

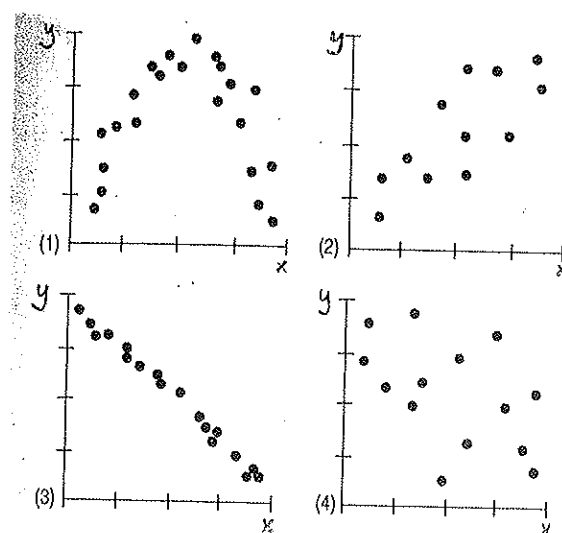
806109 TILASTOTIETEEN PERUSMENETELMÄT I
Taloustieteiden tiedekunnan opiskelijat
Harjoitus 6, viikko 42, syksy 2007

38. Aineistona YRITYS2006
(jatkoa tehtävään 35)

Laske elintarvikealan yrityksillä

- vuosien 2005 ja 2006 TE-arvosanojen välinen Pearsonin tulomomenttikorrelaatiokerroin,
- gearingin ja omavaraisuusasteen välinen Spearmannin järjestyskorrelaatiokerroin. Tulkitse tulokset.

39. Alla on neljä erilaista hajontakuviota.



Missä kuvioista muuttujien x ja y välillä

- ei ole lainkaan tai vain heikkoa riippuvuutta,
- negatiivista riippuvuutta,
- lineaarista riippuvuutta,
- kohtalaista tai melko voimakasta riippuvuutta,
- voimakasta riippuvuutta?

Hajontakuvioita vastaavista aineistoista lasketut Pearsonin tulomomenttikorrelaatiokertoimet ovat -0.923 , -0.487 , 0.006 ja 0.777 . Yhdistä kertoimet oikeisiin kuvioihin.

40. Eräessä suuressa yrityksessä haluttiin selvittää, onko työntekijän sukupuolella yhteyttä siihen, kannattaako yrityksen johdon ehdottamaa uudistusta vai ei. Yrityksen koko henkilökunnan osalta tulokset olivat seuraavat:

	Sukupuoli		Yht.
	Mies	Nainen	
Suhtautuminen kannattaa	60	65	125
uudistukseen ei kannata	170	150	320
Yht.	230	215	445

Koska yrityksen henkilökunta oli jaettavissa kahteen eri organisaatioon tuloksia tarkasteltiin myös erikseen näillä organisaatiotasolla (A ja B) ja saatiin seuraavat taulukot:

Organisaatiotaso A

	Sukupuoli		Yht.
	Mies	Nainen	
Suhtautuminen kannattaa	10	45	55
uudistukseen ei kannata	20	90	110
Yht.	30	135	165

Organisaatiotaso B

	Sukupuoli		Yht.
	Mies	Nainen	
Suhtautuminen kannattaa	50	20	70
uudistukseen ei kannata	150	60	210
Yht.	200	80	280

Tutki ristitulo-suhteen tai ehdollisten prosenttijakaumien avulla sukupuolen ja uudistukseen suhtautumisen välistä riippuvuutta.

- koko aineistossa,
- erikseen eri organisaatiotasolla.

Mitä päätelmiä voit näistä tuloksista tehdä?

41. (jatkoa tehtäviin 35 a) ja 38 a))

- Määrittää elintarvikealan yrityksille regressioyhtälö $y = a + bx$, missä y =TE-arvosana v. 2006 ja x =TE-arvosana v. 2005. Tulkitse saadun yhtälön kertoimet.
- Piirrä regressioyhtälön kuvaaja tehtävän 35a) hajontakuviioon.

- c) Määrää regressioyhtälön determinatiokerroin (= selitysaste) ja tulkitse se.
- d) Laske regressioyhtälön avulla ennustearvo d1) Valion, d2) Pauligin vuoden 2006 TE-arvosanaksi. Mitkä ovat vastaavat Valion ja Pauligin residuaalit?
42. Liitteessä on R:llä saatuja tuloksia aineistosta, joka sisältää seuraavat tiedot 50:stä Oulussa keväällä 2002 myytävänä olleesta rivitaloasunnosta: hintapyyntö (euroa), neliömäärä, ikä (vuosia, väh. yksi) ja etäisyys keskustasta (km).
- a) Laske muuttujien keskihajonnat.
- b) Laske kovarianssimatriisin ja korrelaatiomatriisin peitetty arvot.
- c) Tulkitse c1) hintapyyntön ja neliömäärän, c2) iän ja keskustasta etäisyyden välistä riippuvuutta.
43. (jatkoa tehtävään 42)
- a) Mikä mukana olevista muuttujista on paras selittävä muuttuja hintapyyntölle? Perustele vastauksesi. Määrää kyseinen regressioyhtälö ja tulkitse yhtälön kertoimet. Määrää myös regressioyhtälön determinatiokerroin.
- b) Tulkitse regressioanalyysin tulokset liitteen kohdasta 43 b) (regressioyhtälö, kertoimien tulkinta, determinatiokerroin).
- c) Laske sekä a)- että b)-kohdan yhtälöä käyttäen ennustearvo sellaisen rivitaloasunnon hinnaksi v. 2002, joka on ollut silloin 10 vuotta vanha, kooltaan 100 neliötä ja sijainnut 5 km:n päässä keskustasta.

HUOM!

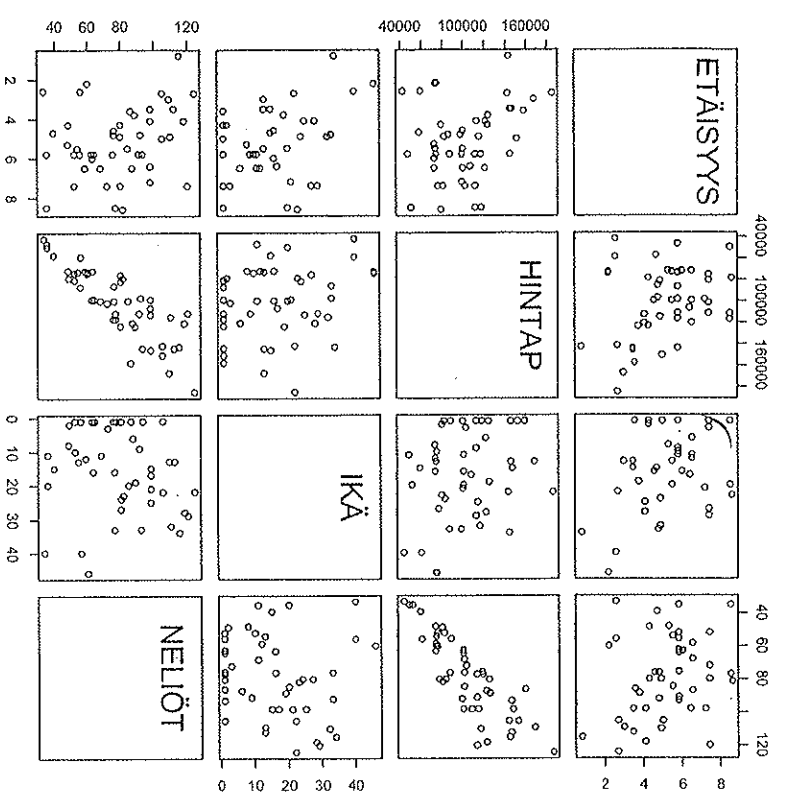
1. välikoe pe 19.10. 8.00-12.00 saleissa IT115 ja IT116.

Saliin IT115 ne, joiden sukunimi alkaa A-K.

Saliin IT116 ne, joiden sukunimi alkaa L-Ö.

Perjantain 19.10 harjoitukset 8-10 siirtyvät pidettäväksi torstaina 18.10.

10-12 salissa M304 (Eeva Vaaramo) ja 14-16 salissa M304 (Hanna-Mari Takkinen).



```
> cov(rivitalot[,c("ETÄISYYS", "HINTAP", "IKÄ", "NELIÖT")], use="complete.obs")
# Korrelationsmatriisi
      ETÄISYYS  HINTAP  IKÄ  NELIÖT
ETÄISYYS  1.0000000 -0.2845631 -0.4725442  0.8499149
HINTAP    -0.2845631  1.0000000 -0.2174021  0.1393607
IKÄ       -0.4725442 -0.2174021  1.0000000  0.1393607
NELIÖT    0.8499149  0.1393607  0.1393607  1.0000000
```

```
> cov(rivitalot[,c("ETÄISYYS", "HINTAP", "IKÄ", "NELIÖT")], use="complete.obs")
# Kovariansiamatriisi
      ETÄISYYS  HINTAP  IKÄ  NELIÖT
ETÄISYYS  3.310759 -17205.71 -11.06931 -9.833363
HINTAP    -17205.708816 1104232163.26 185.74041 701039.512000
IKÄ       -11.069306 185.74041 44.533959 44.533959
NELIÖT    -9.833363 701039.51 44.53396 616.131902
```

```
> runSummary(rivitalot[,c("ETÄISYYS", "HINTAP", "IKÄ", "NELIÖT")],
  statistics="mean")
# Keskiarvot
      mean  n
ETÄISYYS  5.216 50
HINTAP    100603.040 50
IKÄ       16.120 50
NELIÖT    76.756 50
```

436)

```
> malli2 <- lm(HINTAP~ETÄISYYS+IKÄ+NELIÖT, data=rivitalot)
> summary(malli2)
Call:
lm(formula = HINTAP ~ ETÄISYYS + IKÄ + NELIÖT, data = rivitalot)

Residuals:
    Min       1Q   Median       3Q      Max
-17898  -5766  -2740   5179  23055

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  65912.63    7732.13   8.525 5.05e-11 ***
ETÄISYYS    -6085.44    6972.01  -0.872 0.389
IKÄ         -1271.91    121.47  -10.471 9.24e-14 ***
NELIÖT      1132.62    56.89  19.909 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 9639 on 46 degrees of freedom
 Multiple R-squared: 0.921, Adjusted R-squared: 0.9159
 F-statistic: 178.8 on 3 and 46 DF, p-value: < 2.2e-16