

1. Aineisto `yoeng`. Esimerkissä 2.5 merkitsevät kertoimet ovat sukupuolella ja englannin numerolla. Lisää malliin niiden interaktio ja testaa sen merkitsevyyttä.
2. Aineisto `yoruotsi`. a) Sovita kaksisuuntaisen varianssianalyysin malli, jossa faktorit ovat sukupuoli (`sp`) ja koulu (`koulu`). b) Laske kertoimien perusteella ja `tapply()` -funktion avulla sukupuolten keskiarvojen erotukset kouluittain. Milliaista vaihtelua niissä näet? c) Testaa interaktion merkitsevyys.
3. Aineisto `yoruotsi`. a) Lisää ed. tehtävän malliin, jossa on sukupuolen ja koulun interaktio, peruskoulun ruotsin numero `ruotsi`, josta on vähennetty 8. Testaa interaktion merkitsevyyttä. b) Sovita malli ilman interaktiota ja tukitse kertoimet.
4. Aja `library(foreign)` ja lue aineisto komennolla (älä jaa eri riveille):
`read.dta("http://www.stat.columbia.edu/~gelman/arm/examples/earnings/heights.dta")`
Aineisto on peräisin laajasta USA:ssa tehdystä hyvinvointitutkimuksesta v. 1990. Tarkastellaan muuttujia `earn` (vuosiansiot dollareina), `sex` (sukupuoli, 1=mies, 2=nainen), `height` (pituus tuumina), `yearbn` (syntymävuosi, kaksi viimeistä numeroa 1900-luvulla). a) Tee puuttuviksi ilmeisen virheelliset arvot. Muuta pituus senttimetreiksi. b) Rakenna malli, missä pituudella selitetään ansioita. Tee sellainen muunnos, joka tuottaa vakiolle tulkinnan "keskimittaisen ihmisen keskiansiot".
5. Jatkoa edelliseen. Sovita sellaisia regressiomalleja, joiden tavoitteena on ennustaa ansioita joillakin sukupuolen, pituuden ja iän kombinaatioilla. Kokeile erilaisia järkeviä muunnoksia ja interaktioita. Valitse ja perustele suosikkimallisi.