

SAS-ohjelmiston perusteet

Harjoitus 2

torstaina 20.11.08 klo 16-18 ML42

Datamatriisit löytyvät netistä <http://mtl.uta.fi/tilasto/sas/syky08/dataEG/>

1. Poista datamatriisista **h2data1** (pääte aina 'sas7bdat') turhat rivit sekä muuttuja 'turha', järjestä muuttujat järjestykseen eka → toinen → kolmas ja järjestä havaintorivit nousevaan järjestykseen muuttujan 'eka' arvojen mukaisesti.
2. Lisää datamatriisiin **h2data2** tekstitiedostossa h2teksti.txt olevat havainnot niiltä osin kuin ne ovat relevantteja. Ajattele, että kyseinen tekstitiedosto on vain toteutuksesi testausta varten ja varsinaisessa tiedostossa olisi tuhansia vastaavantapaisia rivejä.
3. Hurmeen monisteen tehtävä sivulta 46 ja 47. Datamatriisi **KetutData** on nyt myös kansiossa dataEG ja tehtävänäsi on siis muodostaa monisteen ohjeistuksen mukaiset ketut_kaan ja ketut_takaisin.
4. Hurmeen monisteen tehtävä sirontakuviosta sivulta 51. Datamatriisi **vilj2006_muok** löytyy kurssin datakansioista. Kuvia ei tarvitse parannella jälkikäsitteilyllä. Huom. luennoilla ei ole käsitelty kuvien piirtämistä, joten tarkoitus perehtyä itsenäisesti Hurmeen monisteen lukuun 8. Palauta kirjallisesti (sähköisesti tai paperilla) tekemäsi kuvat.
5. Tee EG:llä (tai SASilla) sama t-testi (Analyze → ANOVA → t test) kuin Leppälän monisteen <http://www.uta.fi/kirjasto/pdf/pdfkirjat/leppala.pdf> esimerkissä 10 (s. 26-27). Vertaile SAS/EG:n ja SPSS:n (monisteessa) tuottamia tulosteita.
6. Lue SASin puutiedosto puut.txt (löytyy mm. datakansioista), jossa siis kolme muuttujaa (voit nimetä ne halkaisija, pituus, tilavuus). Katso mallia verkkojulkaisusta <http://www.helsinki.fi/~mjlaune/sasopas/sasdata00.html> (esim. hakemiston kautta hakusanalla INFILE-lause). Mitä ongelmia tämän tiedoston lukemisessa EG:hen (ja yhtäläillä Exceliin) on? Bonusta: Selvitä, mistä nämä ongelmat johtuvat.
7. Bonustehtävä. Miten regressiomallissa, jossa ei ole vakiota, on tapana laskea selitysaste (voit katsoa esim. Puntasen regressioanalyysin kirjasta). Kokeile parilla valitsemallasi aineistolla, onko EG:n tulosteissa selitysaste laskettu tämän mukaisesti. Palauta vastauksesi kirjallisesti.