

## [MTTTP1] TILASTOTIETEEN JOHDANTOKURSSI, Syksy 2018

<https://coursepages.uta.fi/mtttp1/syksy-2018/>

## HARJOITUS 1 viikko 38

## Ryhmät:

ke	08.30–10.00	LS C6	Paajanen	to	12.15–13.45	LS C8	Korhonen
ke	10.15–11.45	LS C6	Paajanen	to	12.15–13.45	LS C6	Härkönen
ke	12.15–13.45	LS C6	Korhonen	to	14.15–15.45	LS C6	Leppälä
ke	12.15–13.45	LS C8	Männikkö	pe	08.30–10.00	LS C6	Männikkö
ke	14.15–15.45	LS C6	Korhonen	pe	10.15–11.45	LS C6	Korhonen
to	08.30–10.00	LS C6	Männikkö				

Aiheet: *Havaintomatriisi, mittaaminen, mitta-asteikot, otanta, tilastollinen tutkimus*

1. Osoitteessa [http://www.sis.uta.fi/tilasto/mtttp1/syksy2013/MTTTP1\\_pal\\_2013.html](http://www.sis.uta.fi/tilasto/mtttp1/syksy2013/MTTTP1_pal_2013.html) on syksyn 2013 tilastotieteen johdantokurssilla kerätyn opiskelijapalautteen kyselylomake ja palauteyhteenveto. Mikä on tilastoyksikkö? Mikä on populaatio? Mikä on populaation koko? Kuinka monta muuttujaa kyselylomakkeen perusteella muodostuu? Hahmottele vastaavaa havaintomatriisia. Mitä aineiston avulla voidaan tutkia? Onko muuttujien valinnassa mielestäsi onnistuttu? Tutustu tarkemmin palauteyhteenvetoon ja valitse aineistosta muutama muuttuja ja kommentoi saatuja tuloksia. Mitkä seikat voivat vaikuttaa tulosten luotettavuuteen?

2. Mitkä muuttujat ovat kategorisia ja mitkä numeerisia tehtävän 1 aineistossa? Selvitä lisäksi kategoristen muuttujien yhteydessä, ovatko ne luokittelu- vai järjestysasteikkollisia. Olisiko joitain kategorisia muuttujia voinut mitata numeerisesti? Olisiko numeerinen mittaus ollut tämän tutkimuksen kannalta tarpeen?

3. Tutustu yhteen sivulla [http://www.sis.uta.fi/tilasto/mtttp1/syksy2018/MTTTP1\\_aloitusesim\\_6\\_9\\_18.pdf](http://www.sis.uta.fi/tilasto/mtttp1/syksy2018/MTTTP1_aloitusesim_6_9_18.pdf) olevaan tilastollisen tutkimuksen tulokseen. Mikä on tilastoyksikkö? Montako tilastoyksikköä tutkimuksessa on ollut? Mitä muuttujia tutkimuksessa on ollut? Kommentoi lyhyesti tutkimusta, sen tuloksia ja tulostan esitystapaa.

4.

a) Eräällä kurssilla luennot esitettiin toiselle ryhmälle televisioituina ja toiselle ryhmälle tavalliseen tapaan. Osallistujille tehtiin osaamista mittaava testi sekä ennen että jälkeen kurssin. Näiden testien testipistemäärien erotukset olivat:

Naiset	Tavallinen	20.3	23.5	4.7	21.9	15.6	20.3	26.6	21.9	-9.4	4.7	-1.6	25.0
	TV	6.2	15.6	25.0	4.7	28.1	17.2	14.1	31.2	12.6	9.4	17.2	23.4
Miehet	Tavallinen	12.5	7.8	21.9	-7.8	-3.1	3.1	29.7	18.8	28.1	36.0	4.7	-3.3
		45.4	6.3	18.8	9.4	-3.1							
	TV	28.1	-7.8	17.1	14.1	18.8	17.2	1.5	20.3	4.7	15.6	34.4	18.8
		29.7	39.1	9.4	20.3	14.1	18.7	17.2	1.5	25.0	29.7	25.0	23.4
		9.3	1.5	4.7	15.6	26.7	12.5	10.9	32.8	-6.2	3.1	28.1	37.5 20.3

Mikä on tilastoyksikkö ja mitkä ovat muuttujat? Hahmottele vastaava havaintomatriisi. Mitä aineiston avulla voidaan tutkia?

b) Tutkittiin miten tomaattilajike ja istutustiheys vaikuttavat tomaattisatoon. Tutkimustulokset ovat taulukossa 1. Mikähän voisi olla tilastoyksikkö? Mitkä ovat muuttujat? Muodosta vastaava havaintomatriisi. Määritä muuttujien mitta-asteikot. Mitä aineiston avulla voidaan tutkia?

Taulukko 1. Satomäärät (100 kg/ha) eri lajikkeilla ja eri istutustiheyksin.

	Istutustiheys (1000 tainta/ha)			
	10	20	30	40
Lajike 1	7.9, 9.2, 10.5	11.2, 12.8, 13.3	12.1, 12.6, 14.0	9.1, 10.8, 12.5
Lajike 2	8.1, 8.6, 10.1	11.5, 12.7, 13.7	13.7, 14.4, 15.4	11.3, 12.5, 14.5
Lajike 3	15.3, 16.1, 17.5	16.6, 18.5, 19.2	18.0, 20.8, 21.0	17.2, 18.4, 18.9

5. Tutkitaan kuntoliikuntaa vähintään keskimäärin kolme kertaa viikossa harrastavia aikuisia ja tehdään kysely, jossa on mm. seuraava kysymys:

Harrastan (rastita)

- hiihtoa
- pyöräilyä
- uintia
- lenkkeilyä
- yksikään näistä neljästä lajeista ei kuulu harrastuksiini

Tarvitaanko tämän kysymyksen kohdalla viimeinen vaihtoehto? Miksi? Hahmottele vastaava havaintomatriisi.

6. Sivulta <https://coursepages.uta.fi/mttt1/esimerkkiaineistoja/> löydät joitain esimerkkiaineistoja. Tutustu niistä yhteen. Mikä on valitsemassasi aineistossa tilastoyksikkö ja mitkä ovat muuttujat. Määritä myös muuttujien mitta-asteikot. Mitä aineiston avulla voisi tutkia?

7. Etsi pieni (muutama tilastoyksikkö ja muuttuja) aineisto, jossa tilastoyksikkönä on

- a) asunto, auto, kirja, tietokone tai lehti,
- b) urheilija, urheiluseura tai joukkue,
- c) elokuva, TV-uutiset, CD-levy tai www-sivu,

ja muodosta havaintomatriisi. Määrittele aineistosi muuttujien mitta-asteikot.

8. Tarkastele alla kuvattua *ISI*-aineistoa. Mikä on tilastoyksikkö? Mikä voisi olla populaatio? Mitä aineiston avulla voidaan tutkia? Onko muuttujien valinnassa mielestäsi onnistuttu? Mitkä aineiston muuttujista ovat kvantitatiivisia?

### Aineiston *ISI* kuvaus

Havaintoaineistossa on 210 tilastoyksikköä. Aineiston perustana ovat vuosina 1963-1977 Tampereella tapahtuneet synnytykset. Tiedot kerättiin neuvolakorteista Tampereen terveyslautakunnan luvalla. (*Liski & Puntanen, Tilastotieteen peruskurssi I*)

### Muuttujien selitykset

AIDINAMM	Äidin ammatti (koodaus <i>Liski &amp; Puntanen, Tilastotieteen peruskurssi I</i> ) {1,6}
AIDINIK	Äidin ikä täysinä vuosina
ISANAMM	Isän ammatti
ISANIK	Isän ikä täysinä vuosina
AVIOSUH	Aviosuhde: 1=nton,2=nssa,3=eronnut,4=leski
LAPSLKM	Lasten lukumäärä {0,6}
VUOSI	Synnytysvuosi {63,77}
KUUKAUSI	Synnytyskuukausi {1,12}
PAIVA	Synnytyspäivä {1,31}
VIKONPV	Viikonpäivä, jolloin synnytys tapahtui {1,7}
PITUUS	Äidin pituus cm:inä
ALKUPAI	Äidin paino ennen raskautta
VTUTKPAI	Äidin paino viim. tutkimuksessa ennen synnytystä
JTPAINO	Äidin paino jälkitutkimuksessa
SEX	Lapsen sukupuoli, 0=tyttö 1=poika
LAPSPIT	Lapsen pituus cm:inä
LAPSPAI	Lapsen paino grammoina
KUNTO	Lapsen kuntopisteet (0-10)

Aineisto on mikroluokissa hakemistossa p:\pub\pk\data nimellä ISI.SAV.

9. Tutustu sivun <http://www.fsd.uta.fi/menetelmaopetus/otos/otantamenetelmat.html> avulla otantamenetelmiin. Mitä otantamenetelmää käyttäisit, jos tehtävänäsi olisi valita satunnaisotos

- haloista korkeassa pinossa,
- perunoista, jotka ovat säkeissä kuorma-auton lavalla,
- autoista autojonosta
- Suomen kunnista?

10. Opintojaksoon liittyy empiirinen harjoitustyö. Työn teko aloitetaan tutkimusongelmien määrittämisellä ja aineiston hankinnalla. Aineiston sopivuus on syytä varmistaa harjoitustyön ohjaajalta. Lue harjoitustyön teko-ohjeet <http://www.sis.uta.fi/tilasto/mtt1/syksy2018/htyop118.pdf> . Onko ohjeissa epäselviä kohtia? Mitä? Pohdi alustavasti, mistä aiheesta (ja kenen kanssa) voisit tehdä harjoitustyön. Mikä olisi tilastoyksikkö ja mitä muuttujia tarvitsisit? Muuttujien sopivuutta ja riittävyttä voit arvioida ohjeiden sivulla 4 olevan esimerkin kohdan 3 Riippuvuustarkastelut avulla. Millä tavalla saisit hankittua aineiston? Tehtävät 6 ja 7 antavat vinkkejä mahdollisista aiheista ja aineistoista! Huomaa, että tarvitset muuttujien arvot tilastoyksiköittäin.

11. Pohdi näihin harjoituksiin liittyen omaa oppimistasi ja osaamistasi esimerkiksi seuraavien kysymysten avulla. Huomaatko eron populaation ja otoksen välillä? Tiedätkö mikä on tilastoyksikkö ja mikä muuttuja? Osaatko muodostaa havaintomatriisin? Osaatko määrittää muuttujan mitta-asteikon? Onko sinulla käsitys siitä, millainen harjoitustyöaineiston pitää olla? Jäikö joku asia vielä epäselväksi? Mikä vaatii vielä lisäopiskelua?