusaon -0.14192    0.09188    -1.545   0.124
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2588 on 197 degrees of freedom
Multiple R-squared:  0.5141,    Adjusted R-squared:  0.5067
F-statistic: 69.47 on 3 and 197 DF,  p-value: < 2.2e-16
```

Yhteisvaikutuksen  $p=0.124>0.05$ , eli kulmakertoimet eivät poikkea tilastollisesti merkittävästi

10.

```
Call:
lm(formula = tvsaasto ~ evsaasto + omistusa)

Residuals:
    Min       1Q   Median       3Q      Max
-10759.1 -1689.9  -187.1  1633.8  7870.4

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.535e+03   3.907e+02   3.928 0.000118 ***
evsaasto      6.290e-01   4.431e-02  14.196 < 2e-16 ***
omistusaon   -1.553e+03   4.019e+02  -3.863 0.000151 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2597 on 198 degrees of freedom
Multiple R-squared:  0.5082,    Adjusted R-squared:  0.5032
F-statistic: 102.3 on 2 and 198 DF,  p-value: < 2.2e-16
```

$p<0.001$ , joten edellisen vuoden säästöjä kovariaattina käytettäessä tutkimusvuoden säästöissä havaitaan tilastollisesti merkittävä ero

```
contrast estimate SE df t.ratio p.value
ei - on      1553 402 198   3.863  0.0002
```

Kovariaattikorjattujen keskiarvojen välillä on tilastollisesti merkittävä ero  $p<0.001$

Kovariaattikorjatut tutkimusvuoden säästöt ovat käänteiset havaituista keskiarvoista