# Näin tehtiin "**Kyselytutkimuksen mittarit ja menetelmät**" -kirja

**Kimmo Vehkalahti**

yliopistonlehtori, soveltavan tilastotieteen dosentti, VTT
Helsingin yliopisto, matematiikan ja tilastotieteen laitos
`http://www.helsinki.fi/people/Kimmo.Vehkalahti/suomeksi.html`

Tilastollisen tietojenkäsittelyn seminaari
28.10.2008

# Tämän esityksen **autoetnografisesta** luonteesta (1/3)

*"**Autoetnografia.** Tämä jokseenkin omalaatuinen tapa tehdä haastattelututkimusta, on yleistynyt erityisesti etnografisissa teksteissä. [...] Kyseessä on siis **tutkijan tunnustuksellinen itserefleksiivisyys**, jossa hän pyrkii yhdistämään haastattelijan ja haastateltavan roolit tilanteessa, jossa **tutkimuksen aihepiiri liittyy kokemuksellisesti myös tutkijan itsensä elämän rakentumiseen**." (korostukset KV)*

Saastamoinen, Mikko (2003). Tunnustaminen, refleksiivisyys ja representaatiot haastattelututkimuksessa.
`http://www.uku.fi/~msaastam/tunnustaminen.pdf`
Teoksessa Eskola J, Koski-Jännes A, Lamminluoto E, Saaranen A, Saastamoinen M, Valtanen K, (toim.) Tutkimusmenetelmällisiä reflektioita. Kuopio: Kuopion yliopisto, 2003. 234 s.

# Tämän esityksen **autoetnografisesta** luonteesta (2/3)

*"Ellis and Bochner (2000) advocate **autoethnography**, a form of writing that 'make[s] the **researcher's own experience a topic of investigation** in its own right' (p. 733) rather than seeming 'as if they're written from nowhere by nobody' (p. 734). Autoethnography is '**an autobiographical genre of writing**' [...]"*

*"**Some criticize this form of writing as sentimental, unscientific**, and the product of the excesses of postmodernism." (korostukset KV)*

Porter, Noah (2004). CMA Methodology: Autoethnography, `http://anthropology.usf.edu/cma/CMAmethodology-ae.htm`

Ellis, Carolyn and Bochner, Arthur P. (2000). Autoethnography, Personal Narrative, Reflexivity: Researcher As Subject, in N. Denzin and Y. Lincoln (eds.) The Handbook of Qualitative Research (2nd edition). Thousand Oaks, CA: Sage, pp. 733–768.

# Tämän esityksen **autoetnografisesta** luonteesta (3/3)

> *"Tähän kohtaan kuuluu myös **kuvallinen autoetnografia**, jossa tekijä itse luo kuvallisen aineiston esimerkiksi vuoden aikana ja sitten, ikäänkuin seuraavana projektina, alkaa lukea sitä yhdestä tai useasta lähtökohdasta, päämääränään liittää yhteen kuvallinen ja sanallinen käsitteellistäminen suhteessa johonkin ongelmaan, kysymykseen tai tehtävään, jolla saattaa olla myös teoreettista **tai ainakin historiallista ja intersubjektiivista relevanssia**." (korostukset KV)*

Varto, Juha (2005). Tarpeeseen luodut kuvat, `http://vartolog.blogs.fi/2005/09/18/tarpeeseen_luodut_kuvat~188925`

# Kirjan vaiheet lyhyesti

| | 2006 | 2007 | 2008 |
|---|---|---|---|
| **01** | | **kustannussopimus** | — |
| **02** | | mietintää... | keskustelua... |
| **03** | | pohdiskelua... | — |
| **04** | | suunnittelua... | — |
| **05** | | aloittelua... | — |
| **06** | | **kirjoitusviikko**[1] | **kirjoitusviikko**[4] |
| **07** | | **kirjoitusviikko**[2] | paljon kirjoittelua |
| **08** | | **kirjoitusviikko**[3] | PALJON kirjoittelua |
| **09** | | suunnittelua... | **PALJON** kirjoittelua |
| **10** | yllytystä... | kirjoittelua... | **painatus ja julkaisu** |
| **11** | idea ja päätös | kirjoittelua... | |
| **12** | yhteydenotot | **alustava käsikirjoitus** | |

[1] Benalmádena, Espanja
[2] Sastamala, Suomi
[3] Benalmádena, Espanja
[4] Sastamala, Suomi

# Esityksen vaiheet lyhyesti

| Sivut | Sisältö |
|---|---|
| 2–4 | Autoetnografinen johdatus |
| 5–6 | Kirjan ja esityksen vaiheet lyhyesti |
| 7–27 | Suunnittelusta kirjoitustyön alkuun |
| 28–44 | **Näkymiä ”konehuoneesta” eri vaiheissa** |
| 45–67 | Ensimmäinen versio ja puolen vuoden tauko |
| 68–99 | Lopullinen versio kuntoon ja painoon |
| 100–131 | **Näkymiä lopullisen version toimituskentistä** |
| 132–133 | Tietoa kirjasta netissä |

# 4.10.2006 ja 1.11.2006: Tilastokeskuksessa kouluttamassa

- 25.10.2006 jätin **dosenttihakemuksen** tiedekuntaan

Käsillä oleva materiaali on parhaillaan työstövaiheessa. Suurin osa on otettu suoraan vuonna 2002 Metsäntutkimuslaitoksen (METLA) kurssille tekemästäni monisteesta, jonka laitoin samana vuonna verkkoon saataville. Monisteesta on tullut varsin suosittu, ehkä siksi että näistä asioista on edelleen kirjoitettu verraten vähän suomeksi. Onkin ollut mukava huomata, että monet ovat löytäneet monisteeni verkosta ja hyödyntäneet sitä opinnäytetöissään.

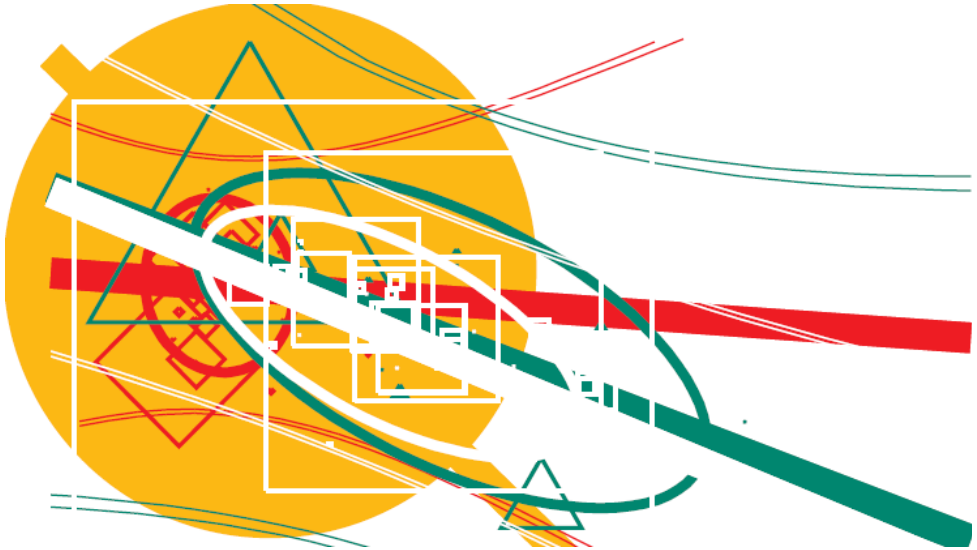
Monisteen mittaamista ja faktorianalyysia koskevan osan olen alustavasti uudistanut, mutta suurin osa (pääkomponenttianalyysista eteenpäin) on yhä entisessä muodossaan aivan sellaisenaan. Tarkoitukseni on monipuolistaa esimerkkejä ja lisätä tarkempia viittauksia eri ohjelmistoihin. Aiemman version esimerkkien painottuminen ekologisiin sovelluksiin juontaa juurensa METLAn ohella vuosina 1995–2001 pitämiini monimuuttujamenetelmien kursseihin Helsingin yliopiston maatalous-metsätieteellisessä tiedekunnassa. Sittemmin kiinnekohdikseni ovat muodostuneet enemmänkin yhteiskunta- ja käyttäytymistieteiden sovellukset, joita aion monisteen uudessa versiossa tuoda enemmän esille.

Olen kiitollinen tähän asti saamastani palautteesta ja pyrin ottamaan vinkit huomioon uutta versiota laatiessani. Tämänhetkisenä tavoitteena on saada uudistettu materiaali verkkoon vuonna 2007. Aikataulu voi kuitenkin muuttua, koska tämä on vain yksi projekti monien muiden joukossa, enkä voi aina asettaa sen prioriteettia niin korkealle kuin haluaisin.

Palautetta kurssilaisilta (anonyymisti):

- "Erittäin hyvä kouluttaja: ulosanti selkeä ja sopivasti huumoria. Asiantuntevuus todella vakuutti."
- "Hyvä, asiallinen, rauhallinen, selkeä ja pätevän oloinen luennoitsija. Kuuntelen mielelläni jossain muussakin yhteydessä."

# 14.11.2006: Ensimmäiset muistiinpanot



Pohdintaa Lahdessa 27.-28.10 ja 10.-11.11.2006 pidetyn kurssin ("Tutkimustulosten graafinen esittäminen") jälkeen:

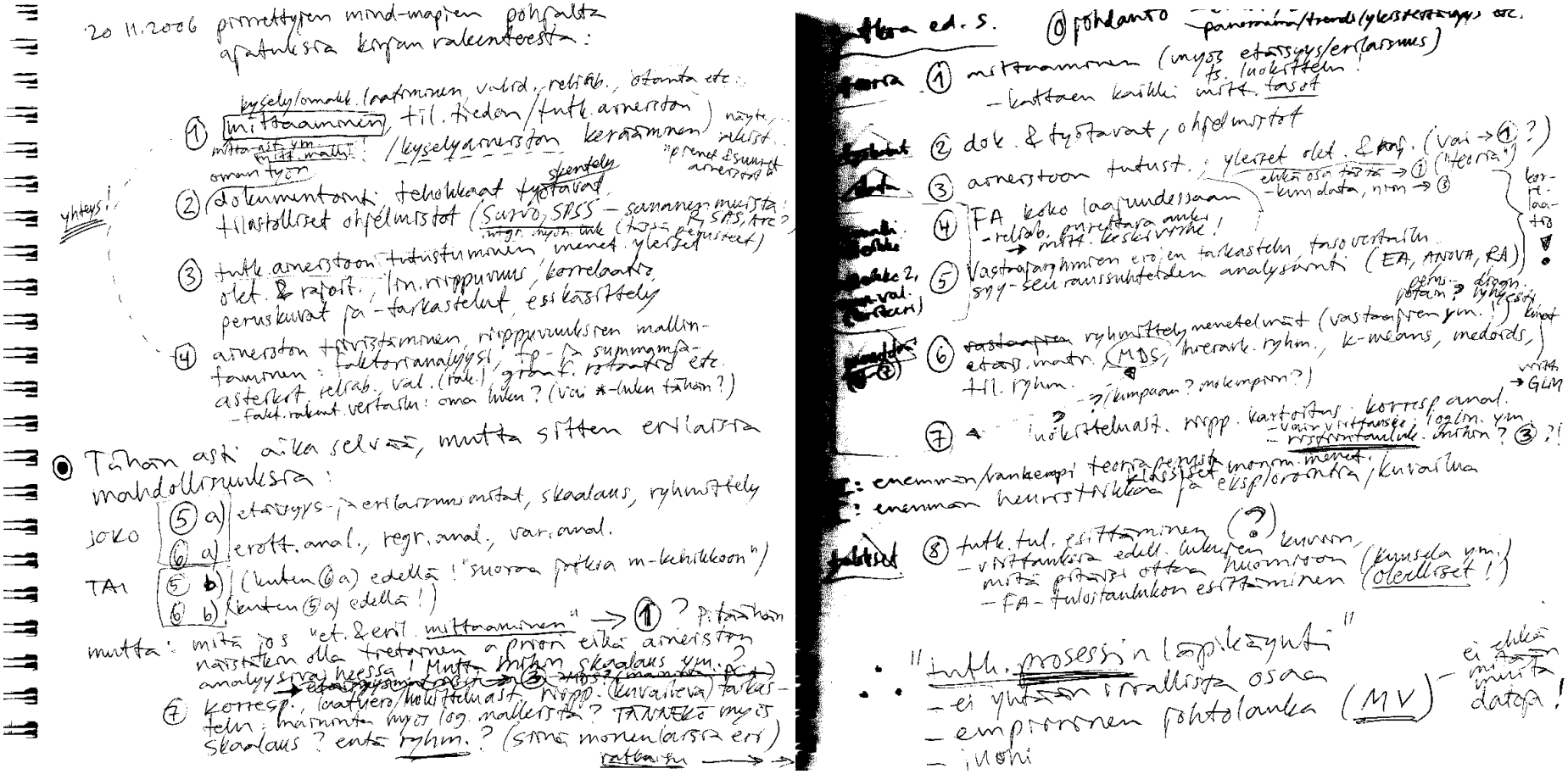
- "**Tutkija Y** kehui; mainitsi kirjoistaan kuulun **opettaja X**:n *'höpöttäneen taululle kirjoittaessaan selkä meihin päin jotain, mistä ei saanut selvää'* – **huoh**".
- "Alkoi syntyä ajatus, että **kirjoitankin monisteen sijasta kirjan**. Tulin ostaneeksi nuoremman Nummenmaan kirjan. Eipä ollut ihmeellinen; kumman paljon kaavaa siinäkin! **Minä voisin tehdä kirjan jossa ei olisi yhtään kaavoja.**"
- "**Maaritin data olisi oiva** esimerkkiaineisto myös kirjaan; siinä olisi silloin samalla vähintään 2 kärpästä, kun ohjaan työtään. Paino on oltava yhteiskunta- (ja käyttäytymis)tieteissä, tässä järjestyksessä."

# 20.11.2006: Muistikirjan perustaminen

Kymmon kirjan luonnostelua
marraskuusta 2006
lähtien...

- ajatuksena tarjot. ideaa WSOY:lle (vrt. Atinula 8-co.)
- tavoitteen = k. poittaa kirja ksi v. 2007
- fokus: kyselytutkim., yht.kunta - pa (siitä kyllä muistikin)
  käytt. tretut, Survo-SPSS, monim. menet., dokumentointi, graaf. esitt.
- Maarit Valtaran vk-aineisto! (kyvettävä vielä; syritty monotteen osalta)
- EI matem. kaavoja. (→ Y:ttäkset teoriaan osalta kyllä) kirjoihin + artikkeleihin!
  vaan tulkintaa, näkemystä.

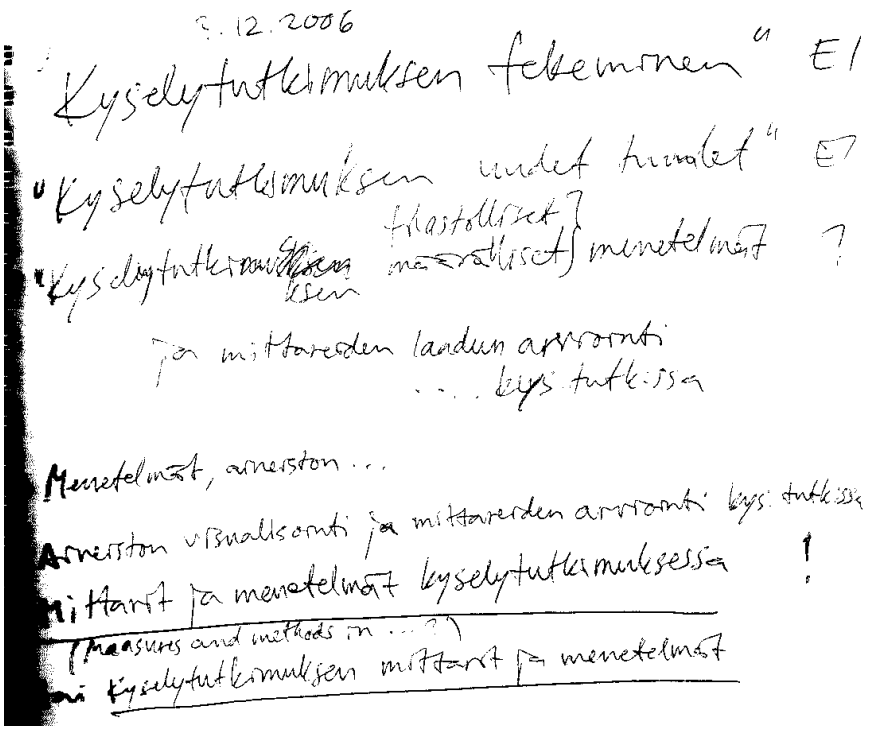
# 26.11.2006: Kirjan rakenteen ensimmäisiä hahmotelmia

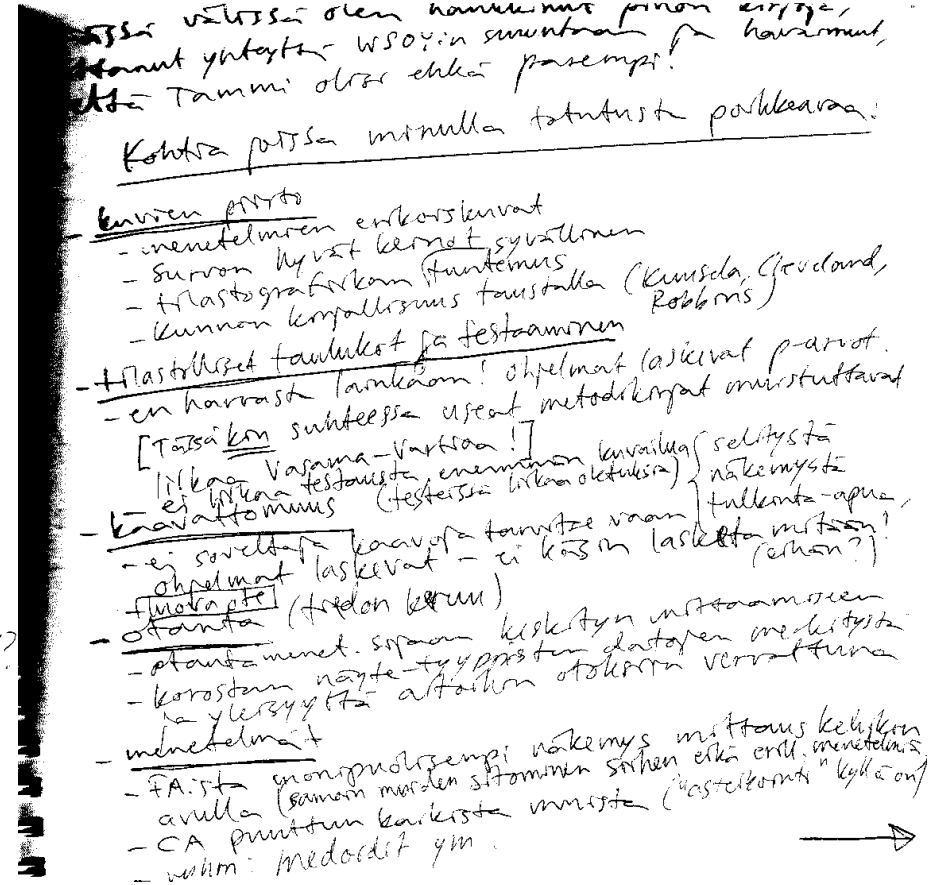


- Kirjan "punainen lanka" oli aika selvä alusta asti.
- Lähetin innoissani viestin asiasta WSOY:lle 9.12.2006.

# 15.12.2006: Kirjan nimen kehittely

- Nimi hahmottui sillä aikaa kun odotin WSOY:n vastausta...




- ...vastausta ei kuulunut, ja **Tammi** alkoi kiinnostaa enemmän.
- Tammi vastasi heti vuoden alussa; sovittiin palaveri 11.1.2007.

# 11.1.2007: Palaveri Tammessa – **kustannussopimus!**

**Kyselytutkimuksen mittarit ja menetelmät**
– rakenteen ja sisällön ideointia uutta oppikirjaa varten

Kimmo Vehkalahti

11.1.2007

- *"Aikataulu: kesään mennessä mahdollisimman pitkälle, syksyllä viimeistely, sitten painoon ja ulos alkuvuodesta 2008."*

  Kirjan **tärkeimmät teemat** ovat:

  - mittaamisen luotettavuuden arviointi ja vaikutukset
  - tilastollisen aineiston tiivistäminen ja analysointi
  - moniulotteisten ilmiöiden visualisointi ja mallintaminen
  - korkeatasoisen tilastografiikan vaatiminen ja laatiminen
  - oman työskentelyprosessin hallinta ja dokumentointi

- *"Tekniikasta: sanoin etten käytä Wordiä ja että teen LaTeXilla."*

# 14.1.2007: Pohdintoja painotuotteesta professoreille

**Jakelu:** Lauri T., Seppo M., Hannu N., Pentti S.

**Viran viimeiset 200 vrk**

Pohdintojani teille, akateemista uraani ohjanneille professoreille:

Pestini "pedagogisena yliopistonlehtorina" päättyy; jäljellä on nyt melko tasan 200 vrk. Yhtenä huippuhetkenä tämän 5-vuotiskauden ajalta mieleen jää **Magister Bonus**, jolla HYY palkitsi minut vuonna 2003.

Huolimatta siitä, että virkani on suunnattu opetuksen kehittämiseen, olen etenkin viime vuosina panostanut myös tutkimukseen; realiteetti kun on, etteivät pelkät (hyvätkään) opetusansiot riitä. Koska mikään ei ole niin varmaa kuin epävarma, **olen alkanut etsiä vihreitä oksia myös muualta**. Kaikesta huolimatta yliopisto tuntuu yhä – 17 vuoden jälkeenkin – houkuttelevammalta ympäristöltä kuin tutkimuslaitokset tai koulutusorganisaatiot, yritysmaailmasta puhumattakaan.

Jätin syksyllä tiedekuntaan hakemukseni *"soveltavan tilastotieteen dosentiksi"*, joten nyt ajattelin vaihteeksi panostaa hieman enemmän opetukseen. **Jo pitkään olen pohtinut oppikirjan tekoa, mutta koska sellaisen "arvo" julkaisuna on melko mitätön, olen malttanut mieleni.**

(jatkuu...)

# 14.1.2007: Pohdintoja painotuotteesta professoreille

(jatkuu...)

Jätin syksyllä tiedekuntaan hakemukseni *"soveltavan tilastotieteen dosentiksi"*, joten nyt ajattelin vaihteeksi panostaa hieman enemmän opetukseen. **Jo pitkään olen pohtinut oppikirjan tekoa, mutta koska sellaisen "arvo" julkaisuna on melko mitätön, olen malttanut mieleni.**

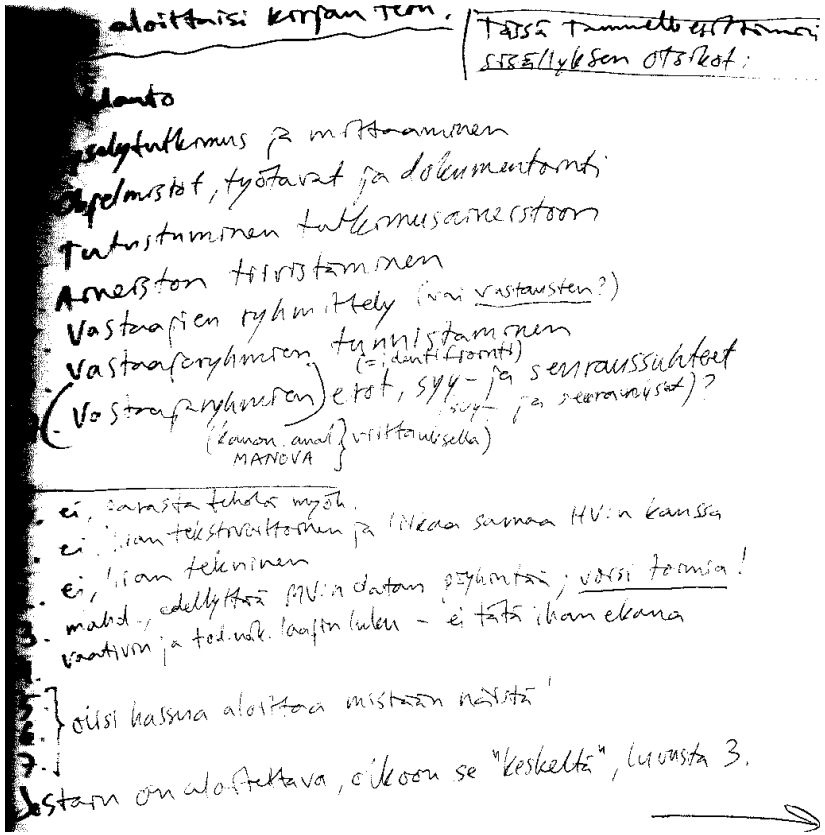
**Syksyn koulutuskeikat mm. Tilastokeskuksessa ja Palmeniassa sytyttivät kirja-ajatuksen uudelleen.** Aika tuntui olevan kypsä, ja eräät muutkin seikat puolsivat päätöstäni ryhtyä "oppikirjailijaksi". Niinpä myin ideani ja **tein viime viikolla sopimuksen kustantajan kanssa.** Kirjani, jonka alustava työnimi on "**Kyselytutkimuksen mittarit ja menetelmät**", ilmestynee vuoden päästä.

Aion käyttää suuren osan jäljellä olevasta pestistäni kirjan kirjoittamiseen. Se vastaa hyvin virkani tavoitteita, "kvantitatiivisen menetelmäopetuksen kehittämistä keskustakampuksella". Näin tästä erikoisesta virasta jää myös jotain konkreettista käteen.

Kiitos jo etukäteen tuesta ja kannustuksesta; hyvää lukukauden alkua!

– Kimmo

# 19.1.2007: ”Mistä kohtaa aloittaisin kirjan teon?”




- *”Jostain on aloitettava, olkoon se ’keskeltä’, luvusta 3.”* – ”Ohjelmistot, työtavat ja dokumentointi”
- Lopulta luvusta 3 tuli Liite A (täysin uudelleen kirjoitettuna)

# 19.1.-19.3.2007: Lahdessa kouluttamassa

Palautetta kollegalta, jolle kerroin kustannussopimuksestani:

- "Paitsi että juttu on mahtava sinun ja kaikkien meidän opettavaista tekstiäsi janoavien kannalta, se rohkaisee myös henk.koht. tavoitteissa, kun huomaa, että unelmaprojektejaan on mahdollista toteuttaa. Wautsi."

Sitten ideoimaani työpajaan takomaan viikonloppuisin:

Palautetta osallistujilta (sähköpostitse):

- "Kiitos myös aivan loistavasti ohjauksesta. Valehtelematta voin sanoa, että tämä prosessi on ollut paras tilastosisältö mihin olen koskaan osallistunut. Myös työpajamalli kannattaa säilyttää jatkossa."
- "Ja palautteena näin myös suoraan, että tämä kvanttityöpaja idea toimi erinomaisesti ja hyödytti ainakin omaa etenemistäni tosi paljon. Opetustapasi, jossa ei lähdetä niistä ainaisista regressiokaavoista, oli oikein hyvä."

# 14.-15.3.2007: Tilastokeskuksessa kouluttamassa

Käsillä oleva materiaali jäi tämän talven aikana lopullisesti työstövaiheeseen—hyvästä syystä: olen kirjoittamassa uutta oppikirjaa, enkä enää jatka tämän monisteen päivittämistä. Kirjani työnimi on **Kyselytutkimuksen mittarit ja menetelmät**, ja sen on tarkoitus ilmestyä alkuvuodesta 2008.

Valtaosa tässä olevasta materiaalista on suoraan vuonna 2002 Metlan (Metsäntutkimuslaitos) kurssille tekemästäni monisteesta, jonka laitoin samana vuonna verkkoon saataville. Monisteesta on tullut varsin suosittu, ehkä siksi että näistä asioista on edelleen kirjoitettu melko vähän suomeksi. Onkin ollut mukava huomata, että monet ovat löytäneet monisteeni verkosta ja hyödyntäneet sitä opinnäytetöissään.

Ehdin syksyllä 2006 aloittaa monisteen mittaamista ja faktorianalyysia koskevan osan uudistamisen, mutta loppuosa on yhä entisessä muodossaan aivan sellaisenaan. Tarkoitukseni on ollut monipuolistaa esimerkkejä ja lisätä tarkempia viittauksia eri ohjelmistoihin.

Palautetta kurssilaisilta (anonyymisti):

- "Kouluttaja varmaan paras mitä olen nähnyt – mihin tahansa aiheeseen liittyen!"
- "Kiva kouluttaja, selkeä ja mukava. Jaksoin hyvin seurata opetusta. Odotellaan kirjaa, tsemppiä sen kirjoittamiseen!"

Maaliskuun lopulla työmatkalle **Puolaan**...

# 21.-26.3.2007: Varsova – Poznan – Bedlewo – Varsova

$$\boldsymbol{\rho_u} = (\boldsymbol{A}'\boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{B}'\boldsymbol{A})_d \times (\boldsymbol{A}'\boldsymbol{\Sigma}\boldsymbol{A})_d \tag{5}$$

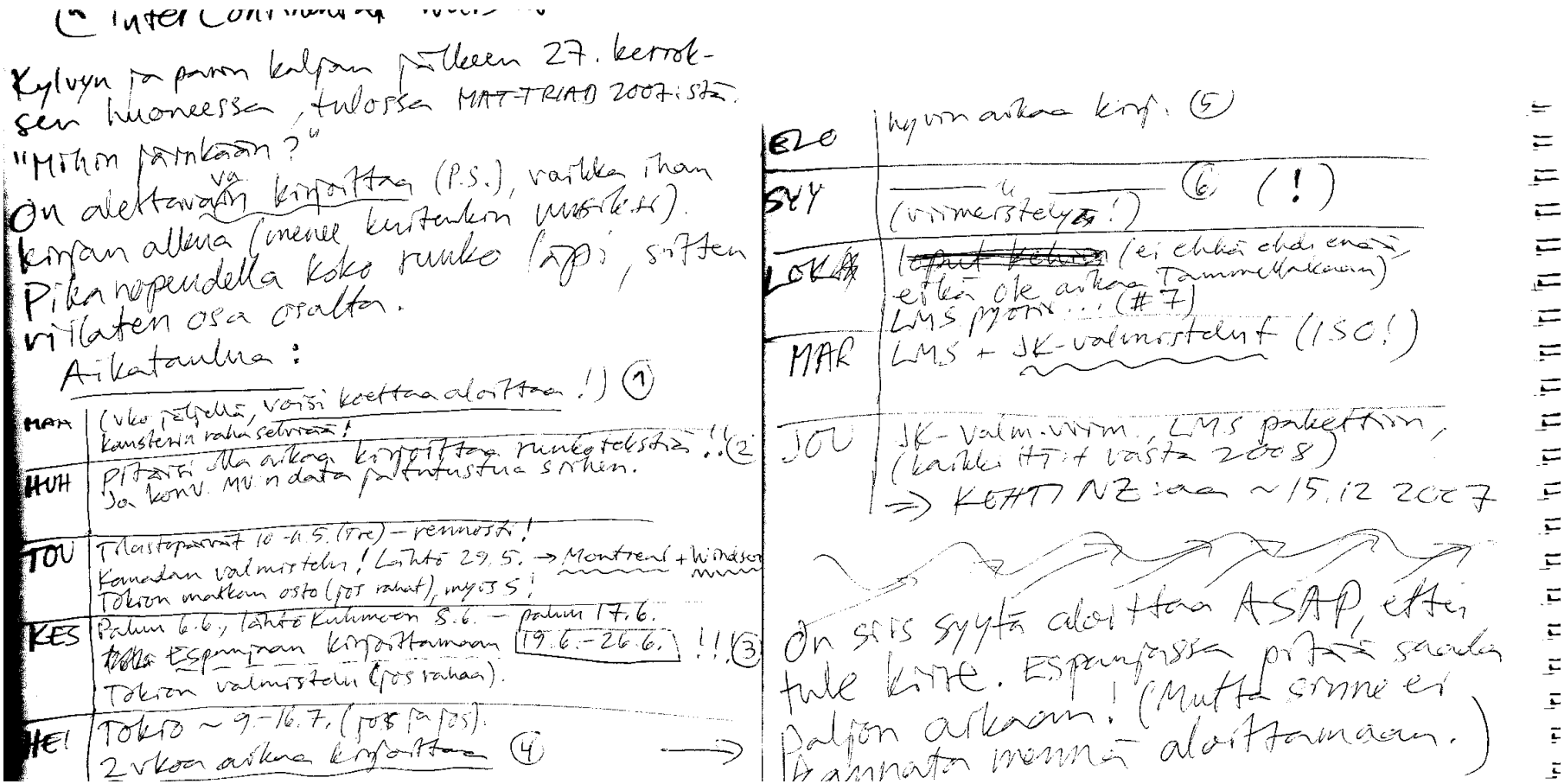
$$= \{\boldsymbol{I}_m + (\boldsymbol{A}'\boldsymbol{\Psi}\boldsymbol{A})_d \times [(\boldsymbol{A}'\boldsymbol{B}\boldsymbol{\Phi}\boldsymbol{B}'\boldsymbol{A})_d]^{-1}\}^{-1}. \tag{6}$$

Cronbach's alpha is a special case of Tarkkonen's rho under a simple model $\boldsymbol{x} = \boldsymbol{1}\tau + \boldsymbol{\varepsilon}$ and with a simple scale $u = \boldsymbol{1}'\boldsymbol{x}$.

Varsovan Inter-Continentalin 27. kerros, 25.3.2007:

- *"Mihin jäinkään?" – "On alettava vain kirjoittaa."*
- *"Pikanopeudella koko runko läpi, sitten viilaten osa osalta."*

# Aikomus oli saada kirja valmiiksi puolessa vuodessa...




- *"On siis syytä aloittaa ASAP, ettei tule kiire."*
- Touko-kesäkuussa odotti jo seuraava työmatka **Kanadaan**...
- Mikäli rahoitus napsahtaisi, menisin heinäkuussa **Japaniin**...
- Kevään 2008 johdantokurssin valmistelutkin olivat mielessä...

# 29.5.-6.6.2007: Montreal – Windsor – Montreal

## Thank you for your attention!

1 Kimmo Vehkalahti, Simo Puntanen, and Lauri Tarkkonen. Effects of measurement errors in predictor selection of linear regression model. *Computational Statistics & Data Analysis*, 2007. http://dx.doi.org/10.1016/j.csda.2007.05.005.

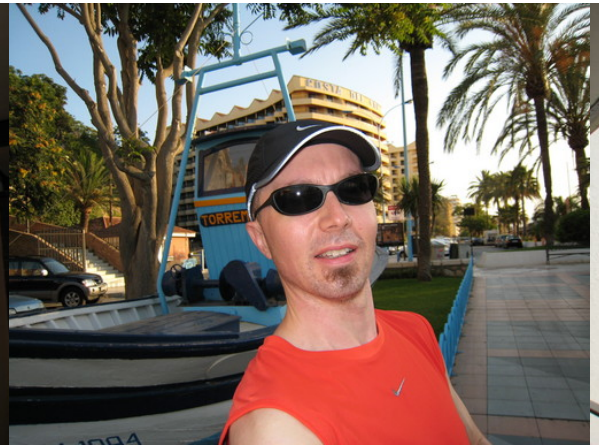
*"Nobody knows the number of essentially different 5 × 5, uniquely solvable open Survo Puzzles"* (Mustonen, 2007, p. 25).

- ...sieltä **Kuhmoon**, josta viikoksi melomaan kohti **Oulua**...

# 8.-17.6.2007: Suomi Meloo (Kuhmo – Oulu)

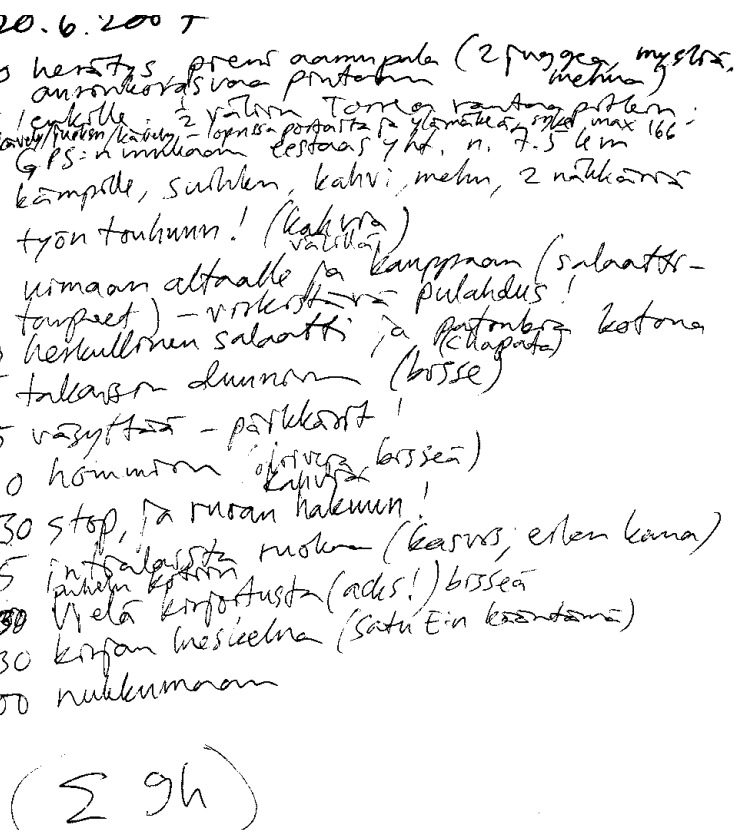
- ...mutta sen jälkeen **Espanjaan kirjoittamaan viikoksi**...

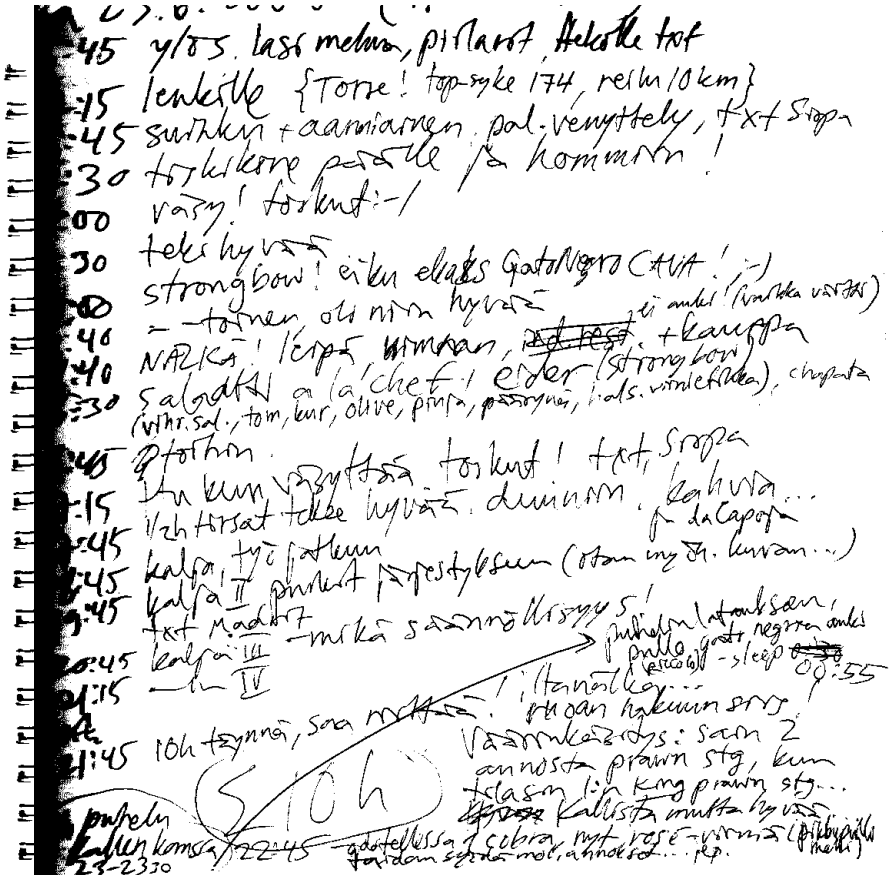
# 19.-26.6.2007: Benalmádena, Espanja



- Aamulenkki rantabulevardia Torremolinokseen ja takaisin...
  - *max-syke:* 174, *keskinopeus:* 7 km/h (GPS, matka 10 km)
  - *asenne:* **loistava**
- ...innoitti kirjoittamaan **syke, nopeus ja asenne** -esimerkin:
  - "Ajatellaan, että haluttaisiin tutkia satunnaisen lenkkeilijän sydämen sykettä, juoksunopeutta ja elämänasennetta. Helppoa, eikö niin? Ei tarvita kuin sykemittari, nopeusmittari ja – asennemittari."

# 19.-26.6.2007: Benalmádena, Espanja



- Lämpimässä ilmanalassa kului hieman virvoitusjuomia.

# 26.-28.6.2007: kotiinpaluu, tilannekatsaus

Malagan lentokentällä paperille kirjattua (26.6.2007):

Aikaansaannokset "10th floorissa":

- kirjoitustyö kunnolla alulle
- alkupään lukuihin hyvää perustekstiä
- alkupään lukujen rakenteen ja "juonen" terävöitystä
- loppuosa toistaiseksi alkutekijöissään, vaatii tiivistystä

Keskustelu kustannuspäällikön kanssa (28.6.2007):

- **LP:** "Tässä vielä ennen lomia kirjoittelen syksyä miettien. Teen elokuun alussa ensi vuoden tuotantobudjettia. **Onko kaikki ok, aikataulu ym.?**"
- **KV:** "Kiitos kysymästä, kaikki OK! Tulin juuri Espanjasta, jossa olin viikon kirjoittamassa. **Sain kirjan mukavasti alulle**, kuten oli tarkoituskin. Ensi viikolla lähden työmatkalle Japaniin, minkä jälkeen pääsenkin syventymään kirjaan pidemmäksi aikaa. Elokuussa minulla on varmasti jo jotain näytettävääkin."
- Rahoitus siis tuli, ja oli vuorossa työmatka **Japaniin**...

# 9.-16.7.2007: Tokio – Nagoya – Kioto – Tokio

- *Survo+LaTeX kuvien käytössä ja näytössä* (Vehkalahti, 2007)
- Ihmisvilinästä **kirjoittamaan** (tulevaan) **Sastamalaan**...

# 23.-27.7.2007: Sastamala, Suomi

(23.7.2007 16:10) "...fillari mukana, pyöräilin Laviaan (12 km, 20 min), pari bissee terassilla, back 30 min (enemmän ylämäkeä). **Nyt luen Espanjassa kirjoittamaani** (printtasin himassa koko roskan) **ja aloittelen kirjoittelua saunamökissä.** Hyvä ilma (vielä), huomiseksi luvattu sadetta (ja koko loppuviikoksi)."

(24.7.2007 07:07) "Klo 6 ylös, ulos ja lenkille! Kylällä hölkäten ja kävellen. Suihku ja petit dejeuner, **nyt kirjan äärellä. Luin eilen tuotokseni** sekä Alastalon kirjan loppuun. Nukkumaan mennessä **idea kirjan nimen perustelusta; miksi ei 'lomaketutkimus'**."

(24.7.2007 22:13) "**Jotain pientä olen saanut aikaan**, mutta jo nyt alkaa vaikuttaa siltä, että **projekti venyy pidemmälle kuin alun perin oli tarkoitus**. Hyvä tavoite olisi saada elo-syyskuun aikana kasaan RAAKILE koko jutusta. Se voisi sitten kypsyä talven sekä itsellä että Tammella. Nyt ei saisi liikaa juuttua yksityiskohtiin eikä muotoiluihin, vaan pitää rakentaa suurempia linjoja."

# 23.-27.7.2007: Sastamala, Suomi

(26.7.2007 11:21) ”Eilen taisin keksiä, että **liitteeseen voisi kerätä kaikkien Survo- ja SPSS-esimerkkien komentokaaviot**. Näin ne eivät sotkisi liikaa tekstiä, jossa on parempi keskittyä menetelmien kuvaamiseen, ohjelmien tulosteisiin jne. Liitteessä voisi kuvata perusteellisesti mistä kussakin komennossa ym. on kyse. Sieltä voisi myös ristiviitata aina sille sivulle, jossa ko. esimerkki esiintyy. LaTeX mahdollistaa!

**Sain illalla pähkäiltyä rakenteenkin paremmaksi!** Tiedonkeruu on aika sekondäärinen juttu tässä; siirrän sen aineiston perustamisen yhteyteen. Näin kirjan ekat luvut (johdannon jälkeen) ovat MITTAAMINEN ja DOKUMENTOINTI.”

(27.7.2007 13:01) ”Nyt kone kiinni; **jotain sain täällä aikaan, mutta ehkä vähemmän varsinaista kirjoittamista**. Työrauha ei ollut ihan niin hyvä. Rakenteellisia ideoita kuitenkin. Tästä voi jatkaa...”

# Näkymiä ”konehuoneesta” eri vaiheissa (1/17)

Survo+LaTeX-yhdistelmä oli jo tullut tutuksi, mutta kirjan teko vaati (LaTeXin osalta) lukuisia uusia asennuksia, masennuksia, säätöjä, vääntöjä, perehtymisiä, kokeiluja, testauksia, yrityksiä, erehdyksiä, viilailuja, ...

```
157 *%\usepackage{makeidx} % (vain yksi hakemisto!!) [toimi tähän asti muttei riitä...]
158 *%\makeindex
159 -
160 - / hirveellä ähellyksellä saatu VIHDOINKIN pelaamaan 5.9.2007!!!
161 - / (tavallinen LaTeX ei tue multi-indeksejä, siksi tämä)
162 - / perusohje d:\simo\trix mutta netistä kaivelin ja asentelin lisää.
163 - / ks. http://www.tex.ac.uk/cgi-bin/texfaq2html?label=multind
164 - / ASENNUS ONNISTUI PARHAITEN MIKTEX-BROWSE-PACKAGES (tms.) Start - Programs---
165 *\usepackage{index}
166 *\makeindex
167 - / Sinänsä toimiva (itse sain kuntoon), mutta laitan sivuun - kokeilen authorindexiä!
168 - / \newindex{authors}{adx}{and}{Henkilöhakemisto}
169 *\newindex{yleiset}{idx}{ind}{Hakemisto}
170 - / äskeinen: oma kokeilu! en saanut "Index"ia millään pois  :)
171 - / sekalaisia yritelmiä, jätän joksikin aikaa:
172 *%\renewcommand*{\indexname}{Hakemisto}
173 *%\newindex{default}{idx}{ind}{Hakemisto}
174 *%\newcommand*{\authorindexname}{Henkilöhakemisto}
175 *
176 - / http://www.tex.ac.uk/cgi-bin/texfaq2html?label=multind
177 - / ärsyttävästi muokkaa myös ulkoasua! ei olekaan chapter enää.
178 *%\usepackage{multind}
179 *%\makeindex{general}
180 *%\makeindex{authors}
```

# Näkymiä ”konehuoneesta” eri vaiheissa (2/17)

Vietin joitain lohduttomia hetkiä mm. hakemistojen kanssa.

```
197 - / latasin tuolta:
198 - / DD E:\LZMA_SDK
199 - / ja paketin jo aiemmin:
200 - / DD E:\AUTHORINDEX
201 - / sain siellä luotua tuon .tarin komennolla
202 - / E:\LZMA_SDK\lzma d authorindex.tar.lzma authorindex.tar
203 - / nyt pitää taas vaihtaa oikeuksille - ttu!!
204 - / Olipa aikaavievää ja hankalaa.  Voi prkl prkl!! Koko iltapäivä meni.
205 - / (enkä tiedä vieläkään, toimiiko...)
206 - / Siis: käsin nämä kaksi tiedostoa paikoilleen (loin dirit):
207 - / C:\PROGRA~1\MIKTEX~1.6\SCRIPTS\PERL\authorindex\authorindex.pl  <- .pl lisätty!
208 - /   sisältöä pätsätty kahdesta kohtaa (!) ->  http://www.nabble.com/Re:-authorindex
209 - / C:\PROGRA~1\MIKTEX~1.6\TEX\LATEX\authorindex\authorindex.sty
210 - / Sitten MikTeX Options auki (start - programs - miktex) ja
211 - / sieltä Refresh FNDB ja Update formats.
212 - / Asensin myös tässä välissä Active Perl v.5.88:n !
213 - / Ja kopsasin tänne uusimman authorindex-pdf:n. Printtasin sen Exassa.
214 - / Vielä natbib.cfg luotu (ks. n. rivi 825 tässä kentässä!).
215 - / Näyttäisi toimivan!! Nyt vielä tälle määrittelyt: (pdf:n mukaan):
216 *\let\cite=\aicite
217 *\let\citet=\aicite
218 *\let\citep=\aicite
```

# Näkymiä ”konehuoneesta” eri vaiheissa (3/17)

Toinen haasteellinen kohde olivat ”kelluvat elementit” (kuvat, taulukot, tulosteet jne.), joita säädin satoja kertoja ennen kuin sain ne toimimaan halutulla tavalla.

```
268 - / Jarmo Niemelä:  ks. LaTeX-jatkokurssi - voisi mainita lähteissä?
269 - / (tämä ohje tosin taitaa olla muuallakin)
270 *\newcommand*{\pvm}{\number\day.\number\month.\number\year}
271 *
272 - / tämä jo tänne eikä vasta BiBTeX-PREAMBLE, esim. MSM-viittaus!
273 *\newcommand{\url}[1]{\texttt{#1}}
274 -
275 - / kokeillaan tehdä tulostuksia varten oma ympäristö: äh, ks. alla!
276 - / \newcounter{tulnro}
277 - /  %\setcounter{tulnro}{0}
278 - / \newenvironment{tuloste}%
279 - / {\begin{center}\refstepcounter{tulnro}}{\end{center}}
280 - / \renewcommand*{\thetulnro}{\thesection.\arabic{tulnro}}
281 - / vielä selvittämättä, miten saan captionin toimimaan näin: (kuten figure!)
282 - / "Tuloste x.y: Perustunnuslukuja"
283 - / s e:\tktltiki.dtx / tästä voisi oppia jotain...
284 -
285 - / Sittenkin (COMPANION, s.292) ihan toisella tavalla: TÄMÄ TOIMII!
286 *\usepackage{float}
287 *\floatstyle{plain}
288 - / {H} ehdoton, katsotaan ensin {h}:lla..
289 *\newfloat{Tuloste}{h}{loo}[chapter]
290 *\newfloat{Työkaavio}{h}{los}[chapter]
291 *\floatstyle{plaintop}
292 *\newfloat{Esimerkki}{h}{lox}[chapter]
293 *
294 *\usepackage{caption}
295 - / ks. Pekan ohjeet:
296 - / S D:\D\PEKKAN2 83 / !!!
297 *\DeclareCaptionLabelSeparator{psep}{.\hspace{0.5em}}
298 *\captionsetup{labelsep=psep,font={sf,bf},textfont=md}
299 - / PJN: font=small,labelfont={sf,bf},margin=3em}
300 *
301 - / näin Esimerkkien caption vasempaan reunaan (!):
302 *\captionsetup[Esimerkki]{singlelinecheck=false,belowskip=0pt,aboveskip=0pt}
303 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (4/17)

Survon PRINT-kaavion sain aika yksinkertaisen näköiseksi – kun lakaisin lähes kaikki LaTeX-sotkut erillisiin tiedostoihin.

```
*\mainmatter
*
*Johdanto
- chapter A in KMM'JK1
*
*Moniulotteinen mittaaminen\label{ch:meas}
- chapter A in KMM'KM1
*
- / tämä otsikko on muuttunut MONEEN kertaan! :)
- / Työskentely ja dokumentointi\label{ch:doc}
- / Ohjelmistot ja työskentely\label{ch:doc}
*Työskentely tietokoneella\label{ch:docs}
- chapter A in KMM'OT1
*
*Tiedonkeruu ja aineiston perustaminen\label{ch:data}
- chapter A in KMM'TP1
*
*Alustavat tarkastelut ja esikäsittely\label{ch:base}
- chapter A in KMM'TT1
*
*Aineiston tiivistäminen\label{ch:FA}
- chapter A in KMM'FA1
*
*Selitysmallien laatiminen\label{ch:RA}
- chapter A in KMM'SM1
*
*Vastaajien ryhmittely\label{ch:G1}
- chapter A in KMM'VR1
*
*Vastaajaryhmien analysointi\label{ch:G2}
- chapter A in KMM'VA1
*
*\appendix
- / comp.text.tex (1997): The real meaning of the LaTeX \appendix command is
- /   \StartTreatingChaptersAsAppendices
-
```

# Näkymiä ”konehuoneesta” eri vaiheissa (5/17)

Lukujen ja liitteiden otsikot muuttuivat **hyvin** moneen kertaan.

```
416 -
417 - / Survo ja SPSS lyhyesti\label{app:SurvoSPSS}
418 *Kirjan ohjelmistot lyhyesti\label{app:SurvoSPSS}
419 - chapter A in KMM'LI1
420 *
421 - / \chapter{Liite: esimerkkien tekniset toteutukset} % (Survo ja SPSS)}
422 - / \chapter{Liite: Esimerkkien Survo- ja SPSS-toteutukset}
423 - / Liite: Esimerkkien toteutukset
424 - / Liite: Kuvien ja esimerkkien toteutukset
425 *
426 - / Liite 2: Survo- ja SPSS-kaaviot
427 - / Liite 2: Kuvien ja tulosteiden työkaaviot
428 *\addtocontents{toc}{\protect\setcounter{tocdepth}{0}}
429 *Kuvien ja tulosteiden työkaaviot\label{app:kaaviot}
430 - / "kuvat ja kaaviot" rimmais, mutta... hmmm...
431 - / kaaviot: kuvista, tulosteita (ja taulukoista?)
432 - chapter A in KMM'LI2
433 - / ei ehkä alakohtia "Survo" ja "SPSS" vaan järjestyksessä
434 - / kaavio kerrallaan; viittaus sille sivulle jossa sen output on
435 - / ja tieto kummalla ohjelmalla kulloinkin tehdään...
436 - /
437 -
438 - / LIITE 2 ?
439 -
440 - / idea 15.8.2007: voisi olla liite, jossa olisi Survon ja SPSS:n
441 - / "viilailu"tietoja, esim. Survon säätäminen mahd. vikkeläksi,
442 - / SPSS:n fontti- ym. säädöt jne. - KV:n suosituksia?
443 - / tämä voisi myös olla osa allaolevaa, jonka keksin myöhemmin
444 - / "Säätösuosituksia" tms. -> Liite 1:n loppuun yksi kohta molemmista?
445 -
446 - / LIITE 3 ? (ehkä jopa 1)
447 -
448 - / idea 15.8.2007: voisi olla liite, jossa olisi Survon ja SPSS:n
449 - / taustat (ks. kohta 3.1.1, vie siellä liikaa tilaa!) ja ehkäpä
450 - / myös peruskäyttöselostukset? 3.1.1:een viittaukset tuollaiseen?
451 - / "Survo ja SPSS lyhyesti" (laitoin jo ekaks!) joo, liite 1:n alkuun
452 -
```

# Näkymiä ”konehuoneesta” eri vaiheissa (6/17)

Toimituskentät ovat täynnä kohtia, jotka on kommentoitu pois, mutta joista näkee, mitä olen milloinkin ajatellut. Näin välttyy aika tehokkaasti kokeilemasta jotain jo hylättyä ideaa pian uudelleen.

```
460 ^
461 - / Sinänsä toimiva (itse sain kuntoon), mutta laitan sivuun - kokeilen authorindexiä!
462 - / \addcontentsline{toc}{chapter}{Henkilöhakemisto}
463 - / %\footnotesize
464 - / %\printindex{authors}{Henkilöhakemisto}
465 - / \printindex[authors]
466 - / %\normalsize
467 *
468 - / tämä siis nojaa authorindex-pakettiin: (automaattiset hlö-viittaukset lähdev.)
469 - / toimii siis, mutta olisiko sittenkin parempi vain YKSI HAKEMISTO (kaikki samaan)?
470 - / Kokeillaan 26.10.2007: tehdään authorindex, mutta yhdistetään kaikki samaan!
471 - / Henkilöhakemisto
472 - / \begin{multicols}{2}
473 - / \printauthorindex
474 - / \end{multicols}
475 - / %\cleardoublepage
476 *
477 - / jos authorindex osoittautuu OK:ksi, tätä voi ehkä vähän keventää?
478 *%\renewcommand{\indexname}{Hakemisto}    % ei pure!!
479 *\addcontentsline{toc}{chapter}{Hakemisto}
480 *%\footnotesize
481 - / yritin tännekin multicols, menee vituiksi
482 *\printindex[yleiset]                      % NYT PUREE!
483 *%\normalsize
484 E
485 *═══════════════════════════════════════════════════════════════════════
486 - / \part{Puitteet} ÄH, nämä ihan turhia (vrt. Heikkilä!)
487 - / \part{Menetelmät} PÖH, turhia otsikoita
488 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (7/17)

Kieliasetusten säätämisessä oli hommia, esimerkiksi lähteiden ja hakemiston osalta. Vaikka LaTeXista löytyy **miljoonittain** verkkosivuja, oli kirjakin usein tarpeen.

```
842 *PVM / Tiistai, 8. toukokuuta 2007 klo 16.29 (vko 19, pvä 128/365)
843 *
844 *Tässä välissä hankin Companionin 2nd ed:in!
845 *Monet noista printatuista sisältyvät siihen.
846 *
847 *Asensin uudemman Miktexin ja laitoin sen kieliasetukset kuntoon.
848 *
849 *PVM / Maanantai, 21. toukokuuta 2007 klo 13.08 (vko 21, pvä 141/365)
850 *
851 *
852 *Luvuille omat kentät, olen lukenut ahkerasti Companionia!
853 *
854 *dd suomek*.*
```

Kesän 2007 lopulla siirryin pdfLaTeXin käyttöön. Alkuvaikeuksia:

```
856 *PVM / Perjantai, 17. elokuuta 2007 klo 17.53 (vko 33, pvä 229/365)
857 *
858 *Päätin siirtyä pdflatex-käyttäjäksi, jotta SPSS-taulukot saa
859 *kivasti pdf-muodossa mukaan. (Hyvä että ne saa niin!)
860 *
861 *tehdään pikakoe:
862 *
863 *>EPSTOPDF KMM-mittauskehikko.eps
864 *>EPSTOPDF sv_bar.eps
865 *>EPSTOPDF xy1.eps
866 *>EPSTOPDF yhdessa_histo.eps
867 *
868 *Ei toimi hyvin (johtunee Survon bound.boxista).
869 *Tehdään mieluummin /GS-PDF:llä:
870 *
871 */GS-PDF KMM-mittauskehikko.eps
872 */GS-PDF sv_bar.eps
873 */GS-PDF xy1.eps
874 */GS-PDF yhdessa_histo.eps
875 *
876 *ja kokeillaan sitten kääntää...
877 *Säädän jo sen verran, että jätän tiedostotyypin POIS
878 *incl.graph:eista; Jarmo Niemelä ohjeisti näin Jori Mäntysalolle
879 *sfnet.atk.tex-ryhmässä.
880 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (8/17)

Montahan tuntia mahdoin käyttää nimihakemiston säätämiseen? Lopulta luovuin siitä, vaikka sainkin sen toimimaan ihan hienosti.

```
966 *Tämä on authorindex.pdf:n sivulta 11:     (10.9.2007)
967 *
968 *SAVEW CUR+1,E-1,natbib.cfg / en tiedä mihin pitäisi tallettaa; löytäiskö täältä?
969 *% natbib.cfg
970 *\AtBeginDocument{%
971 *\@ifpackageloaded{authorindex}{%
972 *\ifNAT@numbers
973 *\let\org@@citex\NAT@citexnum
974 *\else
975 *\let\org@@citex\NAT@citex
976 *\fi
977 *\def\@citex[#1][#2]#3{%
978 *\typeout{indexing: [#1][#2]{#3}}%
979 *\org@@citex[#1][#2]{#3}%
980 *\@aicitey{#3}}%
981 *\renewcommand\NAT@wrout[5]{%
982 *\if@filesw{%
983 *\let\protect\noexpand\let~\relax
984 *\immediate\write\@auxout{\string\aibibcite{#5}{#1}}%
985 *\immediate\write\@auxout{\string\bibcite{#5}{{#1}{#2}{{#3}}{{#4}}}}}%
986 *\fi}}{}}
987 *\endinput
988 E═══════════════════════════════════════════════════════════════════
989 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (9/17)

LaTeX synnyttää paljon tiedostoja; piti tarkistaa mitä kaikkea:

```
18 */acro kmm.pdf
19 *>notepad KMM.lof
20 *dd d:\simo\mitra*.*  / ks. täältä mallia !!!! (Jarmon käsialaa nuo cls:t.)
21 *dd d:\simo\trix      / ja täältä VARSINKIN!
22 *dd d:\art\3\KIMMOL~1 / ja täältä
23 *http://www.tex.ac.uk/tex-archive/help/Catalogue/index.html
24 * - orporiveistä ks. #207: \enlargethispage{2\baselineskip}
25 * HUOM! Ks. miten a) floatit EI-keskelle kpl:tta,  b) \noindent niiden jälkeen?
26 *                 a) tyhjä rivi ennen ja jälkeen!  b) [en laita, katsotaan...]
27 *TIME COUNT START
28 *FILE DEL KMM.pdf / (pdfLaTeX)
29 *FILE DEL KMM.aux / (pdfLaTeX)
30 *FILE DEL KMM.bbl / (BiBTeX)
31 *FILE DEL KMM.ind / (index)
32 *FILE DEL KMM.ain / (authorindex)
33 *FILE DEL KMM.toc / (table of contents)
34 *FILE DEL KMM.lof / (list of FILES) - tähän ahdettu kaikki samaan! TOIMII! :) proud!
35 *PRINT e+1,E TO KMM.tex / tätä ennen olemassa vain KMM.edt KMM.dv2 KMM.bib KMM.ist
36  -> ■>dir KMM.tex
37 *>pdfLaTeX  KMM.tex
38  -> ■>dir KMM.aux, KMM.idx, KMM.toc, KMM.lof, KMM.log
39 *>BiBTeX KMM
40  -> ■>dir KMM.bbl, KMM.blg
41 *>makeindex -c -o KMM.ind -s KMM.ist -t KMM.ilg KMM.idx
42  -> ■>dir KMM.ind, KMM.ilg
43 *>authorindex -i KMM
44  -> ■>dir KMM.ain
45 *>pdfLaTeX  KMM.tex
46  -> ■>dir KMM.aux, KMM.idx, KMM.toc, KMM.lof, KMM.log
47 *>authorindex -i KMM
48  -> ■>dir KMM.ain
49 *>copy KMM.ain KMM.tmp
50 *TXTCONV KMM.tmp KMM.ain
51 *>makeindex -c -o KMM.ind -s KMM.ist -t KMM.ilg KMM.idx KMM.ain
52  -> ■>dir KMM.ind, KMM.ilg
53 *>pdfLaTeX  KMM.tex
54  -> ■>dir KMM.aux, KMM.idx, KMM.toc, KMM.lof, KMM.log
55 *>copy KMM.lof KMM.tmp
56 *TXTCONV KMM.tmp KMM.lof
57 *>pdfLaTeX  KMM.tex
58  -> ■>dir KMM.aux, KMM.idx, KMM.toc, KMM.lof, KMM.log
59 */OPEN KMM.pdf
60 *TIME COUNT END   24.984
61 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (10/17)

```
 62 *s kmm.lof / poistetaan aivan lopuksi Caption:ien loppupisteet luettelosta:
 63 *s kmm.ain / modifioidaan skandeja sisältäviä nimiä (ovelasti jaettu kaavio!)
 64 *s kmm.idx / tämän (mahd.) pätsäys ei taida auttaa, jos tarve tulee... ks. Jarmon ohje
 65 *CONVERSIONS:
 66 *t   ".\relax"   "\relax"                        / lof
 67 *t   "{figure}{" "{figure}{Kuva "               / lof - ovela viritys, ks. myös alla!
 68 *t   "{Esimerkki}{" "{Esimerkki}{Esimerkki "    / lof - ovela viritys, ks. myös alla!
 69 *t   "{Tuloste}{" "{Tuloste}{Tuloste "          / lof - ovela viritys, ks. myös alla!
 70 *t   "{Taulukko}{" "{Taulukko}{Taulukko "       / lof - ovela viritys, ks. myös alla!
 71 *t   "PNTINEN"   "PZNTINEN"                     / ain
 72 *t   "YLSTALO"   "YLZSTALO"                     / ain
 73 *END
 74 *
 75 *s kmm.ain
 76 *s kmm.ind
 77 *s kmm.tex "\index"
 78 *
 79 *>authorindex -h
 80 *>makeindex ?
 81 *>makeindex3 ?   (kokeilin; ei löydä style-filea, katson jos pärjäisin ilman...)
 82 *
 83 *dd  kmm.*
 84 *
 85 *Uusin kokeilu: indexit sittenkin yhteen (makeindexiin lisäparametrit)!
 86 *kysymyksiä: 1) miten henkilöt BOLDilla? (äh) 2) miten Pöntinen oikeaan kohtaan? OK!
 87 *            3) miten SPSS Inc., Survo Systems ym. POIS? (?) [ainakin Survolla...]
 88 *            4) miten eka piste pois (Leena)? NÄIN: (viilaa: Companion s.664) OK.
 89 *                                                 ks. taulukko s.661 !
 90 *SAVEP CUR+1,E-1,KMM.ist / siirrä KMM'TYY-kenttään myöh.!
 91 *% MakeIndex style file for KMM
 92 *heading_prefix "{\\bfseries "
 93 *heading_suffix "\\hfil}\\nopagebreak\n"
 94 *headings_flag 1
 95 *% delim_0 " "
 96 *%delim_0 "\\hspace{5pt}"
 97 *delim_0 "\\ \\ "
 98 *%delim_0 "\\dotfill "
 99 *%delim_1 "\\dotfill "
100 *%delim_2 "\\dotfill "
101 E
102 *% tästä eteenpäin Jarmon (2005) ohjeista:
103 *% http://www.mtl.uta.fi/Opetus/LaTeX/hakemistot/hakemisto_ohje.pdf
104 *  /OPEN LaTeX\hakemisto_ohje.pdf
105 *% mutta mutta: nämä versiota 3.0 varten; minulla versio 2 ! UNOHDA TOISTAISEKSI.
106 *show makeindex3_pohjia.txt / tästä käyttöön, jos tähän palaat
107 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (11/17)

Otsikkotyypit ja sivun mittasuhteet hakusessa:

```
165 - / liitteisiin, sisältöön hakemistoon ym. eri tyyli? (ks. ohjeet!) pienempi?
166 - / myös Companion, s. 39 ym.
167 *
168 *
169 - / NÄMÄ JÄÄVÄTKIN POIS; TITLESEC YM. HOITAVAT ASIAT:
170 - / lukujen tyylit:
171 - / \usepackage[Sonny]{fncychap}
172 - / Leenan kommenttien (31.7) perusteella muutetaan (ei sanaa "LUKU"):
173 - / \usepackage[Bjornstrup]{fncychap} ei näytä hyvältä!
174 - / % /OPEN LATEX/fncychap.pdf
175 - / tarviiko tätä, tarviiko muuten noita makeat...???
176 - / \makeatletter
177 - / \ChTitleVar{\Huge\sffamily\bfseries}
178 - / \makeatother
179 - / ---------------------------------
180 - / ks. Companion, s.36:
181 - / %\usepackage[avantgarde]{quotchap}
182 - / %\renewcommand\chapterheadstartvskip{\vspace*{5\baselineskip}}
183 - / harkitse vielä: (on kyllä aika hyvän näköinen!) - toisaalta...
184 - / \usepackage{helvet}
185 - / %\renewcommand\sectfont{\sffamily\bfseries}
186 - / näitä ei toistaiseksi; pitäisi laittaa joka lukuun...
187 - / \begin{savequote}[10pc]
188 - / \sffamily
189 - / Jep jep!
190 - / \qauthor{Kimmo}
191 - / \end{savequote}
192 *
193 *\usepackage[dvips=false,pdftex=false,vtex=false,
194 *            paperwidth=150mm, paperheight=235mm, twoside,
195 *            textheight=42\baselineskip,
196 *            textwidth=105mm,
197 *            inner=28mm,
198 *            includeheadfoot]{geometry}
199 *% /OPEN LATEX/geometry.pdf
200 *\usepackage[frame,a4,center]{crop}  % tarviiko tosiaan 'cam' (ks. MEMO: Leena)
201 *% /OPEN LATEX/crop.pdf
202 *
203 - / seuraavat vähän turhantuntuisia, koetan pärjätä oletuksilla!
204 - / Sittenkin käyttöön Leenan kommenttien (31.7) perusteella:
205 *\usepackage{fancyhdr}
206 *% /OPEN LATEX/fancyhdr.pdf
207 *\pagestyle{fancy}
208 *\setlength\headheight{28pt}
```

# Näkymiä ”konehuoneesta” eri vaiheissa (12/17)

Kuvien, taulukoiden ym. yhteisluettelosta kirjallisuudessa ei annettu paljoa toivoa, mutta ratkaisin asian Survon ja LaTeXin yhteispelillä:

```
304 *\cleardoublepage
305 *\addcontentsline{toc}{section}{Alkusanat}
306 -   chapter Esipuhe in KMM'AL1
307 -   chapter Kiitos  in KMM'AL1
308 -
309 *\cleardoublepage
310 *\pagestyle{KMM}
311 - / sis.luettelo ehkä vähän tiiviimmin - tai sitten ei...      ks. !
312 *\begin{spacing}{0.95}
313 - / "Sisältö" nyt liian isolla! Pienemmäksi - MITEN?
314 *\tableofcontents
315 *\end{spacing}
316 *
317 - / Companion, s. 54:; oma kokeilu muilla kuin figure: TOIMII! :-))
318 *\makeatletter
319 -
320 - / toimiva, mutta hieman kökkö; kokeilen Survoa paremmin!
321 - / \renewcommand\l@figure[2]{\@dottedtocline
322 - /     {1}{0.0em}{2.3em}{#1~(kuva)}{#2}}
323 - / \renewcommand\l@Tuloste[2]{\@dottedtocline
324 - /     {1}{0.0em}{2.3em}{#1~(tuloste)}{#2}}
325 - / \renewcommand\l@Esimerkki[2]{\@dottedtocline
326 - /     {1}{0.0em}{2.3em}{#1~(esimerkki)}{#2}}
327 - / \renewcommand\l@Taulukko[2]{\@dottedtocline
328 - /     {1}{0.0em}{2.3em}{#1~(taulukko)}{#2}}
329 -
330 - / ks. myös TXTCONV edellä! ainakin pitää makustella - MELKOINEN VIRITYS... :)
331 *\renewcommand\l@figure[2]{\@dottedtocline
332 *    {1}{0.000em}{2.3em}{#1}{#2}}
333 *\renewcommand\l@Esimerkki[2]{\@dottedtocline
334 *    {1}{-2.00em}{2.3em}{#1}{#2}}
335 *\renewcommand\l@Tuloste[2]{\@dottedtocline
336 *    {1}{-0.85em}{2.3em}{#1}{#2}}
337 *\renewcommand\l@Taulukko[2]{\@dottedtocline
338 *    {1}{-1.60em}{2.3em}{#1}{#2}}
339 -
340 *\makeatother
341 -
```

# Näkymiä "konehuoneesta" eri vaiheissa (13/17)

Ja lukujen otsikot elävät jälleen (vanhoja vielä näkyvissä)...

```
352 *Johdanto
353 -   chapter A in KMM'JK1
354 *
355 - / ei kyselyLOMAKKEELLA vaan kyselyTUTKIMUKSESSA! (29.7.2008)!!
356 - / Mittaaminen kyselytutkimuksessa\label{ch:meas}
357 - / ohoh: laitoinkin nämä samaan lukuun! (5.8.2008; idea vkl)
358 - / "Mittaaminen" -> "Mittaus" (6.8.2008) muokattava tekstiä tämän vaikutuksesta!!!!
359 *Mittaus ja tiedonkeruu\label{ch:meas}\label{ch:data}
360 -   chapter A in KMM'KM1
361 - / Tiedonkeruu ja aineiston valmistelu\label{ch:data}
362 -   chapter A in KMM'TP1
363 *
364 - / Perustarkastelut ja esikäsittely\label{ch:base}
365 - / 5.8.2008 meni tämäkin otsikko uusiksi:
366 *Aineistoon tutustuminen\label{ch:base}
367 -   chapter A in KMM'TT1
368 *
369 - / JA tämä (myös 5.8.2008) - koko kirja mullin & mallin!
370 - / yhdistän nämäkin kaksi -katsotaan mitä tulee!
371 - / Tilastollinen mallintaminen\label{ch:stat}
372 *Aineiston tiivistäminen\label{ch:stat}\label{ch:FA}
373 -   chapter A in KMM'MA1
374 - / Aineiston tiivistäminen\label{ch:FA}
375 -   chapter A in KMM'FA1
376 *
377 - / Selitysmallien laatiminen\label{ch:RA}
378 - / otsikko täysin uusiksi (myös loput) 6.8.2008!
379 *Havaintojen vertailu\label{ch:RA}
380 -   chapter A in KMM'RA1
381 - / sittenkin tänne jatkeeksi ANOVAsta jotain...??? (eikä vikaan lukuun?)
382 *
383 - / otsikko muuttui heti kun aloin kirjoittaa (25.7.2008)! (oli: "Vastaajien")
384 - / Vastausten ryhmittely\label{ch:G1}
385 *Havaintojen ryhmittely\label{ch:G1}
386 -   chapter A in KMM'VR1
387 *
388 - / otsikko pitää tässäkin muuttaa... olisiko tuo?
389 - / Vastaajaryhmien analysointi\label{ch:G2}
390 - / Ryhmien kuvaaminen\label{ch:G2}
391 - / Ryhmäerojen visualisointi\label{ch:G2}
392 *Ryhmien visualisointi\label{ch:G2}
393 -   chapter A in KMM'VA1
394 *
395 - / kokeilin "Liitteet"-sivua 5.8.; part kuitenkin tuottaa typerän sis.luettelon. :(
396 - / itse tehty väliotsikko sis.luettelossa riittänee?
397 - / \pagestyle{alkusivut}
398 - / \part*{Liitteet}
399 *\appendix
```

# Näkymiä ”konehuoneesta” eri vaiheissa (14/17)

Liitteitäkin soudin ja huopasin vuoden ajan monta kertaa.
Pahimmillaan niitä oli kolme, useimmiten kaksi, lopulta vain yksi.

```
401 - / \pagestyle{KMM}
402 -
403 *Työvälineet ja dokumentointi\label{app:SurvoSPSS}
404 -   chapter A in KMM'LI1 / nämä paremmin toisiinsa! (31.5.2008) tehty 22.7.2008!
405 -                         / 24.7.2008: "Ohjelmistot" -> "Työvälineet"
406 - / idea 15.8.2007: voisi olla liite, jossa olisi Survon ja SPSS:n
407 - / "viilailu"tietoja, esim. Survon säätäminen mahd. vikkeläksi,
408 - / SPSS:n fontti- ym. säädöt jne. - KV:n suosituksia?
409 - / tämä voisi myös olla osa allaolevaa, jonka keksin myöhemmin
410 - / "Säätösuosituksia" tms. -> Liite 1:n loppuun yksi kohta molemmista?
411 -
412 - / ei tarvita (ks. liite B) \addtocontents{toc}{\protect\setcounter{tocdepth}{0}}
413 -
414 *Kuvien ja tulosteiden työkaavioita\label{app:kaaviot}
415 -   chapter A in KMM'LI2 / työkaaviot vai TEKOkaaviot? vrt. tekoturkis :/ KAAVIOT!!
416 -   /                                                     TYÖkaaviot takas 6.8.2008
417 -
418 - / \backmatter     (lienee aivan turha)
419 -
420 *
421 *\cleardoublepage
422 *\addcontentsline{toc}{chapter}{\bibname}
423 *\bibliography{KMM}
424 *
425 *\cleardoublepage
426 *\addcontentsline{toc}{chapter}{Hakemisto}
427 *\printindex[yleiset]
428 E
```

# Näkymiä ”konehuoneesta” eri vaiheissa (15/17)

```
 1 *SAVE RAKSA3 / Kyselytutkimuksen mittarit ja menetelmät - KV 2008        WINCONV=1
 2 *
 3 *
 4 *
 5 *Tämä on KOLMAS tällainen (tästä poistin authorindexit lopullisesti 1.9.2008)
 6 *
 7 *
 8 *
 9 *
10 *
11 *
12 *PRINT e+1,E TO KMM.tex
13 *>pdfLaTeX  KMM.tex
14 */OPEN KMM.pdf
15 *
16 */ACRO KMM.pdf
17 *s kmm.tex
18 *                 http://www.tex.ac.uk/tex-archive/help/Catalogue/index.html
19 *                  - orporiveistä ks. #207: \enlargethispage{2\baselineskip}
20 *
21 *FILE DEL KMM.pdf / (pdfLaTeX)
22 *FILE DEL KMM.aux / (pdfLaTeX)
23 *FILE DEL KMM.bbl / (BiBTeX)
24 *FILE DEL KMM.ind / (index)
25 *FILE DEL KMM.ain / (authorindex)
26 *FILE DEL KMM.toc / (table of contents)
27 *FILE DEL KMM.lof / (list of FILES)
28 *PRINT e+1,E TO KMM.tex / tätä ennen on: { KMM.edt KMM.dv2 KMM.pre KMM.bib KMM.ist }
29 *>pdfLaTeX  KMM.tex
30 *>BiBTeX KMM
31 *>makeindex -c -o KMM.ind -s KMM.ist -t KMM.ilg KMM.idx
32 *goto 1,l,1
33 *-----------------------[nämä pois, jos nimihakemisto pois!]--- ks. myös preamble!
34 *>authorindex -i KMM
35 *>pdfLaTeX  KMM.tex
36 *>authorindex -i KMM
37 *>copy KMM.ain KMM.tmp
38 *TXTCONV KMM.tmp KMM.ain
39 *>makeindex -c -o KMM.ind -s KMM.ist -t KMM.ilg KMM.idx KMM.ain
40 *-------------------------------------------------------------
41 *
42 l>pdfLaTeX  KMM.tex
43 *>copy KMM.lof KMM.tmp
44 *TXTCONV KMM.tmp KMM.lof
45 *>pdfLaTeX  KMM.tex
46 *
```

# Näkymiä ”konehuoneesta” eri vaiheissa (16/17)

```
 63 - include NULL.DEV / PRINT-määrittelyt
 64 - include KMM.DV2  / Survo-määrittelyt
 65 -    text KMM.PRE  / LaTeX-määrittelyt
 66 -
 67 *\begin{document}
 68 -
 69 *\pagestyle{alkusivut}
 70 -
 71 -      chapter Tittelit  in KMM'0
 72 -      chapter Copyright in KMM'0
 73 -
 74 *\cleardoublepage
 75 *\tableofcontents
 76 -
 77 *\cleardoublepage
 78 *\addcontentsline{toc}{section}{Alkusanat}
 79 -      chapter Alkusanat in KMM'0
 80 -      chapter Kiitokset in KMM'0
 81 -
 82 *\cleardoublepage
 83 *\pagestyle{KMM}
 84 *\reversemarginpar
 85 -
 86 *Johdanto
 87 -       chapter A in KMM'1
 88 *Mittaus ja tiedonkeruu
 89 =ch:meas
 90 -       chapter A in KMM'2
 91 -       chapter B in KMM'2
 92 *Tutustuminen aineistoon
 93 =ch:base
 94 -       chapter A in KMM'3
 95 *Aineiston tiivistäminen
 96 =ch:stat
 97 -       chapter A in KMM'4
 98 -       chapter B in KMM'4
 99 *Havaintojen vertailu
100 =ch:RA
101 -       chapter A in KMM'5
102 *Havaintojen ryhmittely
103 =ch:G1
104 -       chapter A in KMM'6
105 *Ryhmien visualisointi
106 =ch:G2
107 -       chapter A in KMM'7
108 -
109 *\appendix
110 -
111 - / \addtocontents{toc}{\protect\contentsline{chapter}{Liitteet}{~}}
```

# Näkymiä ”konehuoneesta” eri vaiheissa (17/17)

```
113 *Ohjelmistot ja dokumentointi
114 =app:SurvoSPSS
115 -       chapter A in KMM'A
116 -     / chapter B in KMM'A
117 -       chapter C in KMM'A
118 -
119 *\cleardoublepage
120 *\addcontentsline{toc}{chapter}{\bibname}
121 *\bibliography{KMM}
122 -
123 - / tämä hässäkkä päätetty siirtää kirjan LOPPUUN:
124 - / Companion, s. 54:: oma kokeilu muilla kuin figure: TOIMII! :-))
125 *\makeatletter
126 - / ks. myös TXTCONV edellä! ainakin pitää makustella - MELKOINEN VIRITYS... :)
127 *\renewcommand\l@figure[2]{\@dottedtocline
128 *    {1}{0.000em}{2.3em}{#1}{#2}}
129 *\renewcommand\l@Esimerkki[2]{\@dottedtocline
130 *    {1}{-2.00em}{2.3em}{#1}{#2}}
131 *\renewcommand\l@Tuloste[2]{\@dottedtocline
132 *    {1}{-0.85em}{2.3em}{#1}{#2}}
133 *\renewcommand\l@Taulukko[2]{\@dottedtocline
134 *    {1}{-1.60em}{2.3em}{#1}{#2}}
135 *\makeatother
136 -
137 *\cleardoublepage
138 *\listof{Tuloste}{Kuvat, esimerkit, tulosteet ja taulukot}
139 -
140 *\cleardoublepage
141 *\addcontentsline{toc}{chapter}{Hakemisto}
142 *\printindex[yleiset]
143 -
144 E
145 *==========================================================================
146 *
```

- Näin seesteiseen näkymään oli vielä aikaa...
- **Palataan siis kesään 2007.**

# 30.7.2007: Aika tarkistaa aikataulu Tammen kanssa


**1 Johdanto** 1
1.1 Kyselytutkimus 1
1.2 Luotettavuus ja virhemarginaalit 3
1.3 Kirjan rakenteesta ja sisällöstä 4

**2 Moniulotteinen mittaaminen** 5
2.1 Johdatteleva esimerkki 5
2.2 Mittauskehikko 7
2.2.1 Mittausmalli 8
2.2.2 Asteikot 9
2.3 Mittauksen laadun arviointi 9
2.4 Mittaustason valinta ja vaikutukset 9
2.5 Kyselylomakkeen laatiminen 10

**3 Työskentelytavat ja dokumentointi** 11
3.1 Tilastolliset ohjelmistot 11
3.2 Kirjassa käytetyt ohjelmistot 13
3.3 Tehokkaat työskentelytavat 14
3.4 Oman työn dokumentointi 14
3.4.1 Työvaiheiden toistaminen 14
3.4.2 Virheistä toipuminen 15
3.4.3 Nimien keksiminen 15
3.4.4 Kommenttien kirjoittaminen 16

**4 Tiedon keruu ja aineiston perustaminen** 17
4.1 Tiedon keruutavat 17
4.1.1 Kokonaistutkimus ja rekisterit 17
4.1.2 Otokset ja näytteet 18
4.2 Esimerkkiaineiston otanta-asetelma 19
4.3 Kyselylomakkeesta havaintotiedostoksi 20
4.3.1 Työvaiheet SPSS-ympäristössä 21
4.3.2 Työvaiheet Survo-ympäristössä 24


Sisällys näyttää minusta pätevältä ja mielenkiintoiselta. **Tulee hyvä teos, kunhan ehdit kirjoittamaan!** Toivottavasti nyt ei joku muu tulisi markkinoille aivan vastaavalla. Muutoin siirtäminen on ok."

**KV:** "Näkisin realistiseksi aikataulun **lykkäämisen yhdellä vuodella** eteenpäin. Alkuvuodesta 2008 voisi tilannetta arvioida uudelleen ja katsoa olisiko julkaiseminen mahdollista **syksyllä 2008** vai jätetäänkö se **alkuvuoteen 2009**."

**LP:** "Kiitos aikataulutiedoista. Ne ovat nyt tärkeitä, kun teen ensi vuoden budjettia. **Mielellämme tekisimme kirjasi jo syksyksi 2008**, mutta tehdään vaan hyvä kirja ja siirretään aikataulua tarpeen mukaan. Katsotaan tilanne alkuvuodesta.


**5 Aineistoon tutustuminen** 26
5.1 Aineiston selailu 26
5.2 Yhden muuttujan tarkastelut 27
5.2.1 Tilastolliset tunnusluvut 27
5.2.2 Esimerkkiaineiston tunnuslukuja ja jakaumia 28
5.2.3 Pylväs- ja piirakkakuvat 32
5.2.4 Histogrammi 32
5.3 Muunnokset ja uudelleenkoodaukset 34
5.4 Usean muuttujan tarkastelut 35
5.4.1 Ristiintaulukointi 35
5.4.2 Laatikkokuva 35
5.4.3 Hajontakuva 35
5.5 Aineiston muokkaaminen 35

**6 Aineiston tiivistäminen** 37
6.1 Keskeiset oletukset 37
6.1.1 Havaintojen riippumattomuus 37
6.1.2 Muuttujien riippuvuus 38
6.2 Faktorianalyysi 38
6.2.1 Mittausmallin analysointi 38
6.2.2 Rakennevaliditeetin arviointi 38
6.3 Tiivistäminen asteikoiksi 38
6.3.1 Faktoripistemäärät 38
6.3.2 Summamuuttujat 39
6.3.3 Reliabiliteetin arviointi 39

**7 Selitys- tai ennustemallien laatiminen** 40
7.1 Keskeiset oletukset 40
7.2 Regressioanalyysi 40
7.2.1 Selittäjien valinta 40
7.2.2 Regressiodiagnostiikka 40
7.2.3 Ennustevaliditeetin arviointi 40

**8 Vastaajien ryhmittely** 41
8.1 Etäisyyksien mittaaminen 41
8.2 Hierarkinen ryhmittely 41
8.3 Muita ryhmittelymenetelmiä 41

**9 Vastaajaryhmien analysointi** 42
9.1 Erotteluanalyysi 42
9.2 Varianssianalyysi 42
9.3 Korrespondenssianalyysi 42

**Liite: Esimerkkien Survo- ja SPSS-toteutukset** 43

# 30.7.-21.8.2007: Keskusteluja kollegan ja kirjan kanssa

"Niin siinä sitten kävi, kuten puhelimessa viime viikolla jo (puolipiloillaan) juteltiin: päätin lykätä kirjaprojektiani.

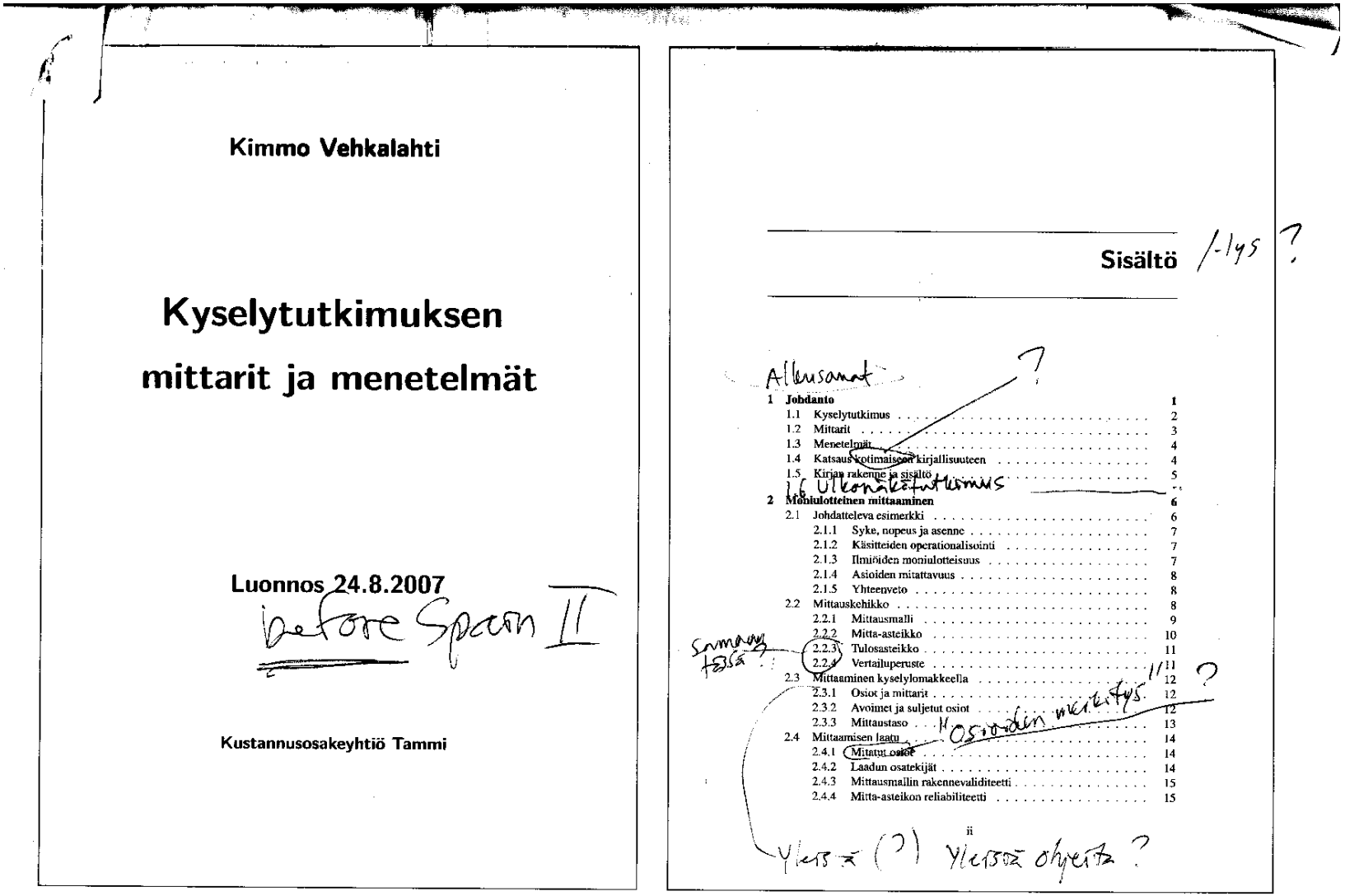
Kirjoitin jo kustannuspäällikölleni. Hän suhtautui erittäin hyvin; palaamme asiaan vuodenvaihteessa ja katsomme mitkä ovat näkymät tuolloin. Ainoa asia josta hän oli huolissaan, oli se että mahtaako joku muu ehtiä markkinoille vastaavan teoksen kanssa ennen minua. Vakuutin hänelle, ettei siitä ole vaaraa." (30.7.2007)

"Monena päivänä Exassa, välillä kotona. Homma edistyy! Eilen teknisiä kuvioita kuntoon; ympäristömäärittelyjä ym. Alkaa näyttää aika kivalta!" (16.8.2007)

"**Miksi kirjoittaminen on välillä niin vaikeaa?** Nytkin olen pykinyt koko aamupäivän, netissä meni liikaakin aikaa, vaikka löysin toki jotain kiintoisaa. Pitäisi saada printti Espanjaan mukaan, jotta siellä saisi tehot irti. Joka lukuun asioita, avainsanoja ym., jotta näkee mitä tulossa mihinkin. Esipuhe lyhennettävä, tosin ei kannata sitä nyt tehdä. Johdantoon kirjan pointteja, vähemmän muista kirjoista!!! Tämä on MINUN kirjani ei mikään kirj.katsaus. Ei liian akateemista/aneemista vaan dynaamista pikemminkin!" (21.8.2007)

"Johan syntyi hauska johdantoaloitus! :) Katsotaan kuinka kauan..." (21.8.2007)

# 24.8.2007: Näkymä sisällysluettelosta



Kimmo Vehkalahti

Kyselytutkimuksen mittarit ja menetelmät

Luonnos 24.8.2007

before Spain II

Kustannusosakeyhtiö Tammi

Sisältö

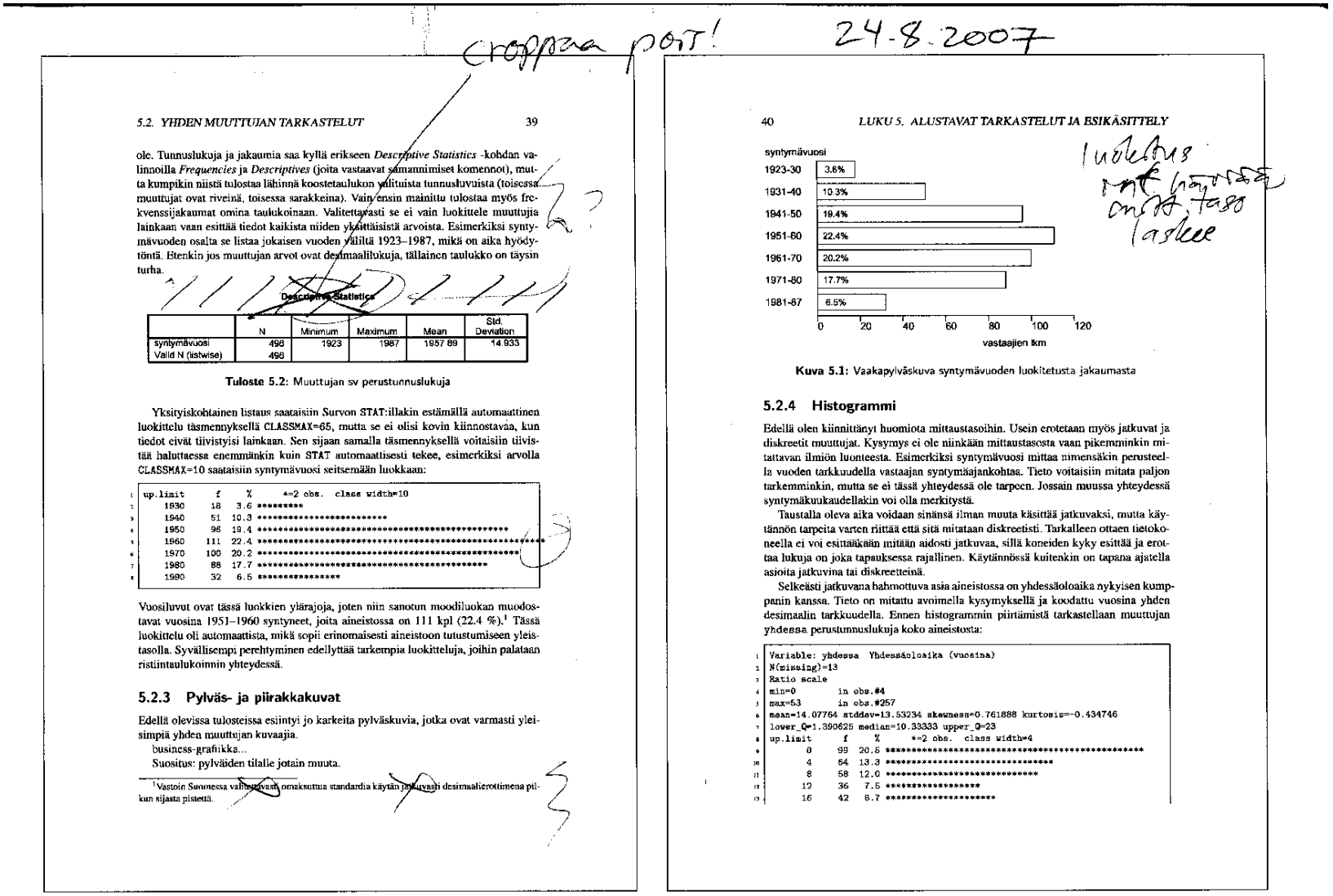
Alkusanat


1 Johdanto 1
1.1 Kyselytutkimus 2
1.2 Mittarit 3
1.3 Menetelmät 4
1.4 Katsaus kotimaiseen kirjallisuuteen 4
1.5 Kirjan rakenne ja sisältö 5
1.6 Ulkonäkötutkimus
2 Moniulotteinen mittaaminen 6
2.1 Johdatteleva esimerkki 6
2.1.1 Syke, nopeus ja asenne 7
2.1.2 Käsitteiden operationalisointi 7
2.1.3 Ilmiöiden moniulotteisuus 7
2.1.4 Asioiden mitattavuus 8
2.1.5 Yhteenveto 8
2.2 Mittauskehikko 8
2.2.1 Mittausmalli 9
2.2.2 Mitta-asteikko 10
2.2.3 Tulosasteikko 11
2.2.4 Vertailuperuste 11
2.3 Mittaaminen kyselylomakkeella 12
2.3.1 Osiot ja mittarit 12
2.3.2 Avoimet ja suljetut osiot 12
2.3.3 Mittaustaso 13
2.4 Mittaamisen laatu 14
2.4.1 Mitatut osiot 14
2.4.2 Laadun osatekijät 14
2.4.3 Mittausmallin rakennevaliditeetti 15
2.4.4 Mitta-asteikon reliabiliteetti 15


ii

- Pari päivää myöhemmin uudelleen **Espanjaan**...

# 24.8.2007: Näkymä tekstistä ennen Espanjaa



croppaa pois!

24.8.2007

5.2. YHDEN MUUTTUJAN TARKASTELUT 39

ole. Tunnuslukuja ja jakaumia saa kyllä erikseen *Descriptive Statistics* -kohdan valinnoilla *Frequencies* ja *Descriptives* (joita vastaavat samannimiset komennot), mutta kumpikin niistä tulostaa lähinnä koostetaulukon valituista tunnusluvuista (toisessa muuttujat ovat riveinä, toisessa sarakkeina). Vain ensin mainittu tulostaa myös frekvenssijakaumat omina taulukoinaan. Valitettavasti se ei vain luokittele muuttujia lainkaan vaan esittää tiedot kaikista niiden yksittäisistä arvoista. Esimerkiksi syntymävuoden osalta se listaa jokaisen vuoden väliltä 1923–1987, mikä on aika hyödytöntä. Etenkin jos muuttujan arvot ovat desimaalilukuja, tällainen taulukko on täysin turha.

**Descriptive Statistics**

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| syntymävuosi | 498 | 1923 | 1987 | 1957.89 | 14.933 |
| Valid N (listwise) | 496 | | | | |

**Tuloste 5.2:** Muuttujan sv perustunnuslukuja

Yksityiskohtainen listaus saataisiin Survon STAT:illakin estämällä automaattinen luokittelu täsmennyksellä CLASSMAX=65, mutta se ei olisi kovin kiinnostavaa, kun tiedot eivät tiivistyisi lainkaan. Sen sijaan samalla täsmennyksellä voitaisiin tiivistää haluttaessa enemmänkin kuin STAT automaattisesti tekee, esimerkiksi arvolla CLASSMAX=10 saataisiin syntymävuosi seitsemään luokkaan:

```
up.limit     f     %      *=2 obs.  class width=10
    1930    18   3.6 *********
    1940    51  10.3 *************************
    1950    96  19.4 ***********************************************
    1960   111  22.4 ****************************************************
    1970   100  20.2 ***************************************************
    1980    88  17.7 ******************************************
    1990    32   6.5 ****************
```

Vuosiluvut ovat tässä luokkien ylärajoja, joten niin sanotun moodiluokan muodostavat vuosina 1951–1960 syntyneet, joita aineistossa on 111 kpl (22.4 %).[1] Tässä luokittelu oli automaattista, mikä sopii erinomaisesti aineistoon tutustumiseen yleistasolla. Syvällisempi perehtyminen edellyttää tarkempia luokitteluja, joihin palataan ristiintaulukoinnin yhteydessä.

### 5.2.3 Pylväs- ja piirakkakuvat

Edellä olevissa tulosteissa esiintyi jo karkeita pylväskuvia, jotka ovat varmasti yleisimpiä yhden muuttujan kuvaajia.

business-grafiikka...

Suositus: pylväiden tilalle jotain muuta.

[1] Vastoin Suomessa valitettavasti omaksuttua standardia käytän jatkuvasti desimaalierottimena pilkun sijasta pistettä.

40 LUKU 5. ALUSTAVAT TARKASTELUT JA ESIKÄSITTELY




luokitus mit häviää mitt. taso laskee

**Kuva 5.1:** Vaakapylväskuva syntymävuoden luokitetusta jakaumasta

### 5.2.4 Histogrammi

Edellä olen kiinnittänyt huomiota mittaustasoihin. Usein erotetaan myös jatkuvat ja diskreetit muuttujat. Kysymys ei ole niinkään mittaustasosta vaan pikemminkin mitattavan ilmiön luonteesta. Esimerkiksi syntymävuosi mittaa nimensäkin perusteella vuoden tarkkuudella vastaajan syntymäajankohtaa. Tieto voitaisiin mitata paljon tarkemminkin, mutta se ei tässä yhteydessä ole tarpeen. Jossain muussa yhteydessä syntymäkuukaudellakin voi olla merkitystä.

Taustalla oleva aika voidaan sinänsä ilman muuta käsittää jatkuvaksi, mutta käytännön tarpeita varten riittää että sitä mitataan diskreetisti. Tarkalleen ottaen tietokoneella ei voi esittääkään mitään aidosti jatkuvaa, sillä koneiden kyky esittää ja erottaa lukuja on joka tapauksessa rajallinen. Käytännössä kuitenkin on tapana ajatella asioita jatkuvina tai diskreetteinä.

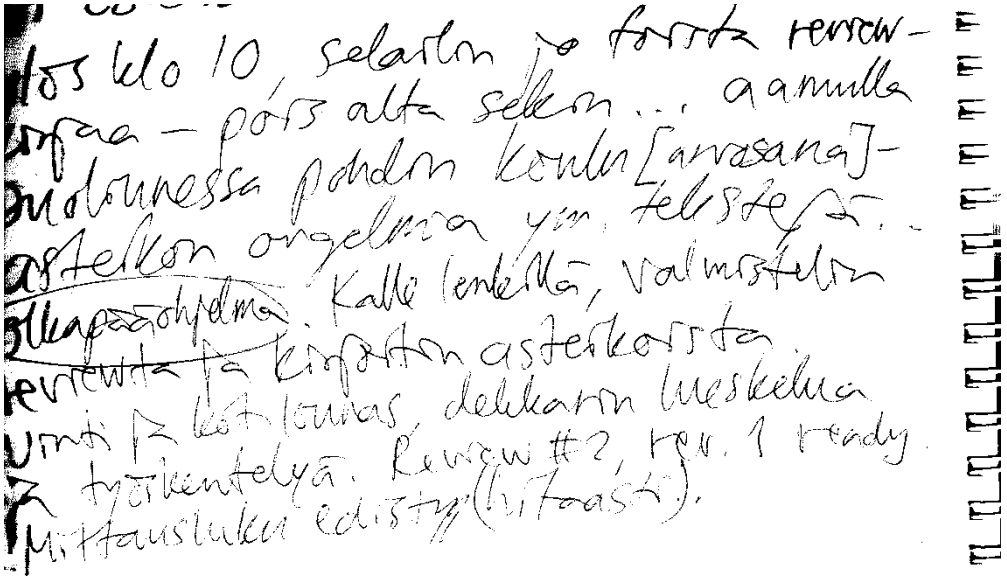
Selkeästi jatkuvana hahmottuva asia aineistossa on yhdessäoloaika nykyisen kumppanin kanssa. Tieto on mitattu avoimella kysymyksellä ja koodattu vuosina yhden desimaalin tarkkuudella. Ennen histogrammin piirtämistä tarkastellaan muuttujan yhdessa perustunnuslukuja koko aineistosta:

```
Variable: yhdessa  Yhdessäoloaika (vuosina)
N(missing)=13
Ratio scale
min=0          in obs.#4
max=53         in obs.#257
mean=14.07764 stddev=13.53234 skewness=0.761888 kurtosis=-0.434746
lower_Q=1.390625 median=10.33333 upper_Q=23
up.limit     f     %      *=2 obs.  class width=4
       0    99  20.5 ***********************************************
       4    64  13.3 ******************************
       8    58  12.0 ****************************
      12    36   7.5 ******************
      16    42   8.7 *********************
```

- Todella alustavaa; vielä "minä"-muodossa kirjoitettua...

# 26.8.-2.9.2007: Benalmádena, Espanja

*"Puoliunessa pohdin kouluarvosana-asteikon ongelmia..."*



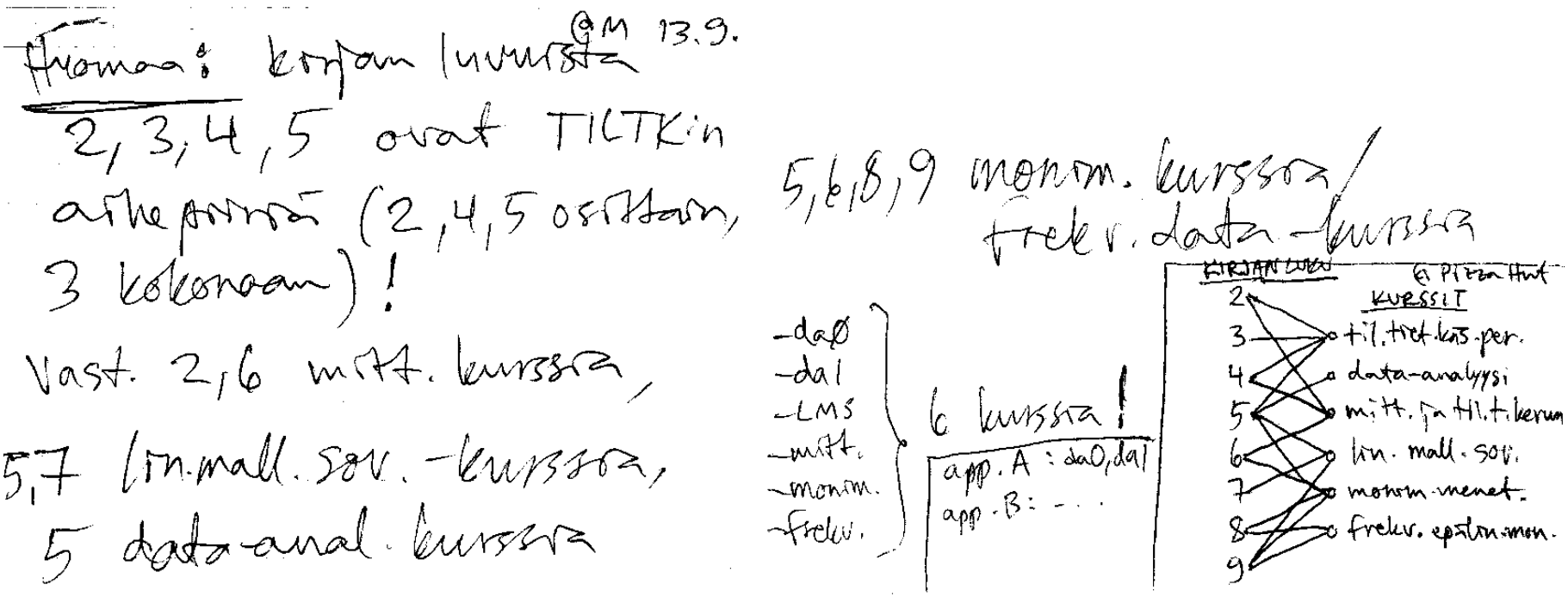

*"Kirjoitin asteikoista" – "Mittausluku edistyy (hitaasti)."*

- Lämpimässä ilmanalassa kului (taas) virvoitusjuomia.

- Paluu Suomeen ja lukuvuosi 2007–2008 alulle.

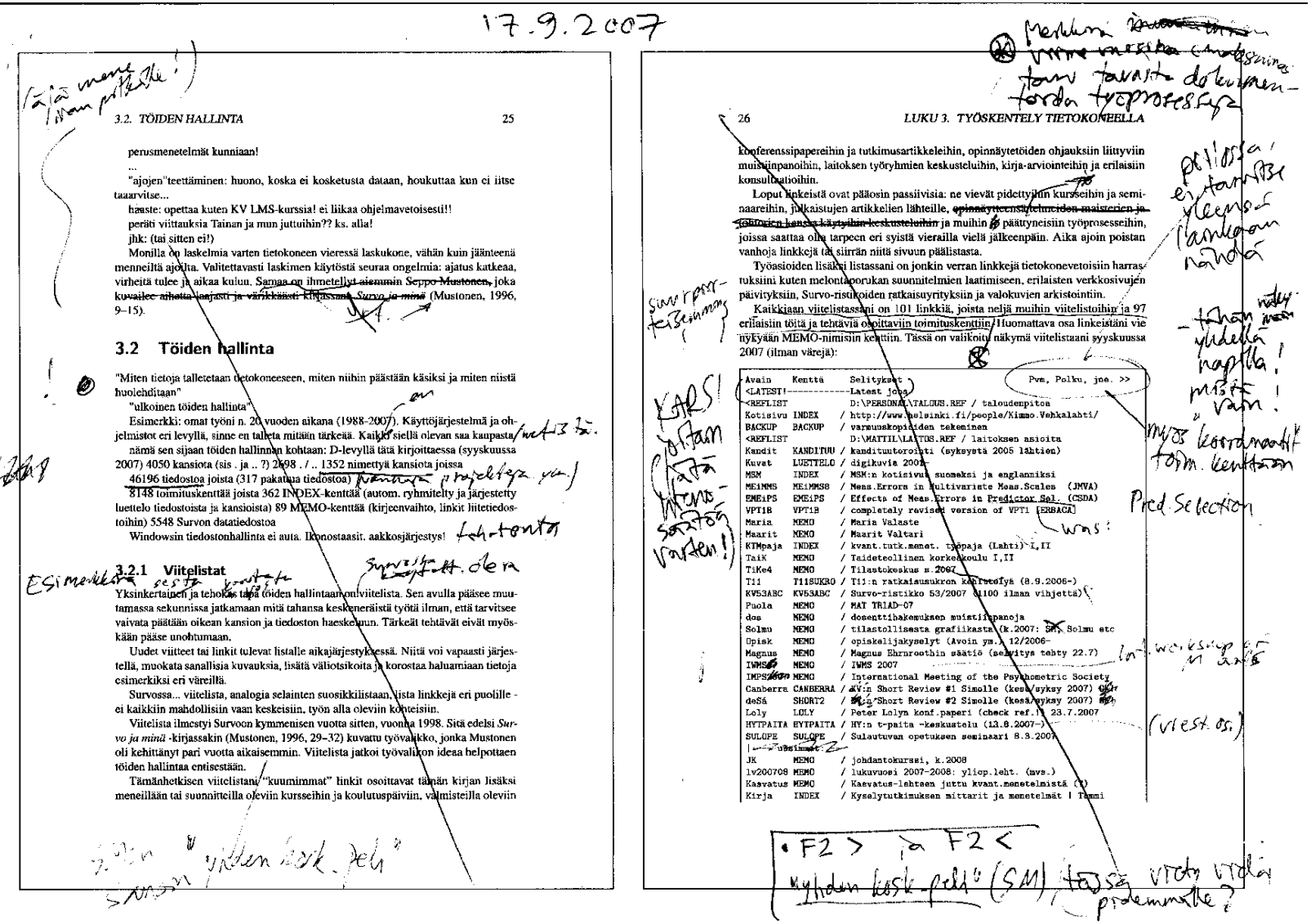
# 13.9.2007: Kirjan ja eräiden kurssien yhteydet

Nämä muistiinpanot (kuten monet muutkin) on tehty metrossa:




- Kirjoitustyö sai jatkua alkusyksyn, koska onnistuin järjestämään I periodin lähes vapaaksi opetuksesta.
- 18.9.2007 kansleri nimitti minut **soveltavan tilastotieteen dosentiksi**.

# 17.9.2007: Töiden "hallintaa" (?!)



3.2. TÖIDEN HALLINTA 25

perusmenetelmät kunniaan!

...

"ajojen"teettäminen: huono, koska ei kosketusta dataan, houkuttaa kun ei iitse taaarvitse...

haaste: opettaa kuten KV LMS-kurssia! ei liikaa ohjelmavetoisesti!!

peräti viittauksia Tainan ja mun juttuihin?? ks. alla!

jhk: (tai sitten ei!)

Monilla on laskelmia varten tietokoneen vieressä laskukone, vähän kuin jäänteenä menneiltä ajoilta. Valitettavasti laskimen käytöstä seuraa ongelmia: ajatus katkeaa, virheitä tulee ja aikaa kuluu. Samaa on ihmetellyt aiemmin Seppo Mustonen, joka kuvailee aihetta lanjasti ja värikkäästi kirjassaan *Survo ja minä* (Mustonen, 1996, 9–15).

## 3.2 Töiden hallinta

"Miten tietoja talletetaan tietokoneeseen, miten niihin päästään käsiksi ja miten niistä huolehditaan"

"ulkoinen töiden hallinta"

Esimerkki: omat työni n. 20 vuoden aikana (1988–2007). Käyttöjärjestelmä ja ohjelmistot eri levyllä, sinne en talleta mitään tärkeää. Kaikki siellä olevan saa kaupasta.

nämä sen sijaan töiden hallinnan kohtaan: D-levyllä tätä kirjoittaessa (syyskuussa 2007) 4050 kansiota (sis . ja .. ?) 2698 . / .. 1352 nimettyä kansiota joissa

46196 tiedostoa (317 pakattua tiedostoa)

8148 toimituskenttää joista 362 INDEX-kenttää (autom. ryhmitelty ja järjestetty luettelo tiedostoista ja kansioista) 89 MEMO-kenttää (kirjeenvaihto, linkit liitetiedostoihin) 5548 Survon datatiedostoa

Windowsin tiedostonhallinta ei auta. Ikonostaasit. aakkosjärjestys!

### 3.2.1 Viitelistat

Yksinkertainen ja tehokas tapa töiden hallintaan on viitelista. Sen avulla pääsee muutamassa sekunnissa jatkamaan mitä tahansa keskeneräistä työtä ilman, että tarvitsee vaivata päätään oikean kansion ja tiedoston haeskelemun. Tärkeät tehtävät eivät myöskään pääse unohtumaan.

Uudet viitteet tai linkit tulevat listalle aikajärjestyksessä. Niitä voi vapaasti järjestellä, muokata sanallisia kuvauksia, lisätä väliotsikoita ja korostaa haluamiaan tietoja esimerkiksi eri väreillä.

Survossa... viitelista, analogia selainten suosikkilistaan, lista linkkejä eri puolille - ei kaikkiin mahdollisiin vaan keskeisiin, työn alla oleviin kohteisiin.

Viitelista ilmestyi Survoon kymmenisen vuotta sitten, vuonna 1998. Sitä edelsi *Survo ja minä* -kirjassakin (Mustonen, 1996, 29–32) kuvattu työvalikko, jonka Mustonen oli kehittänyt pari vuotta aikaisemmin. Viitelista jatkoi työvalikon ideaa helpottaen töiden hallintaa entisestään.

Tämänhetkisen viitelistani "kuumimmat" linkit osoittavat tämän kirjan lisäksi meneillään tai suunniteilla oleviin kursseihin ja koulutuspäiviin, valmisteilla oleviin

26 LUKU 3. TYÖSKENTELY TIETOKONEELLA

konferenssipapereihin ja tutkimusartikkeleihin, opinnäytetöiden ohjauksiin liittyviin muistiinpanoihin, laitoksen työryhmien keskusteluihin, kirja-arviointeihin ja erilaisiin konsultaatioihin.

Loput linkeistä ovat pääosin passiivisia: ne vievät pidettyjihin kursseihin ja seminaareihin, julkaistujen artikkelien lähteille, opinnäytetöiden maisterien ja tohtorien kanssa käytyihin keskusteluihin ja muihin päätyneisiin työprosesseihin, joissa saattaa olla tarpeen eri syistä vierailla vielä jälkeenpäin. Aika ajoin poistan vanhoja linkkejä tai siirrän niitä sivuun päälistasta.

Työasioiden lisäksi listassani on jonkin verran linkkejä tietokonevetoisiin harrastuksiini kuten melontaporukan suunnitelmien laatimiseen, erilaisten verkkosivujen päivityksiin, Survo-ristikoiden ratkaisuyrityksiin ja valokuvien arkistointiin.

Kaikkiaan viitelistassani on 101 linkkiä, joista neljä muihin viitelistoihin ja 97 erilaisiin töitä ja tehtäviä ospittaviin toimituskenttiin. Huomattava osa linkeistäni vie nykyään MEMO-nimisiin kenttiin. Tässä on valikoitu näkymä viitelistaani syyskuussa 2007 (ilman värejä):

```
Avain    Kenttä    Selitykset                                  Pvm, Polku, jne. >>
<LATEST!-----------Latest jobs
<REFLIST            D:\PERSONAL\TALOUS.REF / taloudenpito
Kotisivu INDEX     / http://www.helsinki.fi/people/Kimmo.Vehkalahti/
BACKUP   BACKUP    / varmuuskopiiden tekeminen
<REFLIST            D:\MATTIL\LAITOS.REF / laitoksen asioita
Kandit   KANDITUU  / kanditutkotoimi (syksystä 2005 lähtien)
Kuvat    LUETTELO  / digikuvia 2006
MSM      INDEX     / MSM:n kotisivut suomeksi ja englanniksi
MEIMMS   MEIMMS8   / Meas.Errors in Multivariate Meas.Scales  (JMVA)
EMEiPS   EMEiPS    / Effects of Meas.Errors in Predictor Sel. (CSDA)
VPT1B    VPT1B     / completely revised version of VPT1 [ERBACA]
Maria    MEMO      / Maria Valaste
Maarit   MEMO      / Maarit Valtari
KTMpaja  INDEX     / kvant.tutk.menet. työpaja (Lahti) I,II
TaiK     MEMO      / Taideteollinen korkeakoulu I,II
TiKe4    MEMO      / Tilastokeskus n.2007
T11      T11SUKRO  / T11:n ratkaisusukron kehittely (8.9.2006-)
KV53ABC  KV53ABC   / Survo-ristikko 53/2007 (4100 ilman vihjettä)
Puola    MEMO      / MAT TRIAD-07
dos      MEMO      / dosenttihakemuksen muistiinpanoja
Solmu    MEMO      / tilastollisesta grafiikasta (k.2007: SM, Solmu etc
Opisk    MEMO      / opiskelijakyselyt (Avoin ym.) 12/2006-
Magnus   MEMO      / Magnus Ehrnroothin säätiö (selvitys tehty 22.7)
IWMS     MEMO      / IWMS 2007
IMPS     MEMO      / International Meeting of the Psychometric Society
Canberra CANBERRA  / KV:n Short Review #1 Simolle (kesä/syksy 2007)
deSä     SHORT2    / KV:n Short Review #2 Simolle (kesä/syksy 2007)
Loly     LOLY      / Peter Lolym konf.paperi (check ref.!) 23.7.2007
HYTPAITA HYTPAITA  / HY:n t-paita -keskustelu (13.8.2007-)
SULOPE   SULOPE    / Sulautuvan opetuksen seminaari 8.8.2007
|-----Tulevat:
JK       MEMO      / johdantokurssi, k.2008
lv200708 MEMO      / lukuvuosi 2007-2008: yliop.leht. (mvs.)
Kasvatus MEMO      / Kasvatus-lehtisen juttu kvant.menetelmistä (?)
Kirja    INDEX     / Kyselytutkimuksen mittarit ja menetelmät | Tammi
```

- Tästä piti tulla luku 3, osa siirtyi myöhemmin liitteeksi A...

# 16.-17.10.2007: Tilastokeskuksessa kouluttamassa (taas)

Käsillä oleva materiaali jäi viime talven aikana lopullisesti työstövaiheeseen—hyvästä syystä: olen kirjoittamassa uutta oppikirjaa, enkä enää jatka tämän monisteen päivittämistä. Kirjani työnimi on **Kyselytutkimuksen mittarit ja menetelmät**, ja sen on tarkoitus ilmestyä vuonna 2008.

Valtaosa tässä olevasta materiaalista on suoraan vuonna 2002 Metlan (Metsäntutkimuslaitos) kurssille tekemästäni monisteesta, jonka laitoin samana vuonna verkkoon saataville. Monisteesta on tullut varsin suosittu, ehkä siksi että näistä asioista on edelleen kirjoitettu melko vähän suomeksi. Onkin ollut mukava huomata, että monet ovat löytäneet monisteeni verkosta ja hyödyntäneet sitä opinnäytetöissään.

Ehdin syksyllä 2006 aloittaa monisteen mittaamista ja faktorianalyysia koskevan osan uudistamisen, mutta loppuosa on yhä entisessä muodossaan aivan sellaisenaan. Tarkoitukseni on ollut monipuolistaa esimerkkejä ja lisätä tarkempia viittauksia eri ohjelmistoihin.

Palautetta kurssilaisilta (anonyymisti):

- "Loistava 'pedagogi', on ilahduttavaa kuunnella luennoitsijaa, joka tietää, mistä puhuu ja siten, että sen ymmärtää."
- "Kiitos viehättävistä, tieteen maailmasta kertovista esimerkeistä! Kaikesta näkyi, että sinulla on 'professori-tasoiset', laajat tiedot omasta alastasi. Lämmin kiitos!"

# 26.11.2007: ”Työskentely tietokoneella” – vaiheessa

26.11.2007

LUKU 3

## Työskentely tietokoneella

Olipa kerran aika, jolloin tilastollisia analyyseja tehtiin kynällä ja paperilla. Työskentely oli vaivalloista ja vei tavattomasti aikaa. Tutkijoiden ja kokonaisten tutkimusryhmien aika meni sellaisten tulosten laskemiseen, jotka nyt saa muutaman sekunnissa nappia painamalla. Vähitellen avuksi tuli ensimmäiset laskukoneita, mutta ennen tietokoneiden aikakautta monista menetelmistä voitiin käytännössä vain haaveilla, vaikka ne teoriassa oli tunnettu jo vuosikymmeniä.

Tietokoneet ovat mullistaneet kaiken moneen kertaan, mutta laskemisen perinne heijastuu edelleen muun muassa menetelmäkirjoista, joissa saatetaan neuvoa, millaisella kaavalla korrelaatiokerroin lasketaan. Kaavamaiset neuvot ovat nykyisin jokseenkin turhia. Ei tilastollisia analyyseja tehdä enää laskukoneilla—kynästä ja paperista nyt puhumattakaan.

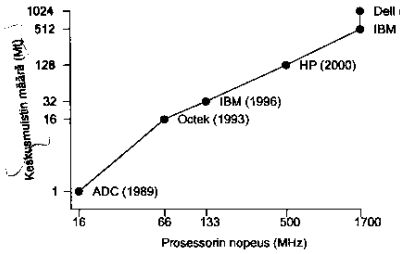
Laskeminen kuuluu tietokoneelle ja ohjelmistoille, eikä niissäkään käytävä menetelmäkirjojen kaavoja. Niiden sijaan tarvitaan tietämystä ohjelmoinnista, numeerisesta laskennasta ja jopa prosessoriden arkkitehtuurista. Näistä asioista ei tarvitse menetelmien soveltajan tietää mitään. Soveltaja voi keskittyä tehtäviin, joihin kone ei kykene: tutkimuksen ja analyysien suunnitteluun sekä tulosten tulkintaan ja arviointiin. Näihin voi toki sisältyä pieniä laskelmia, mutta niissä ei tarvita laskukaavoja vaan käteviä tapoja tehdä laskelmia.

Kunnolliset työvälineet, taito käyttää niitä oikein ja hyvät työskentelytavat tunnuttaan usein yhdistävän lähinnä käsityöläisammatteihin, mutta ne ovat eivät yhtä lailla läsnä tutkimustyyppisessä työssä, ja vaikuttavat niin työn sujuvuuteen ja tehokkuuteen kuin tulosten laatuunkin. Sanonta "työ tekijäänsä opettaa" pätee yhtä lailla, kunhan käytettävät välineet tukevat oppimista ja omien työtapojen kehittämistä.

Aluksi tarkastellaan tutkijan työvälineitä, tietokoneita ja ennen kaikkea ohjelmistoja. Siitä edetään töiden hallintaan, ja lopuksi pohdiltaan, mitä tarkoittaa dokumentoiva työskentelytapa ja mitä hyötyä siitä on erityisesti kyselytutkimuksen kannalta.

29






**Kuva 3.1.** Koneideni prosessoritehon ja keskusmuistin kehitys.

### 3.1 Tietokone ja ohjelmistot

Tietokone on pitkään ollut tutkijan työvälineenä itsestäänselvyys. Kyselytutkimuksenkin työvaiheista siitä voi tehdä melkein kaikki ja paljon enemmän. Koneet ovat tulleet jatkuvasti tehokkaammiksi mutta samalla pienemmiksi ja siirtyneet vaivihkaa työpöydiltä reppuihin ja salkkuihin.

Itselläni on ollut vuodesta 1989 kuusi tietokonetta, kaksi ensimmäistä pöytäkoneita ja loput kannettavia. Uuden koneen myötä laskennalliset tehot kasvavat (katso kuva 3.1), mutta työtehoa ei voi mitata megahertseinä tai gigatavuina. Oleellisempaa on tarkoituksenmukaisten ohjelmistojen käyttö ja omien töiden kätevä hallinta.

Tietokoneen merkillä tai mallilla ei ole juuri mitään merkitystä, ei myöskään käyttöjärjestelmällä. Joissain kohdin viittaan Windowsiin, jonka puitteissa itse enimmäkseen toimin, mutta sinänsä Windowsilla on vain vähäinen merkitys työskentelyssäni. Esittämäni periaatteet pätevät yhtä hyvin Linux-, Mac- ja muissakin järjestelmissä.

#### 3.1.1 Kirjan ohjelmistot

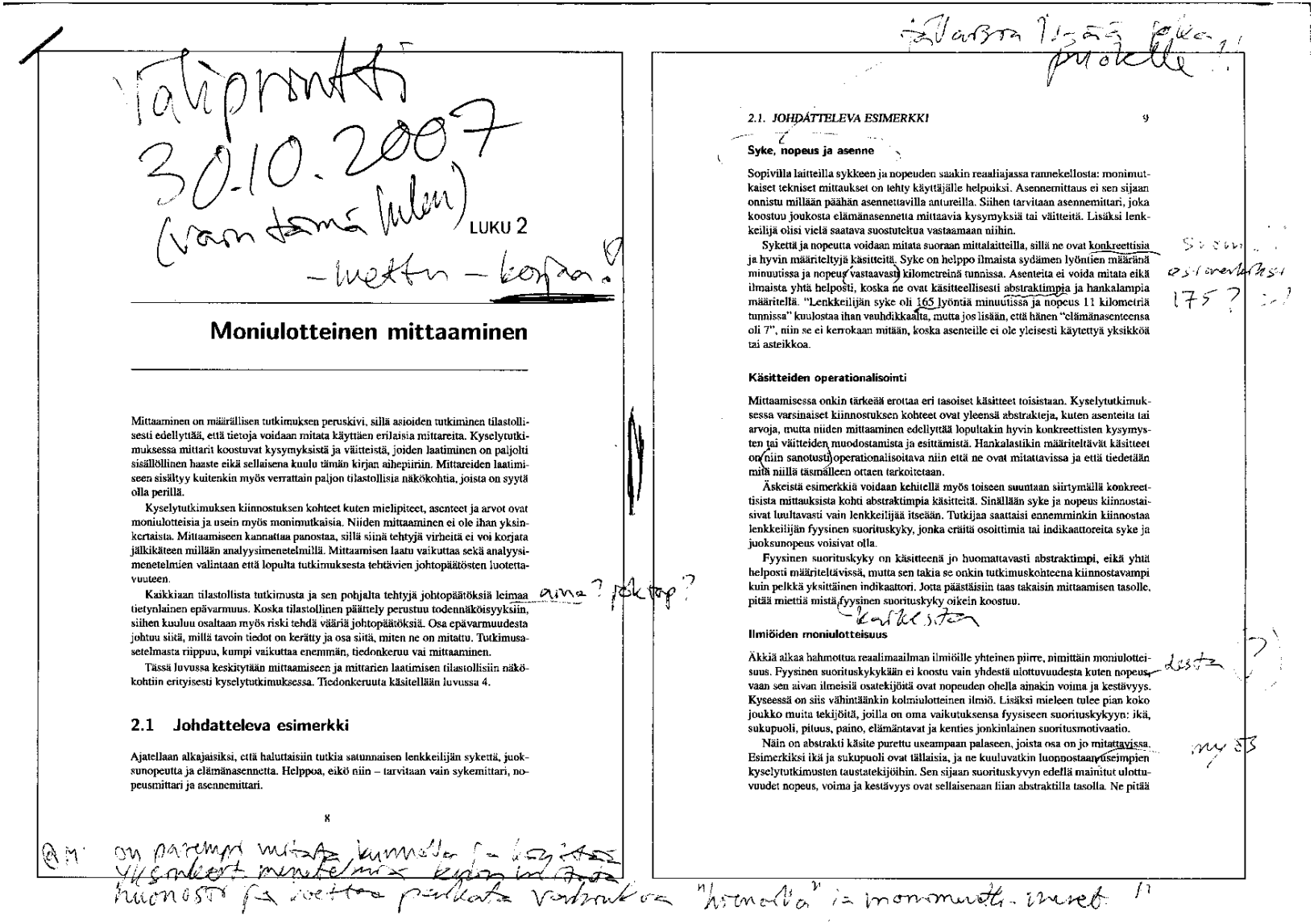
Ohjelmistojen käyttö on oleellinen osa tilastollisten menetelmien soveltamista. Niinpä tällaisen kirjankin käytännön hyöty jäisi vähäiseksi ilman ohjelmistoesimerkkejä. Sopivia ohjelmistoja on runsaasti, joten on pitänyt tehdä valintoja.

Valintani kirjan ohjelmistoiksi ovat Survo ja SPSS. Survo on suomalainen, professori Seppo Mustosen kehittämä tietojenkäsittely-ympäristö, ja SPSS on amerikkalainen, alunperin Norman H. Nien, C. Hadlai Hullin ja Dale H. Bentin kehittämät tilastollisten menetelmien ohjelmapaketti. Molempien ohjelmistojen ensimmäiset versiot ilmestyivät jo 1960-luvulla (Survo Systems, 2007; SPSS Inc., 2007).

Kirjassa käsiteltävien asioiden ymmärtäminen ei edellytä minkään ohjelmiston erityistä osaamista, koska pääosassa ovat menetelmät ja työtavat. Ohjelmistotekniset

- Luvusta 3 tuli myöhemmin liite A ja se lyheni huomattavasti.

# 30.10.2007: ”Moniulotteinen mittaaminen” – vaiheessa



Tahiprontti 30.10.2007 (vain tämä luku) – mutta – koko?

LUKU 2

## Moniulotteinen mittaaminen

Mittaaminen on määrällisen tutkimuksen peruskivi, sillä asioiden tutkiminen tilastollisesti edellyttää, että tietoja voidaan mitata käyttäen erilaisia mittareita. Kyselytutkimuksessa mittarit koostuvat kysymyksistä ja väitteistä, joiden laatiminen on paljolti sisällöllinen haaste eikä sellaisena kuulu tämän kirjan aihepiiriin. Mittareiden laatimiseen sisältyy kuitenkin myös verrattain paljon tilastollisia näkökohtia, joista on syytä olla perillä.

Kyselytutkimuksen kiinnostuksen kohteet kuten mielipiteet, asenteet ja arvot ovat moniulotteisia ja usein myös monimutkaisia. Niiden mittaaminen ei ole ihan yksinkertaista. Mittaamiseen kannattaa panostaa, sillä siinä tehtyjä virheitä ei voi korjata jälkikäteen millään analyysimenetelmillä. Mittaamisen laatu vaikuttaa sekä analyysimenetelmien valintaan että lopulta tutkimuksesta tehtävien johtopäätösten luotettavuuteen.

Kaikkiaan tilastollista tutkimusta ja sen pohjalta tehtyjä johtopäätöksiä leimaa tietynlainen epävarmuus. Koska tilastollinen päättely perustuu todennäköisyyksiin, siihen kuuluu osaltaan myös riski tehdä vääriä johtopäätöksiä. Osa epävarmuudesta johtuu siitä, millä tavoin tiedot on kerätty ja osa siitä, miten ne on mitattu. Tutkimusasetelmasta riippuu, kumpi vaikuttaa enemmän, tiedonkeruu vai mittaaminen.

Tässä luvussa keskitytään mittaamiseen ja mittarien laatimisen tilastollisiin näkökohtiin erityisesti kyselytutkimuksessa. Tiedonkeruuta käsitellään luvussa 4.

### 2.1 Johdatteleva esimerkki

Ajatellaan akkajaisiksi, että haluttaisiin tutkia satunnaisen lenkkeilijän sykettä, juoksunopeutta ja elämänasennetta. Helppoa, eikö niin – tarvitaan vain sykemittari, nopeusmittari ja asennemittari.

8

2.1. JOHDATTELEVA ESIMERKKI 9

**Syke, nopeus ja asenne**

Sopivilla laitteilla sykkeen ja nopeuden saakin reaaliajassa rannekellosta: monimutkaiset tekniset mittaukset on tehty käyttäjälle helpoiksi. Asennemittaus ei sen sijaan onnistu millään päähän asennettavilla antureilla. Siihen tarvitaan asennemittari, joka koostuu joukosta elämänasennetta mittaavia kysymyksiä tai väitteitä. Lisäksi lenkkeilijä olisi vielä saatava suostuteltua vastaamaan niihin.

Sykettä ja nopeutta voidaan mitata suoraan mittalaitteilla, sillä ne ovat konkreettisia ja hyvin määriteltyjä käsitteitä. Syke on helppo ilmaista sydämen lyöntien määränä minuutissa ja nopeus vastaavasti kilometreinä tunnissa. Asenteita ei voida mitata eikä ilmaista yhtä helposti, koska ne ovat käsitteellisesti abstraktimpia ja hankalampia määritellä. “Lenkkeilijän syke oli 165 lyöntiä minuutissa ja nopeus 11 kilometriä tunnissa” kuulostaa ihan vauhdikkaalta, mutta jos lisään, että hänen “elämänasenteensa oli 7”, niin se ei kerrokaan mitään, koska asenteille ei ole yleisesti käytettyä yksikköä tai asteikkoa.

**Käsitteiden operationalisointi**

Mittaamisessa onkin tärkeää erottaa eri tasoiset käsitteet toisistaan. Kyselytutkimuksessa varsinaiset kiinnostuksen kohteet ovat yleensä abstrakteja, kuten asenteita tai arvoja, mutta niiden mittaaminen edellyttää lopultakin hyvin konkreettisten kysymysten tai väitteiden muodostamista ja esittämistä. Hankalastikin määriteltävät käsitteet on (niin sanotusti) operationalisoitava niin että ne ovat mitattavissa ja että tiedetään mitä niillä täsmälleen ottaen tarkoitetaan.

Äskeistä esimerkkiä voidaan kehitellä myös toiseen suuntaan siirtymällä konkreettisista mittauksista kohti abstraktimpia käsitteitä. Sinällään syke ja nopeus kiinnostaisivat luultavasti vain lenkkeilijää itseään. Tutkijaa saattaisi ennemminkin kiinnostaa lenkkeilijän fyysinen suorituskyky, jonka eräitä osoittimia tai indikaattoreita syke ja juoksunopeus voisivat olla.

Fyysinen suorituskyky on käsitteenä jo huomattavasti abstraktimpi, eikä yhtä helposti määriteltävissä, mutta sen takia se onkin tutkimuskohteena kiinnostavampi kuin pelkkä yksittäinen indikaattori. Jotta päästäisiin taas takaisin mittaamisen tasolle, pitää miettiä mistä fyysinen suorituskyky oikein koostuu.

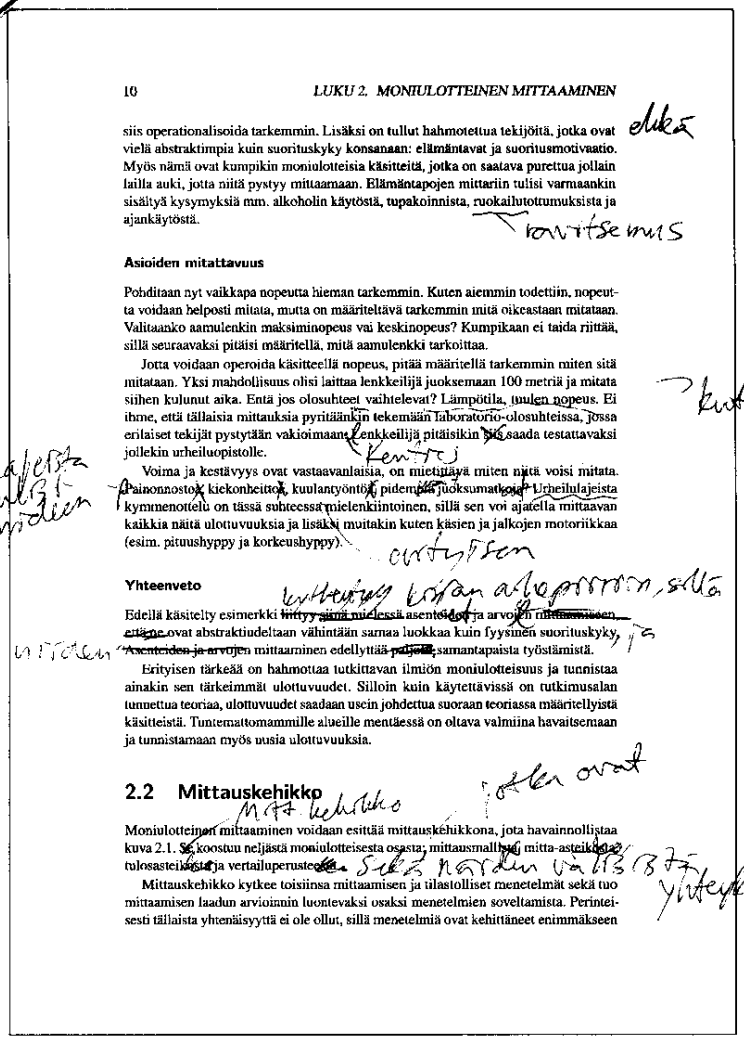
**Ilmiöiden moniulotteisuus**

Äkkiä alkaa hahmottua reaalimaailman ilmiöille yhteinen piirre, nimittäin moniulotteisuus. Fyysinen suorituskykykään ei koostu vain yhdestä ulottuvuudesta kuten nopeus, vaan sen aivan ilmeisiä osatekijöitä ovat nopeuden ohella ainakin voima ja kestävyys. Kyseessä on siis vähintäänkin kolmiulotteinen ilmiö. Lisäksi mieleen tulee pian koko joukko muita tekijöitä, joilla on oma vaikutuksensa fyysiseen suorituskykyyn: ikä, sukupuoli, pituus, paino, elämäntavat ja kenties jonkinlainen suoritusmotivaatio.

Näin on abstrakti käsite purettu useampaan palaseen, joista osa on jo mitattavissa. Esimerkiksi ikä ja sukupuoli ovat tällaisia, ja ne kuuluvatkin luonnostaan useimpien kyselytutkimusten taustatekijöihin. Sen sijaan suorituskyvyn edellä mainitut ulottuvuudet nopeus, voima ja kestävyys ovat sellaisenaan liian abstraktilla tasolla. Ne pitää

- Alun perin luku 2 sisälsi kaikki mittausasiat (!)...

# 30.10.2007: ”Moniulotteinen mittaaminen” – vaiheessa



10 LUKU 2. MONIULOTTEINEN MITTAAMINEN

siis operationalisoida tarkemmin. Lisäksi on tullut hahmotettua tekijöitä, jotka ovat vielä abstraktimpia kuin suorituskyky konsanaan: elämäntavat ja suoritusmotivaatio. Myös nämä ovat kumpikin moniulotteisia käsitteitä, jotka on saatava purettua jollain lailla auki, jotta niitä pystyy mittaamaan. Elämäntapojen mittariin tulisi varmaankin sisältyä kysymyksiä mm. alkoholin käytöstä, tupakoinnista, ruokailutotrumuksista ja ajankäytöstä.

**Asioiden mitattavuus**

Pohditaan nyt vaikkapa nopeutta hieman tarkemmin. Kuten aiemmin todettiin, nopeutta voidaan helposti mitata, mutta on määriteltävä tarkemmin mitä oikeastaan mitataan. Valitaanko aamulenkin maksiminopeus vai keskinopeus? Kumpikaan ei taida riittää, sillä seuraavaksi pitäisi määritellä, mitä aamulenkki tarkoittaa.

Jotta voidaan operoida käsitteellä nopeus, pitää määritellä tarkemmin miten sitä mitataan. Yksi mahdollisuus olisi laittaa lenkkeilijä juoksemaan 100 metriä ja mitata siihen kulunut aika. Entä jos olosuhteet vaihtelevat? Lämpötila, muulen nopeus. Ei ihme, että tällaisia mittauksia pyritäänkin tekemään laboratorio-olosuhteissa, jossa erilaiset tekijät pystytään vakioimaan. Lenkkeilijä pitäisikin ehkä saada testattavaksi jollekin urheiluopistolle.

Voima ja kestävyys ovat vastaavanlaisia, on mietittävä miten niitä voisi mitata. Painonnostoa, kiekonheittoa, kuulantyöntöä, pidempiä juoksumatkoja? Urheilulajeista kymmenottelu on tässä suhteessa mielenkiintoinen, sillä sen voi ajatella mittaavan kaikkia näitä ulottuvuuksia ja lisäksi moitakin kuten käsien ja jalkojen motoriikkaa (esim. pituushyppy ja korkeushyppy).

**Yhteenveto**

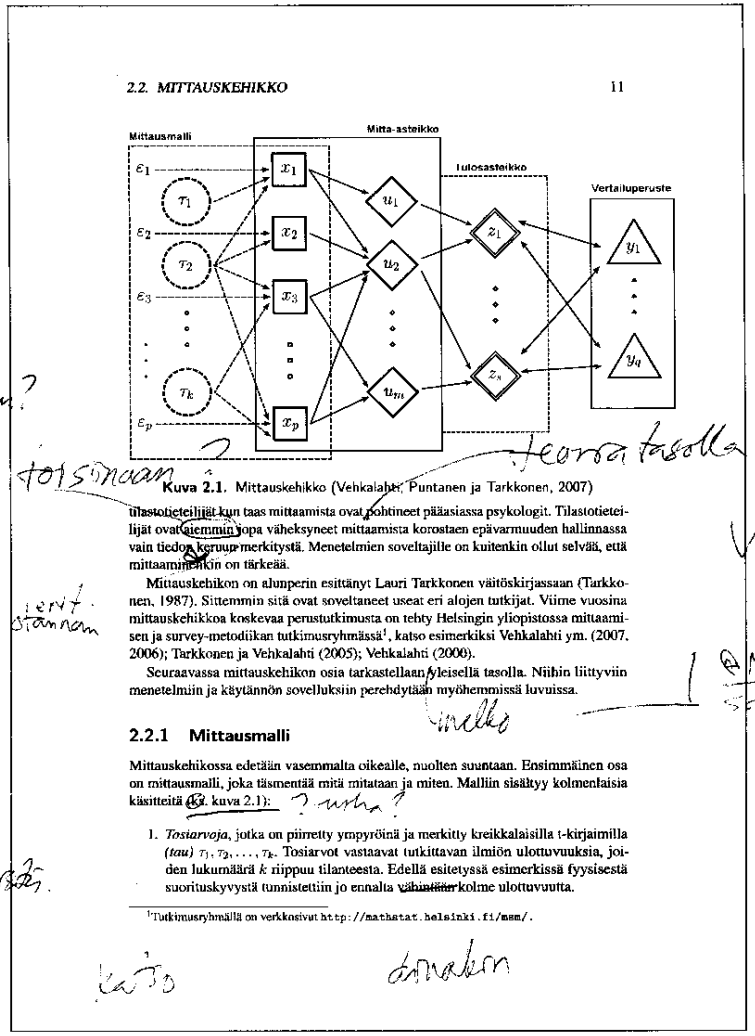
Edellä käsitelty esimerkki liittyy ajatukseen ja arvojen mittaamisen, ettäne ovat abstraktiudeltaan vähintään samaa luokkaa kuin fyysinen suorituskyky, ja Asenteiden ja arvojen mittaaminen edellyttää paljon samantapaista työstämistä.

Erityisen tärkeää on hahmottaa tutkittavan ilmiön moniulotteisuus ja tunnistaa ainakin sen tärkeimmät ulottuvuudet. Silloin kuin käytettävissä on tutkimusalan tunnettua teoriaa, ulottuvuudet saadaan usein johdettua suoraan teoriassa määrittelyistä käsitteistä. Tuntemattomammille alueille mentäessä on oltava valmiina havaitsemaan ja tunnistamaan myös uusia ulottuvuuksia.

## 2.2 Mittauskehikko

Moniulotteinen mittaaminen voidaan esittää mittauskehikkona, jota havainnollistaa kuva 2.1. Se koostuu neljästä moniulotteisesta osasta: mittausmallista, mitta-asteikosta, tulosasteikosta ja vertailuperusteesta.

Mittauskehikko kytkee toisiinsa mittaamisen ja tilastolliset menetelmät sekä tuo mittaamisen laadun arvioinnin luontevaksi osaksi menetelmien soveltamista. Perinteisesti tällaista yhtenäisyyttä ei ole ollut, sillä menetelmiä ovat kehittäneet enimmäkseen



2.2. MITTAUSKEHIKKO 11



**Kuva 2.1.** Mittauskehikko (Vehkalahti, Puntanen ja Tarkkonen, 2007)

tilastotieteilijät kun taas mittaamista ovat pohtineet pääasiassa psykologit. Tilastotieteilijät ovat aiemmin jopa väheksyneet mittaamista korostaen epävarmuuden hallinnassa vain tiedon keruun merkitystä. Menetelmien soveltajille on kuitenkin ollut selvää, että mittaaminenkin on tärkeää.

Mittauskehikon on alunperin esittänyt Lauri Tarkkonen väitöskirjassaan (Tarkkonen, 1987). Siittemmin sitä ovat soveltaneet useat eri alojen tutkijat. Viime vuosina mittauskehikkoa koskevaa perustutkimusta on tehty Helsingin yliopistossa mittaamisen ja survey-metodiikan tutkimusryhmässä[1], katso esimerkiksi Vehkalahti ym. (2007, 2006); Tarkkonen ja Vehkalahti (2005); Vehkalahti (2000).

Seuraavassa mittauskehikon osia tarkastellaan yleisellä tasolla. Niihin liittyviin menetelmiin ja käytännön sovelluksiin perehdytään myöhemmissä luvuissa.

### 2.2.1 Mittausmalli

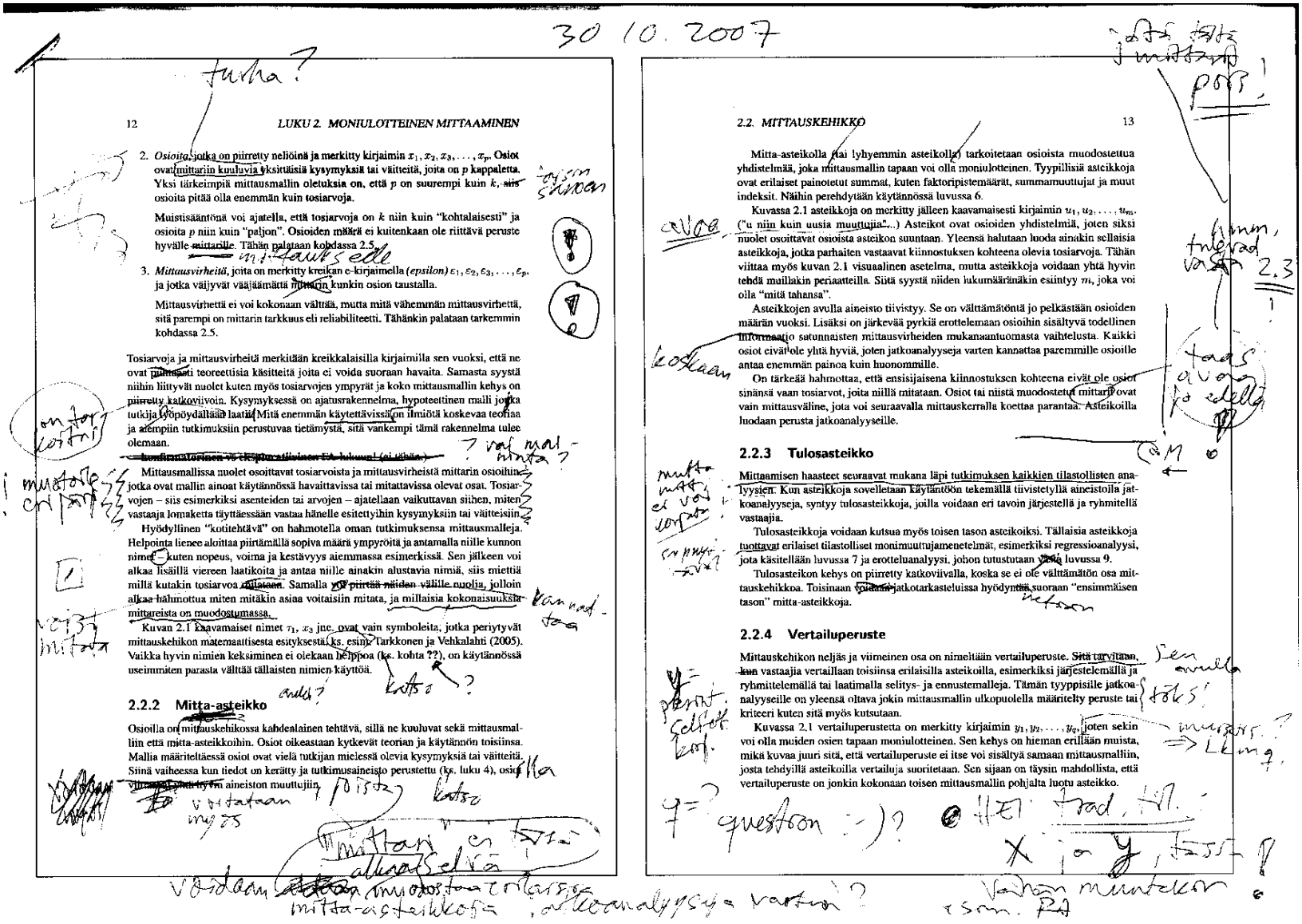
Mittauskehikossa edetään vasemmalta oikealle, nuolten suuntaan. Ensimmäinen osa on mittausmalli, joka täsmentää mitä mitataan ja miten. Malliin sisältyy kolmenlaisia käsitteitä (kt. kuva 2.1):

1. *Tosiarvoja*, jotka on piirretty ympyröinä ja merkitty kreikkalaisilla t-kirjaimilla (*tau*) $\tau_1, \tau_2, \ldots, \tau_k$. Tosiarvot vastaavat tutkittavan ilmiön ulottuvuuksia, joiden lukumäärä $k$ riippuu tilanteesta. Edellä esitetyssä esimerkissä fyysisestä suorituskyvystä tunnistettiin jo ennalta vähintään kolme ulottuvuutta.

[1] Tutkimusryhmällä on verkkosivut http://mathstat.helsinki.fi/msm/.

- ...jopa kehikon sellaisenaan (kirjan sivulla 11!).

# 30.10.2007: "Moniulotteinen mittaaminen" – vaiheessa



12 LUKU 2. MONIULOTTEINEN MITTAAMINEN

2. *Osioita*, jotka on piirretty neliöinä ja merkitty kirjaimin $x_1, x_2, x_3, \ldots, x_p$. Osiot ovat mittarin kuuluvia yksitäisiä kysymyksiä tai väitteitä, joita on $p$ kappaletta. Yksi tärkeimpiä mittausmallin oletuksia on, että $p$ on suurempi kuin $k$, siis osioita pitää olla enemmän kuin tosiarvoja.

   Muistisääntönä voi ajatella, että tosiarvoja on $k$ niin kuin "kohtalaisesti" ja osioita $p$ niin kuin "paljon". Osioiden määrä ei kuitenkaan ole riittävä peruste hyvälle mittarille. Tähän palataan kohdassa 2.5.

3. *Mittausvirheitä*, joita on merkitty kreikan e-kirjaimella *(epsilon)* $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_p$, ja jotka väijyvät vääjäämättä mittarin kunkin osion taustalla.

   Mittausvirhettä ei voi kokonaan välttää, mutta mitä vähemmän mittausvirhettä, sitä parempi on mittarin tarkkuus eli reliabiliteetti. Tähänkin palataan tarkemmin kohdassa 2.5.

Tosiarvoja ja mittausvirheitä merkitään kreikkalaisilla kirjaimilla sen vuoksi, että ne ovat puhtaasti teoreettisia käsitteitä joita ei voida suoraan havaita. Samasta syystä niihin liittyvät nuolet kuten myös tosiarvojen ympyrät ja koko mittausmallin kehys on piirretty katkoviivoin. Kysymyksessä on ajatusrakennelma, hypoteettinen malli jotka tutkija työpöydällään laatii. Mitä enemmän käytettävissä on ilmiötä koskevaa teoriaa ja aiempiin tutkimuksiin perustuvaa tietämystä, sitä vankempi tämä rakennelma tulee olemaan.

Mittausmallissa nuolet osoittavat tosiarvoista ja mittausvirheistä mittarin osioihin, jotka ovat mallin ainoat käytännössä havaittavissa tai mitattavissa olevat osat. Tosiarvojen – siis esimerkiksi asenteiden tai arvojen – ajatellaan vaikuttavan siihen, miten vastaaja lomaketta täyttäessään vastaa hänelle esitettyihin kysymyksiin tai väitteisiin.

Hyödyllinen "kotitehtävä" on hahmotella oman tutkimuksensa mittausmalleja. Helpointa lienee aloittaa piirtämällä sopiva määrä ympyröitä ja antamalla niille kunnon nimet – kuten nopeus, voima ja kestävyys aiemmassa esimerkissä. Sen jälkeen voi alkaa lisäillä viereen laatikoita ja antaa niille ainakin alustavia nimiä, siis miettiä millä kutakin tosiarvoa mitataan. Samalla voi piirtää näiden välille nuolia, jolloin alkaa hahmottua miten mitäkin asiaa voitaisiin mitata, ja millaisia kokonaisuuksia mittareista on muodostumassa.

Kuvan 2.1 kaavamaiset nimet $\tau_1$, $x_3$ jne. ovat vain symboleita, jotka periytyvät mittauskehikon matemaattisesta esityksestä (ks. esim. Tarkkonen ja Vehkalahti (2005). Vaikka hyvin nimien keksiminen ei olekaan helppoa (ks. kohta ??), on käytännössä useimmiten parasta välttää tällaisten nimien käyttöä.

### 2.2.2 Mitta-asteikko

Osioilla on mittauskehikossa kahdenlainen tehtävä, sillä ne kuuluvat sekä mittausmalliin että mitta-asteikkoihin. Osiot oikeastaan kytkevät teorian ja käytännön toisiinsa. Mallia määriteltäessä osiot ovat vielä tutkijan mielessä olevia kysymyksiä tai väitteitä. Siinä vaiheessa kun tiedot on kerätty ja tutkimusaineisto perustettu (ks. luku 4), osiot viittaavat yhtä hyvin aineiston muuttujiin.

2.2. MITTAUSKEHIKKO 13

Mitta-asteikolla (tai lyhyemmin asteikolla) tarkoitetaan osioista muodostetua yhdistelmää, joka mittausmallin tapaan voi olla moniulotteinen. Tyypillisiä asteikkoja ovat erilaiset painotetut summat, kuten faktoripistemäärät, summamuuttujat ja muut indeksit. Näihin perehdytään käytännössä luvussa 6.

Kuvassa 2.1 asteikkoja on merkitty jälleen kaavamaisesti kirjaimin $u_1, u_2, \ldots, u_m$. ("u niin kuin uusia muuttujia"...) Asteikot ovat osioiden yhdistelmiä, joten siksi nuolet osoittavat osioista asteikon suuntaan. Yleensä halutaan luoda ainakin sellaisia asteikkoja, jotka parhaiten vastaavat kiinnostuksen kohteena olevia tosiarvoja. Tähän viittaa myös kuvan 2.1 visuaalinen asetelma, mutta asteikkoja voidaan yhtä hyvin tehdä muillakin periaatteilla. Siitä syystä niiden lukumääränäkin esiintyy $m$, joka voi olla "mitä tahansa".

Asteikkojen avulla aineisto tiivistyy. Se on välttämätöntä jo pelkästään osioiden määrän vuoksi. Lisäksi on järkevää pyrkiä erottelemaan osioihin sisältyvä todellinen informaatio satunnaisten mittausvirheiden mukanaantuomasta vaihtelusta. Kaikki osiot eivät ole yhtä hyviä, joten jatkoanalyyseja varten kannattaa paremmille osioille antaa enemmän painoa kuin huonommille.

On tärkeää hahmottaa, että ensisijaisena kiinnostuksen kohteena eivät ole osiot sinänsä vaan tosiarvot, joita niillä mitataan. Osiot tai niistä muodostetut mittarit ovat vain mittausväline, jota voi seuraavalla mittauskerralla koettaa parantaa. Asteikoilla luodaan perusta jatkoanalyyseille.

### 2.2.3 Tulosasteikko

Mittaamisen haasteet seuraavat mukana läpi tutkimuksen kaikkien tilastollisten analyysien. Kun asteikkoja sovelletaan käytäntöön tekemällä tiivistetyllä aineistolla jatkoanalyyseja, syntyy tulosasteikkoja, joilla voidaan eri tavoin järjestellä ja ryhmitellä vastaajia.

Tulosasteikkoja voidaan kutsua myös toisen tason asteikoiksi. Tällaisia asteikkoja tuottavat erilaiset tilastolliset monimuuttujamenetelmät, esimerkiksi regressioanalyysi, jota käsitellään luvussa 7 ja erotteluanalyysi, johon tutustutaan luvussa 9.

Tulosasteikon kehys on piirretty katkoviivalla, koska se ei ole välttämätön osa mittauskehikkoa. Toisinaan voidaan jatkotarkasteluissa hyödyntää suoraan "ensimmäisen tason" mitta-asteikkoja.

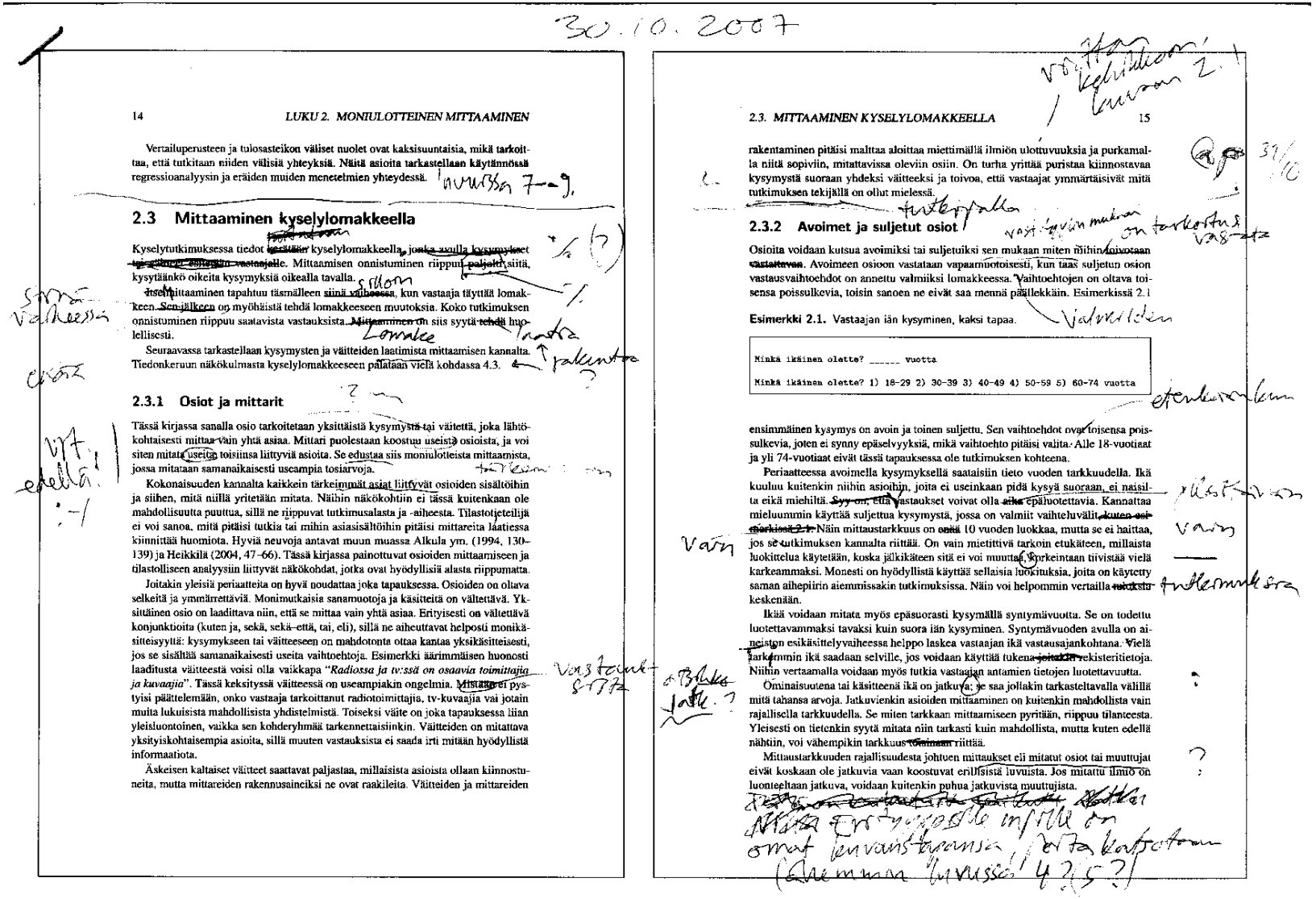
### 2.2.4 Vertailuperuste

Mittauskehikon neljäs ja viimeinen osa on nimeltään vertailuperuste. Sitä tarvitaan, kun vastaajia vertaillaan toisiinsa erilaisilla asteikoilla, esimerkiksi järjestelemällä ja ryhmittelemällä tai laatimalla selitys- ja ennustemalleja. Tämän tyyppisille jatkoanalyyseille on yleensä oltava jokin mittausmallin ulkopuolella määritelty peruste tai kriteeri kuten sitä myös kutsutaan.

Kuvassa 2.1 vertailuperustetta on merkitty kirjaimin $y_1, y_2, \ldots, y_q$, joten sekin voi olla muiden osien tapaan moniulotteinen. Sen kehys on hieman erillään muista, mikä kuvaa juuri sitä, että vertailuperuste ei itse voi sisältyä samaan mittausmalliin, josta tehdyillä asteikoilla vertailuja suoritetaan. Sen sijaan on täysin mahdollista, että vertailuperuste on jonkin kokonaan toisen mittausmallin pohjalta luotu asteikko.

- Kaikki printtini näyttivät jossain vaiheessa tältä...

# 30.10.2007: ”Moniulotteinen mittaaminen” – vaiheessa



14 LUKU 2. MONIULOTTEINEN MITTAAMINEN

Vertailuperusteen ja tulosasteikon väliset nuolet ovat kaksisuuntaisia, mikä tarkoittaa, että tutkitaan niiden välisiä yhteyksiä. Näitä asioita tarkastellaan käytännössä regressioanalyysin ja eräiden muiden menetelmien yhteydessä.

## 2.3 Mittaaminen kyselylomakkeella

Kyselytutkimuksessa tiedot kerätään kyselylomakkeella, joka voulla kysymykset tai väitteet annettiin vastaajelle. Mittaamisen onnistuminen riippuu paljolti siitä, kysytäänkö oikeita kysymyksiä oikealla tavalla.

Itse mittaaminen tapahtuu tämänlaisen siinä vaiheessa, kun vastaaja täyttää lomakkeen. Sen jälkeen on myöhäistä tehdä lomakkeeseen muutoksia. Koko tutkimuksen onnistuminen riippuu saatavista vastauksista. Mittaaminen on siis syytä tehdä huolellisesti.

Seuraavassa tarkastellaan kysymysten ja väitteiden laatimista mittaamisen kannalta. Tiedonkeruun näkökulmasta kyselylomakkeeseen palataan vielä kohdassa 4.3.

### 2.3.1 Osiot ja mittarit

Tässä kirjassa sanalla osio tarkoitetaan yksittäistä kysymystä tai väitettä, joka lähtökohtaisesti mittaa vain yhtä asiaa. Mittari puolestaan koostuu useista osioista, ja voi siten mitata useitakin toisiinsa liittyviä asioita. Se edustaa siis moniulotteista mittaamista, jossa mitataan samanaikaisesti useampia tosiarvoja.

Kokonaisuuden kannalta kaikkein tärkeimmät asiat liittyvät osioiden sisältöihin ja siihen, mitä niillä yritetään mitata. Näihin näkökohtiin ei tässä kuitenkaan ole mahdollisuutta puuttua, sillä ne riippuvat tutkimusalasta ja -aiheesta. Tilastotieteilijä ei voi sanoa, mitä pitäisi tutkia tai mihin asiasisältöihin pitäisi mittareita laatiessa kiinnittää huomiota. Hyviä neuvoja antavat muun muassa Alkula ym. (1994, 130–139) ja Heikkilä (2004, 47–66). Tässä kirjassa painottuvat osioiden mittaamiseen ja tilastolliseen analyysiin liittyvät näkökohdat, jotka ovat hyödyllisiä alasta riippumatta.

Joitakin yleisiä periaatteita on hyvä noudattaa joka tapauksessa. Osioiden on oltava selkeitä ja ymmärrettäviä. Monimutkaisia sanamuotoja ja käsitteitä on vältettävä. Yksittäinen osio on laadittava niin, että se mittaa vain yhtä asiaa. Erityisesti on vältettävä konjunktioita (kuten ja, sekä, sekä–että, tai, eli), sillä ne aiheuttavat helposti monikäsitteisyyttä: kysymykseen tai väitteeseen on mahdotonta ottaa kantaa yksikäsitteisesti, jos se sisältää samanaikaisesti useita vaihtoehtoja. Esimerkki äärimmäisen huonosti laadituista väitteestä voisi olla vaikkapa *“Radiossa ja tv:ssä on osaavia toimittajia ja kuvaajia”*. Tässä keksityssä väitteessä on useampiakin ongelmia. Mittaan ei pystyisi päättelemään, onko vastaaja tarkoittanut radiotoimittajia, tv-kuvaajia vai jotain muita lukuisista mahdollisista yhdistelmistä. Toiseksi väite on joka tapauksessa liian yleisluontoinen, vaikka sen kohderyhmää tarkennettaisiinkin. Väitteiden on mitattava yksityiskohtaisempia asioita, sillä muuten vastauksista ei saada irti mitään hyödyllistä informaatiota.

Äskeisen kaltaiset väitteet saattavat paljastaa, millaisista asioista ollaan kiinnostuneita, mutta mittareiden rakennusainekseksi ne ovat raakileita. Väitteiden ja mittareiden

2.3. MITTAAMINEN KYSELYLOMAKKEELLA 15

rakentaminen pitäisi malttaa aloittaa miettimällä ilmiön ulottuvuuksia ja purkamalla niitä sopiviin, mitattavissa oleviin osiin. On turha yrittää puristaa kiinnostavaa kysymystä suoraan yhdeksi väitteeksi ja toivoa, että vastaajat ymmärtäisivät mitä tutkimuksen tekijällä on ollut mielessä.

### 2.3.2 Avoimet ja suljetut osiot

Osioita voidaan kutsua avoimiksi tai suljetuiksi sen mukaan miten niihin toivotaan vastattavan. Avoimeen osioon vastataan vapaamuotoisesti, kun taas suljetun osion vastausvaihtoehdot on annettu valmiiksi lomakkeessa. Vaihtoehtojen on oltava toisensa poissulkevia, toisin sanoen ne eivät saa mennä päällekkäin. Esimerkissä 2.1

**Esimerkki 2.1.** Vastaajan iän kysyminen, kaksi tapaa.

```
Minkä ikäinen olette? _____ vuotta

Minkä ikäinen olette? 1) 18-29 2) 30-39 3) 40-49 4) 50-59 5) 60-74 vuotta
```

ensimmäinen kysymys on avoin ja toinen suljettu. Sen vaihtoehdot ovat toisensa poissulkevia, joten ei synny epäselvyyksiä, mikä vaihtoehto pitäisi valita. Alle 18-vuotiaat ja yli 74-vuotiaat eivät tässä tapauksessa ole tutkimuksen kohteena.

Periaatteessa avoimella kysymyksellä saataisiin tieto vuoden tarkkuudella. Ikä kuuluu kuitenkin niihin asioihin, joita ei useinkaan pidä kysyä suoraan, ei naisilta eikä miehiltä. Syy on, että vastaukset voivat olla aika epäluotettavia. Kannattaa mieluummin käyttää suljettua kysymystä, jossa on valmiit vaihteluvälit, kuten esimerkissä 2.1. Näin mittaustarkkuus on osaa 10 vuoden luokkaa, mutta se ei haittaa, jos se tutkimuksen kannalta riittää. On vain mietittävä tarkoin etukäteen, millaista luokittelua käytetään, koska jälkikäteen sitä ei voi muuttaa. Korkeintaan tiivistää vielä karkeammaksi. Monesti on hyödyllistä käyttää sellaisia luokituksia, joita on käytetty saman aihepiirin aiemmissakin tutkimuksissa. Näin voi helpommin vertailla tulokset keskenään.

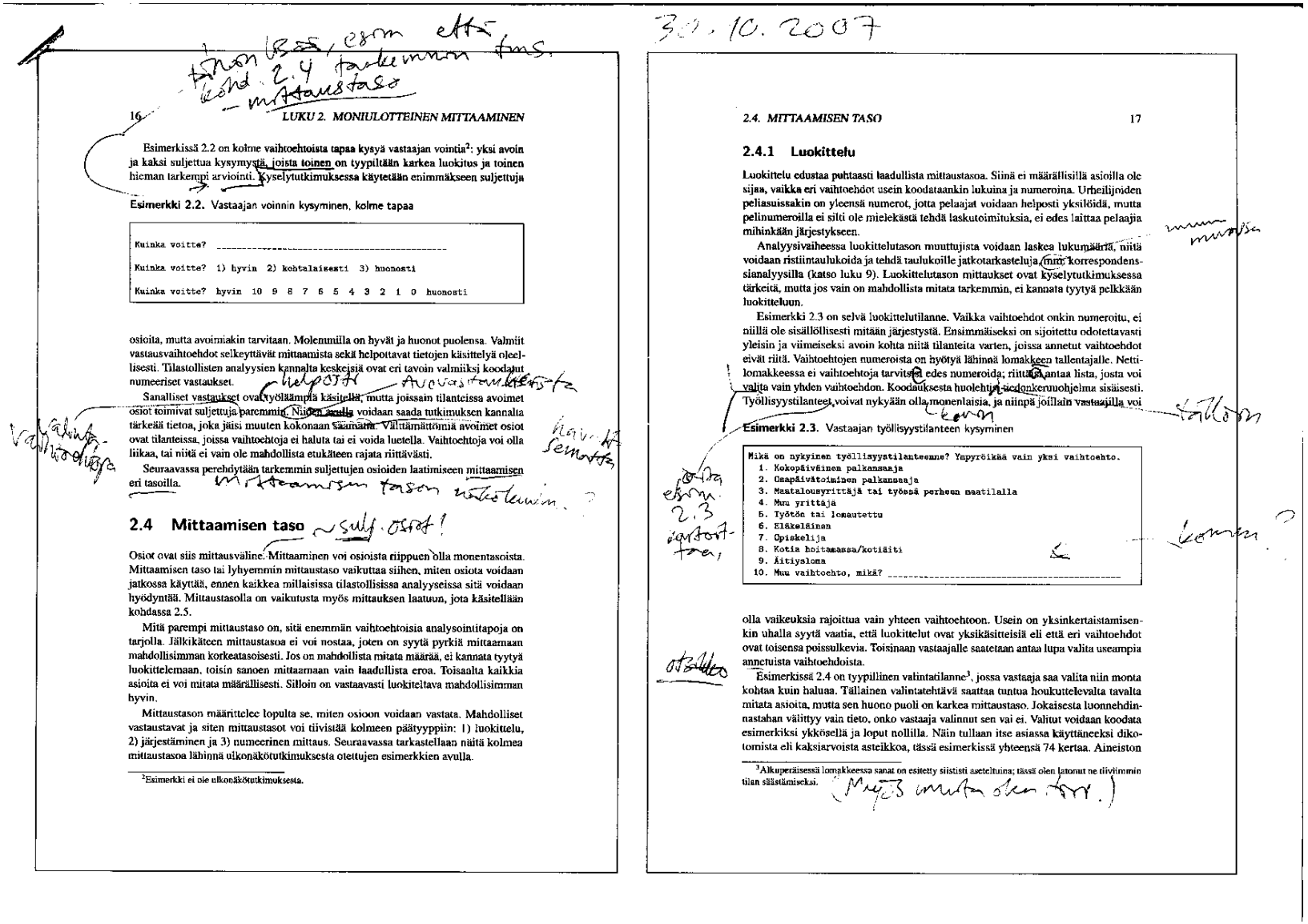
Ikää voidaan mitata myös epäsuorasti kysymällä syntymävuotta. Se on todettu luotettavammaksi tavaksi kuin suora iän kysyminen. Syntymävuoden avulla on aineiston esikäsittelyvaiheessa helppo laskea vastaajan ikä vastausajankohtana. Vielä tarkemmin ikä saadaan selville, jos voidaan käyttää tukena rekisteritietoja. Niihin vertaamalla voidaan myös tutkia vastaajan antamien tietojen luotettavuutta.

Ominaisuutena tai käsitteenä ikä on jatkuva: se saa jollakin tarkasteltavalla välillä mitä tahansa arvoja. Jatkuvienkin asioiden mittaaminen on kuitenkin mahdollista vain rajallisella tarkkuudella. Se miten tarkkaan mittaamiseen pyritään, riippuu tilanteesta. Yleisesti on tietenkin syytä mitata niin tarkasti kuin mahdollista, mutta kuten edellä nähtiin, voi vähempikin tarkkuus olla riittävä.

Mittaustarkkuuden rajallisuudesta johtuen mittaukset eli mitatut osiot tai muuttujat eivät koskaan ole jatkuvia vaan koostuvat erillisistä luvuista. Jos mitattu ilmiö on luonteeltaan jatkuva, voidaan kuitenkin puhua jatkuvista muuttujista.

- ...teksti ei synny tuosta vain, se vaatii editointia...

# 30.10.2007: ”Moniulotteinen mittaaminen” – vaiheessa



16 LUKU 2. MONIULOTTEINEN MITTAAMINEN

Esimerkissä 2.2 on kolme vaihtoehtoista tapaa kysyä vastaajan vointia[2]: yksi avoin ja kaksi suljettua kysymystä, joista toinen on tyypiltään karkea luokitus ja toinen hieman tarkempi arviointi. Kyselytutkimuksessa käytetään enimmäkseen suljettuja

**Esimerkki 2.2.** Vastaajan voinnin kysyminen, kolme tapaa

```
Kuinka voitte?  ____________________________________________

Kuinka voitte?  1) hyvin  2) kohtalaisesti  3) huonosti

Kuinka voitte?  hyvin  10  9  8  7  6  5  4  3  2  1  0  huonosti
```

osioita, mutta avoimiakin tarvitaan. Molemmilla on hyvät ja huonot puolensa. Valmiit vastausvaihtoehdot selkeyttävät mittaamista sekä helpottavat tietojen käsittelyä oleellisesti. Tilastollisten analyysien kannalta keskeisiä ovat eri tavoin valmiiksi koodatut numeeriset vastaukset.

Sanalliset vastaukset ovat työllämpiä käsitellä, mutta joissain tilanteissa avoimet osiot toimivat suljettuja paremmin. Niiden avulla voidaan saada tutkimuksen kannalta tärkeää tietoa, joka jäisi muuten kokonaan saamatta. Välttämättömiä avoimet osiot ovat tilanteissa, joissa vaihtoehtoja ei haluta tai ei voida luetella. Vaihtoehtoja voi olla liikaa, tai niitä ei vain ole mahdollista etukäteen rajata riittävästi.

Seuraavassa perehdytään tarkemmin suljettujen osioiden laatimiseen mittaamisen eri tasoilla.

## 2.4 Mittaamisen taso

Osiot ovat siis mittausvälineitä. Mittaaminen voi osioista riippuen olla monentasoista. Mittaamisen taso tai lyhyemmin mittaustaso vaikuttaa siihen, miten osiota voidaan jatkossa käyttää, ennen kaikkea millaisissa tilastollisissa analyyseissä sitä voidaan hyödyntää. Mittaustasolla on vaikutusta myös mittauksen laatuun, jota käsitellään kohdassa 2.5.

Mitä parempi mittaustaso on, sitä enemmän vaihtoehtoisia analysointitapoja on tarjolla. Jälkikäteen mittaustasoa ei voi nostaa, joten on syytä pyrkiä mittaamaan mahdollisimman korkeatasoisesti. Jos on mahdollista mitata määrää, ei kannata tyytyä luokittelemaan, toisin sanoen mittaamaan vain laadullista eroa. Toisaalta kaikkia asioita ei voi mitata määrällisesti. Silloin on vastaavasti luokiteltava mahdollisimman hyvin.

Mittaustason määrittelee lopulta se, miten osioon voidaan vastata. Mahdolliset vastaustavat ja siten mittaustasot voi tiivistää kolmeen päätyyppiin: 1) luokittelu, 2) järjestäminen ja 3) numeerinen mittaus. Seuraavassa tarkastellaan näitä kolmea mittaustasoa lähinnä ulkonäkötutkimuksesta otettujen esimerkkien avulla.

[2]Esimerkki ei ole ulkonäkötutkimuksesta.

2.4. MITTAAMISEN TASO 17

### 2.4.1 Luokittelu

Luokittelu edustaa puhtaasti laadullista mittaustasoa. Siinä ei määrällisillä asioilla ole sijaa, vaikka eri vaihtoehdot usein koodataankin lukuina ja numeroina. Urheilijoiden peliasuissakin on yleensä numerot, jotta pelaajat voidaan helposti yksilöidä, mutta pelinumeroilla ei silti ole mielekästä tehdä laskutoimituksia, ei edes laittaa pelaajia mihinkään järjestykseen.

Analyysivaiheessa luokittelutason muuttujista voidaan laskea lukumäärä, niitä voidaan ristiintaulukoida ja tehdä taulukoille jatkotarkasteluja, mm. korrespondenssianalyysilla (katso luku 9). Luokittelutason mittaukset ovat kyselytutkimuksessa tärkeitä, mutta jos vain on mahdollista mitata tarkemmin, ei kannata tyytyä pelkkään luokitteluun.

Esimerkki 2.3 on selvä luokittelutilanne. Vaikka vaihtoehdot onkin numeroitu, ei niillä ole sisällöllisesti mitään järjestystä. Ensimmäiseksi on sijoitettu odotettavasti yleisin ja viimeiseksi avoin kohta niitä tilanteita varten, joissa annetut vaihtoehdot eivät riitä. Vaihtoehtojen numeroista on hyötyä lähinnä lomakkeen tallentajalle. Nettilomakkeessa ei vaihtoehtoja tarvitsisi edes numeroida; riittäisi antaa lista, josta voi valita vain yhden vaihtoehdon. Koodauksesta huolehtii tietokonekäyttöohjelma sisäisesti. Työllisyystilanteet voivat nykyään olla monenlaisia, ja niinpä joillain vastaajilla voi

**Esimerkki 2.3.** Vastaajan työllisyystilanteen kysyminen

```
Mikä on nykyinen työllisyystilanteenne? Ympyröikää vain yksi vaihtoehto.
  1. Kokopäiväinen palkansaaja
  2. Osapäivätoiminen palkansaaja
  3. Maatalousyrittäjä tai työssä perheen maatilalla
  4. Muu yrittäjä
  5. Työtön tai lomautettu
  6. Eläkeläinen
  7. Opiskelija
  8. Kotia hoitamassa/kotiäiti
  9. Äitiysloma
  10. Muu vaihtoehto, mikä? ____________________________________
```

olla vaikeuksia rajoittua vain yhteen vaihtoehtoon. Usein on yksinkertaistamisenkin uhalla syytä vaatia, että luokittelut ovat yksikäsitteisiä eli että eri vaihtoehdot ovat toisensa poissulkevia. Toisinaan vastaajalle saatetaan antaa lupa valita useampia annettuista vaihtoehdoista.

Esimerkissä 2.4 on tyypillinen valintatilanne[3], jossa vastaaja saa valita niin monta kohtaa kuin haluaa. Tällainen valintatehtävä saattaa tuntua houkuttelevalta tavalta mitata asioita, mutta sen huono puoli on karkea mittaustaso. Jokaisesta luonnehdinnastahan välittyy vain tieto, onko vastaaja valinnut sen vai ei. Valitut voidaan koodata esimerkiksi ykkösellä ja loput nollilla. Näin tullaan itse asiassa käyttänneeksi dikotomista eli kaksiarvoista asteikkoa, tässä esimerkissä yhteensä 74 kertaa. Aineiston

[3]Alkuperäisessä lomakkeessa sanat on esitetty siististi asetelluina; tässä olen latonut ne tiiviimmin tilan säästämiseksi.

- ...sotkemista ja korjailua...

# 30.10.2007: ”Moniulotteinen mittaaminen” – vaiheessa

30.10.2007



**Esimerkki 2.9.** Pituus, paino ja rahan käyttö:

```
Kuinka pitkä olette ja paljonko painatte? Pituus ______ cm  Paino ____ kg

Kuinka paljon suurin piirtein käytätte rahaa vuosittain vaatetukseenne?
Noin ___________ euroa vuodessa
```

Monet kyselytutkimuksen keskeiset mittaustavat, esimerkiksi asennemittaukset, käsitetään yleensä kirjallisuudessa järjestysasteikoiksi. Ongelma on, että järjestysasteikolle soveltuvia tilastollisia menetelmiä on varsin vähän. Valtaosassa tässäkin kirjassa esitetyistä menetelmistä nojaa keskiarvoihin, hajontoihin ja korrelaatioihin, joiden laskeminen edellyttää vähintään väliasteikollista mittausta. Joidenkin mielestä tässä

**Esimerkki 2.10.** Käsityksiä ulkonäön merkityksestä (?) Likert-asteikolla

```
Ympyröikää seuraavien väittämien kohdalla omaa käsitystänne parhaiten
vastaava vastausvaihtoehto. Vastausvaihtoehdot ovat: 1 = Täysin samaa
mieltä, 2 = Jokseenkin samaa mieltä, 3 = Ei samaa eikä eri mieltä,
4 = Jokseenkin eri mieltä, 5 = Täysin eri mieltä.

  1. Ihmisen ulkonäkö on kulttuurissamme
     liian arvostetussa asemassa.          1   2   3   4   5
  2. Miellyttävästä ulkonäöstä on hyötyä.  1   2   3   4   5
  3. Hyvännäköiset ihmiset pärjäävät
     elämässään paremmin.                  1   2   3   4   5
  4. Hyvännäköiset ihmiset ovat suosituimpia.  1   2   3   4   5
  ...
```

on ristiriita, kun käytännössä kuitenkin lasketaan korrelaatioita vaikka kyseessä olisi järjestysasteikko. -> mittaustarkkuus!

1-2-3-4-5 vs 5-4-3-2-1 ! (selkeys, typografia, suunnat, kääntäminen (ref?))

esimerkkejä: k/e, 1-2-3-4-5 (Likert ym.), kouluarvosana, 1-2-3-4, 1-2-3, 1,...,7

**EOS!!!** pref.skaalat (järj) dikotomiat

Mikä sitten on väliasteikollista mittaamista, on hyvä kysymys. Tyypillisin kirjallisuudessa annettu esimerkki koskee lämpötilan mittaamista, mutta se harvoin kiinnostaa kyselytutkimuksessa. Sen sijaan...

onko Maariitilla semanttisia differentiaaleja??????????

sanan "asteikko" kaksikäsitteisyys (selitä, historiaa Alastalossa!) **minor.**

~~simppeli luokittelu! kohtalokkaita~~

Mikä sitten on ~~väliasteikollista mittaamista,~~ onkin hyvä kysymys.

Valtaosa mittaustasoa koskevista kysymyksistä kulminoituu järjestysasteikkoon. Kirjallisuudessa esitetään usein varsin tiukkoja käsityksiä siitä, mistä saa laskea korrelaatioita ja mistä ei, ja järjestysasteikko rajataan näissä yhteyksissä tylysti ulkopuolelle.

...Likert-asteikko on järjestysasteikko.

Likert (tms. 5-portainen) vs true interval here??

when 2.4 melle. valmis...



Huom! dikotomia...

Esimerkkejä: ikä luokiteltuna (ika1): järj.ast., Likert- ym. asteikot, ... (ikä oli jo edellä, 2 tapaa: ex:ikä2)

mittaustarkkuus muistettava johtopäätöksiä tehdessä: ei liikaa desimaaleja jne.!

## 2.5 Mittaamisen laatu

Ei todellakaan ole samantekevää, miten jotain asioita mitataan. Mittausmallin yhteydessä viitattiin jo "hyvään mittariin" ja "mittarin tarkkuuteen". Tarkkuus onkin yksi mittaamisen laadun osatekijöistä. Mittaamisesta johtuvan epävarmuuden hallinta edellyttää kuitenkin, että mittaamisen laatua lähestytään kokonaisvaltaisemmin. Laatuun vaikuttavat ensinnäkin sekä mittausmallista että mitta-asteikosta tehdyt oletukset. Käytännössä vaikuttaa vielä se, miten hyvin oletukset toteutuvat. Myös mittausten tasolla on oleellinen vaikutus laatuun. Tässä kuvan 2.1 mittauskehikko on taas hyödyksi.

### 2.5.1 Validiteetti ja reliabiliteetti

mm. osioiden avainasema (ks. alla), mutta myös mm. validoidut, kansainväliset ym. mittarit - takaako validointi laadun? mitä laadun osatekijöitä pitäisi tarkastella? -> val, rel, ... Mittauskehikko taas hyvä. (Palataan vielä edellä esitettyyn mittauskehikkoon...)

Avainasemassa ovat osiot, joilla on siis kahdenlainen tehtävä. Ne ovat mittausmallin ainoat mitattavissa olevat osat, tutkimuksen mittausinstrumentti. Niiden suunnittelu, valinta (sanavalinnat?) ja muotoilu vaikuttavat oleellisesti siihen, millaisia tilastollisia menetelmiä tutkimuksessa voidaan käyttää.

**vrt. mittauksen taso! (ks. edellä)**

Toisaalta osiot, sen jälkeen kun ne on mitattu, ovat kaikkien asteikkojen perusta. Koska osioiden taustalla mittausmallissa on sekä tosiarvoja että mittausvirheitä, tulee jokainen asteikko (myös tulosasteikko) perimään näitä molempia jossain sekoitussuhteessa, aivan kuten kuvan 2.1 nuolet osoittavat. Tavoitteena on toimia siten, etteivät mittausvirheet pääsisi liikaa vaikuttamaan, sillä ne heikentävät asteikkojen laatua ja murentavat osaltaan koko tutkimuksen luotettavuutta.

Siitä huolimatta, että osiot ovat keskeisellä sijalla, on kokonaiskäsityksen saamiseksi hyödyllisempää tarkastella erikseen mittaamisen laadun osatekijöitä, validiteettia ja reliabiliteettia. Validiteetti kertoo mitataanko sitä mitä pitikin, kun taas reliabiliteetti kertoo miten tarkasti mitataan. Molemmat ovat mittaamisen laadun kannalta tärkeitä, mutta validiteetti (tai pätevyys) on tärkeämpi. Tarkkuudella ei tee mitään, jos mittaaminen ei ole pätevää.

Yllättävän usein kyselytutkimuksessa mittari saattaa mitata aivan muuta kuin sen luultiin mittaavan. Esimerkiksi jokin aiemmin tehty tutkimus voidaan toistaa toisessa maassa, jolloin kyselylomake käännetään eri kielelle. Jos keskitytään vain osioiden huolelliseen kääntämiseen, voi seurata kohtalokkaita yllätyksiä, sillä osiot saattavatkin mitata toisessa maassa eri asioita.

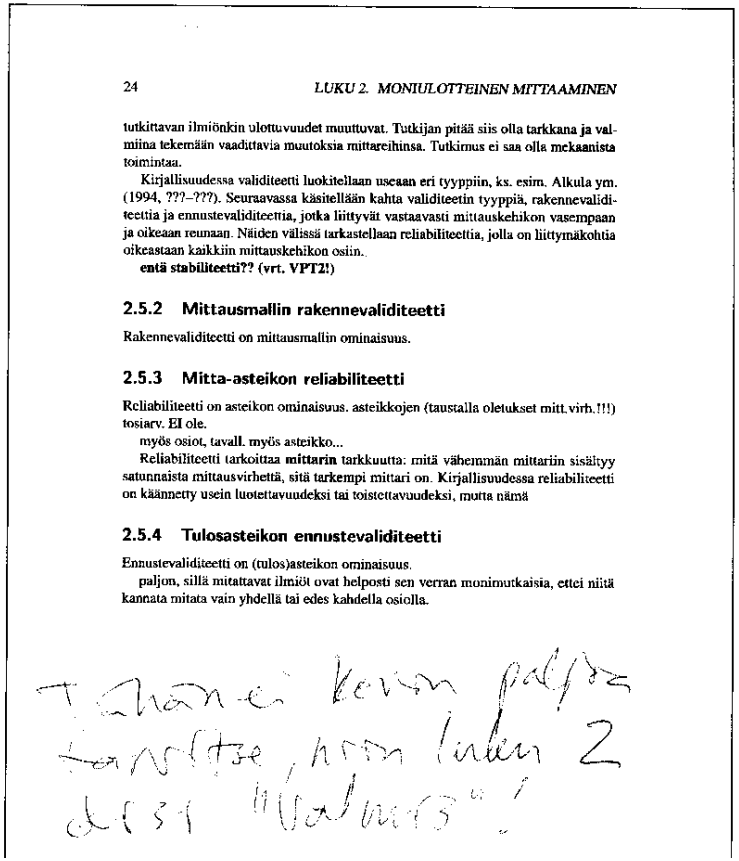
Olisi jälleen ajateltava ensin tosiarvoja (mitä mitataan), ja vasta sitten osioita (miten mitataan). Sama koskee tutkimuksen toistamista yleensäkin, vaikkei kielenkääntäjää tarvittaisikaan, sillä mikään ei takaa mittareiden ajallista pysyvyyttä. Tietenkin

- ...ja taas lisää raakatekstiä...

# 30.10.2007: "Moniulotteinen mittaaminen" – vaiheessa

*"Tähän ei kovin paljoa tarvitse, niin luku 2 olisi 'valmis'!"*

30.10.2007



24 LUKU 2. MONIULOTTEINEN MITTAAMINEN

tutkittavan ilmiönkin ulottuvuudet muuttuvat. Tutkijan pitää siis olla tarkkana ja valmiina tekemään vaadittavia muutoksia mittareihinsa. Tutkimus ei saa olla mekaanista toimintaa.

Kirjallisuudessa validiteetti luokitellaan useaan eri tyyppiin, ks. esim. Alkula ym. (1994, ???–???). Seuraavassa käsitellään kahta validiteetin tyyppiä, rakennevaliditeettia ja ennustevaliditeettia, jotka liittyvät vastaavasti mittauskehikon vasempaan ja oikeaan reunaan. Näiden välissä tarkastellaan reliabiliteettia, jolla on liittymäkohtia oikeastaan kaikkiin mittauskehikon osiin.

**entä stabiliteetti?? (vrt. VPT2!)**

### 2.5.2 Mittausmallin rakennevaliditeetti

Rakennevaliditeetti on mittausmallin ominaisuus.

### 2.5.3 Mitta-asteikon reliabiliteetti

Reliabiliteetti on asteikon ominaisuus. asteikkojen (taustalla oletukset mitt.virh.!!!) tosiarv. EI ole.

myös osiot, tavall. myös asteikko...

Reliabiliteetti tarkoittaa **mittarin** tarkkuutta: mitä vähemmän mittariin sisältyy satunnaista mittausvirhettä, sitä tarkempi mittari on. Kirjallisuudessa reliabiliteetti on käännetty usein luotettavuudeksi tai toistettavuudeksi, mutta nämä

### 2.5.4 Tulosasteikon ennustevaliditeetti

Ennustevaliditeetti on (tulos)asteikon ominaisuus.

paljon, sillä mitattavat ilmiöt ovat helposti sen verran monimutkaisia, ettei niitä kannata mitata vain yhdellä tai edes kahdella osiolla.

Tähän ei kovin paljoa tarvitse, niin luku 2 olisi "valmis"!

- **niin niin... (kirjan ilmestymiseen noin vuosi aikaa...)**
- Tilastokeskuksessa *one more time* 12.-13.11.2007

# 10.12.2007: **Alustava käsikirjoitus** (77 s.)

**Kimmo Vehkalahti**

• jhk (alkuun?) näyttertä palautteista? ... ei aina sia regressiokaa ...
(?) → "takakansi"!

## Kyselytutkimuksen mittarit ja menetelmät

Fil. graf:
• solmu
• JK

**Alustava käsikirjoitus**

**10.12.2007**

Tämä näytetty 11.12.2007 Leenalle korjauksiani merkitty palaverin jälkeen "Hyvältä näyttää, lisää samanlaista! Alku (johdanto) hyvä!!"

Identtinen kopio annettu Katille 10. (ja tarkasteltu yhdess. ...2.2005)

# 10.12.2007: Näkymiä alustavasta käsikirjoituksesta

10.12.2007
→ LP, KT, SL

LUKU 3

## Työskentely tietokoneella

Olipa kerran aika, jolloin tilastollisia analyyseja tehtiin kynällä ja paperilla. Työskentely oli vaivalloista ja vei paljon aikaa. Laskemiseen, jonka nyt saa tehtyä sekunnissa nappia painamalla, saattoi kokonaiselta tutkimusryhmältä kulua päiväkausia. Vähitellen avuksi tuli laskukoneita, mutta ennen tietokoneiden aikakautta monista menetelmistä voitiin edelleen vain haaveilla, vaikka ne teoriassa oli tunnettu jo vuosikymmeniä.

Tietokoneet ovat mullistaneet kaiken moneen kertaan, mutta laskemisen perinne heijastuu edelleen menetelmäkirjoista, joissa neuvotaan millaisella kaavalla korrelaatiokerroin lasketaan. Neuvot ovat paitsi kaavamaisia myös turhia. Ei tilastollisia analyyseja tehdä enää laskukoneilla—kynästä ja paperista nyt puhumattakaan.

Kaavat eivät yleensä edes sellaisenaan sovellu tietokonekäyttöön, vaan tarvitaan ohjelmointitaitoja ja käsitystä numeerisesta laskennasta. Näistä asioista ei menetelmien soveltajan tarvitse tietää mitään. Hän voi keskittyä tehtäviin, joihin kone ei kykene: tutkimuksen ja analyysien suunnitteluun sekä tulosten tulkintaan ja arviointiin. Tehtäviin voi sisältyä pientä laskentaa, mutta siinä ei tarvita niinkään laskukaavoja vaan käteviä tapoja tehdä erilaisia laskelmia.

Kunnolliset työvälineet, taito käyttää niitä oikein ja tehokkaat työskentelytavat yhdistetään usein käsityöläisammatteihin, mutta yhtä lailla ne ovat läsnä tutkimustyyppisessä työssä ja vaikuttavat niin työn sujuvuuteen kuin tulosten laatuunkin. Myös tutkimustyön välineiden tulisi tukea oppimista ja mahdollistaa omien työtapojen mielekäs kehittäminen.

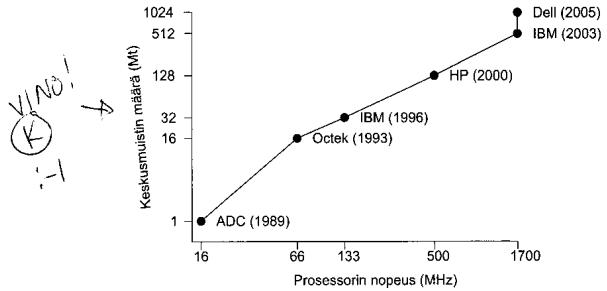
Tässä luvussa perehdytään töiden hallintaan ja oman työn dokumentointiin. Aluksi tarkastellaan tutkijan työvälineitä, tietokoneita ja ennen kaikkea ohjelmistoja. Siitä edetään töiden hallintaan, ja lopuksi pohditaan, mitä tarkoittaa dokumentoiva työskentelytapa ja mitä hyötyä siitä on erityisesti kyselytutkimuksen kannalta.

30

10.12.2007
→ LP, KT, SL



2008?




VINO!

**Kuva 3.1.** Koneideni prosessoritehon ja keskusmuistin kehitys.

### 3.1 Tietokone ja ohjelmistot

Tietokone on tutkijan työvälineistä tärkein. Kaikki kyselytutkimuksen työvaiheet voidaan tehdä tietokoneella. Koneet ovat tulleet jatkuvasti tehokkaammiksi mutta samalla pienemmiksi ja siirtyneet vaivihkaa työpöydiltä reppuihin ja salkkuihin.

Itselläni on ollut vuodesta 1989 kuusi tietokonetta, kaksi ensimmäistä pöytäkoneita ja loput kannettavia. Uuden koneen myötä laskennalliset tehot kasvavat huikein harppauksin (katso kuva 3.1), mutta työtehoa ei voi mitata megahertseinä tai gigatavuina. Oleellisempaa onkin tarkoituksenmukaisten ohjelmistojen käyttö ja omien töiden kätevä hallinta.

Tietokoneen merkillä tai mallilla ei ole juuri mitään merkitystä, ei myöskään käyttöjärjestelmällä. Joissain kohdin viittaan Windowsiin, jonka puitteissa itse enimmäkseen toimin, mutta sinänsä Windowsilla on vain vähäinen merkitys työskentelyssäni. Esittämäni periaatteet pätevät yhtä hyvin Linux-, Mac- ja muissakin järjestelmissä.

#### 3.1.1 Kirjan ohjelmistot

Ohjelmistojen käyttö on oleellinen osa tilastollisten menetelmien soveltamista. Niinpä tällaisen kirjankin käytännön hyöty jäisi vähäiseksi ilman ohjelmistoesimerkkejä. Sopivia ohjelmistoja on runsaasti, joten on pitänyt tehdä valintoja.

Valintani kirjan ohjelmistoiksi ovat Survo ja SPSS. Survo on suomalainen, professori Seppo Mustosen kehittämä tietojenkäsittely-ympäristö, ja SPSS on amerikkalainen, alunperin Norman H. Nien, C. Hadlai Hullin ja Dale H. Bentin kehittämä tilastollisten menetelmien ohjelmapaketti. Molempien ohjelmistojen ensimmäiset versiot ilmestyivät jo 1960-luvulla (Survo Systems, 2007; SPSS Inc., 2007).

Kirjassa käsiteltävien asioiden ymmärtäminen ei edellytä minkään ohjelmiston erityistä osaamista, koska pääosassa ovat menetelmät ja työtavat. Ohjelmistotekniset

- Alkupäässä mukamas valmista (suurin osa lopulta poistettu)...

# 10.12.2007: Näkymiä alustavasta käsikirjoituksesta

10.12.2007
→KT, LP, SL



arvot olisivatkin numeroita (kuten esimerkkiaineiston `maakunta`), niistä ei pidä laskea korrelaatioita, sillä tuloksena on pelkkää roskaa. Hyvä tapa välttää vahinkoja on kertoa ohjelmistolle muuttujien mittaustasot, kuten kohdassa xxx tehtiin ja toivoa, että ohjelmisto osaa ottaa ne huomioon järkevällä tavalla.

### 6.1.4 Muuttujien normaalijakauma

(vrittoa ed. luvun!)

Faktorianalyysin taustalla on myös oletus muuttujien normaalijakaumasta, jonka puitteissa ns. tyhjentävät tunnusluvut (keskiarvot, keskihajonnat ja korrelaatiot) riittävät aineiston oleellisen informaation kuvaamiseen. Tarkemmin sanottuna oletus koskee muuttujien yhteisjakaumaa, jota kutsutaan multinormaalijakaumaksi, katso esimerkiksi Mustonen (1995, luvut 2 ja 3).

Tässä kohtaa teoria ja käytäntö kulkevat vääjäämättä eri reittejä. Vaikka teoriatarkasteluissa multinormaalijakaumaoletus on tavanomainen, on se käytännössä niin vaativa, ettei se oikeastaan edes toteudu todellisilla aineistoilla. Yksittäisten muuttujien normaalisuuskaan ei riitä, joskaan ei se tietenkään haittaakaan.

Tämä teorian ja käytännön ristiriidalta kuulostava seikka ei siis estä faktorianalyysin soveltamista. Käytännössä menetelmä sietää poikkeamia joita empiirisiin aineistoihin joka tapauksessa sisältyy.

### 6.1.5 Muita oletuksia

havaintojen lkm-oletukset...??? n/p

Kuvassa 2.1 esitetty mittauskehikko...

## 6.2 Faktorianalyysi

Faktorianalyysi on kyselytutkimuksen tärkein menetelmä. Sen avulla tiivistetään aineistoa siirtymällä mittauskehikossa (katso kuva 2.1) ennalta pohditusta mittausmallista erilaisiin asteikkoihin, joihin jatkoanalyysit tulevat perustumaan.

Edellä käsiteltyjen oletusten järkevyyttä on syytä pohtia.

### 6.2.1 Mittausmallin analysointi

tehdään siis FA, pohditaan faktorilukua ja mittausvirheitä (kuten me ne määrittelemme, REF!) **VALITTAVA SOPIVA MITTARISTO TÄHÄN!!!**

### 6.2.2 Rakennevaliditeetin arviointi

tarkastellaan a) f-matriisia, b) komm., c) etc., faktorikuvia, d) residuaaleja

rotaatio vasta tänne

## 6.3 Tiivistäminen asteikoiksi

joko "kiinteästi"f-pisteet tai ihan jotain muuta, summat jne. trad.arr.

10.12.2007



### 6.3.1 Faktoripistemäärät

### 6.3.2 Summamuuttujat

### 6.3.3 Reliabiliteetin arviointi

KV:n "ominta"asiaa... jotain alfastakin tänne

- ...luvussa 6 vain keskeneräisiä kappaleita ja otsikoita...

# 10.12.2007: Näkymiä alustavasta käsikirjoituksesta

10.12.2007

LUKU 7

## Selitysmallien laatiminen

Tämä luku on vielä alkutekijöissään. Aion tiivistää siihen yliopistokurssini Lineaaristen mallien sovellukset (ent. Data-analyysi II) keskeiset aiheet sekä kytkeä ne luvussa 2 esitettyyn mittauskehikkoon. Esimerkit tulevat ulkonäköaineistosta. Oheiset alaotsikot ovat vasta alustavia.

**7.1 Tavoitteet ja oletukset**

**7.2 Regressioanalyysi**

**7.2.1 Selittäjien valinta**

**7.2.2 Regressiodiagnostiikka**

**7.2.3 Ennustevaliditeetin arviointi**

61

10.12.2007

LUKU 8

## Vastaajien ryhmittely

Tämä luku on vielä alkutekijöissään. Aion tiivistää siihen oheisten otsikoiden mukaisia asioita samaan tapaan kuin monimuuttujamenetelmien monisteessani, kuitenkin ulkonäköaineistoa käyttäen sekä kytkeä asiat luvussa 2 esitettyyn mittauskehikkoon. Oheiset alaotsikot ovat vasta alustavia.

**8.1 Etäisyyksien mittaaminen**

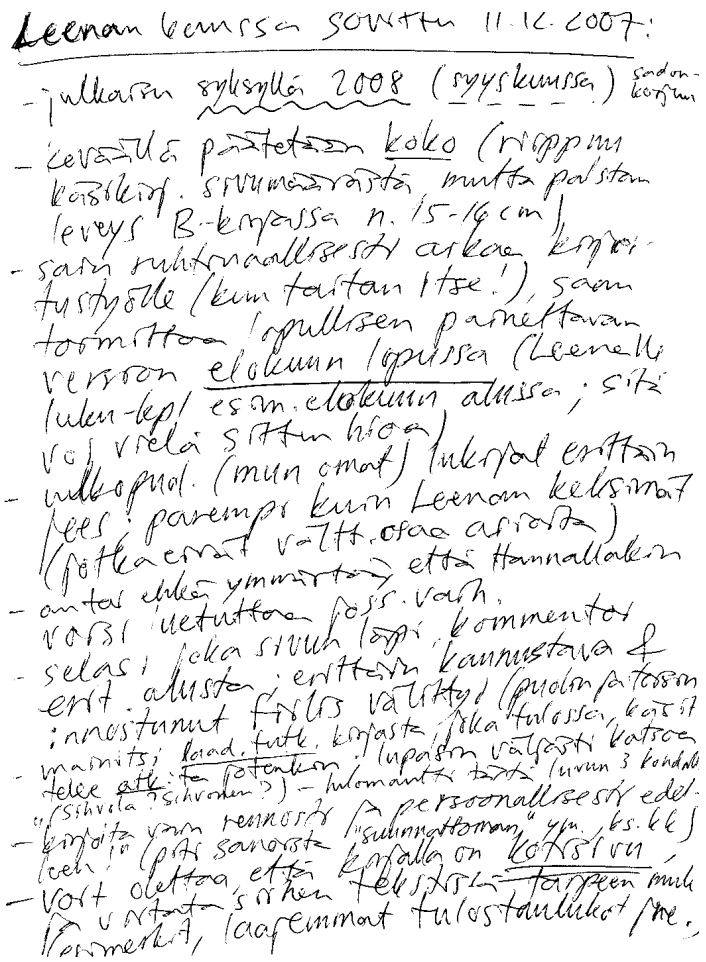
**8.2 Hierarkinen ryhmittely**

**8.3 Muita ryhmittelymenetelmiä**

62

- ...luvuissa 7 ja 8 ei sitäkään! (pelkkiä selityksiä)

# 11.12.2007: Tavoitteena (silti) julkaisu **syksyllä 2008**

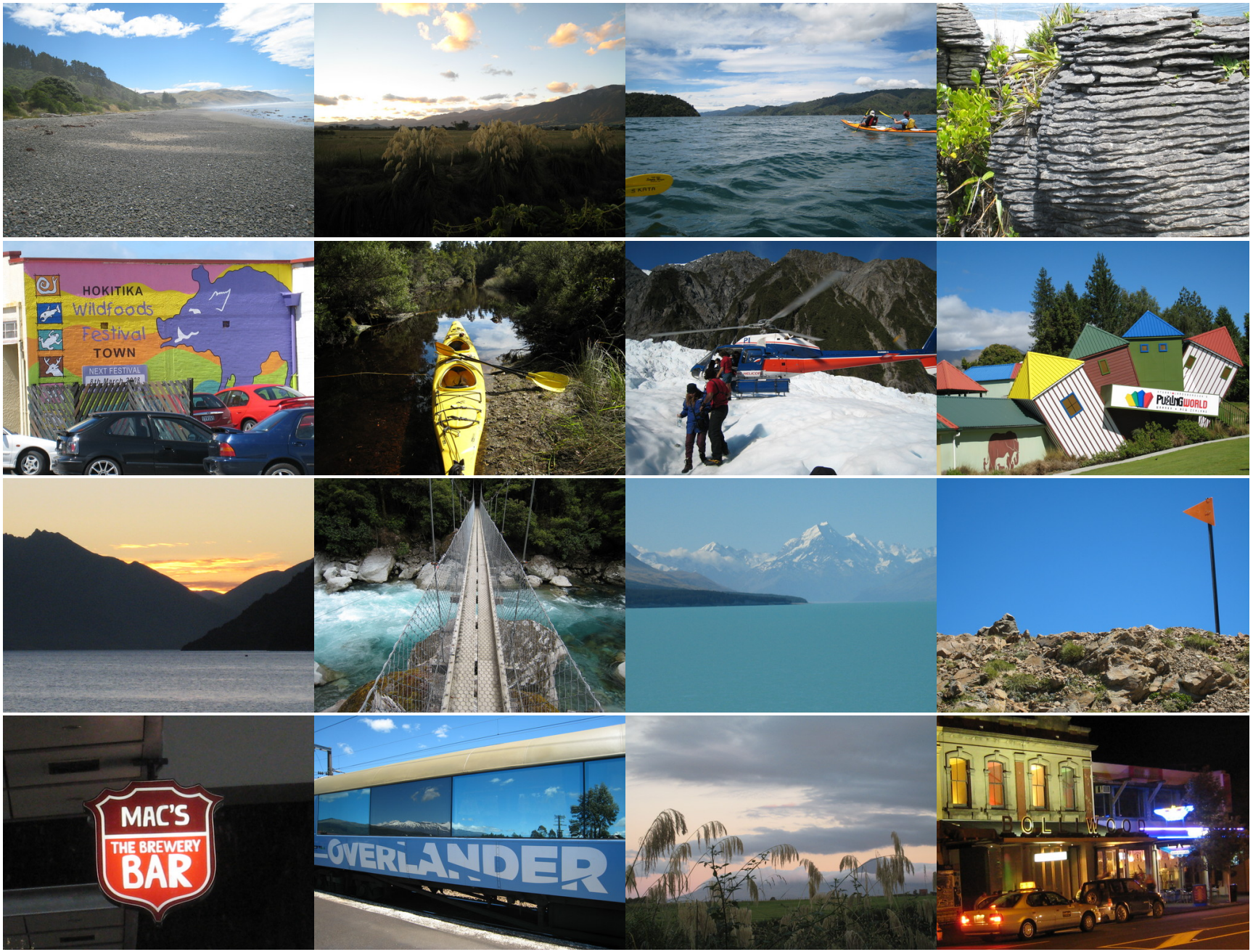


Leenan kanssa sovittu 11.12.2007:

- julkaisu syksyllä 2008 (syyskuussa) sadon-korjuu
- keväällä päätetään koko (riippuu käsikirj. sivumäärästä, mutta palstan leveys B-kirjassa n. 15-16 cm)
- sain ruhtinaallisesti aikaa kirjoitustyölle (kun taitan itse!), saan toimittaa lopullisen painettavan version elokuun lopussa (Leenalle luku-kpl esim. elokuun alussa; sitä voi vielä sittenhioa)
- ulkopuol. (mun omat) lukijat erittäin jees; parempi kuin Leenan keksimät (jotka eivät välttä osaa asioita)
- antai ehkä ymmärtää, että Hannallakin voisi luetuttaa joss. vain.
- selasi joka sivun läpi, kommentoi erit. alusta; erittäin kannustava & innostunut fiilis välittyi (puolin ja toisin
- mainitsi laad. tutk. kirjasta, joka tulossa, käsit telee atk:ta jotenkin; lupasin väljästi katsoa (sihvola? sihvonen?) – hulomantti tästä luvun 3 kondall
- kirjoita vain rennosti ja persoonallisesti edelleen! (piti sanoista "suunnattoman" ym. ks. kk)
- voit olettaa, että kirjalla on kotisivu, ja viitata siihen tekstissä tarpeen mukaan, laajemmat tulostaulukot jne.

- 15.12.2007: lähtö kuukaudeksi **Uuteen-Seelantiin**...

# 16.12.2007–13.1.2008: ei aatostakaan kirjasta kuukauteen



# 14.1.2008: Paluu Suomeen – johdantokurssi (N=325)

Kollegan kullanarvoisia kommentteja katsottu 7.2.2008:

* Otettu melko tiukasti moneen kertaan (laskujeni mukaan ainakin 3 kertaa) kantaa siihen, että faktorianalyysi on tärkein menetelmä, aamen. Onko tämä tarkoituksellista?

+ Mainio ratkaisu sijoittaa liitteisiin työkaaviot, jotka ovat usein ärsyttäviä tekstin lomassa ja häiritsevät lukemista.

* Luku 2.2.2 (Mitta-asteikko): Minusta tämä on aika teoreettista ja hankalasti hahmotettavaa, jos ei ole minkäänlaista aiempaa tietoa asiasta. Kaipaisin asiaan jonkinlaista esimerkkiä heti alkuun, kuten tapanani aina on kaipailla. :-) Aiemmin mainittiin esimerkki lenkkeilijästä, ja pyrkisinkin pitämään esimerkin tiukasti mukana ja vääntämään vähän enemmän rautalangasta, siis mikä on KÄYTÄNNÖSSÄ tulosasteikko ja mikä vertailuperuste yms.

* Asettelusta: esimerkit tulevat joskus keskelle lausetta, esim. s. 20. Nämä kai korjataan myöhemmin siten, ettei keskelle kappalettakaan jäisi mitään häiritseviä laatikoita?

* Sivu 31: minua häiritsee melkein kohtuuttomasti kuvan 3.1 "Keskusmuistin määrä" -tekstin vinous. :-)

Ylipäänsä teksti on todella hyvää ja selkeää ja asia on esitetty valaisevasti. Etenkin sujuva suomen kieli lämmittää lukijan sydäntä ja tekee hampaiden kiristelystä tarpeetonta. :-) (Miten olisi mainoslause *"Tämän menetelmäkirjan kanssa et kiristele hampaitasi!"*? :-))) Myös ulkoasu / taitto on jo nyt siisti ja selkeä.

* *s. 46:* Miten olisi esimerkkinä jokin PIENI AINEISTO, joka esitellään sarkaineroteltuna tekstitiedostona? Asia voisi tulla vielä selvemmäksi.

- Liki ainoa kosketus kirjaan **11.12.2007 – 24.5.2008**
  *(ensimmäinen johdantokurssini vei kaiken ajan ja energian)*

# 26.5.2008: Paluu kirjan äärelle

26-5.2008

## Sisältö


Alkusanat . . . . . iii

**1 Johdanto** . . . . . 1
1.1 Kyselytutkimus . . . . . 2
1.2 Katsaus kirjallisuuteen . . . . . 4
1.3 Kirjan rakenne ja sisältö . . . . . 6

**2 Moniulotteinen mittaaminen** . . . . . 9
2.1 Johdatteleva esimerkki . . . . . 10
2.2 Mittauskehikko . . . . . 12
2.2.1 Mittausmalli . . . . . 12
2.2.2 Mitta-asteikko . . . . . 14
2.2.3 Tulosasteikko ja vertailuperuste . . . . . 15
2.3 Mittaaminen kyselylomakkeella . . . . . 16
2.3.1 Osiot ja mittarit . . . . . 17
2.3.2 Avoimet ja suljetut osiot . . . . . 18
2.4 Mittaamisen taso . . . . . 19
2.4.1 Luokittelu . . . . . 20
2.4.2 Järjestäminen . . . . . 22
2.4.3 Numeerinen mittaus . . . . . 25
2.5 Mittaamisen laatu . . . . . 30
2.5.1 Validiteetti ja reliabiliteetti . . . . . 31
2.5.2 Mittausmallin rakennevaliditeetti . . . . . 32
2.5.3 Mitta-asteikon reliabiliteetti . . . . . 32




26.5.2008




2.5.4 Tulosasteikon ennustevaliditeetti . . . . . 32

**3 Työskentely tietokoneella** . . . . . 33
3.1 Tietokone ja ohjelmistot . . . . . 34
3.1.1 Kirjan ohjelmistot . . . . . 35
3.1.2 Muita ohjelmistoja . . . . . 36
3.1.3 Ohjelmistot vai menetelmät . . . . . 37
3.2 Töiden hallinta . . . . . 38
3.2.1 Töiden tallettaminen . . . . . 38
3.2.2 Töiden tavoittaminen . . . . . 39
3.2.3 Töiden varmistaminen . . . . . 41
3.3 Dokumentoiva työskentelytapa . . . . . 43

**4 Tiedonkeruu ja aineiston perustaminen** . . . . . 47
4.1 Tiedonkeruun vaihtoehdot ja vaikutukset . . . . . 48
4.1.1 Perusjoukko ja otos . . . . . 48
4.1.2 Kokonaistutkimus ja rekisterit . . . . . 49
4.1.3 Harkinnanvarainen näyte . . . . . 50
4.2 Ulkonäkötutkimuksen otanta-asetelma . . . . . 51
4.3 Tiedonkeruu kyselylomakkeella . . . . . 51
4.4 Kyselylomakkeesta aineistoksi . . . . . 51

**5 Alustavat tarkastelut ja esikäsittely** . . . . . 53
5.1 Aineiston selailu . . . . . 53
5.2 Yhden muuttujan tarkastelut . . . . . 54
5.2.1 Tilastolliset tunnusluvut . . . . . 54
5.2.2 Ulkonäköaineiston tunnuslukuja . . . . . 56
5.2.3 Pylväs- ja piirakkakuvat . . . . . 59
5.2.4 Histogrammi . . . . . 59
5.3 Muunnokset ja uudelleenkoodaukset . . . . . 62
5.4 Usean muuttujan tarkastelut . . . . . 62
5.4.1 Ristiintaulukointi . . . . . 63
5.4.2 Laatikkokuva . . . . . 63
5.4.3 Hajontakuva . . . . . 63
5.5 Aineiston muokkaaminen . . . . . 63

**6 Aineiston tiivistäminen** . . . . . 65
6.1 Tavoitteet ja oletukset . . . . . 65
6.1.1 Havaintojen riippumattomuus . . . . . 66
6.1.2 Muuttujien riippuvuus . . . . . 66
6.1.3 Muuttujien mittaustaso . . . . . 67
6.1.4 Muuttujien normaalijakauma . . . . . 67


- Johdantokurssi ohi, samoin yliopistonlehtorin viran haku.

# 26.5.2008: Loppuosasta vain otsikot olemassa

26.5.2008

5/2008: NÄMÄ NYT!




6.1.5 Muita oletuksia . . . . . 67
6.2 Faktorianalyysi . . . . . 67
6.2.1 Mittausmallin analysointi . . . . . 68
6.2.2 Rakennevaliditeetin arviointi . . . . . 68
6.3 Tiivistäminen asteikoiksi . . . . . 68
6.3.1 Faktoripistemäärät . . . . . 68
6.3.2 Summamuuttujat . . . . . 68
6.3.3 Reliabiliteetin arviointi . . . . . 68

**7 Selitysmallien laatiminen** **69**
7.1 Tavoitteet ja oletukset . . . . . 69
7.2 Regressioanalyysi . . . . . 69
7.2.1 Selittäjien valinta . . . . . 69
7.2.2 Regressiodiagnostiikka . . . . . 69
7.2.3 Ennustevaliditeetin arviointi . . . . . 69
7.3 Muunnelmia . . .

**8 Vastaajien ryhmittely** **71**
8.1 Etäisyyksien mittaaminen . . . . . 71
8.2 Hierarkinen ryhmittely . . . . . 71
8.3 Muita ryhmittelymenetelmiä . . . . . 71

**9 Vastaajaryhmien analysointi** **73**
9.1 Erotteluanalyysi . . . . . 73
9.2 Varianssianalyysi . . . . . 73
9.3 Korrespondenssianalyysi . . . . . 73

**Liitteet**

**A Kirjan ohjelmistot lyhyesti** **75**
A.1 Survo lyhyesti . . . . . 75
A.2 SPSS lyhyesti . . . . . 76
A.3 Aineiston perustaminen ja tekninen hallinta . . . . . 77
A.3.1 Työvaiheet SPSS-ympäristössä . . . . . 78
A.3.2 Työvaiheet Survo-ympäristössä . . . . . 81

**B Kuvien ja tulosteiden työkaaviot** **83**

**Kirjallisuutta** **85**

**Hakemisto** **88**


26.5.2008
(1. johd. kurssan jälkeen...)



Ylipäätään tilastotieteen soveltaminen on kaikkea muuta kuin kaavamaista. Kaavat ovat tietenkin tärkeitä tilastotieteen omissa teoriatarkasteluissa, mutta se puoli kuuluu menetelmiä kehittäville tilastotieteilijöille ja matemaatikoille, ei menetelmiä soveltaville tutkijoille.

Tämä kirja on siis soveltavan tilastotieteilijän näkemys siitä, mihin tilastollisia menetelmiä voidaan kyselytutkimuksessa käyttää ja mitä kaikkea on syytä edellyttää mittareilta ja mittaamiselta, jotta tulokset olisivat mahdollisimman luotettavia.

**Kiitokset**

*Kiitoksia kumpuaa käsikirjoituksen keskeneräisyyden kaikotessa.* :-D

*Kaipaan kriittisiä kommentteja kehittääkseni kappaleita, kuvia, ... (kaikkea!)*

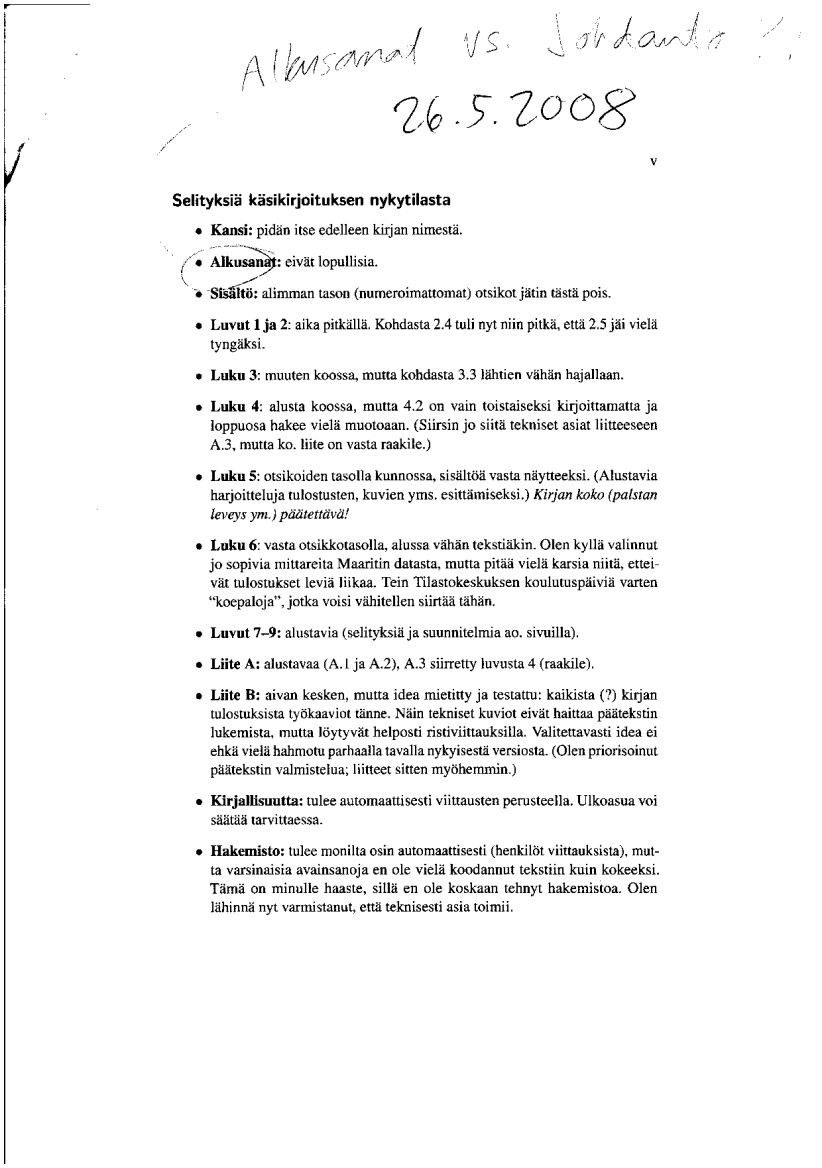
Vuosaaressa joulukuussa 2007,

– Kimmo

PS. Tässä vaiheessa tarjolla on vain **selityksiä** (ks. seuraava sivu!).

- *"Lueskelin ja merkkailin tekstiä kesäparvekkeen auringossa..."*

# 26.5.2008: Selityksiä ja loppuosan luonnostelua



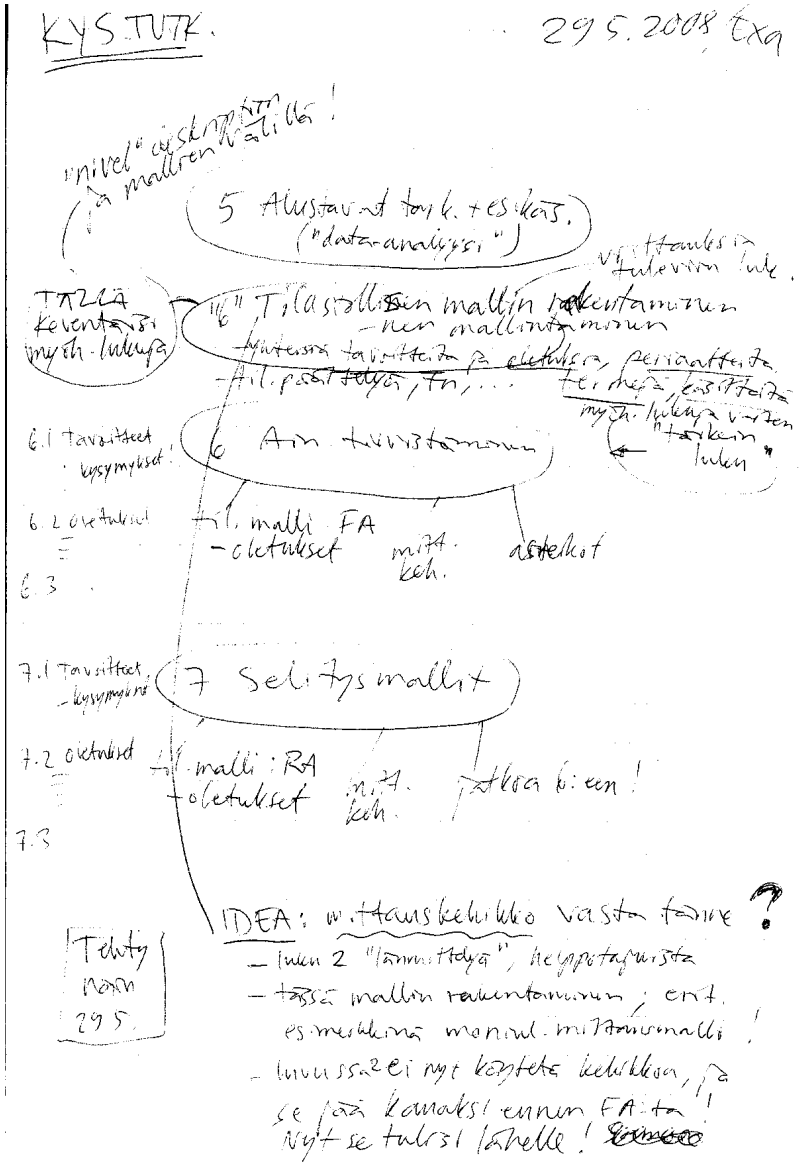
Alkusanat vs. Johdanto ?

26.5.2008

v

**Selityksiä käsikirjoituksen nykytilasta**

- **Kansi:** pidän itse edelleen kirjan nimestä.
- **Alkusanat:** eivät lopullisia.
- **Sisältö:** alimman tason (numeroimattomat) otsikot jätin tästä pois.
- **Luvut 1 ja 2**: aika pitkällä. Kohdasta 2.4 tuli nyt niin pitkä, että 2.5 jäi vielä tyngäksi.
- **Luku 3**: muuten koossa, mutta kohdasta 3.3 lähtien vähän hajallaan.
- **Luku 4**: alusta koossa, mutta 4.2 on vain toistaiseksi kirjoittamatta ja loppuosa hakee vielä muotoaan. (Siirsin jo siitä tekniset asiat liitteeseen A.3, mutta ko. liite on vasta raakile.)
- **Luku 5**: otsikoiden tasolla kunnossa, sisältöä vasta näytteeksi. (Alustavia harjoitteluja tulostusten, kuvien yms. esittämiseksi.) *Kirjan koko (palstan leveys ym.) päätettävä!*
- **Luku 6**: vasta otsikkotasolla, alussa vähän tekstiäkin. Olen kyllä valinnut jo sopivia mittareita Maaritin datasta, mutta pitää vielä karsia niitä, etteivät tulostukset leviä liikaa. Tein Tilastokeskuksen koulutuspäiviä varten "koepaloja", jotka voisi vähitellen siirtää tähän.
- **Luvut 7–9:** alustavia (selityksiä ja suunnitelmia ao. sivuilla).
- **Liite A:** alustavaa (A.1 ja A.2), A.3 siirretty luvusta 4 (raakile).
- **Liite B:** aivan kesken, mutta idea mietitty ja testattu: kaikista (?) kirjan tulostuksista työkaaviot tänne. Näin tekniset kuviot eivät haittaa pääteksti lukemista, mutta löytyvät helposti ristiviittauksilla. Valitettavasti idea ei ehkä vielä hahmotu parhaalla tavalla nykyisestä versiosta. (Olen priorisoinut pääteksti valmistelua; liitteet sitten myöhemmin.)
- **Kirjallisuutta:** tulee automaattisesti viittausten perusteella. Ulkoasua voi säätää tarvittaessa.
- **Hakemisto:** tulee monilta osin automaattisesti (henkilöt viittauksista), mutta varsinaisia avainsanoja en ole vielä koodannut tekstiin kuin kokeeksi. Tämä on minulle haaste, sillä en ole koskaan tehnyt hakemistoa. Olen lähinnä nyt varmistanut, että teknisesti asia toimii.



KYS.TUTK.

29.5.2008 Exa

5 Alustavat tark. + es.käs. ("data-analyysi")

6 Tilastollisen mallin rakentaminen

6.1 Tarvittaet kysymykset

6.2 oletukset

6.3

7 Selitysmallit

7.1 Tarvittaet kysymykset

7.2 oletukset

7.3

IDEA: mittauskehikko vasta tänne ?

Tehty noin 29.5.

- *"Loistavalta näyttää. Nyt vain hommiin. Aikaa on 3 kk. Saa tehdä!"*

# 28.5.2008: Kysyin Tammen faktoreilta kirjan dimensioista

"Olen kirjoittamassa oppikirjaa, ja teen myös taiton itse käyttäen LaTeX-ladontaohjelmaa. Näin saan mahdollisimman paljon arvokasta kirjoitusaikaa itselleni kesäksi (kirjan on tarkoitus ilmestyä syksyllä). Tämä on ensimmäinen kirja, jota teen LaTeXilla ja opettelen tuon systeemin hallintaa samalla. (Olen tehnyt sillä lähinnä artikkeleita, jolloin ulkoasuseikkoihin ei tarvitse kiinnittää huomiota kun ne tulevat lehtien tyyleistä.)

Kirjan sivumäärä ei ole vielä tiedossa, mutta **haluaisin mielelläni säätää jo joitakin asetuksia** lopullista painoasua ajatellen. Voinko tulla käymään jossain vaiheessa Tammitalossa näyttämään tarkemmin mitä olen tekemässä? Tällä ei ole mikään kiire, kesäkuussa ehtii."

# 3.6.2008: Käynti Tammitalossa, aikataulun tarkennus

- **KV:** "Käväisin tänään iltapäivällä Tammitalossa, kun oli asiaa sinne päin.

  **AIKATAULU:** PDF painoon **22.9**, sen jälkeen vedokset (haluan tarkistaa ne myös itse) yms., ja **ilmestyminen lokakuun lopulla**.

  > *Onko siis niin, että minulla olisi ikään kuin efektiivistä kirjoitusaikaa elokuun loppuun, ja vielä viimeistelyaikaa syyskuun alkupuoli? Se olisi hyvän kuuloinen aikataulu minun puolestani. Syyskuun alussa alkaa lukukausi, ja silloin ei enää varsinaista "sisällöntuotantoa" ehtisikään.*

  **KOKO: 150x235** (kuten 'Tutki ja kehitä') tai 150x245 (nämä olivat kuulemma järjestelmässä). Palstan leveys olisi tällöin 10.5 cm, ja sivumäärä jossain **176–192 sivun** paikkeilla.

  Ilmeisesti nämä eivät ihan fiksattuja ole vielä?"

# 3.6.2008: Kustannuspäällikön kommentteja

- **LP:** ”Ne tiedot, joita Jaana katsoi "järjestelmästä", olivat budjetointivaiheessa (viime elokuussa) tehtyjä arviointeja.

  Minusta kirjan koko 150 x 235 mm on hyvä. Tuo 245 korkeus on liiankin pitkulainen.

  Jos pdf tulee sinulta valmiina taittona ja minä olen jo lukenut tekstin jossain vaiheessa (esim. syyskuun alkupuoli), niin silloin tuo **22.9. painoon lähtö on ok**. Kirjoitusjälkesi on niin hyvää, että minulta ei mene kauan lukemiseen.”

# 6.-13.6.2008: Suomi Meloo (Pieksämäki – Lahti)

- ...mutta sen jälkeen **Sastamalaan vajaaksi viikoksi**...

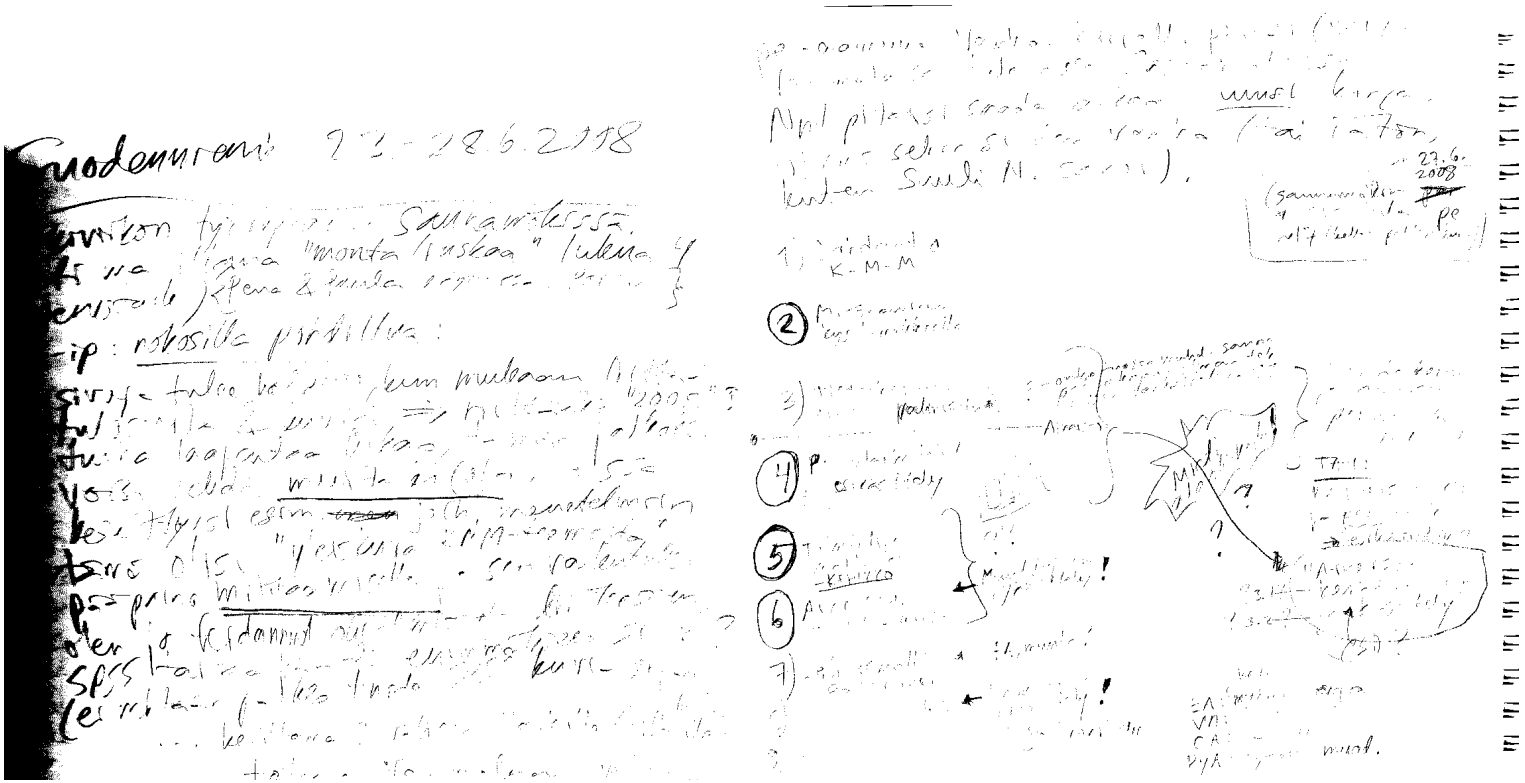
# 23.-28.6.2008: Sastamala, Suomi

"Varsovan Interissä 2007 pohdittujen aikataulujen revisointi:

**Kaikki tapahtuu nyt seuraavan 4 kk aikana!**

- kes: Sastamalaan kirjoittamaan 23.-30.6. Saatava paljon aikaan.
- hei: 2 ekaa viikkoa lomaa, 2 vikaa kirjoittamista.
- elo: Koko kk aikaa kirjoittaa.
- syy: Kuun alussa teksti LP:lle!
    - *JK2 ja LMS8 alkaa.*
    - **Kirjan viimeistely!**
    - **PAINOON 22.9.!!**
- lok: Kirja ilmestyy!"

# 23.-28.6.2008: Sastamala, Suomi



- mm. kohta "Mittauksen luotettavuus" syntyi (vihdoin) tuolla
- ja luvut menivät taas (osin) uuteen järjestykseen...

# 9.7.2008: Luku 5 alkaa hahmottua *(vai alkaako?)*

itsellemi 9.7.2008

LUKU 5

## Tilastollinen mallintaminen

Luvussa 4 käsitellyt perustarkastelut ja aineiston muokkaus ovat minkä tahansa tutkimuksen kannalta välttämättömiä vaiheita. Joskus keskeisten tulosten välittämiseen saattaa riittää hyvä taulukko tai kuva, mutta yleensä niidenkin aikaansaaminen edellyttää pidemmälle meneviä tarkasteluja kuten aineiston tiivistämistä. Näissä tehtävissä tarvitaan tilastollisia malleja ja menetelmiä. Tässä luvussa käydään läpi joitakin yleisiä mallinnusta ~~koskevia~~ tavoitteita ja oletuksia, syvennetään kirjan alkupuolella esitetyn mittausmallin käsitteistöä ja perehdytään ~~sen tilastollisen~~ tarkasteluun faktorianalyysiksi kutsutulla menetelmällä.

( LYH!

### 5.1 Tavoitteet ja oletukset

Tilastollisen mallintamisen yleisenä tavoitteena on ilmaista tiiviissä muodossa jotain kiinnostavaa tutkittavasta ilmiöstä. Mallintamisessa tarvitaan sekä ilmiön sisällöllistä tuntemusta että asianmukaisten tilastollisten menetelmien osaamista. ~~Näiden osa-alueiden tarkempi sisältö ja painotus kuten myös niiden keskinäinen tasapaino poikkeilevat mallintamisen luonteen.~~

(ei sano mitään!)





Lähtöjatus on, että aineisto muodostua satunnaisotoksen jostakin hyvin määritellystä perusjoukosta. Otoksen avulla pyritään tekemään arvioita perusjoukkoa koskevista ominaisuuksista kuten odotusarvosta tai hajonnasta. Ominaisuuksia on tapana kutsua *parametreiksi* ja niiden arviointia *estimoinniksi*. Esimerkiksi perusjouksen tuntematonta odotusarvoa voidaan estimoida otoksesta lasketulla keskiarvolla. Muita estimoitavia parametreja ovat muun muassa ~~korrelaatiot~~, faktorilataukset ja regressiokertoimet, joihin ~~perehdytään~~ tarkemmin myöhemmin. Johtopäätösten tekemistä tällaisten arvioiden tai estimaattien perusteella kutsutaan *tilastolliseksi päättelyksi*.

*Konfirmatorinen* mallintaminen edellyttää vankkaa sisällöllistä teoriapohjaa ja yleensä myös erikoistuneempien menetelmien tuntemusta. *Eksploratiivinen* mallintaminen perustuu enemmänkin aineistolähtöiseen toimintaan ja tavanomaisempien menetelmien käyttöön. ~~Tyypillinen esitelmä jolittuu~~ näiden ääripäiden väliin. Tämän kirjan lähestymistapa vastaa tyypillisimpiä kyselytutkimuksen tilanteita ja on melko eksploratiivinen, maustettuna sopivasti tietyillä konfirmatorisen mallintamisen piirteillä.

#### Epävarmuudet ja riskit

Tilastollisilla malleilla on yksi yhteinen piirre: niiden avulla ei ilmaista tarkkoja totuuksia, ainoastaan todennäköisyyksiä. Mallinnetaan sitten mitä tahansa ilmiöitä kuinka hyvin tahansa, niin mukana on myös huta satunnaisia epävarmuustekijöitä. Osa vaikutuksia voidaan arvioida ja kenties vähentää muttei täysin poistaa. Johtopäätöksiin sisältyy ~~niin~~ epävarmuuksia ja riskejä, joiden kanssa on opittava tulemaan toimeen. Riskien tunnistaminen ja hallinta kuuluu oleellisesti tilastolliseen mallintamiseen ja päättelyyn.

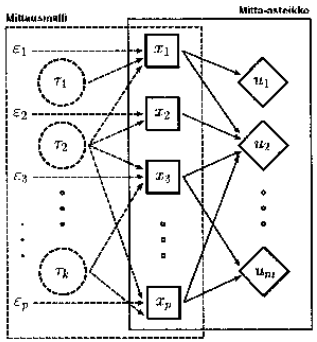
Osa epävarmuuksista voi johtua ilmiön teorian tuntemuksen puutteesta: kaikkia asiaan vaikuttavia tekijöitä ei vain ole osattu ottaa huomioon tai teoria ei kaikilta osin pädekään odotetulla tavalla. Osa voi johtua tiedonkeruusta: otos ei vastauskadon myötä kuvaakaan perusjoukkoa edustavasti tai näyteaineiston keruussa ei saatukaan kiinni tavoiteltuja henkilöitä. Oli tiedonkeruun tapa ja teorian osuus mikä hyvänsä, niin osa epävarmuuksista johtuu mittaamisesta: joidenkin käsitteiden operationalisointi ei ole onnistunut, osaa kysymyksistä ei ole ymmärretty tai kenties kysymyksiä on ollut liikaa, ja vastausväsymys on alkanut vaikuttaa. Epävarmuuden lähteitä on lukuisia.

Joitakin epävarmuuksia voidaan hallita tekemällä sopivia oletuksia, mutta oletukset voivat myös aiheuttaa lisää epävarmuuksia, etenkin jos niitä on liikaa. On helppo tehdä oletuksia, mutta pitäisi myös

# 9.7.2008: Luku 6 alkaa hahmottua *(niinkö luulet?)*

9.7.2008 (itsellens)






**Kuva 6.1.** Mittausmalli ja mitta-asteikko.

Yleensä halutaan tehdä ainakin sellaisia asteikkoja, jotka parhaiten vastaavat teoreettisia tosiarvoja. Tähän viittaa myös kuvan 6.1 visuaalinen asetelma, mutta yhtä hyvin asteikkoja voidaan tehdä muillakin periaatteilla. Siitä syystä niiden lukumääränäkin esiintyy $m$, joka voi olla "mitä tahansa", ~~vaikkapa 1 tai k tai ~~ ~~kuin k~~.

Kuten kuvasta 6.1 näkyy, $x$:illä merkityt muuttujat kuuluvat sekä mittausmalliin että mitta-asteikkoon, joten ne tavallaan kytkevät toisiinsa teorian ja käytännön. Siirtyminen teoriasta käytäntöön tapahtuu faktorianalyysin välityksellä.

Periaatteessa on mahdollista toimia ilmankin mittausmallia, siis ikään kuin "unohtaa" kaikki katkoviivoin esitetty ja tehdä vain suoraan $x$-muuttujista erilaisia yhdistelmiä. Varsinkaan kyselytutkimuksessa tämä ei kuitenkaan ole järkevää, sillä se johtaa helposti liialliseen aineiston ehdoilla etenemiseen ja keinotekoisten ulottuvuuksien "löytämiseen", etenkin kun mittausvirheiden vaikutusta ei ~~näillään tavoin~~ oteta huomioon. Tällaista menettelytapaa vastaa pääkomponenttianalyysiniminen menetelmä, joka historiallisista syistä johtuen on usein sekoitettu faktorianalyysiin. Tässä kirjassa pääkomponenttianalyysilla ei ole sellaisenaan käyttöä, vaikkakin sillä on tiettyjä yhteyksiä eräisiin myöhemmin esiteltäviin menetelmiin.

Mittausmallin ja faktorianalyysin yhdistelmä on paras keino peila-



ta aineiston perusteella nähtyä teorian perusteella pohdittuun. Myös luontevin tapa ~~jatkaa~~ asteikkoihin pohjautuu faktorianalyysiin. Näin ollen faktorianalyysi kattaa kuvan 6.1 kaikki vaiheet, joten sitä voi hyvällä syyllä kutsua kyselytutkimuksen keskeisimmäksi menetelmäksi.

## 6.2 Tiivistäminen asteikoiksi

Asteikkojen tehtävä on aineiston tiivistäminen, joka on välttämätöntä jo pelkästään muuttujien määrän vuoksi. Lisäksi on järkevää ~~pyrkiä~~ erottelemaan muuttujiin sisältyvä todellinen informaatio satunnaisten mittausvirheiden mukanaantuomasta vaihtelusta. (Kaikki ~~osiot~~ eivät muunnostenkaan jälkeen ole yhtä hyviä, joten jatkoanalyyseja varten kannattaa paremmille ~~osioille~~ antaa enemmän painoa kuin huonommille.

### 6.2.1 Faktoripistemäärät

### 6.2.2 Summamuuttujat

### 6.2.3 ~~Reliabiliteetin arviointi~~

#### Mitta-asteikon reliabiliteetti

Reliabiliteetti on asteikon ominaisuus, asteikkojen (taustalla oletukset mitt.virh.!!!) tosiarv. El ole.

myös osiot, tavall. myös asteikko...

Reliabiliteetti tarkoittaa **mittarin** tarkkuutta: mitä vähemmän mittarin sisältyy satunnaista mittausvirhettä, sitä tarkempi mittari on. Kirjallisuudessa reliabiliteetti on käännetty usein luotettavuudeksi tai toistettavuudeksi, mutta nämä

Huomaa, että jos luvusta 2 mainitun tutkijan "kohtelu" on jokseenkin oikea, on vaikea ~~kiistä~~ arvioida, miten hyvin muodostetut ast. vastaavat mitattaan teor. ulottuv.

# 9.7.2008: Luku 7 alkaa hahmottua *(näkisi vain!)*

9.7.2008 (itsellemi)



ja pitävätkö niihin sisältyvät oletukset riittävästi paikkaansa analysoitavan aineiston näkökulmasta.

Nyt laadittavissa malleissa selitetään yhden tai useamman muuttujan avulla jonkin muuttujan kuvaamaa asiaa. Ennen sitä on syytä käydä läpi joukko aineiston havaintoja ja muuttujia koskevia oletuksia. Lisäksi katsotaan, miihin tämä kaikki sijoittuu luvussa 2 käsitellyssä mittauskehikossa.

### 7.2.1 Havainnot ja muuttujat

Aineistosta tehdään erilaisia oletuksia sekä havaintojen että muuttujien suhteen.

s kmm`fal 51

#### Muuttujien roolit

#### Muuttujien väliset yhteydet

Muuttujien välisiä yhteyksiä tutkimalla voidaan selittää jonkin muuttujan kuvaamaa asiaa yhden tai yleensä useamman muun muuttujan avulla. Esimerkiksi voitaisiin tutkia, mitkä tekijät saattaisivat selittää käsityksiä omasta ulkonäöstä.

#### Jakaumaoletukset

Tavallisen regressioanalyysin perustana oleva todennäköisyysjakauma on normaalijakauma. Se ei kuitenkaan tarkoita, että kaikkien muuttujien tulisi olla normaalisti jakautuneita, sillä oletus koskee vain selitettävää $y$-muuttujaa. Koska se on kuitenkin mallintamisen kannalta keskeisin muuttuja, se asettaa heti tiettyjä rajoituksia. Regressioanalyysi perusmuodossaan soveltuukin vain sellaisten ilmiöiden mallintamiseen, joissa selitettävän muuttujan jakauma on tyypiltään normaalijakauman kaltainen, symmetrinen ja yksihuippuinen.

### 7.2.2 Mallit mittauskehikossa

Mallintaminen sijoittuu kuvan 7.1 mittauskehikkoon luontevasti. Kuvan oikeassa reunassa vertailuperuste edustaa selitettävää, varsinaisena mielenkiinnon kohteena olevaa asiaa. Kaaviossa se on kuvattu moniulotteisena, mutta tavallisessa regressiomallintamisessa käyttöön tulee vain yksi selitettävä muuttuja kerrallaan. Kaaviossa on käytetty symbolina kirjainta $y$, joka on myös regressiomalleissa tyypillisin tunnus selitettävälle muuttujalle.



Vastaava tunnus selittäville muuttujille on usein $x$, ja tavanomainen regressioanalyysin asetelma onkin, että $y$:tä selitetään yhdellä tai useammalla $x$:llä. Nyt kuitenkin mittauskehikon merkitys koko asetelman pohjustamisessa näkyy selkeästi: sen sijaan että regressiomallintamista tehtäisiin alkuperäisillä $x$:illä, niistä tiivistetäänkin vähempiulotteisia yhdistelmiä, $a$:lla merkittyjä asteikkoja. Mikään ei sinänsä estä $x$:ien käyttämistä suoraan "ohi" varsinaisten asteikkojen, mutta ilman tiivistystä selitysmallien rakentamisesta aiheutuu ongelmia: $x$:lä on liikaa, niihin sisältyy mittausvirhettä ja ne mittaavat monelta osin samoja asioita.

Mallintamisen tulokset tiivistyvät tulosasteikoiksi, joita kuvassa 7.1 on merkitty kirjaimella $z$. Regressioanalyysin osalta tätä asteikkoa kutsutaan mallin sovitteeksi tai ennustemalleissa ennusteeksi. Kuvan kaksisuuntainen nuoli vertailuperusteen ja tulosasteikon välillä kuvaa arviota siitä, miten hyvä selitysmalleja on onnistuttu löytämään.

## 7.3 Regressioanalyysi

### 7.3.1 Selittäjien valinta

### 7.3.2 Ennustevaliditeetin arviointi

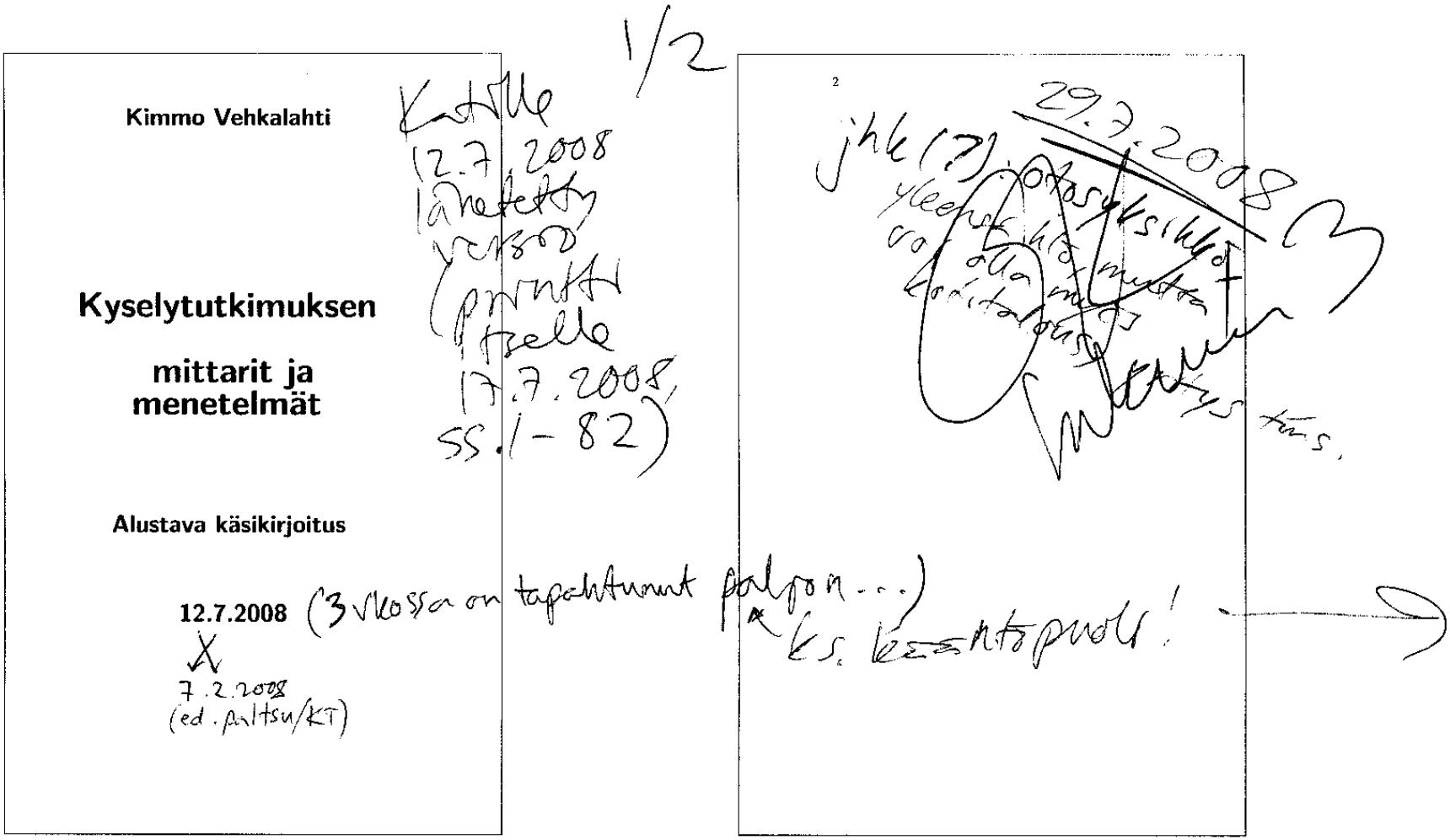
Mikäli aineistoon ja tutkimusasetelmaan sisältyy ajallisia ulottuvuuksia, voidaan puhua myös ennustamisesta tai ennusteiden laatimisesta.

## 7.4 Regressiodiagnostiikka

~ da 2 !

kehikon pala.

# 12.7.2008: Uusi käsikirjoitusversio (114 s.) *(Ohoh!)*



1/2

Kimmo Vehkalahti

Katille 12.7.2008 lähetetty versio (printti itselle 17.7.2008, ss. 1–82)

**Kyselytutkimuksen**

**mittarit ja menetelmät**

**Alustava käsikirjoitus**

**12.7.2008** (3 vkossa on tapahtunut paljon...)

7.2.2008 (ed. paltsu/KT)

2

29.7.2008

jhk (?): otosyksikkö yleensä klo, mutta voi olla mitä koditalous

tys tms.

ks. kääntöpuoli!

# 12.7.2008: Uusi käsikirjoitusversio: jotain valmistakin

12.7.2008




3.1.3 Harkinnanvarainen näyte . . . 41
3.2 Kyselylomake . . . 42
3.2.1 Nettilomake vai paperi . . . 42
3.2.2 Saatekirje ja kato . . . 42
3.2.3 Selkeys ja tiiviys . . . 43
3.3 Lomakkeelta aineistoksi . . . 43

**4 Perustarkastelut ja esikäsittely** **44**
4.1 Aineiston selailu . . . 44
4.1.1 Havainnot ja muuttujat . . . 45
4.2 Yhden muuttujan tarkastelut . . . 46
4.2.1 Jakaumat . . . 46
4.2.2 Tunnusluvut . . . 47
4.2.3 Kuvat . . . 52
4.3 Muunnokset ja uudelleenkoodaukset . . . 55
4.4 Usean muuttujan tarkastelut . . . 58
4.4.1 Taulukot . . . 58
4.4.2 Kuvat . . . 60
4.4.3 Tunnusluvut . . . 62
4.5 Aineiston muokkaaminen . . . 62

**5 Tilastollinen mallintaminen** **63**
5.1 Tavoitteet ja oletukset . . . 63
5.1.1 Mallit ja menetelmät . . . 65
5.2 Mittausmalli . . . 65
5.2.1 Mallin laatiminen käytännössä . . . 67
5.3 Faktorianalyysi . . . 68
5.3.1 Käsitteet ja parametrit . . . 68
5.3.2 Havainnot ja muuttujat . . . 69
5.3.3 Faktorointi . . . 71
5.3.4 Rakennevaliditeetti . . . 71
5.4 Faktorirakenteiden vertailu . . . 71

**6 Aineiston tiivistäminen** **74**
6.1 Mittausmalli ja mitta-asteikko . . . 74
6.2 Tiivistäminen asteikoiksi . . . 76
6.2.1 Faktoripistemäärät . . . 76
6.2.2 Summamuuttujat . . . 76
6.2.3 Mitta-asteikon reliabiliteetti . . . 76





**7 Selitysmallien laatiminen** **77**
7.1 Mittauskehikko . . . 77
7.1.1 Tulosasteikko ja vertailuperuste . . . 78
7.1.2 Mallit mittauskehikossa . . . 81
7.2 Regressioanalyysi . . . 81
7.2.1 Havainnot ja muuttujat . . . 81
7.2.2 Selittäjien valinta . . . 82
7.2.3 Tulosasteikon ennustevaliditeetti . . . 82
7.3 Regressiodiagnostiikka . . . 82

**8 Vastaajien ryhmittely** **83**
8.1 Etäisyyksien mittaaminen . . . 83
8.2 Hierarkinen ryhmittely . . . 83
8.3 Muita ryhmittelymenetelmiä . . . 83

**9 Vastaajaryhmien analysointi** **84**
9.1 Erotteluanalyysi . . . 84
9.2 Varianssianalyysi . . . 84
9.3 Korrespondenssianalyysi . . . 84

**Liitteet**

**A Ohjelmistot ja dokumentointi** **85**
A.1 Tietokone ja ohjelmistot . . . 86
A.1.1 Kirjan ohjelmistot . . . 86
A.1.2 Muita ohjelmistoja . . . 88
A.1.3 Ohjelmistot varmenetelmät . . . 90
A.2 Töiden hallinta . . . 91
A.2.1 Töiden tallettaminen . . . 91
A.2.2 Töiden tavoittaminen . . . 92
A.2.3 Töiden varmistaminen . . . 95
A.3 Dokumentoiva työskentelytapa . . . 96
A.4 Survo lyhyesti . . . 99
A.4.1 Aineiston selailu . . . 100
A.5 SPSS lyhyesti . . . 100
A.5.1 Aineiston selailu . . . 101
A.6 Aineiston perustaminen ja ~~töiden~~ hallinta . . . 101
A.6.1 Työvaiheet SPSS-ympäristössä . . . 102
A.6.2 Työvaiheet Survo-ympäristössä . . . 105

**B Kuvien ja tulostelden työkaavioita** **108**

**Kirjallisuutta** **108**


edelleen aivan auki! (ei enää :-)

aivan sekaisin! Ei enää!

OK!

B Kuvien työt.

Survo/SPSS

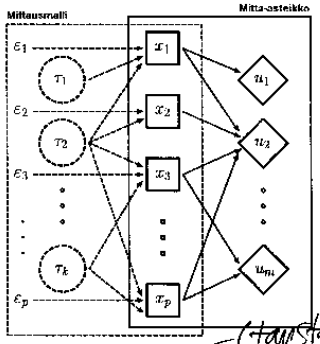
tarve ei tietoa

(Survo) linkki/ref? \pageref

# 12.7.2008: Uusi käsikirjoitusversio: taas mennään...

12.7.2008 → KT






**Kuva 6.1.** Mittausmalli ja mitta-asteikko.

ne itse asiassa ovatkin. Koska asteikot ovat osioiden yhdistelmiä, niin nuolet osoittavat osioista eteenpäin. Kaaviossa liikutaan siis kokoajan vasemmalta oikealle.

Yleensä halutaan tehdä ainakin sellaisia asteikkoja, jotka sisältöltään parhaiten vastaisivat teoreettisia tosiarvoja. Tähän viittaa myös kuvan 6.1 visuaalinen asetelma, mutta yhtä hyvin asteikkoja voidaan tehdä muillakin periaatteilla. Siitä syystä niiden lukumääränäkin esiintyy $m$, joka voi olla "mitä tahansa" – riippumatta $k$:sta ja $p$:stä. Huomaa, että jos luvussa 2 mainittu tutkijan "kotitehtävä" on jäänyt tekemättä, on aika vaikea arvioida, miten hyvin muodostetut asteikot vastaavat itään teoreettisia ulottuvuuksia.

Kuten kuvasta 6.1 näkyy, $x$-muuttujat kuuluvat sekä mittausmalliin että mitta-asteikkoon, joten ne tavallaan kytkevät toisiinsa teorian ja käytännön. Siirtyminen teoriasta käytäntöön tapahtuu faktorianalyysin välityksellä, kuten edellisessä luvussa nähtiin.

Periaatteessa on mahdollista toimia ilmankin mittausmallia, siis ikään kuin "unohtaa" kaikki katkoviivoin esitetty ja tehdä vain suoraan $x$-muuttujista erilaisia yhdistelmiä. Varsinkaan kyselytutkimuksessa tämä ei kuitenkaan ole järkevää, sillä se johtaa helposti liialliseen aineiston ehdoilla etenemiseen ja keinotekoisten ulottuvuuksien "löytämiseen", etenkin kun mittausvirheiden vaikutustakaan ei tällöin mi-



tenkään oteta huomioon. Tällaista menettelytapaa vastaa menetelmä nimeltään *pääkomponenttianalyysi*, joka historiallisista syistä johtuen on usein sekoitettu faktorianalyysiin. Tässä kirjassa pääkomponenttianalyysilla ei ole sellaisenaan käyttöä, vaikkakin sillä on tiettyjä teknisiä yhteyksiä eräisiin myöhemmin esiteltäviin menetelmiin.

Mittausmallin ja faktorianalyysin yhdistelmä on paras keino peilata aineiston perusteella nähtyä teorian perusteella pohdittuun. Myös seuraavaksi tarkasteltavista vaihtoehdoista luontevin tapa siirtyä mallista asteikkoihin pohjautuu faktorianalyysiin. Näin ollen faktorianalyysi kattaa kuvan 6.1 kaikki vaiheet, joten sitä voi hyvällä syyllä kutsua kyselytutkimuksen keskeisimmäksi menetelmäksi.

## 6.2 Tiivistäminen asteikoiksi

Asteikkojen tehtävä on aineiston tiivistäminen, joka on välttämätöntä jo pelkästään muuttujien määrän vuoksi. Lisäksi on järkevää pyrkiä erottelemaan muuttujiin sisältyvä todellinen informaatio satunnaisten mittausvirheiden vääjäämättä mukanaantuomasta vaihtelusta. Kaikki muuttujat eivät muunnostenkaan jälkeen ole yhtä hyviä, joten jatkoanalyyseja varten kannattaa paremmille muuttujille antaa enemmän painoa kuin huonommille.

### 6.2.1 Faktoripistemäärät

yleistä + MV:n asetelma

### 6.2.2 Summamuuttujat

yleistä + MV:n asetelma

### 6.2.3 Mitta-asteikon reliabiliteetti

yleistä + MV:n asetelma

Reliabiliteetti on asteikon ominaisuus. asteikkojen (taustalla oletukset mitt.virh.!!!) tosiarv. EI ole.

myös osiot, tavall. myös asteikko...

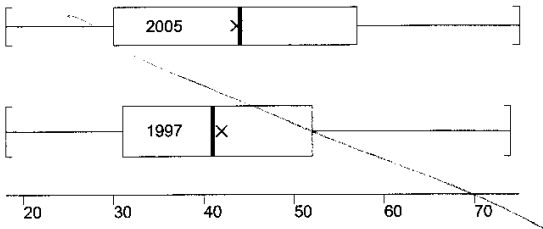
Reliabiliteetti tarkoittaa **mittarin** tarkkuutta: mitä vähemmän mittariin sisältyy satunnaista mittausvirhettä, sitä tarkempi mittari on. Kirjallisuudessa reliabiliteetti on käännetty usein luotettavuudeksi tai toistettavuudeksi, mutta nämä

PCA vain 1:n. luku?

# 24.7.2008: Luku 4 hahmottuu (vuorostaan) uudelleen



24.7.2008




(VRT. ed. taul!)

**Kuva 4.4.** Laatikkokuva

KESKEN!!

### Hajontakuva

Tilastollisista kuvista ainakin tutkimustyössä tärkein lienee hajontakuva, joka esittää kahden ~~enemmän tai vähemmän~~ jatkuvan muuttujan yhteisjakaumaa pisteparvena. Tämän kuvatyypin lukuisat muunnelmat ovat käytössä menetelmien erikoiskuvina. Myös laatikkokuvan on

yhteys korrelaatioon (2 mjan tunnusluku, ALLA!), FA-luvussa lisää . . .

### 4.4.3 Tunnusluvut

korrelaatio lähinnä!

chi2? (voisi mainita?)

US?

## 4.5 Aineiston muokkaaminen

Puuttuvat tiedot ovat ongelmallisia, mutta niitä ei voi ~~empiirisessä~~ tutkimuksessa välttää. Taustamuuttujien osalta niille ei myöskään voi paljoa tehdä. On vain pidettävä puuttuvat tiedot mukana ikään kuin yhtenä luokkana ja raportoidessa päätettävä erikseen millä tavoin ne kulloinkin otetaan huomioon. Tunnuslukuihinhan ne eivät millään lailla sisälly.

~~luokittelujä? muunnoksia? ei vaan~~ **~~yhden ja useamman mjan tark. väliin~~!**

~~virheiden korjaus (=eldokun.!)~~

kato? puuttuvien korvaus? ... Vakavan lisähaasteen asettavat joko osittain tai kokonaan puuttuvat tiedot. vastauskato... Tiedon puuttumisen ongelmaan tilastotiede on pyrkinyt vastaamaan kehittämällä tapoja paikata eli imputoida aineistoa, jotta se mahdollisimman hyvin vastaisi täydellistä. Mikään ei tietenkään voi ihan oikeasti korvata

Tee kuntoon pstm! Tikc:n matsku! → kaavio mol. ohj.

ellei tehty postit täydennys rekisteriä (ref. luku 3?)

tänne pitäen EOS:stä ??

# 24.7.2008: Luku 5 hahmottuu (jälleen) uudelleen

SC 24.7.2008 (4-7)

Tännekö jotain (hypot.) testauksesta? (ora mallintamista yhtys johtop. perusj. ine.!) VAIKO EI? (ks. JK08K!)

LUKU 5

## Tilastollinen mallintaminen

Perustarkastelut ja muokkaukset ovat välttämättömiä vaiheita kyselytutkimusaineistojen analysoinnissa. Joskus keskeisten tulosten välittämiseen saattaa riittää hyvä taulukko tai kuva, mutta yleensä niidenkin aikaansaaminen edellyttää pidemmälle meneviä tarkasteluja. Näissä tehtävissä tarvitaan tilastollisia malleja ja menetelmiä.

Tässä luvussa käydään läpi yleisiä mallintamisen tavoitteita ja oletuksia, tarkennetaan kirjan alkupuolella esitetyn mittausmallin käsitteistöä ja perehdytään mallin tarkasteluun faktorianalyysiksi kutsutulla tilastollisella menetelmällä.

### 5.1 Tavoitteet ja oletukset

Tilastollisen mallintamisen yleisenä tavoitteena on ilmaista tiiviissä muodossa jotain kiinnostavaa tutkittavasta ilmiöstä. Tehtävässä tarvitaan sekä ilmiön sisällöllistä tuntemusta että asianmukaisten tilastollisten menetelmien osaamista.

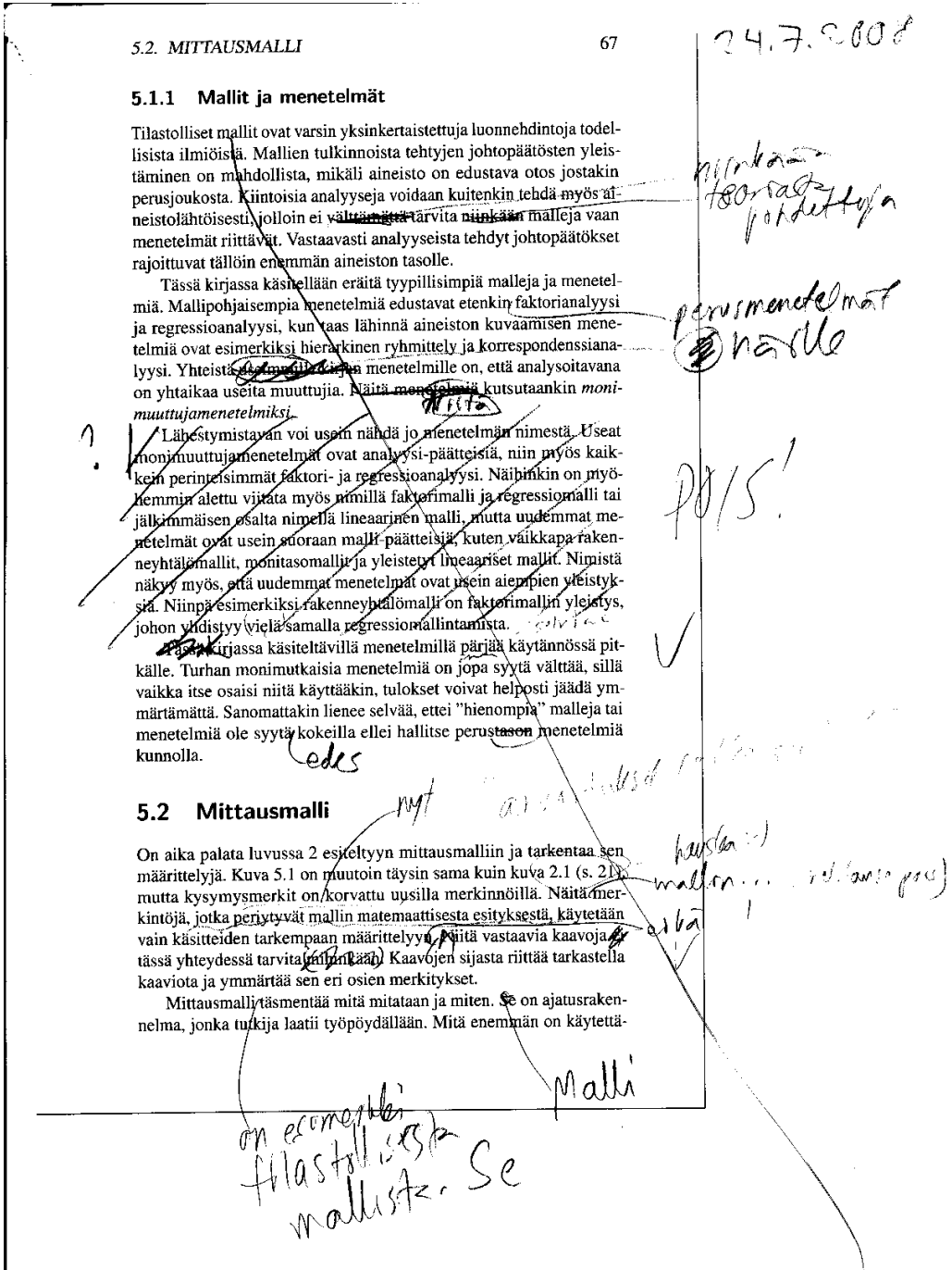
Yleisenä oletuksena on, että aineisto muodostaa satunnaisotoksen jostakin hyvin määritellystä perusjoukosta (vrt. luku 3). Otoksen avulla pyritään tekemään arvioita perusjoukkoa koskevista ominaisuuksista kuten odotusarvoista, hajonnoista, korrelaatioista tai todennäköisyyksistä. Tällaisia ominaisuuksia kutsutaan *parametreiksi* ja

65

Vaikke rumec nataa, pus :-)

Tää on hyvä :)

# 24.7.2008: Luku 5 muotoutuu – insert, delete...



5.2. MITTAUSMALLI 67

### 5.1.1 Mallit ja menetelmät

Tilastolliset mallit ovat varsin yksinkertaistettuja luonnehdintoja todellisista ilmiöistä. Mallien tulkinnoista tehtyjen johtopäätösten yleistäminen on mahdollista, mikäli aineisto on edustava otos jostakin perusjoukosta. Kiintoisia analyyseja voidaan kuitenkin tehdä myös aineistolähtöisesti, jolloin ei välttämättä tarvita niinkään malleja vaan menetelmät riittävät. Vastaavasti analyyseista tehdyt johtopäätökset rajoittuvat tällöin enemmän aineiston tasolle.

Tässä kirjassa käsitellään eräitä tyypillisimpiä malleja ja menetelmiä. Mallipohjaisempia menetelmiä edustavat etenkin faktorianalyysi ja regressioanalyysi, kun taas lähinnä aineiston kuvaamisen menetelmiä ovat esimerkiksi hierarkinen ryhmittely ja korrespondenssianalyysi. Yhteistä useimmille kirjan menetelmille on, että analysoitavana on yhtaikaa useita muuttujia. Näitä menetelmiä kutsutaankin *monimuuttujamenetelmiksi*.

Lähestymistavan voi usein nähdä jo menetelmän nimestä. Useat monimuuttujamenetelmät ovat analyysi-päätteisiä, niin myös kaikkein perinteisimmät faktori- ja regressioanalyysi. Näihinkin on myöhemmin alettu viitata myös nimillä faktorimalli ja regressiomalli tai jälkimmäisen osalta nimellä lineaarinen malli, mutta uudemmat menetelmät ovat usein suoraan malli-päätteisiä, kuten vaikkapa rakenneyhtälömallit, monitasomallit ja yleistetyt lineaariset mallit. Nimistä näkyy myös, että uudemmat menetelmät ovat usein aiempien yleistyksiä. Niinpä esimerkiksi rakenneyhtälömalli on faktorimallin yleistys, johon yhdistyy vielä samalla regressiomallintamista.

Tässä kirjassa käsiteltävillä menetelmillä pärjää käytännössä pitkälle. Turhan monimutkaisia menetelmiä on jopa syytä välttää, sillä vaikka itse osaisi niitä käyttääkin, tulokset voivat helposti jäädä ymmärtämättä. Sanomattakin lienee selvää, ettei "hienompia" malleja tai menetelmiä ole syytä kokeilla ellei hallitse perustason menetelmiä kunnolla.

## 5.2 Mittausmalli

On aika palata luvussa 2 esiteltyyn mittausmalliin ja tarkentaa sen määrittelyjä. Kuva 5.1 on muutoin täysin sama kuin kuva 2.1 (s. 21), mutta kysymysmerkit on korvattu uusilla merkinnöillä. Näitä merkintöjä, jotka periytyvät mallin matemaattisesta esityksestä, käytetään vain käsitteiden tarkempaan määrittelyyn. Niitä vastaavia kaavoja ei tässä yhteydessä tarvita (niinkään). Kaavojen sijasta riittää tarkastella kaaviota ja ymmärtää sen eri osien merkitykset.

Mittausmalli täsmentää mitä mitataan ja miten. Se on ajatusrakennelma, jonka tutkija laatii työpöydällään. Mitä enemmän on käytettä-

# 24.7.2008: Luku 5 *"kesken – iisi biisi jatkaa"*



24.7.2008

(yhteenv.)

Tämä teorian ja käytännön ristiriidalta kuulostava seikka ei estä faktorianalyysin soveltamista. Käytännössä menetelmä sietää poikkeamia joita empiirisiin aineistoihin joka tapauksessa sisältyy.

**Aineiston koko**

Yksi yleisimpiä tilastotieteilijälle esitettyjä kysymyksiä niin faktorianalyysin kuin muidenkin analyysien yhteydessä koskee aineiston kokoa, siis havaintojen lukumäärää. Kysymykseen ei missään tapauksessa ole yksiselitteisiä vastauksia, sillä asiaan vaikuttavat monet seikat, joista sitäpaitsi vain osa on tilastollisia. Aineiston tavoitekoko voi määräytyä tutkimusasetelman perusteella, mutta lopullisen koon ratkaisevat usein tiedonkeruun kustannukset ja aikataulu, lopulta myös vastausaktiivisuus ja vastausten taso. On helppo neuvoa keräämään niin paljon havaintoja kuin mahdollista, mutta taaskaan laatua ei voi korvata määrällä. Pienempienkin aineistojen analysointi on mahdollista tietyt järkevyysehdot huomioonottaen.

vastausprosentti laatu

Mittausmallin yhteydessä todettiin jo, että muuttujia on oltava enemmän kuin tosiarvoja eli faktoreita. Tämä ei kuitenkaan vielä riitä. Faktorianalyysi edellyttää käytännössä melko paljon havaintoja, koska siinä estimoidaan runsaasti parametreja. Vaikka faktoreiden määrä olisikin kohtuullinen eikä muuttujia olisi paljonkaan, on estimoitavia faktorilatauksia äkkiä useita kymmeniä tai satoja. Lisäksi tulevat vielä mittausvirheiden varianssit, joita on saman verran kuin muuttujia.

Esimerkkitilanteessa jota seuraavaksi käsitellään, rajoitutaan alukseen pelkästään aineiston ensimmäiseen keräysvuoteen. Tällöin liikutaan jo hieman arveluttavilla rajoilla, sillä 53 muuttujan estimoitavia parametreja kolmen faktorin asetelmassa on 212, havaintoja on 265. Kun asetelmaa lavennetaan neliulotteiseksi, housee parametrien lukumäärä jo 265:een, siis miltei samaan kuin osa-aineiston koko. Estimoinnin vakauden ja tulosten luotettavuuden kannalta olisi parempi olla useita havaintoja kutakin parametria kohden. Sama pätee muuhunkin tilastolliseen mallintamiseen.

(1997) koska ei oletettaa etta tilanne olisi rakenteellisesti samanlaista kahdeksan vuoden nykyhetkeen.

### 5.3.3 Faktorointi

KESKEN, iisi biisi jatkaa...

Katsastellaan faktorimatriisia... huom! ML:n "hyvyys": valittaa herkästi jos ollaan rajoilla! (voi mennä läpikin, mutta... "suuntaa-antava"??)

perinteisen tilastollisen päättelyn ydin: uskottavuuspäättely, suurimman uskottavuuden menetelmä, pyrkii löytämään parametreille aineiston valossa uskottavimmat estimaatit

**idea:** 1) ML-ratkaisu, EI rotaatiota, ei järjestelyä, ei korostuksia! mjat aakkosjärjestyksessä, faktorit numerojärj.; "vaikeaselkoinen"; kuitenkin havaittavissa ettei 3-ulott. idea toimi 2) otetaan yksi faktori

selkä tauty muutoksen 31.7. vaksi

# 24.7.2008: Liite B:n kaaviotyöt ja valoisat (?) yöt

24.7.2008

LIITE B

## Kuvien ja tulosteiden kaavioita

laadintaan ? tekemisen käytettyjä pätevät annakin versiosta 15 (?)

1.8.2008

Tässä liitteessä on kirjan kuvien ja tulosteiden Survo- ja SPSS-kaavioita. Joihinkin on tässä yhteydessä lisätty rivinumerot viittauksia varten. Kaaviot on ryhmitelty luvuittain.

**Luku 4** → 4.1 jne ?

**Kaavio 1.** Tuloste 4.1, sivu 47 (Survo).

```
STAT MV2007A CUR+1 / VARS=Vuosi
```

Kaaviolla 1 lasketaan jakaumia ja tunnuslukuja Survon STAT-operaatiolla. Tulokset tulevat komentorivin alapuolelle (CUR+1). Täsmennys VARS luettelee käytettävät muuttujat (tässä vain Vuosi).

**Kaavio 2.** Tuloste 4.2, sivu 48 (Survo).

```
STAT MV2007A CUR+1 / VARS=sv CLASSMAX=30
```

Kaaviossa 2 täsmennys CLASSMAX määrää luokkien maksimimäärän sv-muuttujan automaattisessa luokittelussa. Oletusarvo on 30,

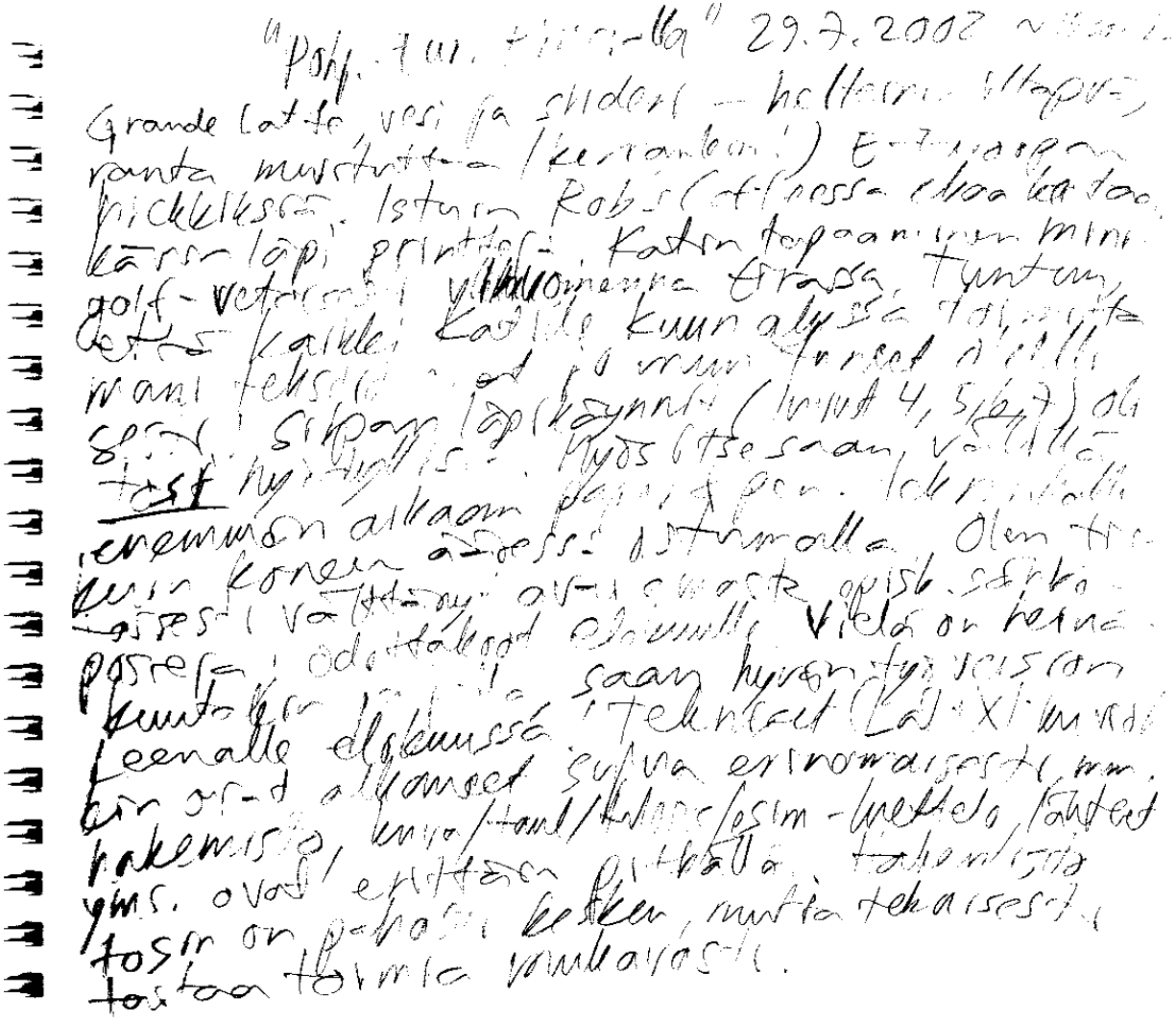
113

ginen idea:
Näihin viittauksel tekstissä MARG/MAR-LIN kohdalta, esim
Kaavio sivulla 113
TA1 (kaavio s. 113)
TA1 (ks. kaavio s. 113)
mutta jos margi-naalia alkaa käyttäy sitä pitäisi käyttää jkk merkittäräpäänkin ... ?!

tässä liitteessä tarvitaan näille nimel, jotta: "... kaaviossa xx lasketaan ..." mutta globaalisti niitä ei tarvita siis "vakaali" viittaus riittäisi.

• KT kuitenkin oikeasa: tekstin lukiessa [illegible] kaavio

# 29.7.2008: ”Pohjois-Euroopan Rivieralla”



- *”Vielä on heinäkuutakin jäljellä, saan hyvän version LP:lle elokuussa...”*

# 29.7.2008: "Pohjois-Euroopan Rivieralla" (Aurinkolahti)

Heinäkuun loppupuoliskolla 2008:
(sekalaisessa järjestyksessä mieleen tulevia 30.7.2008)

- johdanto kirjoitettu ja järjestetty (pari kertaa) uudelleen
- hakemisto saatu eräiltä yks.kohdilta toimimaan paremmin (skandit lähteiden nimissä, pilkut kuten Leena toivoi 12/07, siistimpi ulkoasu)
- luettelo kuvista, tulosteista ym. aikaan omin nenvoin (yhdistetty, ei triviaali!)
- luvut 8 ja 9 alulle
- lukujen 5–7 kokonaisuus selkeämmäksi
- luvut 2 ja 3 lähes "valmiiksi" (puuttuuko vielä kohtia?)
- luvun 4 työnjako liitteiden kanssa: periaate selvä, toteutusvielä kesken (ohjelmisto spesifit asiat liitteisiin)
- lukujen ja liitteiden otsikoiden tarkennuksia
- liitteen B toteutustavan määrittely ja testaus
- liitteen A organisointi ex-luvusta 3 yhdistellen
- rivivälin säätäminen (teksti, sis.luett., tulosteet ym.)
- floatien — " — (uusi paketti, tarkemmat mahd.)
- ? (mitä muuta?)

TAVOITE:

Leenalle, Katille, Maaritille, Sirpalle
KOKO TEKSTI
aika pian!
(paperina?) • ENSIN Leenalle valikoima näytesivuja? (LEVEYS!)

Mitä seuraavaksi?

- luv. 4 aukkoja (4.4.2, 4.4.3)
- luv. 5, 6, 7 —"—
- luvut 8, 9 alkutek.
- liite A kesken + tiivistämättä
- liite B alussa
- hakemisto (lopuksi?)
- kotisivu??
- kiitokset

# 3.8.2008: Sisällysluettelon editointia

KMM 3.8.2008

## Sisältö


Alkusanat . . . . . 3

**1 Johdanto** **13**
1.1 Kyselytutkimus . . . . . 13
1.1.1 Mittarit ja menetelmät . . . . . 14
1.2 Kirjan rakenne ja sisältö . . . . . 15

**2 Mittaaminen** ~~kyselytutkimuksessa~~ ja tiedonkeruu
2.1 Johdatteleva esimerkki . . . . . 19
2.2 Kyselylomake mittausvälineenä . . . . . 22
2.2.1 Moniulotteisuus . . . . . 22
2.2.2 Osiot ja mittarit . . . . . 24
2.2.3 Avoimet ja suljetut osiot . . . . . 25
2.3 Mittauksen taso . . . . . 27
2.3.1 Luokittelu . . . . . 28
2.3.2 Järjestäminen . . . . . 30
2.3.3 Numeerinen mittaus . . . . . 33
2.4 Mittauksen luotettavuus . . . . . 39
2.4.1 Validiteetti ja reliabiliteetti . . . . . 40

Tiedonkeruun (-tavan ?)

~~3 Tiedonkeruun ja sisältön vaikutuksia~~ ~~43~~
~~3.1 Vaihtoehdot ja~~ vaikutukset . . . . . 44
2.5 3.1.1 Perusjoukko ja otos . . . . . 44
3.1.2 Kokonaistutkimus ja rekisterit . . . . . 46
3.1.3 Harkinnanvarainen näyte . . . . . 47
2.6 Kyselylomake tiedonkeruuvälineenä . . . . . 48
2.7 Lomakkeelta aineistoksi . . . . . 49 → ?

3 **Perustarkastelut ja esikäsittely** **51** ← ?
4.1 Aineistoon tutustuminen . . . . . 51
4.1.1 Havainnot ja muuttujat . . . . . 51
4.2 Yhden muuttujan tarkastelut . . . . . 52
4.2.1 Jakaumat . . . . . 53
4.2.2 Tunnusluvut . . . . . 54

Tilastollinen aineisto ?
Kyselytutkimuksen ?


LAAJENNA: kehikkeiksi samaan kuin kaikkeen + kerrassa esit. menetelmät vol sivu...

OK

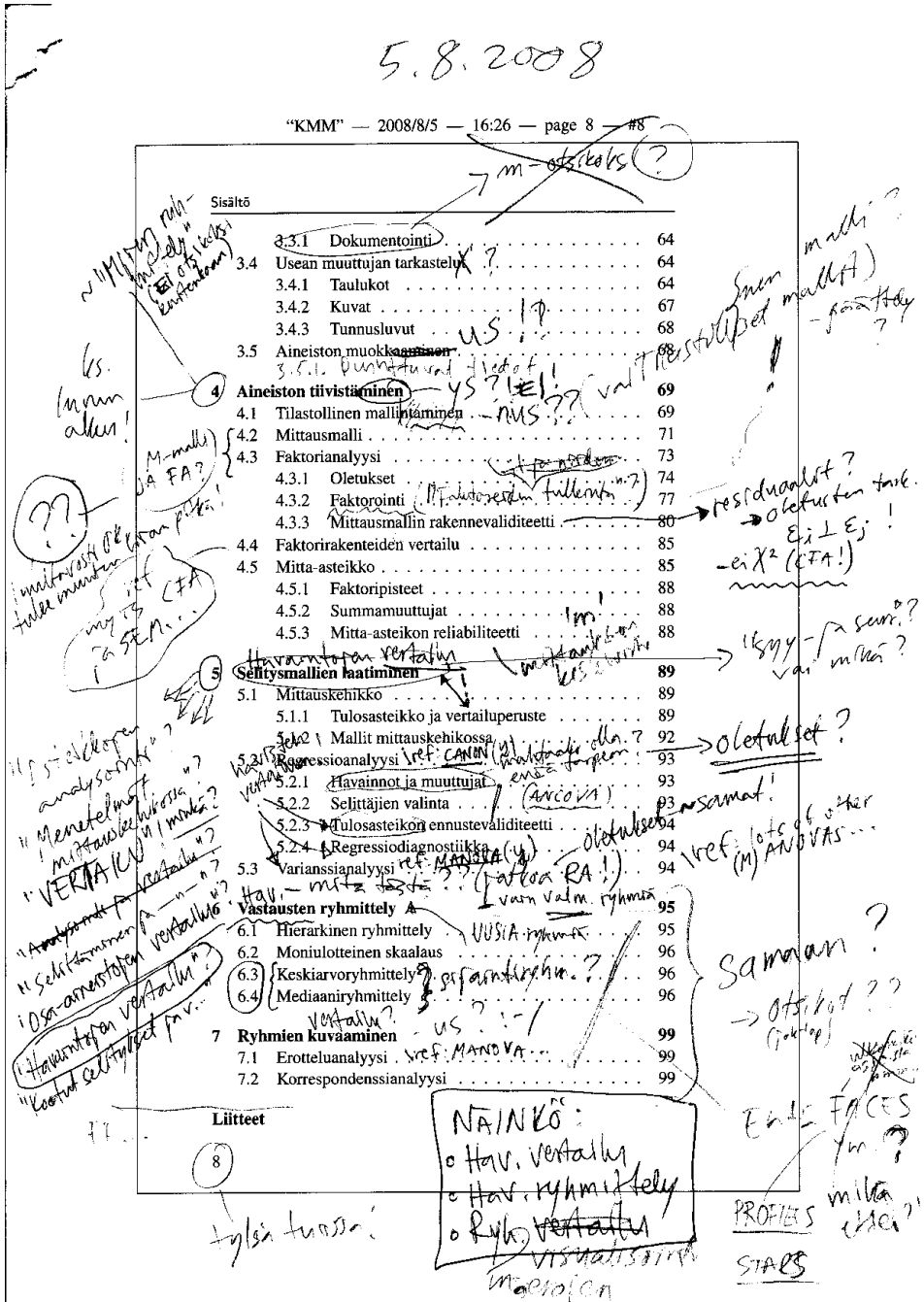
VAI vain Aineisto ?

tai perehtyminen

vai Aineistoon tutustuminen ?
(jollain 1. alaotsi. jot. muuta)

vai Aineiston esikäsittely ?

# 5.8.2008: Sisällysluettelon editointia (!)



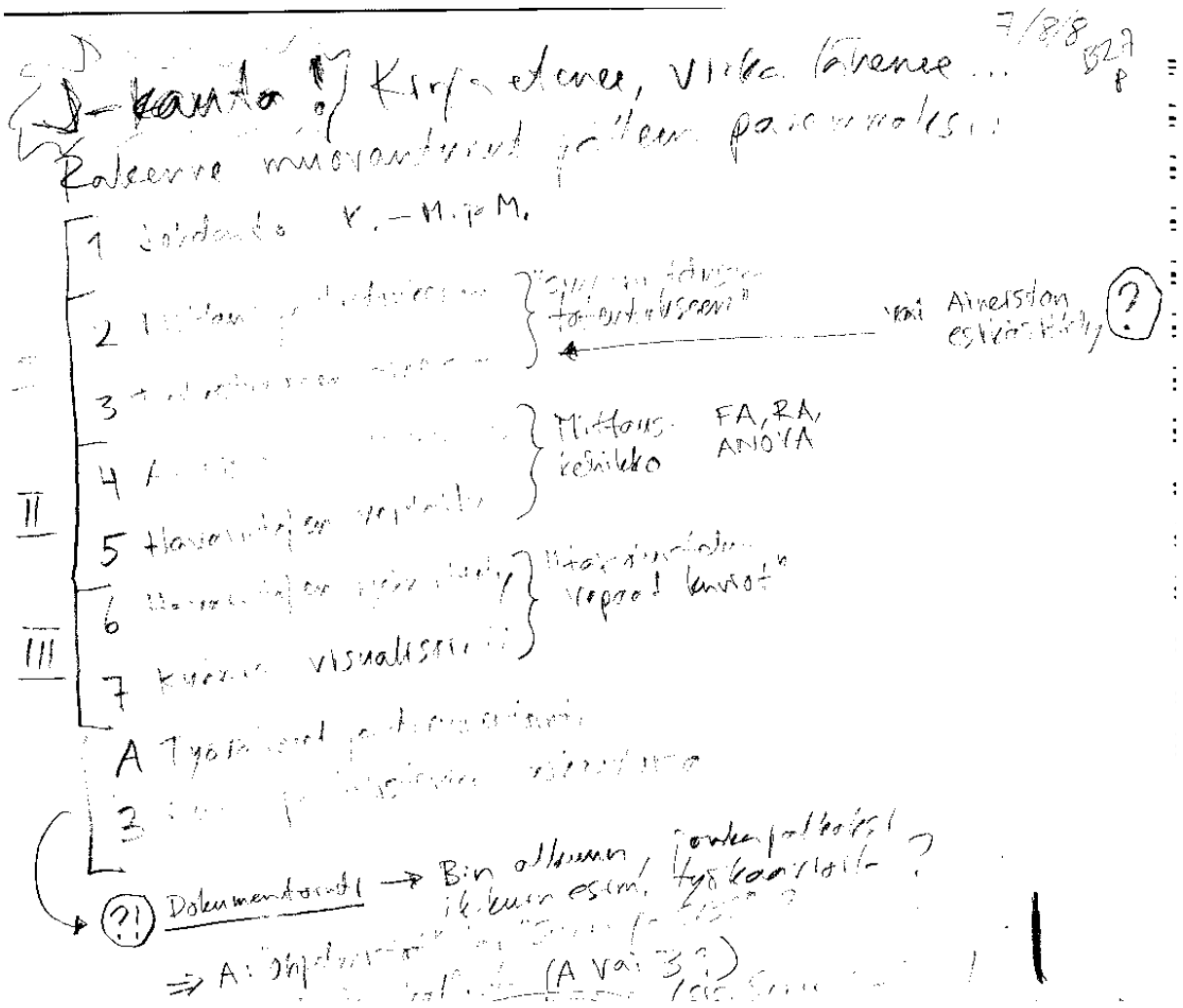
5.8.2008

"KMM" — 2008/8/5 — 16:26 — page 8 — #8

Sisältö

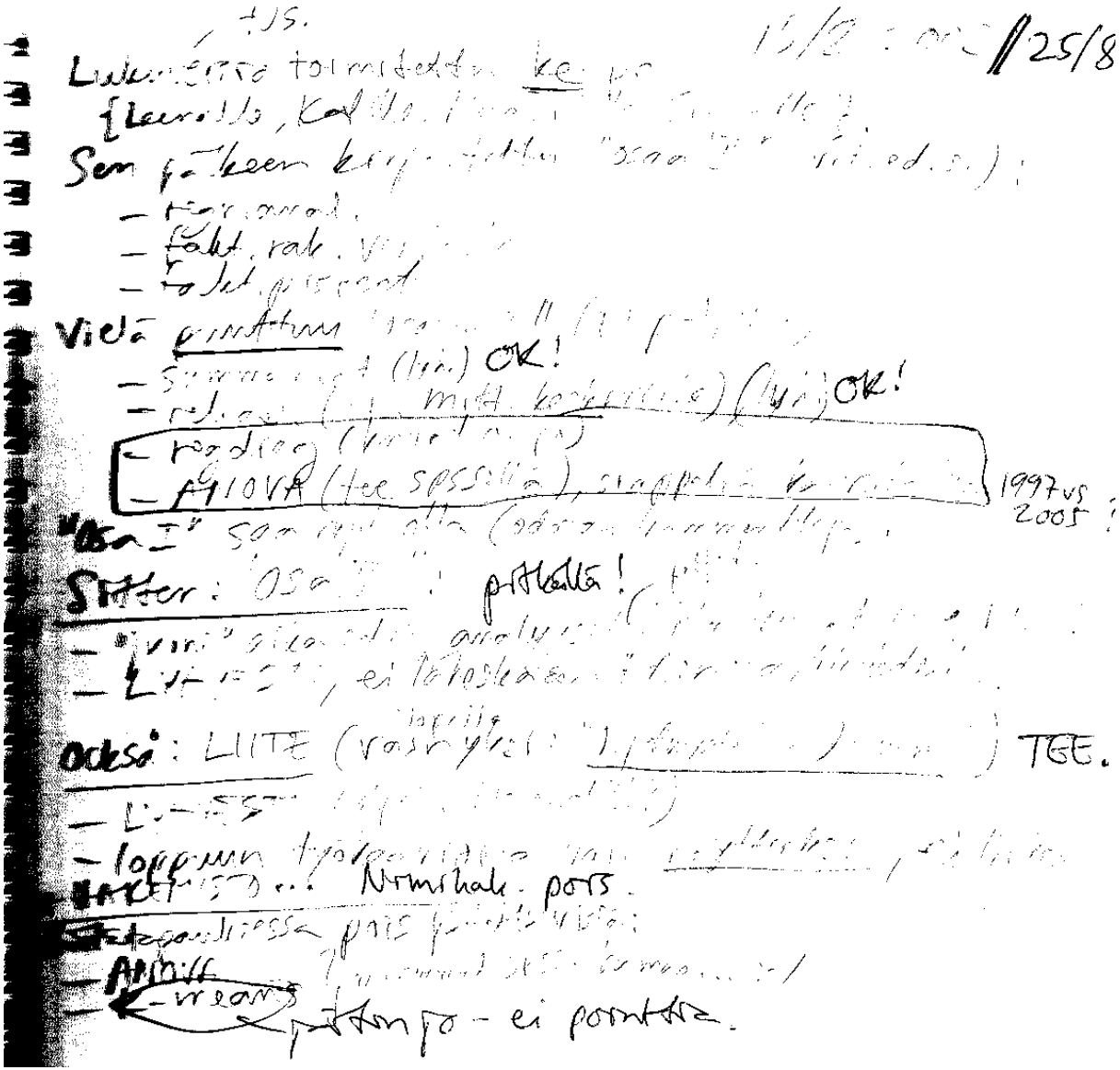

3.3.1 Dokumentointi . . . . . 64
3.4 Usean muuttujan tarkastelu . . . . . 64
3.4.1 Taulukot . . . . . 64
3.4.2 Kuvat . . . . . 67
3.4.3 Tunnusluvut . . . . . 68
3.5 Aineiston muokkaaminen . . . . . 68
**4 Aineiston tiivistäminen . . . . . 69**
4.1 Tilastollinen mallintaminen . . . . . 69
4.2 Mittausmalli . . . . . 71
4.3 Faktorianalyysi . . . . . 73
4.3.1 Oletukset . . . . . 74
4.3.2 Faktorointi . . . . . 77
4.3.3 Mittausmallin rakennevaliditeetti . . . . . 80
4.4 Faktorirakenteiden vertailu . . . . . 85
4.5 Mitta-asteikko . . . . . 85
4.5.1 Faktoripisteet . . . . . 88
4.5.2 Summamuuttujat . . . . . 88
4.5.3 Mitta-asteikon reliabiliteetti . . . . . 88
**5 Selitysmallien laatiminen . . . . . 89**
5.1 Mittauskehikko . . . . . 89
5.1.1 Tulosasteikko ja vertailuperuste . . . . . 89
5.1.2 Mallit mittauskehikossa . . . . . 92
5.2 Regressioanalyysi . . . . . 93
5.2.1 Havainnot ja muuttujat . . . . . 93
5.2.2 Selittäjien valinta . . . . . 93
5.2.3 Tulosasteikon ennustevaliditeetti . . . . . 94
5.2.4 Regressiodiagnostiikka . . . . . 94
5.3 Varianssianalyysi . . . . . 94
**6 Vastausten ryhmittely . . . . . 95**
6.1 Hierarkinen ryhmittely . . . . . 95
6.2 Moniulotteinen skaalaus . . . . . 96
6.3 Keskiarvoryhmittely . . . . . 96
6.4 Mediaaniryhmittely . . . . . 96
**7 Ryhmien kuvaaminen . . . . . 99**
7.1 Erotteluanalyysi . . . . . 99
7.2 Korrespondenssianalyysi . . . . . 99
**Liitteet**

# 7.8.2008: ”Kirja etenee, virka lähenee...”



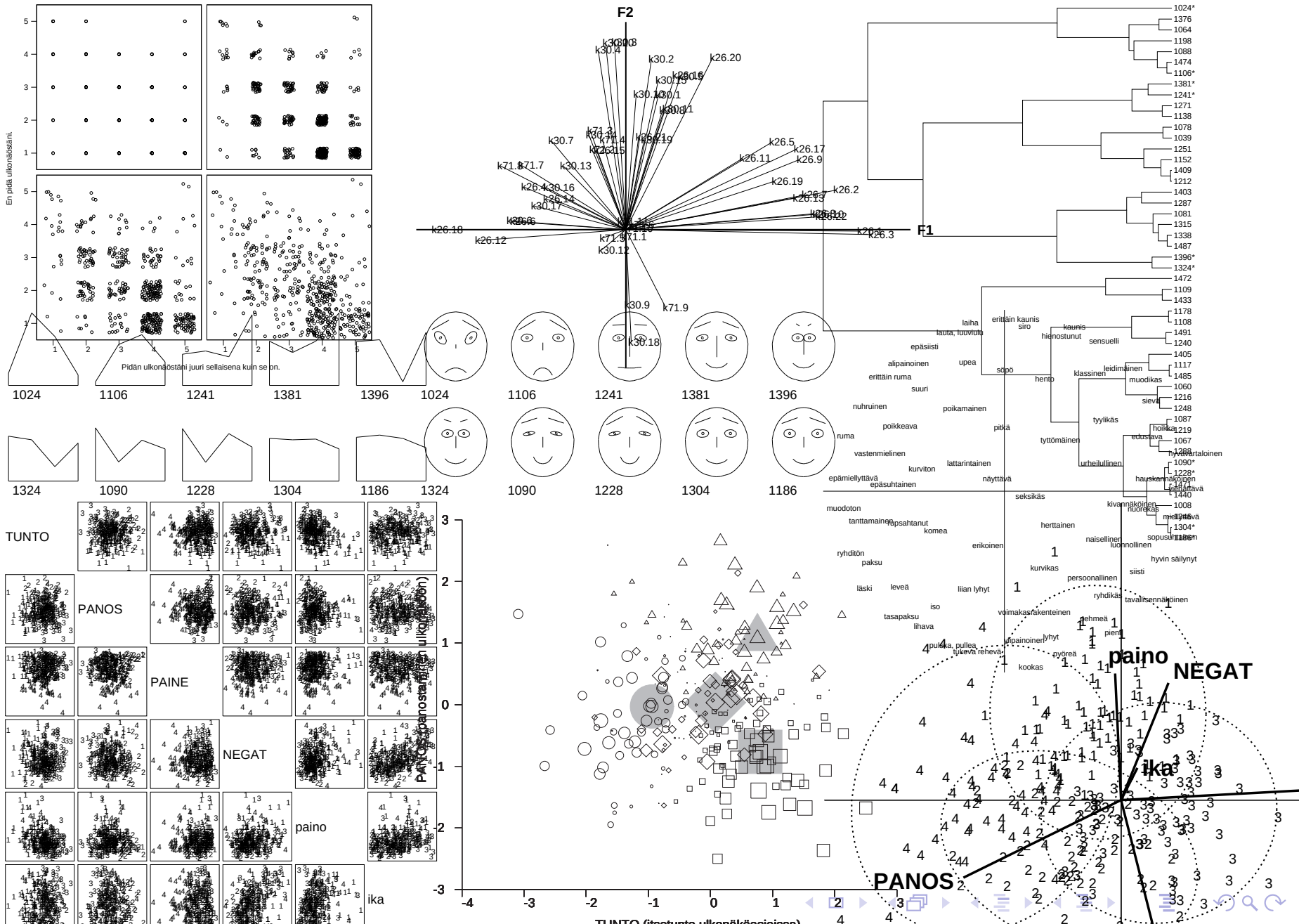

- *”Rakenne muovautunut jälleen paremmaksi...”*

# 15.8.2008: ”Vielä puuttuu...”




- *”Liite: lopulta vain yksi”*

# 25.8.2008: Kuvakooste Tammen kansigraafikolle



# 4.9.2008: Kansigraafikon luomus nähtäväksi

# 9.9.2008: Käsikirjoitus (211 s.) jakoon lähipiirille



Kimmo Vehkalahti

**Kyselytutkimuksen mittarit ja menetelmät**

- Deadlineen 2 viikkoa (lopulta tarvittiin vielä lisäviikko).

# 18.9.2008: 15 min esitys Tammessa opettajille

## Kyselytutkimuksen mittarit ja menetelmät

– etukäteistietoa uudesta oppikirjasta

**Kimmo Vehkalahti**
Soveltavan tilastotieteen dosentti, VTT, yliopistonlehtori
Matematiikan ja tilastotieteen laitos, Helsingin yliopisto

18.9.2008

- Ennakkotilauksia oli kuulemma alkanut jo tulla.

# 26.9.2008: Kustannustoimittajan korjausten vienti taittoon





perusjoukkoa tai näyteaineistoon ei saatu tavoiteltuja henkilöitä. Oli tiedonkeruun tapa ja teorian osuus mikä hyvänsä, niin osa epävarmuuksista johtuu mittauksesta: joidenkin käsitteiden operationalisointi on epäonnistunut, osaa kysymyksistä ei ole ymmärretty, tai ne on ymmärretty eri tavalla kuin tutkija on tarkoittanut, tai eri vastaajat ovat ymmärtäneet ne eri tavoin. Kenties kysymyksiä on ollut liikaa, ja vastausväsymyskin on alkanut vaikuttaa. Lisäksi vastauksia voi hämärtää *sosiaalinen suotavuus*, joka tarkoittaa sitä, että vastaaja pyrkii vastauksillaan välittämään itsestään myönteisen vaikutelman. Epävarmuuden lähteitä on siis lukuisia.

Joitakin epävarmuuksia voidaan hallita tekemällä sopivia oletuksia, mutta oletukset voivat myös aiheuttaa lisää epävarmuuksia, etenkin jos niitä on liikaa tai ne ovat epärealistisia. Oletuksia on helppo tehdä, mutta pitäisi myös tutkia missä määrin ne pitävät paikkansa. Mallit ovat yleiskäyttöisempiä ja tulokset luskottavampia mitä vähemmillä oletuksilla toimitaan.

### Mallit ja menetelmät

Tilastollisista malleista tehtyjen johtopäätösten yleistäminen on mahdollista, mikäli aineisto muodostaa edustavan otoksen jostakin perusjoukosta. Kiintoisia analyyseja voidaan tehdä myös aineistolähtöisesti, jolloin ei tarvita teoriasta johdettuja malleja vaan menetelmät riittävät. Analyysien perusteella tehdyt päätelmät rajoittuvat tällöin enemmän aineiston tasolle. MUUTOS

Tässä kirjassa käsitellään eräitä tyypillisimpiä malleja ja menetelmiä. Mallipohjaisempia menetelmiä edustavat etenkin perusmenetelmät *faktorianalyysi* ja *regressioanalyysi*, kun taas puhtaasti aineiston kuvaamisen menetelmiä ovat esimerkiksi *hierarkinen ryhmittely* ja *korrespondenssianalyysi*. Yhteistä näille menetelmille on, että analysoitavana on yhtaikaa useita muuttujia. Siksi niitä kutsutaan monimuuttujamenetelmiksi.

Tässä kirjassa käsiteltävillä menetelmillä selviää käytännössä pitkälle. Turhan monimutkaisia menetelmiä on jopa syytä välttää. Vaikka itse osaisi niitä soveltaa, voivat tulokset jäädä muilta ymmärtämättä. MUUTOS "Hienompia" malleja tai menetelmiä ei ole syytä edes kokeilla, ellei hallitse perusmenetelmiä kunnolla.







POISTO Burtin matriisi muistuttaa korrelaatiomatriisia (ks. tuloste 3.11, s. 79). Lävistäjän ykkösiä vastaavat frekvenssijakaumat ja korrelaatioita ristiintaulukot. LISÄYS Toisaalta se muistuttaa hajontakuvamatriisia (ks. kuva 7.1, s. 172), jossa hajontakuvien tilalla ovat ristiintaulukot.

MUUTOS Burtin matriisin haittapuolena voidaan pitää sen laajuutta: tulosteen 7.6 matriisissa on kaikkiaan 1 444 lukua, mikä on yli 20-kertainen määrä verrattuna kahdeksan muuttujan korrelaatiomatriisin 64 lukuun. POISTO Edut menevät kuitenkin haittojen edelle, koska Burtin matriisin avulla voidaan analysoida muitakin kuin lineaarisia riippuvuuksia. LISÄYS Ryhmäkoodi ja työllisyystilanne ovat luokittelutasoisia muuttujia, joiden analysointiin hajontakuvat ja korrelaatiot soveltuvat huonosti, jos lainkaan.

### 7.3.4 Usean muuttujan kuva

Burtin matriisiin sisältyviä yhteyksiä voidaan kuvata *usean muuttujan korrespondenssianalyysilla*. Tämä korrespondenssianalyysin yleistys soveltuu etenkin laajojen, luokitteluihin perustuvien kyselyaineistojen kuvaamiseen. Menetelmää on kehitetty ja sovellettu erityisesti Ranskassa, jossa sitä toisinaan kutsutaan "luokiteltujen muuttujien faktorianalyysiksi".

Yhteiskuntatieteissä analyysia on tehnyt tunnetuksi ranskalainen sosiologi Pierre Bourdieu, esimerkiksi teoksessaan *Distinction* (Bourdieu, 1984). Menetelmän perusteisiin johdattaa muun muassa Greenacre (2007). Pidemmälle meneviin mahdollisuuksiin ja muunnelmiin perehdyttää laaja teos *Multiple Correspondence Analysis and Related Methods* (Greenacre & Blasius, 2006).

POISTO Usean muuttujan korrespondenssianalyysissa numeeriset tulokset kiinnostavat yleensä vielä vähemmän kuin tavallisessa korrespondenssianalyysissa. Tavoitteena on jälleen visualisoida olennaiset tiedot kaksiulotteisena kuvana. Akseleiden lukuarvot on tapana jättää pois, sillä useiden luokiteltujen muuttujien samanaikaisessa visualisoinnissa kuvaan tulee helposti "ruuhkaa". Kiinnostavia ovat vain luokkien ja luokittelijoiden väliset suhteet ja niiden sijoittuminen "korrespondenssiavaruuteen".

# 30.9.-28.10.2008: homma pakettiin ja jakeluun

Kirjaprojektin viimeiset vaiheet lyhyesti:

- 28.9.2008: **U2 3D** -leffa *(syksyn ensimmäisiä vapaapäiviä!)* `http://en.wikipedia.org/wiki/U2_3D`
- 29.9.2008: viimeiset viimeistelyt
- 30.9.2008: **pdf (223 s., 1.8 MB) Vammalan kirjapainoon!**
- 2.10.2008: markkinointipalaveri Tammessa
- 3.10.2008: kilistelybileet @ Sling In
- 6.10.2008: ensimmäisen painovedoksen tarkistus Tammessa
  - muuten OK mutta SPSS-taulukot pielessä! (palautus painoon)
- 9.10.2008: vedos 2 (kust.toim.) – yhä yksi taulukko pielessä!
- 14.10.2008: ennakkomyyntiesitteet; vedos 3 (koko kirja!) lähettipalveluna Kumpulaan – tarkistus ja takaisin: **OK!**
- 17.10.2008: sisällysluettelo Tammen nettisivuille
- 20.10.2008: lopullisen kustannussopimuksen allekirjoitus
- 21.10.2008: lisää materiaalia Tammen nettisivuille
- 28.10.2008: **kirja ilmestyy!**

# Näkymiä lopullisen version toimituskentistä (1/32)

Pois kommentoituna melkein toinen mokoma, ns. turhia töitä:

```
1239 * ruokak          'Ruokakunnan koko'
1240 * raskaus         'Raskaana'
1241 * lapset          'Lapsia'
1242 * lluku           'Lasten lukumäärä'
1243 * <loput 483 muuttujaa vastaavasti> .
1244 ]
1245 *Vastaavasti voitaisiin lisätä tiedot muuttujien mittaustasosta. Vaihtoehtoja on
1246 *vain kolme: luokittelu-, järjestys- ja suhdeasteikko. SPSS ei tosin
1247 *osaa hyödyntää näitä tietoja juuri millään lailla.
1248 *
1249 - / MISSING-asiat myöhemmin...
1250 *
1251 - / Näin se homma etenee... olipa prkl hankalaa - koko juhannusaatto (2007)
1252 -_/ meni tämän kanssa painiesssa!!!!!!
Shadow:                                    JA lopulta kaikki tuo turhaa... 15.9.2008
1254 *VALUE LABELS \ldots
1255 *
1256 *
1257 *Survo
1258 *
1259 - / Survolla ei ole annetussa tehtävässä mitään vaikeuksia.
1260 *Havaintotiedosto perustetaan FILE SAVE -komennolla käyttäen
1261 *täsmennystä DELIMITER=TAB.
```

# Näkymiä lopullisen version toimituskentistä (2/32)

Joitain osia on rajattu kokonaan tulostusalueen ulkopuolelle:

```
1444 *Jokaiselle työlle on syytä perustaa kansio,
1445 *johon talletetaan kaikki siihen liittyvä materiaali: aineistot,
1446 *dokumentit, kirjeenvaihto, analyysit, kuvat, taulukot, raportit,
1447 *muistiot, artikkelit jne. Näin työ pysyy paremmin koossa, ei
1448 *sotkeennu muihin töihin eikä jää lojumaan työpöydälle.
1449 *
1450 *Äskeinen kappale voisi olla arkistointikansion käyttöohjeesta
1451 *ajalta ennen tietokoneita. Ohje pätee kuitenkin yhtä hyvin
1452 *tietokoneen kansioihin ja tiedostoihin. Erona on vain, että
1453 *materiaali talletetaan sähköisessä muodossa. Periaate on täysin sama.
1454 *
1455 *Tuskinpa kukaan pitäisi kovin järkevänä työskentelytapaa, jossa kaikkien
1456 *projektien dokumentit sullottaisiin samaan kansioon. Kuitenkin
1457 *Windows ohjaa tämäntyyppiseen toimintaan ehdottamalla tiedostojen
1458 *talletuspaikaksi ''My Documents'' -kansiota. Jos siihen suostuu, voi pian ainoa
1459 *keino löytää töitään ollakin hakukone, jota Windows todella
1460 *tarjoaakin töiden ''hallintaan''.
1461 *
1462 - / Ajan myötä töiden hallintaa on Windows-ympäristössäkin koetettu kehittää, ja
1463 - / niinpä mainitun ''My Documents'' -kansion alle on ilmestynyt valmiiksi sellaisia
1464 - / kansioita kuin ''My Music'' ja ''My Pictures''.
1465 - / Oleellisesti nämä eivät auta paljoakaan, kuten ei sekään, että
1466 - / uusimmassa Windowsissa on taas tarjolla kokonaan uudet kansiot ja työtavat.
1467 -
```

# Näkymiä lopullisen version toimituskentistä (3/32)

Lisää vaivalla kirjoitettuja, sittemmin poistettuja kohtia:

```
1567 *Viitelista
1568 *
1569 *Viitelista on Survolla hallittava luettelo täsmälinkkejä missä tahansa
1570 *kansioissa sijaitsevien tiedostojen, yleensä Survon toimituskenttien
1571 *eri kohtiin. Viitelistan kehitti Seppo Mustonen vuonna 1998. Sitä
1572 *edelsi Survo ja minä -kirjassakin \citep[29--32]{Mustonen1996} kuvattu
1573 *staattisempi työvalikko. Viitelista jatkoi ideaa ''yhden kosketuksen
1574 *pelistä'', ja sen tarkoitus oli helpottaa edelleen töiden hallintaa
1575 *kokoamalla yhteen paikkaan viittaukset kaikkiin työn alla oleviin kohteisiin.
1576 *
1577 *Edellä mainittu muistiotiedosto MEMO on juuri tällainen toimituskenttä.
1578 *Se toimii työn tarttumakohtana, josta pääsen nopeasti jatkamaan
1579 *aiemmin keskeytynyttä työskentelyä käyttämättä aikaani turhaan
1580 *haeskeluun ja selailuun. Samalla kun palaan työn pariin, tulee
1581 *MEMO-kenttään myös kirjattua mitä olen tekemässä ja milloin.
1582 *
1583 *Tämänhetkisen
1584 *viitelistani ''kuumimmat'' linkit osoittavat tämän kirjan lisäksi meneillään tai
1585 *suunnitteilla oleviin kursseihin ja koulutuspäiviin, valmisteilla oleviin
1586 *konferenssiesitelmiin ja tutkimusartikkeleihin, opinnäytetöiden ohjauksiin
1587 *liittyviin muistiinpanoihin, laitoksen työryhmien keskusteluihin,
1588 *kirja-arviointeihin ja erilaisiin konsultaatioihin.
1589 *Loput linkeistä ovat pääosin passiivisia: ne vievät pidettyihin
1590 *kursseihin ja seminaareihin, julkaistujen artikkelien
1591 *rakennustelineille ja muihin päättyneisiin työprosesseihin, joissa
1592 *saattaa olla tarpeen eri syistä vierailla vielä jälkeenpäin.
1593 *
```

# Näkymiä lopullisen version toimituskentistä (4/32)

Ja lisää *(olisikohan syytä kirjoittaa toinen kirja?)*:

```
1666 *
1667 - / oli: Varmistaminen
1668 *Töiden varmistaminen
1669 *
1670 *Tietokone voi hajota milloin tahansa. Tiedostot voivat mennä rikki
1671 *tai tulla vahingossa päällekirjoitetuksi. Kannettava tietokone
1672 *voi unohtua johonkin tai tulla varastetuksi.
1673 *Kaikissa näissä tapauksissa tietoa häviää, joko osittain tai kokonaan.
1674 *Uusia koneita saa kaupasta ja ohjelmia verkosta, mutta
1675 *ilman varmuuskopioita tärkeät tiedot ovat mennyttä -- ja työtunnit sen mukana.
1676 *
1677 *On todella syytä miettiä, kuinka paljon olisi valmis tekemään uudelleen.
1678 *Vaivalla mietityn lauseen? Viiden minuutin työn? Koko päivän aherruksen?
1679 *Viikon vaivannäön? Kaikki aineiston koodaukset ja tarkistukset?
1680 *Kuukausien tiedonkeruun? Ei kai?
1681 *
1682 *Järkevä käyttäjä huolehtii varmuuskopioista, eikä anna aihetta
1683 *iltapäivälehtien ''tietokone söi väitöskirjan'' -tyyppisille
1684 *otsikoille. Levytila ei enää maksa paljoakaan, ja varsinkin
1685 *irtokovalevyt ja muistipalikat ovat todella käteviä. Myös DVD- ja
1686 *CD-levyt soveltuvat varmuuskopiointiin.
1687 *
1688 *Siinä vaiheessa kun kone häviää tai tiedosto hajoaa,
1689 *varmuuskopioita ei voi enää tehdä. Niitä pitää siis tehdä jonkinlaisella
1690 *rutiinilla jatkuvasti. Omassa viitelistassani (katso esimerkki ex:viitelista)
1691 *onkin ihan alussa kohta BACKUP, johon suuntaan melko säännöllisesti.
1692 *Siellä odottavat valmiit yksinkertaiset XCOPY-komennot, joilla kopioin
1693 *haluamani tiedot talteen.
1694 *
```

# Näkymiä lopullisen version toimituskentistä (5/32)

Täällähän on tarinaa (ja narinaa) vaikka kuinka:

```
1822 *
1823 *Virheistä toipuminen
1824 *
1825 *Virheet ovat luonnollinen osa tutkimustyyppistä työskentelyä.
1826 *Jokainen löytää virheitä niin aineistostaan kuin omista tekemisistään,
1827 *kyse onkin enemmän siitä miten virheisiin suhtautuu.
1828 *Ainakaan niiden tekemistä ei kannata yrittää välttää, sillä se
1829 *johtaa vain helposti liian varovaiseen työskentelyyn,
1830 *joka seurailee totuttuja uria. Virheistä voi sitäpaitsi oppia paljonkin.
1831 *Tutkimustyö ei ole mitään
1832 *rutiinien toistamista, vaan siinä pitää lähteä uusille
1833 *urille, kyseenalaistaa aiempia ratkaisuja ja poiketa rohkeasti sivupoluille.
1834 - / Sanonta aina roiskuu kun rapataan pätee myös tutkimustyöhön.
1835 *
1836 *Jotta sivupoluille uskaltaisi lähteä, sieltä
1837 *löytäisi helposti takaisin ja löytämistään virheistä saisi vielä
1838 *kaiken hyödyn irti, on kyettävä kohtaamaan virhetilanteita
1839 *milloin tahansa. Pitää vain havaita ja paikallistaa virhe, tehdä
1840 *tarvittavat korjaukset ja lopuksi toipua virhetilanteesta.
1841 *Toipuminen tarkoittaa, että ymmärtää mihin kaikkeen havaittu virhe
1842 *on siihen mennessä vaikuttanut, mitkä työvaiheet on tehtävä
1843 *uudelleen ja mistä voi taas jatkaa. Mitä tehokkaammiksi saa kehitettyä
1844 *työskentelytapansa, sitä nopeammin toipuminen käy ja sitä vähemmän
1845 *tarvitsee pelätä virheitä.
1846 *
```

# Näkymiä lopullisen version toimituskentistä (6/32)

Jossain vaiheessa kirjaan oli tulossa kaksi SPSS-kuvaakin:

```
2095 *
2096 *Kuva piirretään kahdella tavalla. Rivin~16 GRAPH-komennon
2097 *tuottaa valinta Graphs -- Legacy Dialogs -- Bar ja rivin~19
2098 *IGRAPH-komennon valinta Graphs -- Interactive -- Bar.
2099 *Vaihtoehdot edustavat SPSS:n eri kehitysvaiheita. Vanhempi
2100 *GRAPH osaa hyödyntää luokkien nimiä, mutta sitä ei voi ohjata
2101 *komennoilla paljoa pidemmälle. Uudempi IGRAPH tarjoaa
2102 *enemmän mahdollisuuksia, muttei oletuksena hyödynnä
2103 *nimiä, vaan sijoittaa pylväiden alle luvut 1--5.
2104 - / (kahdella desimaalilla).
2105 *
2106 *Molemmissa tapauksissa kuvan viimeistely edellyttäisi käsin tapahtuvaa
2107 *muokkausta. Koska sellaisten työvaiheiden dokumentointimahdollisuudet ja
2108 *toistettavuus ovat olemattomia, ei tässä kirjassa ole yhtään SPSS:llä
2109 *piirrettyä kuvaa.
2110 *
2111 *
2112 - / EN HALUA näitä tänne:
2113 -  /  F
2114 -  /  CSPSS GRAPH ja IGRAPH
2115 -  /  =fig:SPSS-kuvat
2116 -  /  *\includegraphics[width=0.55\textwidth]{KMM-GRAPH}      % .45
2117 -  /  *\includegraphics[width=0.45\textwidth]{KMM-IGRAPH}     % .37
2118 -  /  f
2119 -
```

# Näkymiä lopullisen version toimituskentistä (7/32)

Liitteet menivät lopulta niin uusiksi, että jätin ne välillä pois tulosteista kokonaan (kumma kyllä; joskus vaikuttivat niin hyviltä!):

```
2535 *═══════════════════════════════════════════════════════════════════════
2536 *
2537 - / shadow k: \section*{ }   /
2538 *
2539 - / Alun tietokonejorinat siirretty pois; ota jotain jos otat! 8/8/8 (ks. END)
2540 *
2541 *Liitteiden uudelleenorganisointi (8/8/8) kesken:
2542 <
2543 Lvain yksi liite (aiemman kahden sijasta)
2544 Ltiiviisti ohjelmistoista (Survo ja SPSS)
2545 Lyleistä dokumentoinnista, esimerkkeinä työkaavioita
2546 Lalustavat otsikot alla (tekstit remontissa)
2547 >
2548 *═══════════════════════════════════════════════════════════════════════
2549 *
```

# Näkymiä lopullisen version toimituskentistä (8/32)

Hakemistossa olisi selvästi kehittämisen varaa; sen teko osoittautui aika haastavaksi:

```
29 *Kyselytutkimus
30 *
31 *kyselytutkimus
32 *Kyselytutkimus on tärkeä tapa kerätä ja tarkastella tietoa
33 *muun muassa erilaisista yhteiskunnan ilmiöistä, ihmisten toiminnasta,
34 *mielipiteistä, asenteista ja arvoista. Tämäntyyppiset kiinnostuksen
35 *kohteet ovat sekä moniulotteisia että monimutkaisia.
36 *
37 *Kyselytutkimuksessa tutkija esittää vastaajalle kysymyksiä
38 *kyselylomake
39 *kyselylomakkeen välityksellä. Kyselylomake on mittausväline,
40 *jonka sovellusalue
41 - / KILL BILL vol 2: (siirto alakohtaan!)
42 - / tosiasioiden, käsitysten, mielipiteiden, asenteiden ja arvojen
43 - / mittaamiseen erilaisilla mittareilla. Sen
44 *ulottuu yhteiskunta- ja
45 - [M!] / "kattaa" -> "ulottuu" 23.9.2008
46 *käyttäytymistieteellisestä tutkimuksesta
47 *mielipidetiedusteluihin, katukyselyihin, soveltuvuustesteihin ja
48 *palautemittauksiin.
49 *
50 *haastattelututkimus
51 *Haastattelututkimuksessa
52 *tutkija tai haastattelija esittää kysymyksiä suoraan vastaajalle, esimerkiksi
53 *puhelimitse tai kasvotusten.
```

# Näkymiä lopullisen version toimituskentistä (9/32)

Varhaisia näkymiä johdantoon, jonka kirjoitin ainakin kolme kertaa...

```
164 - /        |           |           |           |           |
165 - /        |           |           |           |           |
166 - /       ---         ---         ---         ---         ---
167 - /      / | \       / | \       / | \       / | \       / | \
168 - /     o  o  o     o  o  o     o  o  o     o  o  o     o  o  o
169 - /
170 - / *Mittarit
171 - / *
172 - / *Tilastollisessa tutkimuksessa epävarmuudet aiheutuvat tiedonkeruusta ja
173 - / *mittaamisesta. Mittaaminen on ollut perinteisesti käyttäytymistieteiden,
174 - / *etenkin psykologian alaa. Pohjan sille loivat Charles Spearmanin
175 - / *Spearman, Charles
176 - / *varhaiset tutkimukset ihmisen älykkyydestä. Spearman määritteli jo 1900-luvun
177 - / *alussa psykologisen mittaamisen keskeiset käsitteet sekä kehitti tällaisten
178 - / *mittausten analysointiin soveltuvan menetelmän, jotka kutsutaan
179 - / *faktorianalyysiksi.
180 - / *
181 - / - / tässä harhautuu: pitäisi puhua MITTAREISTA eikä menetelmistä!
182 - / *
183 - / *Tilastotieteilijät suhtautuivat faktorianalyysiin epäillen vuosikymmenien ajan.
184 - / *Yleensäkin tilastotieteessä on keskitytty tiedonkeruun, eritoten otannan
185 - / *haasteisiin ja jopa oletettu, että mittaukset on tehty virheettömästi. Oletus
186 - / *on epärealistinen, sillä mittauksia on lähes mahdotonta tehdä virheettömästi
187 - / *edes tarkoilla teknisillä mittalaitteilla, puhumattakaan kyselylomakkeiden mittar
188 - / *
189 - / *Tilastotieteen ja psykologian vastakkainasettelu saattaa jo nykyään
190 - / *ihmetyttää, tuovathan sekä tiedonkeruu että mittaaminen vääjäämättä osansa
191 - / *kaikkien tutkimusten virhemarginaaleihin. Tavallaan ''kädenvääntö'' alkoi jo
192 - / *psykologi Spearmanin ja tilastotieteilijä Karl Pearsonin
193 - / *Pearson, Karl
194 - / *kiistoista sata vuotta sitten. Heidän värikkäät keskustelunsa jatkuivat
195 - / *tieteellisten lehtien palstoilla vuosien ajan, ja yhteisymmärrystä tuntui
196 - / *olevan vaikea saavuttaa.
```

[Spearman ja Pearson katosivat tästä myöhemmin]

# Näkymiä lopullisen version toimituskentistä (10/32)

...muokaten jokaista vähintään kymmeneen otteeseen:

```
297 -
298 - / Nämäkin alkusanoihin: (16.6.2008) (oli: johdantoa)
299 - / sittenkin johdantoon!?!
300 - / otsikko sittenkin: "Tämä kirja"?  vrt. 1.2! EI, 1.2 hävisi jo... 8.7.2008
301 - /    tänne mukaan "Muut kirjat" -yleisosa?
302 *Kirjan rakenne ja sisältö
303 *
304 - / onpa hauska tiivistelmä! (30.7.2008) Kestääkö? ;)
305 - / ... kesti 5 min huomata että tämä EI kuulu kohdan 1.1.1
306 - / vaan kohdan 1.2 alkuun! :-) (jossa se korvaa 2 kpl:tta!)
307 -
308 *Rakenteellisesti kirja seuraa kyselytutkimuksille yhteisiä vaiheita.
309 - [M!]
310 *Sisällöllisesti se yhdistää kaksi kokonaisuutta:
311 <
312 Lmittauksen tilastolliset näkökohdat kyselytutkimuksessa
313 Lkyselyaineiston tilastollinen mallintaminen ja kuvailu.
314 >
315 - [L!] / siirsin alempaa tähän 19.9.2008:
316 *Olennaista käsiteltävien asioiden ymmärtämisessä ovat erilaiset tulosteet ja kuvat,
317 *joita kirjassa on runsaasti.
318 - / Niihin viitataan tekstistä lukukohtaisilla
319 - / numeroilla kuten esimerkiksi kuva~fig:FACES.
```

# Näkymiä lopullisen version toimituskentistä (11/32)

Vihonviimeinen muutos 29.9.2008 klo 23:29 (aamulla painoon):

```
378 *
379 *\newpage
380 -
381 *Liitteessä~app:SurvoSPSS kuvataan lyhyesti
382 - [M!]
383 *Survo
384 *SPSS
385 *kirjassa käytettyjä Survo- ja SPSS-ohjelmistoja sekä tarkastellaan
386 *kuvien ja tulosteiden työkaavioita esimerkkeinä dokumentoivasta työskentelytavasta.
387 *Työkaavioihin viitataan luvuista~ch:base--ch:G2
388 *marginaaliin sijoitetulla sivunumerolla.
389 *Sekä tekstissä että tulosteissa on yhdenmukaisuuden vuoksi
390 - [BOX(wrk:FACES)] / piruuttani laitan tuon (29.9.2008 23:29) <- viimeinen muutos! :)
391 *käytetty desimaali\-erottimena pilkun sijasta pistettä.
392 *
393 *Kirjalla on myös kotisivu (ks. www.tammi.fi).
394 *
395 *Ulkonäkötutkimus
396 *
397 *ulkonäkötutkimus
398 *Sekä rakenteellisesti että sisällöllisesti tärkeä kokonaisuus muodostuu
399 - [M!]
```

# Näkymiä lopullisen version toimituskentistä (12/32)

”Sopua itseni kanssa” oli joskus vaikea saavuttaa:

```
223 - / valmiina havaitsemaan ja tunnistamaan myös uusia ulottuvuuksia.
224 *
225 - / mittauskehikko siirretty 29.5.2008 uuteen mallintamislukuun!
226 - / MUTTA MUTTA: TÄNNE PELKKÄ MALLI? SEHÄN SE KOTITEHTÄVÄ ON!
227 - / sitten viittauksia eteenpäin, jossa käsitellään tarkemmin kehikkoa
228 -
229 - / siirretty ylöspäin:
230 - / haa: 9.8.2007 keksin vihdoin otsikon, vrt. myös luvun 4 samanalkuinen!!
231 - / \section{Kyselylomake mittaamisen välineenä}
232 - / vähän kömpelö kuitenkin... (tuntia myöhemmin jo viilaamassa...)
233 - / olisko parempi: "Mittaaminen kyselylomakkeella"
234 - / ja vast.        "Tiedonkeruu kyselylomakkeella"    JEP!
235 - / 29.5.2008 on taas aika tehdä muutoksia...
236 - / Mittaaminen kyselylomakkeella\label{sec:meas-questionnaire}
237 *
238 - / lisätty otsikkoon jälk.sana 29.7.2008 (vrt. tiedonkeruu!) JA LUVUN NIMI (muuttui)!
239 *Kyselylomake mittausvälineenä
240 =sec:meas-questionnaire
241 *
242 *Kyselytutkimuksessa mittaus tapahtuu kyselylomakkeella. Kun
243 *kyselylomake
```

# Näkymiä lopullisen version toimituskentistä (13/32)

Myöskään sopua ennakkolukijoiden kanssa ei aina syntynyt:

```
537 *Välttämättömiä avoimet osiot ovat tilanteissa, joissa vaihtoehtoja ei haluta
538 *tai ei voida luetella. Vaihtoehtoja voi olla liikaa, tai niitä ei
539 *vain ole mahdollista etukäteen rajata riittävästi.
540 *
541 *\newpage
542 - / Jatkuva vai diskreetti?
543 - / hirvee ähellys, ehkä vihdoin selvempi! (KAIKKI huomauttivat tästä!!)
544 *Osioista jatkuviin ja diskreetteihin muuttujiin
545 - / (tämän viittausnumeroksi tulee edellinen ]-otsikko - OK.) [pari viittausta 3:ssa]
546 =sec:jatkdiskr
547 *
548 *Esimerkissä ex:ikä? käsitelty ikä on ominaisuutena tai käsitteenä
549 *jatkuva, koska se liittyy ajan jatkuvaan kulumiseen. Jatkuvien
550 *ilmiöiden mittaus ja esittäminen tietokoneella on kuitenkin mahdollista
551 *vain rajallisella tarkkuudella. Ikää voidaan mitata paljon tarkemmin
552 *kuin vuosina, mutta kyselytutkimuksessa se on harvoin tarpeellista.
553 *
```

# Näkymiä lopullisen version toimituskentistä (14/32)

Usein oli helpompi kirjoittaa suunnitelmia siitä, mitä pitäisi kirjoittaa kuin varsinaista asiaa:

```
587 *tapauksissa tutkita jatkuvaksi. Kaikki kysymykset mittaavat kuitenkin
588 *vastaajan vointia. Selvimmin ne erottaa toisistaan mittauksen taso, johon
589 *seuraavaksi perehdytään.
590 *
591 - / S-niemessä (2007) kirjoitettua...
592 - / tukeudu esim. Alkula (130-139), Heikkilä (47-66), ...
593 - / tuo til.tieteilijän NÄKÖKULMIA ("konsultointi")
594 - / EOS, 5-portainen ("järj.ast."), ...
595 - / tänne EI kuitenkaan esimerkkiä "ulkonäkötutkim.kyselylomake" tms.,
596 - / tämä luku olkoon enemmän "teoriaa". ne asiat voi laittaa tuonne datalukuun
597 - / PALUU tänne sateisena syyspäivänä 21.9.2007! Vihdoin alkaa sujua. Illalla Tammi...
598 - / TÄSTÄ KOHDASTA YLEMMÄN TASON OTSIKKO -> luokittelu ym. sis.luetteloon!!
599 - / ja samalla tänne saadaan pieniä otsikoita kuten "En osaa sanoa" ym.
600 *\newpage
601 *Mittauksen taso
602 =sec:meas-level
603 *
604 *Osiot ovat siis mittausväline. Seuraavassa rajoitutaan suljettuihin
605 *osioihin, joiden mittaus voi edelleen
606 *olla monentasoista. Mittauksen taso
```

# Näkymiä lopullisen version toimituskentistä (15/32)

Yksi lukuisista paikoista, joissa rakentelin kirjan säätöjä: juna Tampereelle 7.9.2007:

```
*PVM / Perjantai, 7. syyskuuta 2007 klo 10.20 (vko 36, pvä 250/365)
*Treen junassa
*
*     !"$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ
*     !"$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ
*
*     [\]^_`abcdefghijklmnopqrstuvwxyz{|}~□ÇüéâäàåçêëèïîìÄÅÉæÆô
*     [\]^_ abcdefghijklmnopqrstuvwxyz{|}~□ÇüéâäàåçêëèïîìÄÅÉæÆô
*
*     öòûùÿÖÜ¢£¥₧ƒáíóúñÑªº¿⌐¬½¼¡«»
*     öòûùÿÖÜ¢£¥₧ƒáíóúñÑªº¿⌐¬½¼¡«»
*
*     ßΓπΣσµτΦΘΩδ∞φε∩≡±≥≤⌠⌡÷≈°∙·√ⁿ²
*     αßΓπΣσµτΦΘΩδ∞φε∩≡±≥≤⌠⌡÷≈°∙·√ⁿ²■
*
*
*s kmm'tyy
*
*
*
*Vastenmielinen
*
*s kmm
*\newfloat{Tuloste}{h}{los}[chapter]
*\newfloat{Työkaavio}{h}{los}[chapter]
*\newfloat{Esimerkki}{h}{los}[chapter]
*
*Tuloste      O,o   / output
*Työkaavio    S,s   / scheme
*Esimerkki    X,x   / example
*
* control O: \begin{Tuloste}    / output
* control o: \end{Tuloste}
* control S: \begin{Työkaavio}  / scheme
* control s: \end{Työkaavio}
* control X: \begin{Esimerkki}  / example
* control x: \end{Esimerkki}
* control C: \caption{
* control =: \label{
* shadow T: }                   / terminate (caption or label)
*
```

# Näkymiä lopullisen version toimituskentistä (16/32)

Kelluvien elementtien hallinta: Tammi vaati **ILMAA**, LaTeX-ohjeet neuvoivat vain tyhjän tilan *minimoinnissa*:

```
896 - / "leijuvaisten" hallinta: (Jarmo: "kelluvien elementtien") :)
897 -
898 - define [ILMAA] \vspace*{2.0em} / 26.8.2008 (toteutetaan Leenan toive tällä tavalla!)
899 - /                                              VAATII KÄSITYÖTÄ LOPPUVAIHEESSA!!!
900 - / ei sittenkään tämäkään ole hyvä, ainakaan näin. "ILMAA ennen floatia" johtaa
901 - / floatin siirtyessä tyhjään tilaan kpl:iden välissä! ILMAA floatin sisään?
902 - / koetan saada ILMAA jollain muulla konstilla, joten:
903 - define [ILMAA] [%=0]
904 -
905 - / no nyt Jarmon ohjeilla (31.8.2008):  ks. myös floatrow-ohje, kohta 2.7
906 -
907 - / define [FBa] \renewcommand*{\FBaskip}{10mm}
908 - / define [FBb] \renewcommand*{\FBbskip}{10mm}
909 - / define [ILMAA] [FBa][LF][FBb][LF]
910 - / parametroidaan: (7.9.2008)
911 - define [FBa(x)] \renewcommand*{\FBaskip}{xmm}
912 - define [FBb(x)] \renewcommand*{\FBbskip}{xmm}
913 - define [ILMAA(x)] [FBa(x)][LF][FBb(x)][LF] / (sama tila ylös ja alas)
914 -
915 - / 29.7.2008: spacing-paketilla tulosteet tiiviimmiksi (linesp pienemmäksi)!
916 - define [TIIVIS] [LF]\begin{spacing}{0.9} / 25.8.2008 (oli 0.85)
917 - define [/TIIVIS] \end{spacing}[LF]
918 - /
919 - control O: [ILMAA(7)]\begin{Tuloste}[LB]htb[RB][TIIVIS]
920 - control o: [/TIIVIS]\end{Tuloste}
921 - control S: [ILMAA(3)]\begin{Kaavio}[LB]htb[RB]
922 - control s: \end{Kaavio}
923 - control X: [ILMAA(7)]\begin{Esimerkki}[LB]htb[RB]
924 - control x: \end{Esimerkki}
925 - control F: [ILMAA(10)]\begin{figure}[LB]htb[RB]
926 - control f: \end{figure}
927 - control D: [ILMAA(7)]\begin{Taulukko}[LB]htb[RB]
928 - control d: \end{Taulukko}
929 - control Z: [ILMAA(1)]\begin{NΣkymΣ}[LB]htb[RB] / Kuva -> Näkymä 27.9.2008
930 - control z: \end{NΣkymΣ}
931 - control Y: [ILMAA(3)]\begin{Listaus}[LB]htb[RB] / Esimerkki -> Listaus 27.9.2008
932 - control y: \end{Listaus}
933 - control C: \caption{
934 - control =: \label{
935 - shadow T: }                    / terminate (caption or label)
936 - shadow =: \ref{ }
937 - shadow #: \pageref{ }
938 -
939 - control \: \noindent[S]  / AIVAN LOPPUVAIHEESSA LISÄILLÄÄN NÄITÄ...
```

# Näkymiä lopullisen version toimituskentistä (17/32)

Viimeiseksi luulemastani versiosta tulikin toiseksi viimeinen, ja jouduin merkkaamaan kaikki vielä tekemäni muutokset marginaaleihin (jottei kustannustoimittajan tarvitsisi käydä läpi koko kirjaa toistamiseen):

```
940 -
941 - shadow 0: \texttt{\underline{ }} / keskeneräiset kohdat (31.7.2008-)
942 -
943 - shadow ?: [%=0] / varmuuden vuoksi; käytetty visuaalisena korostuksena!
944 - shadow \: [%=0] / varmuuden vuoksi; käytetty visuaalisena korostuksena!
945 - shadow ^: [%=0] / varmuuden vuoksi; käytetty visuaalisena korostuksena!
946 - shadow 5: [%=0] / varmuuden vuoksi; käytetty visuaalisena korostuksena!
947 - shadow _: [%=0] / varmuuden vuoksi; käytetty visuaalisena korostuksena!
948 - shadow 2: [%=0] / varmuuden vuoksi; käytetty visuaalisena korostuksena!
949 -
950 - / 22.8.2008 kopioitu PS.DEV:istä, johon Seppo lisäsi ne ehdotuksestani :)
951 - / ■ SHOW KMM'4 "ei-PRINT"
952 - define [PRINT_OFF] [%printing_on=0]
953 - define [PRINT_ON] [%printing_on=1]
954 -
955 - / Työkaavioviittaukset:
956 - define [BOX(KAAVIO)] \vspace*{-7mm}\reversemarginpar\marginnote{\footnotesize{\fbox
957 -
958 - / muutosten merkkaus kustannustoimittajaa varten (16.9.2008):
959 - define [L!] \reversemarginpar\marginnote{\fbox{\textbf{LIS-YS}}}\normalmarginpar%[L
960 - define [P!] \reversemarginpar\marginnote{\fbox{\textbf{POISTO}}}\normalmarginpar%[L
961 - define [M!] \reversemarginpar\marginnote{\fbox{\textbf{MUUTOS}}}\normalmarginpar%[L
962 -
963 - / muutoskohtien poisto poistamatta enää koodeja (27.9.2008):
964 - define [L!] [%=0]
965 - define [P!] [%=0]
966 - define [M!] [%=0]
```

# Näkymiä lopullisen version toimituskentistä (18/32)

Välillä oli suuria vaikeuksia kelluvien elementtien sijoittelussa:

```
1112 -
1113 -%/ SEURAAVAT, ks. LaTeX Companion s.284- (kaikki koskevat float:eja)
1114 -%/ ks. myös http://www.eng.cam.ac.uk/help/tpl/textprocessing/squeeze.html (23.7.2008
1115 -%/ ks. myös fancyhdr:n helppi, s.22 (!) ----------------------------------------
1116 *%  Leena: "ILMAA lisää!!!" joten vaihdetaan näitä ihan toiseen suuntaan... :)
1117 *%  Lähdetään oletuksista uudelleen:
1118 *%\setcounter{topnumber}{4}
1119 *%\setcounter{bottomnumber}{4}
1120 *%\setcounter{totalnumber}{10}
1121 *%\renewcommand{\textfraction}{0.15}
1122 *%\renewcommand{\topfraction}{0.85}
1123 *%\renewcommand{\bottomfraction}{0.70}
1124 *%\renewcommand{\floatpagefraction}{0.66}
1125 -%/ \setlength{\floatsep}{6pt plus 1pt minus 1pt}      % puolet oletuksista
1126 -%/ \setlength{\textfloatsep}{10pt plus 1pt minus 2pt} % puolet oletuksista
1127 -%/ astetta rajummat - johan alkoi tiivistyä! (22.7.2008)
1128 -%/ löysin vielä intextsep:in (ks. Companion s.285)  --- aika hienovaraista...
1129 *%\setlength{\floatsep}{8pt plus 1pt minus 2pt}
1130 *%\setlength{\textfloatsep}{8pt plus 1pt minus 2pt}
1131 *%\setlength{\intextsep}{8pt plus 1pt minus 2pt}
1132 *% otetaas nää takaisin: (ISOT floatit ei meinaa asettua!!)
1133 *% näiden kolmen muuttamisessa ei ole pointtia:
1134 *%\setcounter{topnumber}{4}
1135 *%\setcounter{bottomnumber}{4}
1136 *%\setcounter{totalnumber}{10}
1137 *% näitä säädetty hartaasti Companionin avulla 28.8.2008:
1138 *\renewcommand{\textfraction}{0.00}       % tekstin  min-osuus sivusta, oletus 20 %
1139 *\renewcommand{\floatpagefraction}{0.50}  % floatien min-osuus sivusta, oletus 50 %
1140 *\renewcommand{\topfraction}{1.00}        % max float-osuus at the top, oletus 70 %
1141 *\renewcommand{\bottomfraction}{0.50}     % max float-osuus at the bot, oletus 30 %
1142 *% nyt lisätään ILMAA: :) suuria vaikeuksia saada kuvat tekstin sekaan! :-/
1143 *\setlength{\floatsep}{10pt plus 1pt minus 2pt}
1144 *\setlength{\textfloatsep}{10pt plus 1pt minus 2pt}
1145 *\setlength{\intextsep}{10pt plus 1pt minus 2pt}
1146 *% SEURAAVAKSI: ks. Companion, s. 286-287-... sijoittelun ohjaamisesta!
1147 -
```

# Näkymiä lopullisen version toimituskentistä (19/32)

Riviväli ja sivujen taitto – alaosiin saa jäädä tyhjää, tekstikappaleiden väliin ei:

```
-%/ kokeeksi vähän tiheämpi riviväli: (23.7.2008) - sis.luett.erikseen, ks. alla!
*% setspace-paketti on tähän parempi; käytetään sitä!
*% http://web.mit.edu/answers/latex/formatting/latex_spacing.html
*\usepackage{setspace}
-%/ \setstretch{0.95}
-%/ Leenan kommenttien (31.7) perusteella vedetään takaisin!
*%  Leenan kommenttien (20.8) perusteella KASVATETAAN välistystä! ILMAA!
*% 1.2 oli ihan liikaa, 1.1 tuntuisi olevan ~ OK, vähennän vielä vähän...
*\setstretch{1.05}
-
*% Leena: "sivut mieluummin vajaiksi alhaalta kuin täytetyiksi sieltä täältä!!"
*% http://magic.aladdin.cs.cmu.edu/2008/05/12/using-raggedbottom-to-identify-where-to-
*\raggedbottom
*% matemaattiset merkit Times-tyyliin (\chi oli LIIAN ISO oletuksena!)
*% eipä juuri auta; koeta pienentää \chi niissä harvoissa kohdissa...
*% \usepackage{mathptmx}
*
*% 4.10.2008 (Tammen sivuille menevää pdf:ää varten); jotta web-osoitteet OIKEIN!
*
*\usepackage{hyperref}
*\hypersetup{
*    bookmarks=false,
*    pdftitle={Kyselytutkimuksen mittarit ja menetelmät},
*    pdfauthor={Kimmo Vehkalahti},
*    pdfsubject={Kyselytutkimuksen mittarit ja menetelmät},
*    pdfcreator={Kimmo Vehkalahti},
*    pdfkeywords={kyselytutkimus, mittarit, menetelmät},
*    colorlinks=true,
*    linkcolor=black,
*    citecolor=black,
*    filecolor=black,
*    urlcolor=black
*}
*
```

# Näkymiä lopullisen version toimituskentistä (20/32)

Lähteiden tyylitiedosto rakennettu jo 24.5.2007 (onneksi):

```
1238 *- _ Lähteiden tyylitiedosto (suomeksi.bst)
1239 *
1240 *Alkujaan (5/2007):
1241 *show c:\progra~1\miktex~1.6\tex\latex\custom-bib\makebst.tex
1242 *>latex c:\progra~1\miktex~1.6\tex\latex\custom-bib\makebst.tex
1243 *Vastailin kysymyksiin, aluksi tarkasti, sitten vain enteriä pohjassa
1244 *pitäen (niitä on satoja!)
1245 *Tuloksena syntyi tällainen (verbose-optioilla), tätä voi editoidakin:
1246 *
1247 * - _ dbj:n kääntäminen
1248 *
1249 *SAVEW CUR+1,E-1,suomeksi.dbj / dbj = docstrip batch job
1250 *%% Driver file to produce suomeksi.bst from merlin.mbs
1251 *%% Generated with makebst, version 4.1 (2003/09/08)
1252 *%% Produced on 2007/05/24 at 9:49
1253 *%% Edited a few times since that...
1254 *\input docstrip
1255 *
1256 *\preamble
1257 *----------------------------------------
1258 *** Kimmon kirjaa (2007-8) varten ***
1259 *
1260 *\endpreamble
1261 *
1262 *\postamble
1263 *End of customized bst file
1264 *\endpostamble
1265 *
1266 *\keepsilent
1267 *
1268 *\askforoverwritefalse
1269 *\def\MBopts{\from{merlin.mbs}{%
1270 *%EXTERNAL FILES:
1271 *%Name of language file: \cfile=suomeksi.mbs.
1272 *%No included files.
```

# Näkymiä lopullisen version toimituskentistä (21/32)

Joissain kohdin on näkyvissä enemmän kommentteja kuin tekstiä:

```
 4 *DEF A A+1,E-1
 5 *
 6 *>zip -j9 h:\bck\kmm_20080624a kmm*.edt
 7 *
 8 *PVM / Maanantai, 23. kesäkuuta 2008 klo 16.55 (vko 26, pvä 175/366)
 9 *Alusta uudelleen Suodenniemellä...
10 *PVM / Tiistai, 24. kesäkuuta 2008 klo 01.42 (vko 26, pvä 176/366)
11 *ukku ihan kohta...
12 *PVM / Maanantai, 28. heinäkuuta 2008 klo 23.47 (vko 31, pvä 210/366)
13 *korjailuja Sirpan merkintöjen pohjalta...
14 *PVM / Perjantai, 1. elokuuta 2008 klo 18.00 (vko 31, pvä 214/366)
15 *koko päivän revisoinut näitä juttuja... ja vielä... nallea...
16 *
17 *muista (TUlosteissa):
18 - / label pitää olla heti captionin jälkeen, muuten ref antaa section-nron!
19 *
20 *
21 A
22 *Kun aineisto on koossa, siihen päästään tutustumaan piirtämällä kuvia,
23 *tekemällä taulukoita ja tutkimalla tunnuslukuja. Tämä on vaihe, jossa
24 *ei pidä hätäillä. Vaikka tekisi jo mieli rynnätä regressioanalyyseihin,
25 *on viisainta malttaa mielensä. Perusteellinen aineiston esikäsittely
26 *luo pohjan varsinaisille analyyseille. Samalla se
27 *auttaa löytämään virheitä; niitä ei voi kokonaan välttää.
28 *Virheet kannattaa yrittää korjata
29 *saman tien, sillä myöhemmin ne aiheuttavat enemmän vaikeuksia.
30 *
31 - / oli: -n selailu (30.7.2008)
32 - / Aineistoon tutustuminen
33 - / uusi otsikko 5.8.2008, kun koko luvun otsikko uusiksi
34 - / Kyselyaineiston rakenne
35 - / uusi otsikko 8.9.2008, kun koko luvun otsikko uusiksi (vielä kerran!)
36 *Aineistoon tutustuminen
37 - / jhk: "aikaa ja tilaa säästämättä", tässä vain esimerkkejä!! tärkeä vaihe.
38 *
```

# Näkymiä lopullisen version toimituskentistä (22/32)

(Aika monissakin kohdin.)

```
445 *Syntymävuotta kiinnostavampaa olisi tarkastella vastaajan ikää -- etenkin
446 *kun tiedot on kerätty eri vuosina. Tähän palataan
447 *kohdassa~sec:codings (s.~sec:codings).
448 *
449 - / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / /
450 - / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / /
451 - / ----- HUH - KOKO PERJANTAI 1.8.2008 MENI TÄHÄN ASTI PUNNERTAESSA ------------
452 - / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / /
453 - / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / / /
454 -
455 - / Myös SPSS tulostaa tunnuslukuja (ks. tuloste~out:descriptives_sv)
456 - / ja jakaumia, mutta automaattista luokittelua siinä ei jostain
457 - / syystä ole, vaikka automatiikkaa moneen muuhun kohtaan sisältyykin.
458 - / SPSS listaa kaikki havaitut arvot yksittäin, mikä ei ole
459 - / kovin kiinnostavaa etenkään jatkuvien muuttujien osalta, koska tiedot
460 - / eivät tiivisty lainkaan ja tulostaulukotkin kasvavat helposti valtavan kokoisiksi.
461 -
462 -
463 - / Survon STAT-komentoon sisältyvä
464 - / automaattinen luokitus sopii hyvin aineistoon tutustumiseen, kun
465 - / käydään läpi kaikki aineiston muuttujat yleisellä tasolla.
466 - / ?
467 - / Yksityiskohtaisempi perehtyminen edellyttää toki tarkemmin mietittyjä
468 - / luokitteluja, joihin palataan ristiintaulukoinnin yhteydessä (ks.
469 - / kohta~sec:crosstabs).
470 -
```

# Näkymiä lopullisen version toimituskentistä (23/32)

Aika pitkään roikkui mukana SPSS:llä tehty (ruma) runko–lehti-kuva:

```
720 *Seuraavassa perehdytään erikseen
721 *jatkuville ja diskreeteille muuttujille soveltuviin
722 *yhden muuttujan pylväskuviin.
723 -
724 - / Samaan sarjaan kuuluu niin
725 - / sanottu runko--lehti -kuva, josta on esimerkki tulosteessa~fig:stemleaf.
726 - / Sen tulkintatapa auennee parhaiten kun vertailee sitä
727 - / tulosteeseen~out:Svuosi (s.~out:Svuosi). EI AUKENE (KT&KV)
728 -
729 - / muutettu 7.8.2008 "kuvaksi" (AINOA SPSS-kuva tässä kirjassa :)) EI sittenkään!!!
730 - / antaa olla 'fig:', mutta vaihdetaan kontrollikirjaimet (ja tekstiviitteet)
731 -
732 - / O
733 - / CSyntymävuoden jakauman runko--lehti -kuva.
734 - / =fig:stemleaf
735 - / *\medskip
736 - / [
737 - / *sv Stem-and-Leaf Plot
738 - / -
739 - / * Frequency    Stem &  Leaf        Stem width: 10, Each leaf: 2 case(s)
740 - / -
741 - / *     2.00      192 .  &                   & denotes fractional leaves.
742 - / *    11.00      192 .  6889&
743 - / *    17.00      193 .  0012444&
744 - / *    34.00      193 .  5556667777888899
745 - / *    40.00      194 .  001111222333444444
746 - / *    51.00      194 .  555556666777778888899999
747 - / *    51.00      195 .  000001112222222333334444
748 - / *    57.00      195 .  555566666677788888888999999
749 - / *    60.00      196 .  000000111111122223333333444444
750 - / *    48.00      196 .  555555666677788899999999
751 - / *    47.00      197 .  001111122223333344444444
752 - / *    40.00      197 .  555666777778888899
753 - / *    29.00      198 .  0001122223344
754 - / *     9.00      198 .  5567
755 - / -  / säästän tilaa, siirrän selitykset kuvan yläreunaan!
756 - / - / Stem width:        10
757 - / - / Each leaf:       2 case(s)
758 - / -
```

# Näkymiä lopullisen version toimituskentistä (24/32)

Jälleen kommentteja on enemmän kuin varsinaista tekstiä:

```
799 *Kuvan mukaan yleisin yhdessäoloaika olisi noin 20 vuotta.
800 *Kuva on tässä muodossa hyödyksi lähinnä
801 *aineistoon tutustumisessa; tietojen raportoimiseen se on turhan
802 - [P!]
803 *yksityiskohtainen.
804 - / hemmetti, että on hankalia nämä histogrammit tällaisista mjista!
805 - / Kuvan mukaan yleisin yhdessäoloaika olisi noin 20 vuotta,
806 - / mutta osa tästä selittynee karkealla mittaustarkkuudella. Mediaanihan oli 15
807 - / vuotta. Osa ''tasaluvuista'' 5, 10, 15 jne. on suositumpia
808 - / kuin esimerkiksi 29, 31 tai 39, jota ei esiinny yhtään. Arvo
809 - / 32 esiintyy jostain syystä useammin kuin tasan 30.
810 -
811 - / jätän nämä jorinat pois; tein kuvat uudelleen vähän eri tavalla (19.9.2008)
812 -
813 - / Kuvasta fig:histo_yhdessa_1 on jätetty pois edellä mainitut 99 vastaajaa,
814 - / ja pylväiden alle on merkitty luokkakeskukset.
815 - / HISTO-komennon yhteydessä annettu luokittelu muodostaa siis luokat
816 - / 0.1--10, 10.1--20, \ldots, 50.1--60 vuotta. Viimeiseen luokkaan kuuluu
817 - / vain kaksi vastaajaa, joten joissain tilanteissa sen voisi yhdistää
818 - / edelliseen luokkaan ja antaa näin saadulle luokalle esimerkiksi nimi
819 - / ''40 vuotta tai kauemmin''.
820 -
821 - / vain 11 luokkaan. Nyt luokkakeskuksiksi on valittu tasaluvut, joten
822 - / esimerkiksi toinen pylväs vasemmalta sisältää kaikki
823 - / 5.1--10 vuotta yhdessä olleet. Luokkaväli on siis viisi vuotta.
824 - / Nollaluokkaan sisältyvät vain arvot 0.1--2.5, mutta sen pylväs on
825 - / hieman hämäävästi piirretty tässä samanlevyisenä kuin muut.
826 *
827 F
828 *\includegraphics[width=\textwidth]{KMM-histo_yhdessa_2}
829 CYhdessäolon histogrammi viiden vuoden välein.
830 =fig:histo_yhdessa_2
831 f
```

# Näkymiä lopullisen version toimituskentistä (25/32)

Isot revisiot syyskuussa 2008: useampi peräkkäinen, (yli)pitkä viikonloppu (pe–ma, n. 10–16 h/vrk):

```
923 *Luokitusten valinta vaikuttaa huomattavasti kuvien ja tulosten tulkintaan
924 *ja on erinomainen muistutus siitä,
925 *että määrällinen tutkimus ei ole sen objektiivisempaa kuin mikään
926 *muukaan tutkimus, vaikka näin toisinaan näkee väitettävän.
927 - / en saa nyt kuulaa tähänkään:
928 - / Tutkijan on vain pyrittävä kuvaamaan aineistoa mahdollisimman
929 - / totuudenmukaisesti.
930 - / Tutkimusetiikkaan kuuluu, että tutkija kuvaa aineistoa mahdollisimman
931 *
932 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
933 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
934 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
935 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
936 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
937 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
938 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
939 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
940 - / TÄHÄN ASTI ISO REVISIO VALMIS 5.9.2008 KLO 21:30. JATKUU HUOMENNA (LA)...
941 -
942 -
943 - / tässä ollaan taas, 6.9.2008 klo 8:10...
944 - / tässä ollaan taas, 6.9.2008 klo 8:10...
945 - / tässä ollaan taas, 6.9.2008 klo 8:10...
946 - / tässä ollaan taas, 6.9.2008 klo 8:10...
947 - / tässä ollaan taas, 6.9.2008 klo 8:10...
948 - / tässä ollaan taas, 6.9.2008 klo 8:10...
949 -
950 -
951 *
952 - / lyh. vain Muunnokset (6.8.2008)
953 - / Muunnokset ja uudelleenkoodaukset
954 *Muunnokset
955 =sec:codings
```

# Näkymiä lopullisen version toimituskentistä (26/32)

Lopullisessa versiossa on **99** pakotettua sivunvaihtoa sekä **87** pakotettua sisennyksen poistoa (kuvien ym. jälkeen):

```
56 *Mittauskehikon kaksi viimeistä osaa, tulosasteikko ja vertailuperuste,
57 *täsmentävät, miten asteikkoja käytetään tilastollisissa analyyseissa.
58 *
59 *\newpage
60 \Seuraavassa vertailuperustetta ja tulosasteikkoa kuvaillaan yleisellä
61 - [M!] / tohon laitoin jo \noindent:in kokeeksi!
62 *tasolla. Konkreettisempi käsitys niistä muodostuu regressioanalyysin
63 *yhteydessä, kohdassa~sec:REGRAN (s.~sec:REGRAN).
64 *
65 *Vertailuperuste
66 =sec:criterion
67 -
68 - / olkoon "historialliset" kommentit tässä tallessa: :)
69 - / vai vain "Kriteeri"?? vai ULKOINEN kriteeri      vaihda kuvateksti!!!
70 - / 9.8.2007: vaihdoin kuvaan tekstit "Tulosasteikko" ja "Vertailuperuste" !!
71 - / s d:\msm\logo2007 235 / [(vuosi myöh.) Hyvät suomennokset; tulivat jäädäkseen!]
72 *
73 *Vertailuperuste muodostaa pohjan havaintojen vertailulle.
74 - [M!]
75 *Se määritellään yleensä erillään mallista ja asteikoista.
```

# Näkymiä lopullisen version toimituskentistä (27/32)

Myös uusi [PRINT_OFF] -koodi on käytössä (sain idean ehdottaa sitä Sepolle 21.8.2008 – toteutus tuli jo samana päivänä!):

```
846 *Käytännössä selitysastetta enemmän kiinnostavat selittäjät, joihin
847 *seuraavaksi perehdytään tarkemmin.
848 *Tulosteen~out:ipainoREG2 mallin luotettavuuteen palataan
849 *vielä kohdassa~sec:regdiag (s.~sec:regdiag),
850 *jossa tutkitaan, miten hyvin mallista tehdyt oletukset pitävät paikkansa.
851 *
852 *
853 - [PRINT_OFF]  / tämä on hieno ominaisuus (vaikka itse sanonkin)! :)
854 *
855 *Regressioanalyysin antamaa ennustevaliditeetin arviota painaa kuitenkin
856 *alaspäin ...
857 *
858 *dd d:\k\mitt
859 *
860 *Attenuaatio menee KAAVOIKSI, ei sovi tähän kirjaan.
861 *
862 *Reliabiliteetti 2.asteen asteikolle, jossa ylimääräisiä mjia?? (ikä,paino)
863 *
864 *Pitäisi ilmeisesti olla jo ikä,paino mittausmallissa mukana! Graaf.rotaatio etc...
865 *
866 *Ehkä parempi jättää suosiolla tähän; alkaa mennä liian vaikeaksi!
867 *
868 *Jos tätä tulee selviteltyä, niin ehkä "2nd, revised edition" sisällyttää... :)
869 *
870 - [PRINT_ON]
871 *
872 *
873 - / Luokitellut selittäjät ja varianssianalyysi vain sittenkin ANOVA pois otsikosta?
874 - / ei sittenkään k-tason otsikkoa tätä varten! REGDIAG sen sijaan saa olla.
875 *Luokitellut selittäjät
876 =sec:ANCOVA
877 -
878 - / italic: kovarianssianalyysi   (dummyt!)
879 -
880 - / tämän yhdeksi alakohdaksi Varianssianalyysi !
881 -
882 *Edellä laadituissa malleissa selittäjät olivat jatkuvia -- joko
```

# Näkymiä lopullisen version toimituskentistä (28/32)

Jo pelkät ristiviittaukset (lukuihin, kohtiin, sivuihin, lähteisiin) ovat painava peruste LaTeXin käytölle. Survon avulla viittailu käy jopa vielä yksinkertaisemmin:

```
1049 *frekvenssiaineisto
1050 *log-lineaarinen malli
1051 *lineaarinen malli!yleistetty
1052 *frekvenssiaineistojen analyysimenetelmillä.
1053 *Niitä ovat muun muassa
1054 - [P!]
1055 *kohdassa~sec:REGoletukset (s.~ref:GLM)
1056 *mainittuihin, yleistettyihin lineaarisiin
1057 *malleihin kuuluvat log-lineaariset mallit.
1058 *Johdatuksen näihin malleihin esittää \citet[220--232]{Alkula&al}.
1059 *Mallien yhteyksistä korrespondenssianalyysiin kertoo
1060 *esimerkiksi \citet{Greenacre&Blasius2006}.
1061 *
1062 - / kaneetti... 23.9.2008 11:32 Jes! SL approved! :) ..ttu mikä *neetti!
1063 *
1064 *Kirjan menetelmällinen sisältö päättyy tähän.
1065 *Liitteessä~app:SurvoSPSS perehdytään ohjelmistoihin
1066 *ja dokumentointiin sekä tutustutaan kirjassa esitettyjen kuvien ja
1067 *tulosteiden työkaavioihin.
1068 *
1069 *
1070 *
1071 - / kirja loppuu tähän! miten lopetat??
1072 - / NÄIN:
1073 *\newpage
1074 *\thispagestyle{empty}
1075 E
```

Tuossa kohtaa tajusin, että kirjan tekstiosuus päättyy! :-D

# Näkymiä lopullisen version toimituskentistä (29/32)

Monet Survo-tulosteet on tehty PRINT:illä ja upotettu kuvina dokumenttiin pdf-muodossa, niin myös Burtin matriisin näkymä:

```
*PVM / Maanantai, 22. syyskuuta 2008 klo 08.55 (vko 39, pvä 266/366)
*riviväliä lisätty; tilaa löytyi lopulta enemmän
*
*PVM / Tiistai, 23. syyskuuta 2008 klo 10.41 (vko 39, pvä 267/366)
*sittenkin "luokka 5" näkyviin!
*
*/GS-PDF K.PS
*PRINT CUR+1,E-1 TO K.PS  /  käsin korjattu säädetyt nimet! NIMET OIKEALLE TASATTU!
- include PS.DEV
- [Courier(9)][line_spacing(11.5)][margin(10)] / kuten MV'FA2 (linesp oli 9.5)
- shadow 7: [BOLD][REVERSE] [NORMAL]
- shadow 4: [BOLD][DOTTED] [NORMAL]
- shadow 5: [%=0] / [UNDERLINE] [NORMAL]
- shadow Z: [%=0] / [BOLD] [NORMAL]
-     / ■ i cur+1 334
- /      Class 1a 1b 1c  1d  1e 2a 2b 2c  2d 2e 2f 2g  3a 3b 3c 3d 3e ...
*             1                 2                     3                 4
*1          1 51  0  0   0   0  0  0  0   0  8 29 14  10  2  4 35  0 ...
*           2  0 82  0   0   0 12 26 44   0  0  0  0  39  4 20  0 19 ...
*           3  0  0 31   0   0  0  0  4   6 13  6  2  12  7  7  5  0 ...
*           4  0  0  0 100   0  0  0  8  63 29  0  0  81  3 15  0  1 ...
*           5  0  0  0   0 232 10 30 39  49 42 43 19 118 12 39 39 24 ...
-
*2       18-20  0 12  0   0  10 22  0  0   0  0  0  0   1  0  7  0 14 ...
*        21-25  0 26  0   0  30  0 56  0   0  0  0  0  22  2 11  0 21 ...
*        26-35  0 44  4   8  39  0  0 95   0  0  0  0  62  4 23  0  6 ...
*        36-45  0  0  6  63  49  0  0  0 118  0  0  0  88  9 16  2  3 ...
*        46-55  8  0 13  29  42  0  0  0   0 92  0  0  65  9 18  0  0 ...
*        56-65 29  0  6   0  43  0  0  0   0  0 78  0  22  4 10 42  0 ...
*        66-74 14  0  2   0  19  0  0  0   0  0  0 35   0  0  0 35  0 ...
-
*3      työssä 10 39 12  81 118  1 22 62  88 65 22  0 260  0  0  0  0 ...
*    yrittäjä  2  4  7   3  12  0  2  4   9  9  4  0   0 28  0  0  0 ...
*  työtön/muu  4 20  7  15  39  7 11 23  16 18 10  0   0  0 85  0  0 ...
* eläkeläinen 35  0  5   0  39  0  0  0   2  0 42 35   0  0  0 79  0 ...
```

# Näkymiä lopullisen version toimituskentistä (30/32)

Faktorimatriisin selitysten pätkiminen oli aikamoista touhua...

```
480 * 358 NQ_   1 k71.11   Naisten ulkonäkö merkitsee kulttuurissamme enemmän kuin miesten
481 *END
482 *
483 *lyhennys # 2: max 40 merkkiä (sis. "..." jos katkaistu)
484 *
485 *                                        <- katkaisukohta!
486 *         1         2         3         4         5         6         7         8
487 *12345678901234567890123456789012345678901234567890123456789012345678901234567890...
488 *Olen tyytyväinen ulkonäköön.
489 *Vaatteet näyttävät hyvältä päälläni.
490 *Pidän ulkonäöstäni juuri sellaisenaan.
491 *En pidä kehostani.
492 *Olen naisellinen.
493 *En ole fyysisesti viehättävä.
494 *Olen aina hyvännäköinen ajankohdasta...
495 *Laittautumattakin näytän hyvälle.
496 *Kehoni on seksuaalisesti viehättävä.
497 *Olen aina ollut tyytyväinen ulkonäköön.
498 *Ulkonäkö kertoo millainen ihminen olen.
499 *Olen ruma.
500 *Ulkonäköni vastaa sisäistä minääni.
501 *Suhtautumiseni ulkonäköön vaihtelee.                HUOM! Tämän jälkeen tehty
502 *Meikattuna olen tyytyväisempi ulkonäköön.                muutoksia vielä alla
503 *Ulkonäköni on tärkeä osa minua.                          olevaan UPDATE-kaavioon!
504 *Olen kaunis nainen.
505 *En pidä ulkonäöstäni.
506 *Olen fyysisesti hyvässä kunnossa.
507 *Minulle on tärkeätä, että näytän hyvälle.
```

# Näkymiä lopullisen version toimituskentistä (31/32)

...mutta niin oli koko faktorimatriisin esittämisenkin pohdinta; halusin ehdottomasti koko roskan samalle sivulle:

```
*FACTA R1997,4
*MAT F1997=FACT.M    / *F1997~F 53*4
*ROTATE F1997,4 / ROTATION=ORTHO_CLF,0.3  GRAPHICAL    !!
*MAT B1997=AFACT.M   / *B1997~A 53*4
*/LOADFACT UN2007FA B1997
* LIMITS=-0.6,-0.3,0.3,0.599,1   muutettu 0.7 -> 0.6, jotta enemmän B&W!
* SHADOWS=7,1,0,1,7                          + pieniä säätöjä kuten edellä!
* SUMS=2 WIDE=1 POSDIR=1 COLUMNS=SORT      (EI sittenkään välityhjiä!)
*LOADM B1997,12.12 ,CUR+8 / SORT=-UN2007FA,0.1
*
*/GV-SHOW K.PS            / leveystesti neljällä faktorilla ensin!
*PRINT CUR+1,E-1 TO K.PS / blankkoja voi vielä poistaa... ja on poistettava!
- include PS.DEV          / menee silti liian pieneksi; lyhennettävä/katkottava txt...
- [Courier(9)][line_spacing(9.0)][margin(10)] / sopiva fonttikoko+riviväli löydettävä!
- shadow 7: [BOLD][REVERSE] [NORMAL] / lisäoptio: 'k' pois? EI. tyhjät rivivälit!
-  /   d cur+1 E-2 2 / mjien nimien perästä 2 kpl, EI MUITA!
-  / HUOM! HUOM!  KORVASIN LOPULTA GRAAFISELLA ROTAATIOLLA, KS. PALJON ALEMPANA!!!!
*           F2     F4     F3     F1 Sumsqr
-
*k26.3   0.81 -0.07 -0.11  0.02  0.68 Pidän ulkonäöstäni juuri sellaisenaan.
*k26.1   0.78 -0.06 -0.09 -0.01  0.62 Olen tyytyväinen ulkonäköönі.
*k26.2   0.70  0.11 -0.19 -0.05  0.54 Vaatteet näyttävät hyvältä päälläni.
*k26.10  0.68 -0.05  0.21 -0.12  0.52 Olen aina ollut tyytyväinen ulkonäköönі.
*k26.7   0.65  0.03  0.21 -0.07  0.47 Olen aina hyvännäköinen ajankohdasta...
*k26.8   0.64 -0.01  0.05  0.02  0.41 Laittautumattakin näytän hyvälle.
*k26.22  0.63  0.02 -0.20 -0.05  0.44 Pidän siitä mille näytän ilman vaatteita.
*k26.13  0.59  0.06  0.02 -0.04  0.35 Ulkonäköni vastaa sisäistä minääni.
*k26.17  0.56  0.29 -0.24  0.05  0.46 Olen kaunis nainen.
*k26.9   0.56  0.27 -0.37 -0.01  0.52 Kehoni on seksuaalisesti viehättävä.
*k26.19  0.52  0.13 -0.02 -0.11  0.30 Olen fyysisesti hyvässä kunnossa.
*k26.5   0.48  0.32 -0.22  0.05  0.39 Olen naisellinen.
*k26.11  0.43  0.20  0.14 -0.05  0.25 Ulkonäkö kertoo millainen ihminen olen.
*k26.14 -0.29  0.17 -0.10  0.18  0.16 Suhtautumiseni ulkonäköönі vaihtelee.
*k71.8  -0.41  0.26  0.20  0.32  0.38 Koetan olla kauneusihanteiden mukainen.
*k26.18 -0.58 -0.03  0.48 -0.01  0.57 En pidä ulkonäöstäni.
-
```

# Näkymiä lopullisen version toimituskentistä (32/32)

Myös analyysikaavioissa on valtavasti osia, jotka eivät päätyneet kirjan lopulliseen versioon (tässä on transformaatioanalyysia tehtynä kokeeksi eri dimensioisista faktorimatriiseista):

```
3426 *k30.3     -0.07  0.65 -0.14  0.18  0.48 Käytän aikaa itseni "laittamiseen"...
3427 *k30.4     -0.15  0.64 -0.08  0.20  0.48 Yritän aina parantaa ulkonäköäni.
3428 *k30.20    -0.10  0.64  0.10  0.19  0.46 Käytän aikaa ulkonäköni tutkimiseen.
3429 *k30.2     -0.04  0.57 -0.15  0.15  0.38 Tarkastan ulkonäköni peilistä aina,...
3430 *k26.20     0.03  0.55 -0.08  0.13  0.32 Minulle on tärkeätä, että näytän hyvälle.
3431 *k30.15     0.17  0.54  0.05  0.10  0.33 Pyrin herättämään huomiota ulkonäölläni.
3432 *k26.16     0.11  0.53 -0.28  0.10  0.37 Ulkonäköni on tärkeä osa minua.
3433 *k30.8      0.02  0.51 -0.34  0.12  0.39 Ostan vaatteita, joissa näytän hyvälle.
3434 *k30.5      0.20  0.50 -0.13  0.07  0.31 Nautin kun ihmiset katsovat minua.
3435 *k26.15    -0.25  0.43 -0.43  0.16  0.45 Meikattuna olen tyytyväisempi ulkonäköön.
3436 *k30.10     0.27  0.40 -0.03  0.03  0.23 Pukeudun mielelläni seksikkäästi.
3437 *k71.4     -0.06  0.37  0.32  0.12  0.26 Hyvännäköiset ihmiset ovat suositumpia.
3438 *k71.3     -0.10  0.36  0.34  0.13  0.27 Hyvännäköiset pärjäävät elämässään...
3439 *k30.14    -0.12  0.36 -0.22  0.12  0.21 Viehättävänä olen myös halukkaampi...
3440 *k30.13    -0.13  0.36  0.06  0.13  0.17 Olen aikonut mennä plastiikkakirurgille.
3441 *k26.21    -0.20  0.33 -0.31  0.13  0.26 Tiedän, jos olen "huonosti laitettu".
3442 *k71.2     -0.05  0.32 -0.02  0.09  0.11 Miellyttävästä ulkonäöstä on hyötyä.
3443 *k30.11     0.20  0.29 -0.17  0.02  0.15 Liikun pitääkseni vartaloni "kunnossa".
3444 *k30.19     0.05  0.29 -0.23  0.05  0.14 Kiinnitän erityistä huomiota hiuksiini.
3445 *k71.11    -0.09  0.22 -0.07  0.08  0.07 Naisten ulkonäkö merkitsee enemmän...
3446 *k30.18     0.11 -0.40  0.21 -0.13  0.23 Käytän vähän kauneudenhoitotuotteita.
3447 *k30.17    -0.21  0.12  0.51  0.10  0.32 En osallistu iltamenoihin ulkonäköni...
3448 *k30.16    -0.28  0.20  0.45  0.14  0.34 Ulkonäköni takia jätän osallistumatta...
3449 *k30.12    -0.27 -0.23  0.39  0.02  0.28 Yritän olla huomaamattoman näköinen.
3450 *k30.9      0.08 -0.26  0.37 -0.08  0.22 En välitä miltä vaatteeni näyttävät...
3451 *k71.9      0.17 -0.16 -0.26 -0.09  0.13 Elämässä pärjää ulkonäöstä riippumatta.
3452 *k26.1      0.00  0.00  0.00  0.00  0.00 Olen tyytyväinen ulkonäköön.
3453 *k30.1      0.00  0.00  0.00  0.00  0.00 Katson aina miltä näytän kun lähden...
3454 *k71.1      0.00  0.00  0.00  0.00  0.00 Ulkonäkö on yliarvostetussa asemassa.
3455 *k71.5     -0.10  0.08 -0.07  0.04  0.02 Hoikkuuden ihannointi asettaa paineita.
3456 *k71.10    -0.06  0.09  0.08  0.04  0.02 Nuoria ja hoikkia naisia ihannoidaan.
3457 *Sumsqr     6.87  5.14  2.56  0.83
3458 *
3459 *SHADOW BLOCK o
3460 *
3461 *OK!
```

# 17.10.2008: Kirjan kotisivu ja paloja sisällöstä

Tammi - Tieto- ja käsikirjat - Kyselytutkimuksen mittarit ja menetelmät - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.tammi.fi/kirjat/oppi/1/ISBN/9789512657605/

Sähköposti :: Inbox | Helsingin yliopisto - Valtiotieteell... | Tammi - Tieto- ja käsikirjat...

LUKIO
AIKUISKOULUTUS
SOSIAALI- JA TERVEYSALA
YRITYSKIRJAT
TIETO- JA KÄSIKIRJAT

Ostoskori

: : Tieto- ja käsikirjat : :

» Sisällysluettelo

## Kyselytutkimuksen mittarit ja menetelmät

Tilastotiedettä innostavasti!

Kirja pureutuu sellaisiin tutkimuksen tekemisen vaiheisiin, joihin liittyy tilastotieteen asioita ja käsitteitä. Näitä ovat mittaaminen ja tiedonkeruu sekä aineiston hallinta, muokkaus, kuvaaminen ja analysointi tilastollisilla menetelmillä. Pääpaino on kyselytutkimuksessa, jota nykyisin hyödynnetään eri muodoissa useilla aloilla.

Kirja sopii yliopistojen ja ammattikorkeakoulujen opiskelijoille sekä tutkimuslaitosten ja yritysten asiantuntijoille. Kirjasta on hyötyä erityisesti lomakepohjaisten kyselyjen suunnittelussa ja toteuttamisessa, mutta myös näissä käytettävien tilastollisten menetelmien ymmärtämisessä ja tulosten tulkinnassa.

Kirjassa korostuu käytäntö: miten näitä asioita oikeasti tehdään, mitä kannattaa tietää ja mitä on kulloinkin järkevää olettaa. Selkeä lähestymistapa, konkreettiset esimerkit ja kuvat auttavat ymmärtämään monesti vaikeiksi koettuja asioita uudella, tuoreella tavalla.

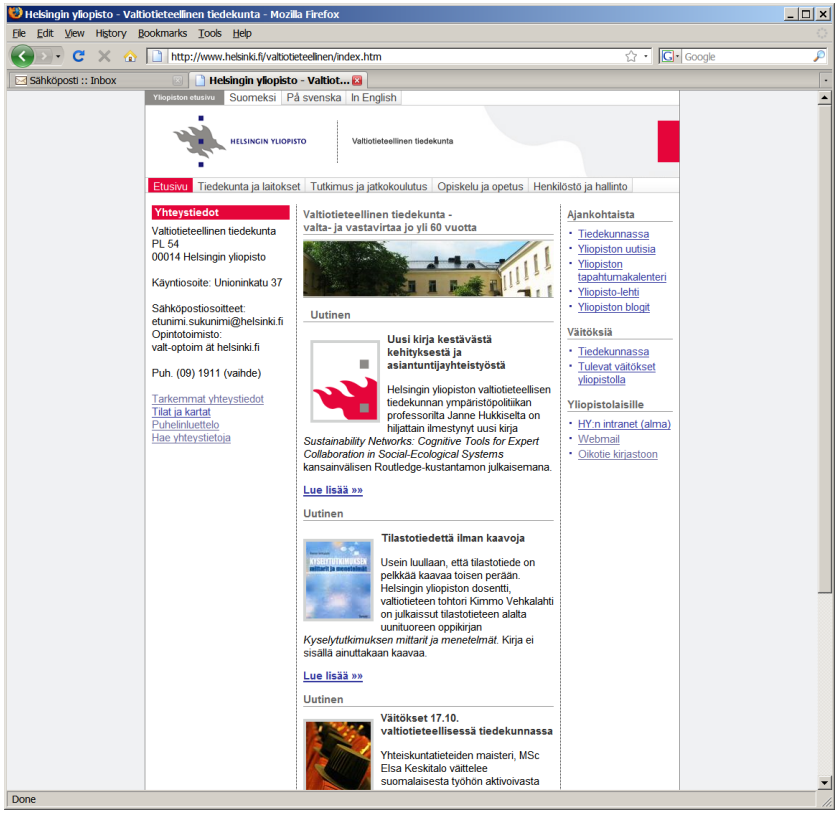
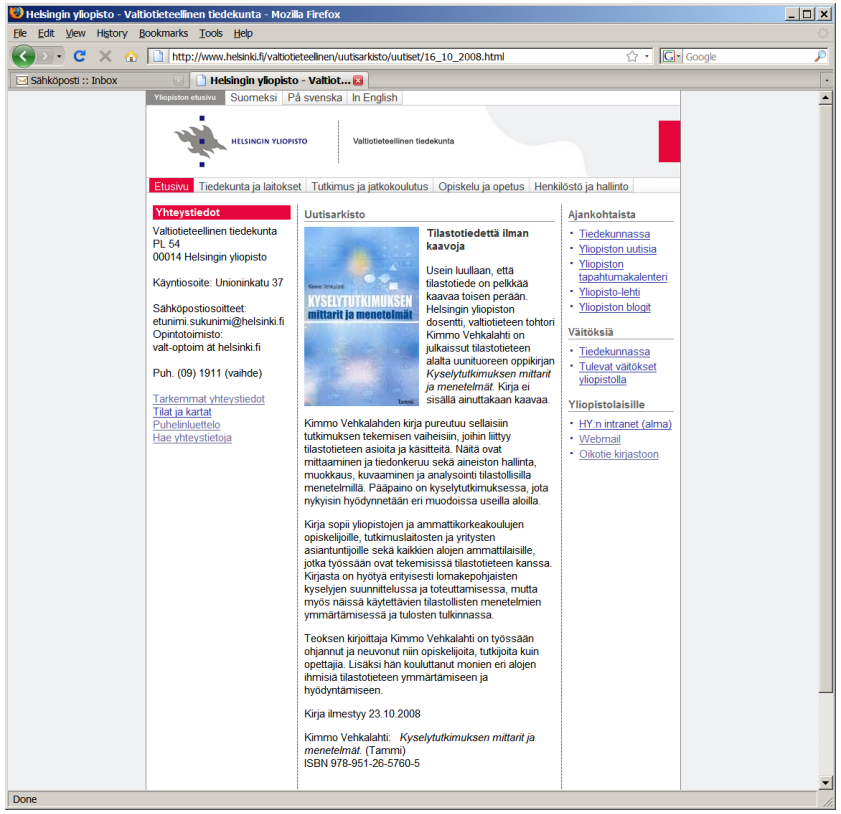
Tekijä:
dosentti, VTT Kimmo Vehkalahti

**ISBN 978-951-26-5760-5**
**Sivuja 223**
**Hinta 29,75 €**
ALV 0% 27,55 €
Lisää ostoskoriin

Ennakkotilattava tuote.
Ilmestyy 28.10.2008

Ilmestyy 28.10.2008

Done

# 17.10.2008: ”Tilastotiedettä ilman kaavoja”

Näin otsikoi valtiotieteellisen tiedekunnan tiedottaja.