Christina Gustafsson

Tilastollinen tietojenkäsittely STAT2100 IBM SPSS Statistics 20 for Windows Osa 1

Kevät 2013

SISÄLLYSLUETTELO

| 1. OHJELMAN PERUSKÄYTÖSTÄ |
|---|
| 1.1. Yleistä |
| 1.2. Ohjelman käynnistäminen2 |
| 1.3. SPSS:n ikkunointi |
| 1.4. Esimerkki tehtävän määrittelystä5 |
| 1.5. Asetukset |
| 1.6. Apua? |
| 2. HAVAINTOAINEISTON MUODOSTAMINEN |
| 2.1. Muuttujan määrittely9 |
| 2.2. Havaintoaineiston syöttäminen 13 |
| 2.3. Aineistotiedoston tallettaminen |
| 2.4. Havaintoaineiston avaaminen ja lukeminen |
| 2.4.1. Excel-taulukon avaaminen |
| 2.4.2. Tekstitiedoston lukeminen |
| 2.5. SPSS-aineistojen yhdistäminen |
| 3. HAVAINTOAINEISTON MUOKKAAMINEN |
| 3.1. Muuttujan arvot lausekkeiden avulla |
| 3.2. Muuttujan arvojen vaihtaminen |
| 3.3. Tilastoyksiköiden järjestäminen |
| 3.4. Tilastoyksiköiden painottaminen |
| 3.5. Osajoukon valitseminen |
| 3.6. Tilastoyksikköryhmien muodostaminen |
| 3.7. Muuttujajoukon muodostaminen |
| 3.8. Aineiston rakenteen vaihtaminen |
| 3.9. Yhteenvetoaineiston muodostaminen |
| 4. GRAAFISET ESITYKSET |
| 4.1. Kuviotyypit |
| 4.2. Esimerkkikuvioita Chart Builder –vaihtoehdon avulla |
| 4.3. Esimerkkikuvioita Legacy Dialogs –vaihtoehdon avulla |
| 4.4. Kuvion muokkaaminen ja mallikuvio |

1. OHJELMAN PERUSKÄYTÖSTÄ

1.1. Yleistä

SPSS (Statistical Product and Service Solutions) on laaja ja joustava tilastollisen analysoinnin ohjelma. SPSS 20 koostuu useista eri moduuleista, joista vain pieni osa on käytössä Vaasan yliopistossa. Ohjelmaan on lisäksi olemassa erillistuotteita, mm. rakenneyhtälömallitarkasteluun sopiva AMOS. Lisäpakettina ohjelmaan sisältyy myös mahdollisuus hyödyntää algoritmeja, jotka kirjoitetetaan Rohjelmistolla tai Python-kielellä.

Suomenkielisiä oppaita (tosin aika varhaisille versioille) SPSS-ohjelman käyttöön löytyy verkosta mm.

- SPSS-opas osoite http://myy.haaga-helia.fi/~taaak/k/SPSS.pdf
- SPSS-opas 12.0-versiolle osoite http://www.valt.helsinki.fi/blogs/stat/opas.htm
- SPSS-oppimisympäristö osoite http://www.fsd.uta.fi/menetelmaopetus/SPSS/SPSS.html

1.2. Ohjelman käynnistäminen

SPSS-ohjelmaa käytetään Windows-ympäristössä muiden Windows-ohjelmien tapaan pääasiassa hiiren avulla.

SPSS-ohjelma käynnistetään Windows-ympäristössä seuraavasti:

- 1. Näpäytetään **Start**-painiketta kuvaruudun vasemmasta alakulmasta.
- 2. Valitaan All Programs.
- 3. Valitaan **IBM SPSS Statistics**-kansio.
- 4. Näpäytetään vaihtoehtoa **IBM SPSS Statistics 20,** jolloin saadaan näkyviin SPSS-ohjelman pikavalintaikkuna (kuvio 1).

Pikavalintaikkuna ehdottaa valitsemaan jonkin aikaisemmin käytetyn SPSS-aineiston (kohta **Open an existing data source**). Muitakin vaihtoehtoja on valittavana kuten aloittelijan opetusohjelma (kohta **Run the tutoria**). Ikkunan voi sulkea painamalla **Cancel**-painiketta. Ikkunan avautuminen jatkossa voidaan estää valitsemalla kohta: **Don't show this dialog in the future**.

SPSS-ohjelman käyttö lopetetaan valitsemalla valikosta **File** kohta **Exit**. Ohjelmasta poistuttaessa voi päättää, haluaako tallettaa niiden ikkunoiden sisällöt, joihin on tullut muutoksia edellisen talletuskerran jälkeen.



Kuvio 1. Pikavalintaikkuna

1.3. SPSS:n ikkunointi

SPSS:ssä on kaksi perusikkunaa: (**IBM SPSS Statistics**) **Data Edito**r - ja (**IBM SPSS Statistics**) **Viewer** -ikkuna.

SPSS-ohjelman **Data Editor**- eli **aineisto**ikkuna (kuvio 2) saadaan näkyviin automaattisesti, kun SPSSohjelma käynnistetään. Tässä ikkunassa on valikoita ja työkalupainikkeita, joiden avulla havaintoaineistoa voidaan muokata ja analysoida. Ikkuna on jaettu kahteen välilehteen: **Data View** ja **Variable View**. **Data View** -välilehdellä on käsiteltävä havaintoaineisto taulukkomuodossa ja **Variable View** –välilehdellä on havaintoaineiston muuttujien määrittelyt. **Data Editor** -ikkunan sisältö molempine välilehtineen voidaan tallettaa tiedostoksi **File** -valikosta valinnalla **Save** tai **Save As...** tai **Save All Data** tai työkalulla **B**. Havaintoaineistotiedoston oletustalletustyyppi on **SPSS Statistics** (.sav). **Data Editor** -ikkunan havaintoaineisto-osuus voidaan myös tallettaa esimerkiksi Excel-taulukoksi. Samanaikaisesti avoinna voi olla useita aineistoikkunoita, suositeltavaa on kuitenkin pitää auki vain yhtä aineistoa kerrallaan.

| ta kys | elyaine | isto.sav [DataSet2 | 2] - IBM SP | SS Statistics | Data Editor | | OTSIKKORIVI |) | | _ | _ | | | x |
|--|---------|---------------------------|-------------------|---------------|----------------|-------------------|-------------------------------|------------------|----------|----------------|-------------|--------|----------|---|
| File | Edit | <u>V</u> iew <u>D</u> ata | Transform | Analyze | <u>G</u> raphs | <u>U</u> tilities | Add- <u>o</u> ns <u>W</u> ind | low <u>H</u> elp | 0 | VALIKK | ORIVI | | | |
| | | | | | | | | | | | | | | |
| SOLUEDITORI TYÖKALURIVI Visible: 17 of 1 | | | | | | | | of 17 Varia | bles | | | | | |
| | | nimi | tdk | paaaine | kurssi | ika | sukupuol | avoavio | ansiotyo | tyomaara | edistys | opetus | osallist | |
| | 1 | Mirja | 1 | 3 | 3 | 24 | 2 | 2 | 2 | | 2 | 16 | 16 | 4 |
| | 2 | Brit | t <mark>(1</mark> | 1 | 2 | 21 | 2 | 2 | 2 | | 3 | 16 | 16 | |
| | 3 | Anna | 2 | 6 | 1 | 20 | 2 | 2 | 2 | 24 | 3 | 19 | 19 | |
| | 4 | Liisa | 1 | 4 | 4 | 22 | 2 | 2 | 1 | 20 | 4 | 20 | 16 | |
| | 5 | Ritva | 2 | 7 | 1 | 20 | 2 | 2 | 1 | 2 | 4 | 22 | 18 | |
| | 6 | Kristiina | 2 | 6 | 1 | 19 | 2 | 1 | 2 | 22 | 4 | 30 | 27 | |
| | 7 | Helena | 1 | 2 | 4 | 23 | 2 | 1 | 1 | 15 | 3 | 8 | 12 | |
| | 8 | Jarkko | 1 | 5 | 4 | 27 | 1 | 1 | 1 | 12 | 2 | 14 | 14 | |
| | 9 | Marko | 1 | 1 | 2 | 22 | 1 | 2 | 2 | | 3 | 22 | 20 | |
| 1 | 0 | Stina | ı <mark>1</mark> | 3 | 1 | 19 | 2 | 2 | 2 | | 2 | 22 | 18 | - |
| | | 1 | | | | | | - min | | | | | | |
| Data | View | Variable View | VÄLI | EHDET | | | | | | | | | | |
| | | | CALI | | _ (| TILARIVI |) | | IBM SPS | S Statistics P | rocessor is | ready | | |

Kuvio 2. Aineistoikkuna ja sen Data View -välilehti

| 🔚 *Output1 [Document1] - IBM SPSS Stat | istics Viewer OTSIKKORIVI | | | | | _ _ X |
|--|------------------------------|---------------|---------------------------|--------------------|-----------------------|--------------------|
| <u>File Edit View Data Transform</u> | Insert Format Analyze Graphs | Utilities Add | - <u>o</u> ns <u>W</u> in | idow <u>H</u> elp | VALIKKORIVI | J |
| | 🛄 🗠 🛥 🧱 🎬 | | 0 | | 护 🖻 🐌 | |
| • • + - 🛍 | | | TYÖKALU | RIVIT | | |
| E E Output | TULOKSET | Pääaine | (| | | <u> </u> |
| Descriptives Frequencies | | Frequency | Percent | Valid Percent | Cumulative Percent | |
| Title | Valid englanti | 14 | 4,3 | 4,4 | 4,4 | |
| Notes | nykysuomi | 33 | 10,2 | 10,3 | 14,6 | |
| Active Dataset | ruotsi | 68 | 21,1 | 21,2 | 35,8 | |
| Statistics | saksa | 49 | 15,2 | 15,3 | 51,1 | |
| | viestintätieteet | 55 | 17,1 | 17,1 | 68,2 | |
| Correlations | hallintotiede | 58 | 18,0 | 18,1 | 86,3 | |
| | julkisoikeus | 32 | 9,9 | 10,0 | 96,3 | |
| SISÄLLYSLUETTELO | · | | | 2.5 | | 4 |
| | TILARIVI | | IBM SPS | S Statistics Proce | essor is ready | H: 310, W: 559 pt. |

Kuvio 3. Tulosikkuna

Tulosikkuna on jaettu pystysuunnassa kahteen osaan: vasemmalla puolella on sisällysluettelo ja oikealla puolella varsinaiset tulokset. Suurinta osaa tulosikkunan tuloksista voidaan järjestellä ja muokata. Tulosikkunassa on myös valikoita ja työkalupainikkeita. Samanaikaisesti voi olla avoinna useita tulosikkunoita.

Jotkin SPSS-ohjelman ikkunat avautuvat vain, jos niin erikseen halutaan. Chart Editor- eli kuvanmuokkausikkunassa voidaan muokata muodostettuja tilastokuvioita. SPSS-ohjelmaa voidaan hallita valikoiden lisäksi myös Syntax Editor- eli komentoikkunaan kirjoitettavilla SPSS:n komentokielisillä käskyillä. SPSS Table Editor- eli taulukonmuokkausikkunassa voidaan muokata tulosikkunan taulukkojen ulkoasua. Ikkunassa Script Editor eli skriptieditorilla voidaan automatisoida monia SPSS:n toimintoja.

Ikkunoissa näkyy ylimpänä otsikkorivi, joka kertoo ikkunan nimen tai/ja sisällön.

Otsikkorivin alapuolella on **valikkorivi**. Valikot vaihtelevat hiukan ikkunoittain. Seuraavia valikoita löytyy mm. aineisto- ja/tai tulosikkunasta:

File-valikon valinnoilla voidaan mm. avata tiedostoja, tallettaa ikkunan sisältöä, tulostaa, pysäyttää SPSS-prosessori ja poistua SPSS:stä.

Edit-valikon valinnoilla voidaan mm. siirtää tekstiä, kuva tai taulukko Windowsin leikepöydälle, liittää leikepöydällä oleva osuus SPSS:ään, muuttaa SPSS:n oletusarvoja ja lisätä muuttujia ja tilastoyksiköitä.

View-valikon valinnoilla voidaan näyttää tai piilottaa ruudun osia.

Data-valikon valinnoilla voidaan mm. määritellä muuttujia, lajitella havaintoja, yhdistellä havaintoaineistoja, valita osajoukkoja sekä ryhmitellä tilastoyksiköitä.

Transform-valikon valinnoilla voidaan tehdä muuttujamuunnoksia ja uusia muuttujia luokittelun avulla.

Analyze-valikon valinnoilla suoritetaan tilastolliset analyysit sekä muodostetaan raportit ja listaukset. Graphs-valikon valinnoilla muodostetaan tilastokuviot.

Utilities-valikon avulla saa tietoja yksittäisistä muuttujista ja koko aineistosta, ja voidaan määrittää muuttujaryhmät.

Add-ons-valikosta löytyy SPSS-lisämoduleista ja -erillisohjelmista

Window-valikon valinnoilla voidaan mm. minimoida ikkunoiden koko ja siirtyä ikkunasta toiseen.

Help-valikossa on SPSS:n avustusjärjestelmä, linkki SPSS:n kotisivulle ja tilastomenetelmän valinnan opastus.

Insert-valikon (vain tulosikkunassa) avulla voidaan lisätä haluttuun kohtaan esim. otsikoita.

Format-valikon (vain tulosikkunassa) avulla voidaan tulostus sijoittaa haluttuun kohtaan tulostusikkunassa.

Useissa ikkunoissa on valikkorivin alapuolella **työkalupainikkeita**, joilla voidaan toteuttaa käskyjä, jotka sisältyvät myös valikoihin. Työkalun toiminnosta saadaan lyhyt kuvaus painikkeen alapuolelle ja myös tilarivin vasempaan reunaan, kun työkalua osoitetaan hiirellä.

SPSS-ohjelman näkyvät osat ovat käyttöliittymä varsinaiseen laskentaohjelmaan (SPSS Statistics Processor), joka toimii taustalla. Aktiivisen (= otsikkorivi yleensä kirkkaansininen) ikkunan alareunassa on tilarivi, joka kertoo, mikä on laskentaohjelman tilanne (esim. IBM SPSS Statistics Processor is ready, Running GET, Running CROSSTABS, ...). Aineistoikkunan tilarivillä on tietoa siitä, onko aineistosta käytössä vain osajoukko (esim. Filter On, Use On), onko havaintoaineisto jaettu ryhmiin (Split by ...), tai onko painomuuttuja käytössä (Weight On). Tulosikkunan tilarivi kertoo mm. valittuna olevan Pivot-taulukon koon.

1.4. Esimerkki tehtävän määrittelystä

Kukin SPSS-tehtävä muodostuu yleensä analyysikokonaisuudesta, johon tarvittavista määrityksistä osa on sellaisia, ettei niihin välttämättä tarvitse puuttua, kun taas osa määrityksistä vaatii toimenpiteitä jokaisella tehtäväkerralla. Näistä edellisiin kuuluvat mm. erilaisten tunnuslukujen valinnat ja jälkimmäisiin esim. tehtävän muuttujien valinnat.

Valinnaisille toiminnoille on usein määritelty oletusarvot, joita käytetään, ellei käyttäjä muuta määrittele. Vaikka erilaiset analysoinnit ja muunnokset vaativat hyvin erilaisia valintoja, ovat tehtävämääritykset rakenteiltaan samantyyppisiä. Esimerkkinä tehtävämäärityksestä on frekvenssijakaumataulukon muodostaminen (kuvio 4), jolloin on valikosta **Analyze** valittu alivalikko **Descriptive Statistics** ja vielä valittu proseduuri **Frequencies**.

| | | Variable(s): | Statistics |
|--|------|------------------|----------------------------------|
| nimi Tiedekunta [tdk] Vuosikurssi [kurs Ikä vuosina [ika] Sukupuoli [sukup Onko avo- tai avio Onko ansiotyöss Ansiotyön määrä | | Räaine [paaaine] | <u>C</u> harts <u>F</u> ormat |
| Display frequency tal | oles | | |

Kuvio 4. Frequencies-proseduurin päämäärittelyikkuna

Kunkin tehtävän valinnat on jaettu ensimmäisenä avautuvaan **päämäärittely-** eli **valintaikkunaan**, jonka valinnat on määriteltävä jokaiselle tehtävälle erikseen. Tässä ikkunassa valitaan yleensä käytettävät muuttujat sekä tehdään tarvittaessa muita tehtävään liittyviä valintoja. Ikkunan alalaidassa olevat (tai oikeassa reunassa) olevat painikkeet ovat

- OK= suorittaa määritellyn tehtävänPaste= näyttää SPSS:n komentokieliset käskyt SPSS:n komentoikkunassaReset= asettaa kaikki määritykset alkutilaanCancel= keskeyttää määritykset ja ikkuna sulkeutuu
- **Help** = avaa tehtävästä kertovan aputekstin

määrittelyikkunan Muuttuja valitaan tehtävään vasemman reunan muuttujaluettelosta kaksoisnäpäytyksellä, jolloin muuttuja siirtyy ko. tehtävässä käsiteltävien muuttujien listaan. Muuttujan voi siirtää muuttujaluettelosta myös laahaamalla sen hiiren avulla käsiteltävien muuttujien listaan. Kolmas tapa on näpäyttää muuttujan nimeä kerran ja painaa sitten käsiteltävien muuttujien listan vasemmalla puolella olevaa oikealle osoittavaa nuolipainiketta. Useita muuttujia voi siirtää valitsemalla joko Ctrl- tai Shift -näppäimen avustuksella ne muuttujat, jotka halutaan siirtää, ja sitten laahataan muuttujaryhmä hiiren avulla. Jos valituksi tuli väärä muuttuja, poistetaan muuttuja joko laahaamalla se hiiren avulla takaisin muuttujaluetteloon tai näpäyttämällä hiirellä ko. muuttujan nimeä ja painamalla nyt vasemmalle osoittavaa nuolipainiketta. Kustakin muuttujasta näkyy muuttujaluettelossa ensimmäisenä kuvake, jota kertoo muuttujalle valitun mitta-asteikon (ja tietyissä tilanteissa näkyy tyyppitietoa), seuraavana muuttujan selitys/nimike (, jos sellainen on muuttujalle annettua) ja viimeisenä muuttujan lyhyt nimi.

Ikkunan oikeassa reunassa (tai alalaidassa) on painikkeita, joilla voidaan tarvittaessa avata erilaisia tehtävän **lisämäärittelyikkunoita**. Lisämääritysten oletusarvot näkyvät valmiiksi valittuina kohdissa, joissa oletusarvoja on. Esimerkkinä lisämäärityksestä on frekvenssijakaumakuvion muodostaminen **Charts**-option avulla (kuvio 5).

Lisämääritysikkunoissa on yleensä painikkeet

| Continue | = määritykset vahvistetaan ja ikkuna sulkeutuu |
|----------|--|
| Cancel | = määritykset keskeytetään ja ikkuna sulkeutuu |
| Help | = avaa tehtävästä kertovan aputekstin |

| P F | requencies: Charts |
|------------|--------------------------------|
| FC | hart Type |
| 0 | None |
| (| Bar charts |
| 0 | <u>P</u> ie charts |
| 0 |) <u>H</u> istograms: |
| | Show normal curve on histogram |
| FC | chart Values |
| 0 | Erequencies O Percentages |
| | |

Kuvio 5. Frequencies-proseduurin Charts-lisämäärittelyikkuna

1.5. Asetukset

SPSS:n perusasetukset riittävät ohjelman käytön aloittamiseen, mutta joskus tarvitaan käyttöön hieman erilaisia ominaisuuksia. Asetuksia voidaan muuttaa pysyvästi valitsemalla **Edit**-valikosta kohta **Options**. Asetuksiin liittyviä määrittelyvälilehtiä on useita (kuvio 6). Seuraavia muutoksia ominaisuuksissa tarvitaan joskus.

General-välilehdellä voidaan valita, halutaanko valintaikkunoiden muuttujaluetteloissa (Variable Lists) muuttujista näyttää nimike[lyhyt nimi] (Display labels) tai pelkkä lyhyt nimi (Display names), muuttujat voidaan esittää aakkosjärjestyksessä (Alphabetical) tai aineistonmukaisessa järjestyksessä (File) tai tilastollisen mitta-asteikon mukaan (Measurement level). Välilehdellä saa myös valittua tulosikkunan tulostuskielen kohdassa Language. File Locations -välilehdellä voidaan tehdä muutoksia SPSS:n automaattisesti luoman päiväkirjan (Session Journal) määrityksiin.

Viewer-lehdellä voidaan kullekin tuloskohtatyypille määritellä erikseen mm. millaisella tekstityylillä se esitetään, ja näytetäänkö sitä tulosikkunassa lainkaan. Output Labels -lehdellä voidaan määrittää, näytetäänkö tuloksissa muuttujista nimi, nimike vai molemmat ja muuttuja-arvoista arvo, arvon nimike vai molemmat.

Charts-lehdellä voidaan määritellä mm. kuvien leveys/korkeussuhde, tekstityyli ja kehykset sekä tyyli mallien joukosta. **Pivot Tables** -lehdellä voidaan valita taulukoiden esitystyyli valmiiden mallien joukosta.

Data-lehdellä voidaan uusille numeerisille muuttujille määrätä formaatti, joka poikkeaa SPSS:n oletusarvosta. **Currency**-lehdellä voidaan muuttaa lukuarvojen esittämiseen liittyviä määrittelyjä.

7

| Serieral Viewer Data Currency Output Labels Cr | arts Pivot Tables File Locations Scripts Syntax Editor | | | | |
|---|--|--|--|--|--|
| ariable Lists | Coutput | | | | |
| Display la <u>b</u> els | No scientific notation for small numbers in tables | | | | |
| 🕽 Display <u>n</u> ames | Apply locale's digit grouping format to numeric values | | | | |
| ⊘ Alp <u>h</u> abetical | Measurement system: Centimeters | | | | |
| <u>File</u> | - | | | | |
| 🛇 Measuremen <u>t</u> level | Language: English 👻 | | | | |
| Roles | Notification: | | | | |
| To save time, some dialogs allow the use of predefined | | | | | |
| field roles to automatically assign variables (fields) to list in dialoos. | | | | | |
| Use predefined roles | | | | | |
| 🛇 Use custom assignments | Character Encoding for Data and Syntax | | | | |
| | Locale's writing system | | | | |
| Vindows | Itpianda (upiuszad obergatar aat) | | | | |
| | G Onicode (universal character set) | | | | |
| Look and reel: SPSS Standard | Character encoding cannot be changed when any non-empty datasets are open | | | | |
| 🔲 Open syntax window at startup | | | | | |
| Open only one dataset at a time | User Interface | | | | |
| | Language: English | | | | |

Kuvio 6. Options-ikkuna

1.6. Apua?

SPSS:n käyttöön saa apua mm. **Help**- eli ohjevalikon valinnoilla. **Topics**-vaihtoehdon avulla avautuu ohjemäärittelyikkuna, johon voidaan esim. kirjoittaa sana tai käsky, josta halutaan lisätietoa. Tämä valinta sisältää myös luettelot SPSS:n käskyistä ja tilastotermeistä. **Tutorial**-valinnan avulla voi itsekseen opetella SPSS:n peruskäyttöä. **Case Studies** -valinnalla saa esimerkkejä tilastoanalyysityypeistä tulostulkintoineen. **Statistics Coach** -valinnalla saadaan ohjeita tilastomenetelmän valintaan ja tulosten tulkintaan. **Command Syntax Reference** -valinnalla saa avattua oppaan SPSS:n komentokielen rakenteesta ja käskyistä. **Algorithms**-valinnalla saa avattua kuvauksen tilastollisen menetelmän laskenta-algoritmista. **IBM SPSS Products Home** -valinta avaa SPSS:n kotisivut. **About**-valinnalla saa tietoa esim. ohjelmiston laitteistovaatimuksista.

Valintaikkunoiden valintoihin saa apua ikkunan **Help**-painikkeella, jolloin avautuu ikkuna, jossa kuvataan eri valintavaihtoehdot lyhyesti.

2. HAVAINTOAINEISTON MUODOSTAMINEN

2.1. Muuttujan määrittely

Uuden SPSS-istunnon käynnistyttyä avautuu aineistoikkuna. **Data View** -välilehden (kuvio 2) soluihin voidaan suoraan kirjoittaa havaintoarvoja, mutta aineiston muodostaminen kannattaa yleensä aloittaa muuttujien määrittelyllä **Variable View** -välilehdellä (kuvio 7). **Variable View** -välilehti valitaan aineistoikkunan alareunasta näpäyttämällä välilehteä hiirellä. Tällä välilehdellä yksi muuttuja vastaa yhtä vaakariviä ja sarakkeet edustavat muuttujien ominaisuuksia. Välilehdellä voidaan määritellä mm. muuttujan nimi, tyyppi ja nimike.

| Until | tled1 [Datas | Get0] - IB | M SPSS Statisti | cs Data Edito | r | | | | | | | |
|-------------|--------------------------|------------|-----------------|-----------------|----------------------------|----------------------|-----------------|-----------------|--------------|--------------|----------------|------|
| <u>File</u> | <u>Edit</u> <u>V</u> iew | Data | Transform | <u>A</u> nalyze | <u>G</u> raphs <u>U</u> ti | lities Add- <u>c</u> | ns <u>W</u> ind | ow <u>H</u> elp | | | | |
| | | | | 2 | | H | | | | | | AB6 |
| | N | ame | Туре | Width | Decimals | Label | Values | Missing | Columns | Align | Measure | Role |
| 1 | s jj | | | | | | | | | | | 4 |
| 2 | | | | | | | | | | | | |
| 3 |) <u> </u> | | | | | | | | | | | |
| 4 | | | | | | | | | | | | |
| 5 | | | | | | | | | | | | |
| 6 | | | | | | | | | | | | |
| 7 | a []. | | | | | | | | | | | |
| 8 | | | | | | | | | | | | |
| 9 | | | | | | | | | | | | _ |
| 36 | 4 | | | | | | | | | | | 4 |
| Data V | /iew Varia | able Vie | w | | | | | | | | | |
| | | | | | | | | IE | BM SPSS Stat | istics Proce | essor is ready | |

Kuvio 7. Variable View -välilehti

Name-sarakkeessa voidaan muuttujalle kirjoittaa nimi (oletuksena VAR00001 jne.). Muuttujan nimi saa olla korkeintaan 64 merkkiä pitkä. Nimi kannattaa aloittaa kirjaimella a-z. Muuten nimessä voi käyttää kirjaimia a-z, numeroita sekä merkkejä _, # ja @ . Muita merkkejä joko ei saa käyttää lainkaan (esim. välilyönti) tai sitten kannattaa välttää (esim. piste, å,ä,ö). Ison ja pienen kirjaimen välillä ei ole eroa. Muuttujaniminä ei saa käyttää sanoja ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO ja WITH, koska ne ovat ohjelman ns. varattuja sanoja. Kahta samannimistä muuttujaa ei voi määritellä.

Type-sarakkeessa voidaan määrätä muuttujan esitystyyppi. Näpäyttämällä solun oikean reunan värillistä

painiketta **e** avautuu muuttujatyypin määrittelyikkuna (kuvio 8).

| o <u>N</u> umeric | | |
|---------------------|------------------------|---|
| © <u>C</u> omma | <u>W</u> idth: | 8 |
| © <u>D</u> ot | Decimal Places | 0 |
| Scientific notation | Decimar <u>naces</u> . | 2 |
| © D <u>a</u> te | | |
| 🛇 Do <u>l</u> lar | | |
| ◯ Custom currency | | |
| 🛇 String | | |

Kuvio 8. Muuttujatyypin määrittelyikkunan yläosa.

Vaihtoehdoista

Numeric on numeerinen muuttuja, joka on oletusarvo. Kaikki sellainen tieto, joka on mahdollista esittää reaalilukuina, kannattaa esittää tässä muodossa. Arvojen esittämiseen ja tallettamiseen tarvittavaa merkkimäärää voidaan muuttaa valinnoilla/sarakkeilla Width (oletus 8, maksimi 40) ja Decimal Places/Decimals (oletus 2, maksimi 16). Merkkimäärään lasketaan mukaan myös mahdolliset etumerkit ja desimaalierotin. Desimaalierotin on yleensä pilkku (,) (tai piste (.) riippuen käyttäjätilin kieli- ja aluemäärityksistä).

Comma on numeerinen muuttuja, jossa tuhannet erotetaan pilkuilla ja desimaalierottimena on piste.

Dot on numeerinen muuttuja, jossa tuhannet erotetaan pisteillä ja desimaalierottimena on pilkku.

Scientific notation on numeerinen muuttuja, joka esitetään eksponenttimuotoisena E-lisäyksen ja potenssin avulla, esim. 123 = 1,23E+002 ja 0,004 = 4,00E-003

- Date on numeerinen muuttuja, joka voi sisältää kalenteri- ja kellonaikatietoja useassa eri muodossa.
- **Dollar** on numeerinen muuttuja, johon on lisätty dollarimerkki (\$) ja desimaalierottimena on piste (.). Esitysmuodon voi valita useasta vaihtoehdosta.
- Custom currency on numeerinen valuuttamuuttuja, jonka laadun saa määrättyä itse (esim. €). Laadun merkinnän määrittäminen tehdään Edit-valikon Options-valinnan kortissa Currency.
- String on alfanumeerinen eli merkkijonomuuttuja, jonka pituuden saa määrättyä haluamakseen valinnalla Characters. Pienet ja isot kirjaimet tarkoittavat muuttuja-arvoina eri merkkejä.
- **Restricted Numeric** on kokonaislukuarvoinen muuttuja siten, että ensimmäiset merkit vasemmalta ovat nollia.

Width- ja Decimals-sarakkeissa voidaan määrittää muuttuja-arvojen merkkimäärät, jos niitä ei antanut jo esitystyyppimäärittelyn yhteydessä.

Label-sarakkeessa voidaan muuttujalle antaa selväkielinen kuvaus eli nimike. Nimike voi olla 256 merkkiä pitkä, ja siinä ei ole rajoituksia käytettävistä merkeistä. Jos nimikkeet annetaan, tulevat ne näkyviin tulostuksissa haluttaessa, mikä helpottaa tulosten tulkintaa.

Values-sarakkeessa voidaan muuttujan arvoille antaa nimike eli selitys, joka helpottaa usein laadullisten muuttujien käyttöä ja tulosten tulkintaa. Näpäyttämällä solun oikean reunan värillistä painiketta avautuu määrittelyikkuna (kuvio 9), jossa voidaan määrittää muuttuja-arvojen nimikkeet

| -Value Labels Val <u>u</u> e: 2 Label: nainen | TEKEILLÄ OLEVA SELITYS | Spelling |
|---|--------------------------------|----------|
| Add Change Remove | 'mies" JO TEHDYT SELITYKSET | |
| | OK Cancel Help | |

Kuvio 9. Muuttuja-arvojen nimikkeiden määrittelyikkunan yläosa

Muuttuja-arvo kirjoitetaan **Value**-kohtaan ja sitä vastaava selitys eli nimike **Label**-kohtaan. Muuttujaarvon nimike voi olla 120 merkkiä pitkä. Tulosikkunassa nimikkeiden isot kirjaimet säilyvät isoina ja pienet pieninä. Määrittelyikkunan painikkeista

Add lisää muuttuja-arvonimikkeen muiden ko. muuttujan arvonimikkeiden listaan

- Change muuttaa valitun muuttuja-arvonimikkeen
- **Remove** poistaa valitun muuttuja-arvonimikkeen listasta.

Spelling voidaan tarkistaa esim. englanninkielinen oikeinkirjoitus.

Määritellyt muuttuja-arvojen nimikkeet saadaan näkyviin lukujen sijasta aineisto-ikkunaan, kun View-

valikosta valitaan Value Labels tai työkalupainikepalkista valitaan painike

Systeeminmäärittelemäksi puuttuvaksi tiedoksi (system-missing values) tulevat kaikki numeromuotoisten muuttujien tyhjät havaintosolut, johon SPSS automaattisesti lisää pisteen puuttuvan tiedon merkiksi. Jos puuttuva tieto halutaan merkitä erityistä koodia käyttäen, voidaan tällainen oma koodi (user-missing values) määrittää Missing-sarakkeessa. (Erityinen puuttuvan arvon koodi (esim. E) kannattaa määrittää ainakin String-tyyppiselle muuttujalle, ja kun havaintoaineistoa syötetään, laitetaan puuttuvan tiedon paikalle ko. koodi (esim. E). Mielipideväittämissä esiintyy usein yhtenä vastausvaihtoehtona "ei osaa sanoa". Jos tällä vastausvaihtoehdolla on oma koodinsa, ja ko. vastauksen antaneita tilastoyksiköitä ei haluta mukaan analysointiin, voidaan tällöinkin koodiarvo määrittää puuttuvaksi tiedoksi.) Näpäyttämällä solun oikean reunan painiketta avautuu puuttuvien havaintojen määrittelyikkuna (kuvio 10).

| Jes |
|---------|
| |
| |
| High |
| <u></u> |
| |

Kuvio 10. Puuttuvien havaintojen määrittelyikkuna

Avautuvassa ikkunassa valinnat ovat

No missing valueskäyttäjä ei määrittele itse puuttuvan tiedon merkintää (oletusarvo)Discrete missing valueskäyttäjä valitsee muuttujalle enintään kolme erillistä puuttuvan tiedon
koodiaRange plus of ...käyttäjä valitsee muuttujalle puuttuvan tiedon arvoiksi reaalilukuvälin
(Low-High) sekä halutessaan vielä yhden erillisen muuttuja-arvon
(Discrete value)

Columns-sarakkeessa voidaan määrätä aineistoikkunan sarakkeen leveys merkkien lukumäärän avulla. Sarakkeen leveyden määrittely vaikuttaa vain siihen, millaisena muuttuja näkyy aineistoikkunassa. **Align**-sarakkeessa voidaan määrittää muuttuja-arvojen sijoittelu aineistoikkunaan (**Left/Right/Center**). Sijoittelu vaikuttaa vain aineistoikkunan ulkoasuun.

Measure-sarakkeessa voidaan valita alasvetovalikon avulla muuttujan mitta-asteikko

| Scale | välimatka- (eli intervalli-) tai suhdeasteikoksi, joka näkyy mm. muuttujalistassa muuttujan |
|---------|---|
| | nimen yhteydessä merkkinä 🦑 (numeerinen) tai 🔏 (päivämäärä/kellonaikatietoa) |
| Ordinal | järjestys- eli ordinaaliasteikoksi, merkkinä 🚅 (numeerinen) tai 🛛 📠 (merkkijono) |
| Nominal | luokittelu- eli nominaaliasteikoksi, merkkinä joko 💑 (numeerinen) tai 💑 (merkkijono) |

HUOM. Välimatka- tai suhdeasteikollinen muuttuja ei voi olla merkkijonomuuttuja. Mitta-asteikolla on merkitystä mm. kuvioiden muodostamisessa.

Role-sarakkeessa voidaan valita muuttujan rooli tilastoanalyyseissä. Perustilastoaineistoissa tällä valinnalla ei ole merkitystä, joten oletusarvoa **Input** ei kannata vaihtaa.

Luokittelu- ja järjestysasteikollisten muuttujien muuttuja-arvojen määrittelyä ja jo syötetyn aineiston tarkistuksia voidaan suorittaa myös **Data**-valikon valinnalla **Define Variable Properties**. Avautuvassa määrittelyikkunassa valitaan ruutuun **Variables to Scan** ne muuttujat, joiden arvomäärittelyjä halutaan tehdä, tarkistaa ja/tai korjata. Sen jälkeen painetaan **Continue**-painiketta, jonka jälkeen avautuu muuttujamäärittelyjen tarkistusikkuna (kuvio 11).

| Define Variable Properties | | 14 A.S | | 1 m m | × |
|--|-----------------------------------|---------------|--------------------|--------------------------|----------------------------------|
| Scanned Variable List | Current Variable: | sukupuol | Label: | Sukupuoli | |
| UnMeRole Variable | Measurement Level: | 💑 Nomi 💌 | Suggest | Type: Numeric | T |
| | Rol <u>e</u> : | 🔪 Input 🔍 | | Width: 1 | Decimals: 0 |
| Sukupuol edistys | Unlabeled values: | 0 | | Att | ri <u>b</u> utes |
| | Value Lahel orid: | Enter or edit | lahels in the | e orid. You can enter ad | ditional values at the holtom |
| | Changed | Missing | Count | Value | Label |
| | 1 | | 51 | 1 | mies |
| | 3 | | 2/1 | 2 | nainen |
| 1 Cases scanned: 322 Value list limit: 200 | Copy Properties From Another V | fariable | To Other V Help | /ariables) | beled Values Automatic Labels |

Kuvio 11. Muuttujamäärittelyjen tarkistusikkuna

Tämä tehtävä käy automaattisesti läpi kaikki havainnot muuttujista ja esittää aktiiviseksi valitun muuttujan (nyt *sukupuol*) määrittelytiedot ja frekvenssijakauman tarkistusikkunaan. Tarkistusikkunassa näkyvät mm. ne muuttujan arvot, joille ei ole määritelty muuttujan arvonimikettä. Uudet muuttujaarvonimikkeet voidaan määritellä, ja vanhoja voidaan korjata Label-sarakkeessa. Nimiketekstejä voidaan kopioida (Copy Properties) toisesta muuttujasta (From Another Variable) tai toisiin muuttujiin (To Other Variables). Myös muita muutoksia voidaan tehdä.

Muuttujamäärittelyjä voidaan kopioida muille muuttujille myös Variable View -kortissa. Valitaan näpäyttämällä hiiren oikeanpuoleisella painikkeella se solu, jossa on tarvittava määrittely, ja valitaan avautuvasta valikosta Copy, sitten näpäytetään hiiren oikeanpuoleisella painiketta sen solun kohdalla, johon määrittelyä halutaan soveltaa ja valitaan Paste. Jos yhden muuttujan kaikki määrittelyt sopivat

sellaisenaan toiselle muuttujalle, voidaan kaikki määrittelyt siirtää kerralla. Tällöin näpäytetään yhden solun sijasta määritellyn muuttujan rivinumeroa, mutta muuten toimitaan kuten edellä.

Muuttujamäärittelyt voidaan kopioida myös esim. toisesta havaintoaineistosta ohjatun toiminnon avulla **Data**-valikon valinnalla **Copy Data Properties**.

2.2. Havaintoaineiston syöttäminen

Aineistoikkunan **Data View** -välilehteen voidaan syöttää havaintoaineisto. Taulukon yksi rivi vastaa yhtä tilastoyksikköä, yksi sarake yhtä muuttujaa. Kussakin solussa on yhden tilastoyksikön yhden muuttujan arvo. Kortissa voi kirjoittaa ja muokata havaintoarvoja, mutta mm. kaavoja ei voi suoraan kirjoittaa soluun tai solueditoriin (vrt. Excel).

Aineiston syöttäminen on helppoa: ensin valitaan solu ja sen jälkeen kirjoitetaan muuttuja-arvo. Muuttuja-arvo näkyy solussa ja solueditorissa, ja valitun solun rivinumero ja valitun muuttujan nimi näkyvät ikkunan vasemmassa ylälaidassa. Painetaan vielä **Enter**- tai **Tab**-näppäintä tai osoitetaan seuraava solu, jolloin muuttujan arvo tallentuu soluun. **Enter**-näppäin aktivoi saman sarakkeen seuraavan rivin solun ja **Tab**-näppäimellä siirrytään määriteltyjen muuttujien osalta aina seuraavaan soluun.

Numeromuotoisten muuttujien kohdalla tyhjäksi jätetyt solut tulkitaan puuttuviksi tiedoiksi. Näihin soluihin SPSS automaattisesti lisää pisteen (.). Merkkijono- eli **String**-tyyppisen muuttujan kohdalla tyhjäksi jätettyä solua SPSS ei tulkitse puuttuvaksi tiedoksi, vaan tyhjät solut SPSS tulkitsee välilyönneiksi, jotka ovat kuitenkin merkkejä muiden merkkien joukossa. Tällöin siis kannattaa määrittää erityinen koodi puuttuvan tiedon merkiksi.

Jos syötettävässä muuttujan arvossa on tyyppiin sopimattomia merkkejä, ei arvo tallennu. Numeerisen muuttujan liian pitkät lukuarvot tallentuvat annetussa muodossa, mutta solussa näkyy usein tässä tilanteessa luvun eksponenttiesitysmuoto, vaikka solueditorissa näkyvä esitys on sama kuin millä luku tuli syötettyä. Kokonaislukumuuttujaan tallennettavat desimaaliluvut pyöristyvät lähimpään kokonaislukuun soluesityksessä, mutta solueditorissa näkyy desimaaliesitys. Huomaa, että tietyt tilastoproseduurit käyttävät juuri näitä solussa näkyviä arvoja, ja toiset proseduurit taas käyttävät aitoja muuttuja-arvoja.

Jos muuttuja-arvoille on määrätty nimikkeet, ja ne ovat näkyvissä aineistoikkunassa, voidaan solun havaintoarvo valita solun oikean reunan alasvetovalikon avulla.

Tallennettua aineistoa voidaan muokata monella eri tavalla aineistoikkunassa tai muunnosproseduureilla. Aineistoikkunassa voidaan muuttujan arvoa vaihtaa näpäyttämällä korjattavaa solua, jolloin korvattava arvo näkyy solueditorissa. Kirjoitetaan uusi arvo ja painetaan **Enter**-näppäintä. Yhden solun arvon voi kopioida tai siirtää muihin soluihin, yhden rivin arvoja voidaan siirtää muille tilastoyksiköille, yhden sarakkeen arvot voidaan kopioida muille muuttujille sekä alueen arvot toiselle alueelle. Näissä toimenpiteissä käytetään apuna leikepöytää: merkitään siirrettävä/kopioitava alue, valitaan **Edit**–valikosta **Cut** tai **Copy**, merkitään alue, johon kopioidaan ja valitaan **Edit**–valikosta **Paste.** Arvot näkyvät uudella alueella, jos muuttujien tyypit vastaavat toisiaan tai muuntaminen tyypistä toiseen on mahdollinen, muutoin tuloksena on puuttuva tieto.

Jos tilastoyksiköiden/muuttujien väliin halutaan lisätä uusia, merkitään ensin rivi/sarake, jonka ylä-/vasemmalle puolelle uusin tilastoyksikkö/muuttuja halutaan lisätä. Sen jälkeen valitaan **Edit**-valikosta **Insert Cases** (työkalu) tai **Insert Variable** (työkalu). Uuden rivin/sarakkeen kaikissa soluissa on puuttuvaa tietoa.

Tilastoyksikön tai muuttujan saa poistettua näpäyttämällä sen rivinumeroa tai nimeä ja painamalla **Delete**-näppäintä tai valitsemalla **Edit**-valikosta **Clear**.

2.3. Aineistotiedoston tallettaminen

Syötetty aineisto ja aineistoon tehdyt muutokset säilyvät vain SPSS-istunnon ajan, joten myöhempää käyttöä varten ne täytyy muistaa tallettaa tiedostoon. Kun aineistoikkuna on aktiivisena, valitaan ensimmäisellä talletuskerralla valikosta File joko Save Data (työkalu 🕒) tai Save As... tai Save All Data.

Avautuvassa ruudussa tiedoston nimi kirjoitetaan **File name** -ruutuun. Talletusmuoto valitaan **Save as type:** -luettelosta. Mahdollisia talletusmuotoja on lukuisia, mm.

- SPSS Statistics (*.sav) on SPSS-muoto, joka on myös oletusarvo. Tässä muodossa talletettuja aineistoja voi käyttää SPSS/SPSS-ohjelmaversiosta 7.5 lähtien
- SPSS 7.0 (*.sav) on SPSS 7.0 for Windows-muoto, jota voi käyttää aiemmissa SPSS-ohjelmaversioissa
- **Portable** (*.por) on SPSS-siirtomuoto. Tässä muodossa talletettua aineistoa voidaan käyttää muissa käyttöjärjestelmissä (esim. Macintosh ja Unix)
- **Tab delimited** (*.dat) on ASCII-tekstiä, jossa muuttuja-arvot on erotettu tabulaattorilla (sarkaimella) ja muuttujien nimet saadaan halutessa talletettua
- Excel 2007 through 2010 (*.xlsx) on Excel 97 -työkirja. Excel-taulukkoon voidaan tallettaa muuttujien nimet ja muuttuja-arvonimikkeet numerokoodien sijasta. Muuttujista tallettuu korkeintaan ensimmäiset 16000. Jos havaintorivien määrä ylittää Excel-taulukon maksimimäärän, muodostuu Excel-tiedostoon automaattisesti useita Excel-taulukoita.

Talletettaessa voidaan **Variables**-optiolla valita ne muuttujat, jotka aineistoon halutaan tallettaa. Oletusarvoisesti tiedostoon tallettuu aineistoikkunan kaikki muuttujat.

Myöhemmillä talletuskerroilla, jos talletusmuotoa ja aineiston nimeä ei haluta muuttaa, riittää talletusvalinnaksi File-valikon Save Data tai työkalu

2.4. Havaintoaineiston avaaminen ja lukeminen

SPSS-muodossa talletetun havaintoaineiston saa avautumaan SPSS:n aineistoikkunaan, kun valitaan työkalu 🔁 tai File-valikosta Open ja sen alivalikosta Data ja vielä kaksoisnäpäytetään aineiston nimeä.

Jos havaintoaineisto on muodostettu tekstinkäsittely, taulukkolaskenta- tai tilasto-ohjelmalla, voidaan se yleensä avata tai lukea SPSS:n aineistoikkunaan. Avaamis- ja lukemistapa riippuu aineiston tallennus- ja talletusmuodosta. Kun aineisto avataan tai luetaan ulkoisesta lähteestä, tehdään muuttujien määrittelyt ja tarkennukset vasta sen jälkeen, kun aineisto on näkyvissä **Data Editor**-ikkunassa.

2.4.1. Excel-taulukon avaaminen

Jos havaintoaineisto on muodostettu Excel-ohjelmalla (talletusmuoto tyyppiä .xls, .xlsx, .xlsm), voidaan aineisto suoraan avata SPSS:n aineistoikkunaan valitsemalla SPSS-ohjelmassa työkalu itai Filevalikosta **Open** ja sen alivalikosta **Data.** Tiedoston tyypin muodoksi vaihdetaan **Excel** ja edelleen kaksoisnäpäytetään aineiston nimeä sopivasta alihakemistosta.

Jos luettavan Excel-tiedoston ensimmäiselle riville on kirjoitettu muuttujien nimet, valitaan avautuvasta **Opening Excel Data Source** -ikkunasta (kuvio 12) kohta **Read variable names from the first row of data**. Samassa ikkunassa myös määritetään, miltä Excel-työkirjan taulukosta (**Worksheet**-

alasvetovalikko) aineisto luetaan. Jos aineistoksi halutaan valita vain tietyt sarakkeet (= muuttujat) ja/tai rivit (= tilastoyksiköt), voidaan käytettävät rivit ja sarakkeet määrittää **Range**-kohtaan.

| E:\STAT2070 | SPSS luentoihin liittyvää\Kulutustilasto_excel.x | ls |
|-------------|--|----|
| Read varia | ble names from the first row of data | |
| Worksheet: | Kulutustilasto_excel [A1:128] | |
| Range: | 1 | |
| | th for string columns: | |

Kuvio 12. Excel-aineistotiedoston avausikkuna

Kun tietoja luetaan Excel-taulukosta, SPSS arvioi ensimmäisen havaintorivin perusteella muuttujien tyypit. Esimerkiksi numeeriset muuttujat SPSS lukee Excel-tiedostosta 11 merkillä, joten aineiston lukemisen jälkeen kannattaa SPSS-tiedoston vielä muokata. Jos ensimmäisellä rivillä on jollakin muuttujalla arvona tyhjä ruutu (=puuttuva tieto), niin ko. muuttujan määrittely täytyy suorittaa Variable View -kortissa. Jos Excel-ohjelmassa annetut muuttujien nimet vastaavat SPSS-ohjelman muuttujanimien käytäntöä, siirtyvät ne suoraan SPSS-ohjelmassa muuttujanimiksi (Name), muutoin Excel-nimistä tulee SPSS-ohjelmaan muuttujien selitykset (Label).

2.4.2. Tekstitiedoston lukeminen

Tilastollista aineistoa esittävä tekstitiedosto voi olla muodoltaan kiinteäkenttäinen (fixed width) tai sellainen, jossa muuttuja-arvot on erotettu toisistaan jollain merkillä (delimited).

Kiinteäkenttäisessä aineistossa jokaiselle tiedolle (eli muuttujan arvolle) on varattu määrätty määrä merkkejä. Jos muuttujan arvo on lyhyempi kuin varattujen merkkien määrä, käytetään välilyöntiä täyttämään "tyhjä osuus". Yhden tilastoyksikön muuttujan arvot voivat olla sijoitettu usealle perättäiselle riville. Usein numeerisen muuttujan arvot ovat oikealta tasattuja (tällöin "tyhjä osuus" on ennen numeroita) ja merkkijonomuuttujan arvot vasemmalta tasattuja ("tyhjä osuus on merkkien jälkeen"). Puuttuva tieto on esitetty välilyöntien avulla. Aineistotiedostoa kannattaa tutkiskella jollain tekstinkäsittelyohjelmalla niin, että kirjoitusfonttina on sellainen, jossa merkit ovat yhtä leveitä (esim. Courier). Tällöin aineiston rakenne on helpompi hahmottaa. Aineisto voi olla esim. tämän näköinen

| 11 | Iikko | 3 | 2,345 | 9 | 234,232 |
|----|-------|---|--------|----|---------|
| 21 | Kirsi | 4 | 12,234 | 12 | 9 |
| 3r | natti | 2 | 121,23 | 10 | 222,12 |
| 4 | Jörö | 4 | 12 | 12 | 222,24 |
| 51 | Marja | 1 | | 8 | 22,2 |
| б | Make | | 5,2 | | 23,4 |
| 7 | Ida | 3 | 12,2 | 7 | 234 |

Jos aineisto on sellainen, että **muuttuja-arvot on erotettu** toisistaan jollain **merkillä**, on aineistossa muuttuja-arvojen välissä on aina yksi erotinmerkki, joka ei (yleensä) saa muuten esiintyä aineistossa. Erotinmerkkinä voi olla esim. tabulaattori, puolipiste, välilyönti tai/ja pilkku. Pilkku on huono vaihtoehto

varsinkin silloin, jos desimaalierottimen on pilkku! Jos erotinmerkki voi olla osa muuttuja-arvoa, on erotinmerkki kirjoitettava esim. heittomerkkien sisään. Samalla rivillä voi olla usean tilastoyksikön arvoja tai saman tilastoyksikön arvoja voidaan sijoittaa usealle perättäiselle riville, muuttujien järjestyksen eri tilastoyksiköillä täytyy kuitenkin olla sama. Puuttuvaa tietoa ei esitetä välilyöntien avulla, vaan se on esitettävä jollakin sopivalla merkillä, esim. pilkulla. Esimerkkiaineistomme alkuosa voisi vapaamuotoisena olla vaikkapa seuraavanlainen

```
1 Mikko 3 2,345 9 234,232 2 Kirsi 4 12,234 12 9 3 matti 2 121,23 10 222,12 4 Jörö 4 12 12 222,24
```

Tekstinkäsittelyohjelmassa pitää havaintoaineisto tallettaa ASCII-muotoisena tekstitiedostona, jotta aineisto voidaan lukea SPSS-ohjelmaan, esimerkiksi **Word**-ohjelmassa talletusmuoto on **Plain text**.

Tekstitiedostoksi talletettu havaintoaineisto avataan File-valikon Read Text Data-valinnalla. Avattavan tiedoston tyyppi on oletusarvoisesti Text. Aineiston nimeä kaksoisnäpäytetään sopivasta alihakemistosta, jolloin avautuu ensimmäinen Tekstivelho-ikkuna (Text Import Wizard) (kuvio 13). Ensimmäisessä vaiheessa voidaan valita sovellettavaksi jo aiemmin määritelty lukuformaatti valitsemalla Yes kohdassa Does your text file match a predefined format? Jos valmista lukuformaattia ei ole, valitaan No ja näpäytetään Next-painiketta ikkunan alareunasta.

Ruudun alareunan kohdassa **Text file** on esikatselutaulukko havaintoaineiston esitysmuodosta. Jokaisen Tekstivelho-vaiheen jälkeen kannattaa katsoa ruudun alareunasta, millaiseksi SPSS-aineisto on muodostumassa.

| 8 840 1 2400 1 1020 | 81 28.5 0 73 40.33 | Welcome to the text import wizard! | 1 1 5 1 |
|---------------------------|--|---|-------------------|
| 3 870 0 5 1740 | 933117 18341.91 | specify information about the variables. | our text file and |
| 1 | wart war2 war3 w | Does your text file match a predefined for | ormat? |
| 2 | | © <u>Y</u> es | Browse |
| 3 | | (No | 1 |
| 4 | | | |
| - | and a second sec | | |
| 10 | | ESIKATSELU | |
| xt file | E:\STAT2070 SPSS I | ESIKATSELU uentoihin liittyvää\pelkkatekstipilkuilla.txt 20 30 40 3 | 50 . |
| ext file | E:\STAT2070 SPSS I 010 1Mikko 3 2,345 2Kirsi 4 12,234 | ESIKATSELU uentoihin liittyvää\pelkkatekstipilkuilla.txt 20 | 50 . 60. |
| at file | E:\STAT2070 SPSS I 010 1Mikko 3 2,345 2Kirsi 4 12,234 2 3matti 2 121.23 | ESIKATSELU uentoihin liittyvää\pelkkatekstipilkuilla.txt 20 | 50 1 |
| 1 2 3 4 | E:\STAT2070 SPSS I 010 1Mikko 3 2,345 2Kirsi 4 12,234 3matti 2 121,233 4 Jörö 4 12 | ESIKATSELU uentoihin liittyvää\pelkkatekstipilkuilla.txt 20 | 50 . 60. |
| 1 2 3 4 5 | E:\STAT2070 SPSS I 010 1Mikko 3 2,345 2Kirsi 4 12,234 : 3matti 2 121,23 : 4 Jörö 4 12 : 5Marja 1 | ESIKATSELU uentoihin liittyvää\pelkkatekstipilkuilla.txt 20 | 50 . 60. |
| 1 2 3 4 5 | E:\STAT2070 SPSS I 010 1Mikko 3 2,345 2Kirsi 4 12,234 : 3matti 2 121,23 : 4 Jörö 4 12 : 5Marja 1 | ESIKATSELU uentoihin liittyvää\pelkkatekstipilkuilla.txt 20 | jo |

Kuvio 13. Tekstivelho, ensimmäinen vaihe

Tekstivelhon toisessa vaiheessa (kuvio 14) valitaan tekstitiedoston muoto. Kohdassa **How are your variables arranged?** valinnat ovat

Delimited, jota käytetään, kun muuttuja-arvot on erotettu toisistaan tietyllä erotinmerkillä **Fixed width**, jota käytetään silloin, kun aineisto on kiinteäkenttäinen

Kohdassa **Are variable names included at the top of your file?** valinnat ovat **Yes**, jota käytetään silloin, kun tekstitiedoston alkuun on tallennettu muuttujien nimet **No**, jota käytetään silloin, kun tekstitiedostossa ei ole muuttujien nimiä

| Delimited Fixed width | - Variables are delimited by a specific character (i.e., comma, tab). - Variables are aligned in fixed width columns. TITIEDOSTON MUODON VALINTA | |
|-----------------------|--|---|
| Are variable nar | nes included at the top of your file? | ٦ |
| O Yes | | |

Kuvio 14. Tekstivelho, toinen vaihe (ikkunan yläosa)

Tekstivelhon kolmannessa vaiheessa (kuvio 15) määritetään aineiston rakenne. Kohdassa **The first case** of data begins on which line number? voidaan määrittää tekstitiedoston sen rivin numero, josta havainnot alkavat.

| e first case of (| data begins on w | hich lin | ne nur | mber? | 1 | * | |
|-------------------------|-------------------|----------|--------|-------|---|---|--|
| ow many <u>l</u> ines i | represent a case | ? | 1 | 4 | | | |
| ow many case | es do you want to | import | t? | | | | |
| All of the c | ases | | | | | | |
| ◎ <u>T</u> he first | 1000 | cas | ses. | | | | |
| | | | | | | | |

Kuvio 15. Tekstivelho, kolmas vaihe (ikkunan yläosa), kiinteäkenttäinen aineisto

Kohdassa How many cases do you want to import? voidaan valita

All of the cases, jolloin tekstitiedoston kaikki havainnot muodostavat SPSS-aineiston,

The first ## cases, jolloin ## ensimmäistä tilastoyksikköä muodostavat SPSS-aineiston,

A percentage of the cases ## %, jolloin tekstitiedoston tilastoyksiköistä valitaan satunnaisesti ## % muodostettavaan SPSS-aineistoon

Jos aineistomuodoksi on on toisessa vaiheessa valittu **Fixed Width** (kiinteäkenttäinen), kohdassa **How many lines represent a case?** kerrotaan, kuinka monta vaakariviä on käytetty yhden tilastoyksikön muuttuja-arvojen esittämiseen.

Jos aineistomuodoksi on valittu **Delimited** (erotinmerkki käytössä), on edellisen valinnan sijasta kysymys **How are your cases represented?**, jonka vastausvaihtoehdot ovat

Each line represents a case, jolloin yksi rivi edustaa yhden tilastoyksikön tietoja,

A specific number of variables represents a case, jolloin määrätään, kuinka monta muuttujaa yksittäiseen tilastoyksikköön liittyy.

Kiinteäkenttäisen aineiston käsittelyssä Tekstivelhon neljännessä vaiheessa (kuvio 16) on ruudun alareunassa havaintoaineiston alustava esitysmuoto, mutta kuvassa on lisäksi SPSS:n sijoittamia pystysuoria viivoja osoittamassa kunkin muuttujan arvojen (mahdollista) alkusaraketta. Näitä viivoja voi siirtää, lisätä ja poistaa. Jos viiva on väärässä paikassa, voidaan se siirtää oikealle kohdalle laahaamalla sitä hiiren avulla. Uuden viivan lisääminen tapahtuu näpäyttämällä haluttua kohtaa. Viivan poistaminen tapahtuu laahaamalla se hiiren avulla kuva-alueen ulkopuolelle. Näitä toimintoja voi suorittaa myös kirjoittamalla **Column Number** –ruutuun sen sarakkeen numeron, johon viivan haluaa lisätä (**Insert Break**) tai poistaa (**Delete Break**)

| Text Import Wizard - Fixed Width Step 4 of 6 |
|---|
| Specify where each variable begins. The first column is column 0. |
| To INSERT a variable break line, click at the desired position in the ruler or data area. Alternatively, move to the position using the arrow keys or by typing the column number; then press the Insert Break button. |
| To MOVE a variable break line, drag it to the new position. |
| To DELETE a variable break line, select it or type its position. Then press the Delete key or the Delete Break button. |
| Buler: 10 20 30 40 50 60 1 1Mikko 3 2,345 9 234,232 2 2Kirsi 4 12,234 12 9 3 4 12 12 222,12 4 Jörö 4 12 12 222,24 |
| Column Number: 1 Insert Break Delete Break Current Variable Width: 1 |
| < Back Next > Finish Cancel Help |

Kuvio 16. Tekstivelho, neljäs vaihe, kiinteäkenttäinen aineisto

Jos aineistomuodoksi on valittu **Delimited**, Tekstivelhon neljännessä vaiheessa (kuvio 17) kerrotaan tekstitiedoston erotinmerkit kohdassa **Which delimiters appear between variables?** Vaihtoehdot ovat

| Tab | tabulaattori | Space | välilyönti |
|-------|------------------|-----------|------------|
| Comma | pilkku | Semicolon | puolipiste |
| Other | jokin muu merkki | | |

Kohdassa **What is the text qualifier?** määritetään se merkkipari, jonka väliin jäävä mahdollinen erotinmerkki tulkitaan muuttuja-arvoon sisältyväksi merkiksi.

| Vhich delimiters | appear between variables? | What is the text qualifier? |
|------------------|---------------------------|-----------------------------|
| 🗸 <u>T</u> ab | Space | One |
| Comma | Semicolon | Single guote |
| Othe <u>r</u> : | | ◎ <u>D</u> ouble quote |
| | | © Ot <u>h</u> er: |
| | | |

Kuvio 17. Tekstivelho, neljäs vaihe (ikkunan yläosa), aineistomuoto Delimited

Tekstivelhon viidennessä vaiheessa (kuvio 18) voidaan yksittäistä muuttujaa tarkastella näpäyttämällä esikatseluruudussa (**Data preview**) sen sarakeotsikkoa. Kohtaan **Variable name** voidaan muuttujalle kirjoittaa haluttu nimi ja kohdassa **Data format** valitaan muuttujan tyyppi. Muuttujien tyypit kannattaa tarkistaa, koska ohjelma tunnistaa ne joskus eri tavalla kuin käyttäjä on tarkoittanut. Jos muuttujaa ei haluta mukaan muodostuvaan SPSS-aineistoon, voidaan **Data format** -kohdasta valita **Do Not Import**.

| - | | 200 | | | | |
|---------------------|--|-----------------------------|---|--------------------------------------|--|---|
| <u>V</u> aria V1 | able name: | M | UUTTUJAN NIN | MUUTTUJA | : V1 VALITTU, N VAIHTAA TÄSS | Ä |
| Data Nun | ı format: neric | _ | ENSIMM TYYPPI:N TARKIST/ | ÄISEN MUU IUMERIC. A, ETTÄ TYY | JTTUJAN V1 VPPI ON SOPIVA | |
| ata prev | view | | ESIKATSELU | | | 5 |
| ata prev | view | 110 | ESIKATSELU | 1.00 | | 7 |
| ata prev V1 | view V2 Mikke | V3 | ESIKATSELU | V5 | V6 | |
| ata prev V1 | view V2 Mikko Kirci | V3 3 | ESIKATSELU V4 2,345 | V5 9 | V6 234,232 | |
| ata prev /1 | View V2 Mikko Kirsi matti | V3 3 4 2 | ESIKATSELU V4 2,345 12,234 121 23 | V5 9 12 | V6 234,232 9 222,12 | |
| ata prev V1 | view V2 Mikko Kirsi matti Jörö | V3 3 4 2 4 | ESIKATSELU 2,345 12,234 121,23 12 | V5 9 12 10 12 | V6 234,232 9 222,12 222,24 | |
| ata prev /1 | view V2 Mikko Kirsi matti Jörö Maria | V3 3 4 2 4 1 | ESIKATSELU 2,345 12,234 121,23 12 | V5 9 12 10 12 8 | V6 234,232 9 222,12 222,24 22.2 | |

Kuvio 18. Tekstivelho, viides vaihe

Tekstivelhon kuudennessa vaiheessa (kuvio 19) voidaan muodostettu tekstitiedoston lukuformaatti tallettaa kohdassa **Would you like to save this file format for future use?** Kohdassa **Would you like to paste the syntax?** voidaan valita, halutaanko edellä suoritetut komennot näkyviin SPSS-komentoeditoriin.

| | at for future use? | like to save this file f | rou | | | | | |
|---------|----------------------|--------------------------|-----|---|------|--------------|------------|---|
| ls. | Save As | | C | - | var3 | var2 | var1 | |
| (CSSND) | (<u>—</u> | | (0 | | 1 | 840 | 628 | 1 |
| | | | | | 0 | 2400 | 638 | 2 |
| | - | like to paste the syn | ΓW | | 8 | 10200 | 632 | 3 |
| | | | 0 | | 0 | 870 | 633 | 4 |
| locally | Cache data lo | | | • | | | | T |
| | ☑ <u>C</u> ache data | like to paste the syn | C C | • | 0 | 10200 870 | 632 633 | 3 |

Kuvio 19. Tekstivelho, kuudes vaihe (ikkunan yläosa)

Aineistoikkunan esitysmuoto ei heti aineiston lukemisen jälkeen ole aina sopiva, esimerkiksi desimaaliluvuista voi näkyä vain kaksi desimaalia. Esitysmuotoa voidaan muokata sopivaksi mm. muuttujien tyyppimäärittelyjen avulla.

2.5. SPSS-aineistojen yhdistäminen

SPSS:ssä voidaan yhdistää kaksi havaintoaineistoa yhdeksi isoksi aineistoksi tilastoyksiköitä tai muuttujia lisäämällä.

Jos havaintoaineistoa laajennetaan lisäämällä tilastoyksiköitä eli jos yhdistettävissä aineistoissa muuttujat ovat (ainakin osittain) samoja, mutta tilastoyksiköt eivät, avataan ensimmäinen yhdistettävistä aineistoista aktiiviseksi SPSS-aineistoikkunaan ja valitaan **Data**-valikosta **Merge Files** ja edelleen **Add Cases.**

Nyt avautuvasta määrittelyruudusta valitaan se havaintoaineisto, joka halutaan liittää aineistoikkunassa avoinna olevan aineiston tilastoyksiköiden perään, jolloin edelleen avautuu **Add Cases from** -ikkuna (kuvio 20).

Kohdassa Variables in New Active Dataset: näkyvät ne muuttujat, joiden nimet ja tyypit ovat samanlaisia yhdistettävissä aineistoissa. Jos jotakin näistä muuttujista ei haluta uuteen aineistoon, voi muuttujan poistaa kaksoisnäpäytyksellä. Kohdassa Unpaired Variables ovat ne muuttujat, joiden nimet ja/tai tyypit eivät vastaa toisiaan. Muuttujan nimen perässä oleva merkki kertoo, kummassa aineistossa muuttuja esiintyy: *-merkki viittaa aineistoikkunan aktiiviseen aineistoon ja +-merkki lisättävään aineistoon. Jos erinimisistä muuttujista halutaan muodostaa yksi muuttuja, näpäytetään ensimmäistä muuttujaa ja sen jälkeen painetaan Ctrl-painiketta ja näpäytetään samanaikaisesti toista muuttujaa ja vielä näpäytetään Pair-painiketta. Unpaired Variables -kohdassa olevan muuttujan voi myös nimetä uudelleen Rename-valinnalla. Indicate case source as variable -valinnalla voidaan aineistoon lisätä muuttuja, joka kertoo, kummassa aineistossa tilastoyksikkö alun perin oli.

| npaired Variables: | Variables in New Active Dataset: | Unpaired Variables: | Variables in New Active Dataset: |
|-------------------------|-----------------------------------|-------------------------------|-----------------------------------|
| tunimi>(*) | | | sukunimi & Surname> |
| ioe1(*) | | | etunimi & Firstname> |
| loe2(*) | (| | Koe1 & Exam1 |
| ukunimi>(*) | ↔ | | Koe2 & Exam2 |
| ukupuoli(*) | | | sukupuoli & gender |
| ge(+) | Pair | | Pair Age(+) |
| xam1(+) | | | |
| xam2(+) | | | |
| irstname>(+) | | | |
| ender(+) | Indicate case source as variable: | | 📝 Indicate case source as variabl |
| Rename | source01 | R <u>e</u> name | source01 |
|)=Active dataset | | (*)=Active dataset | |
| =E1STAT2070 SPSS luento | ihin liittwää\lisättävät tvt sav | (+)=E:\STAT2070 SPSS luentoih | in liittwää\lisättävät tvt.sav |

Kuvio 20. Määrittelyikkuna tilastoyksiköiden lisäämiselle ennen muokkausta ja sen jälkeen

Jos havaintoaineistoa laajennetaan lisäämällä muuttujia eli jos yhdistettävissä aineistoissa tilastoyksiköt ovat samoja, mutta muuttujat eivät, avataan ensimmäinen havaintoaineisto SPSS-aineistoikkunaan ja valitaan sitten **Data**-valikosta **Merge Files** ja edelleen **Add Variables**.

Nyt avautuvasta määrittelyruudusta valitaan aktiiviseen ikkunaan yhdistettävä havaintoaineisto, jolloin avautuu Add Variables from –ikkuna. Ikkunan kohdassa New Active Dataset näkyvät ne muuttujat, joista uusi aineisto on muodostumassa. Kohdassa Excluded Variables ovat ne yhdistettävän aineiston muuttujat, joilla on sama nimi aktiivisena aineiston jonkin muuttujan kanssa. Jos kahdella eri ominaisuudet sisältävällä muuttujalla on sama nimi, valitaan muuttuja Exclude Variables -kohdasta ja annetaan sille uusi nimi Rename-painikeen avulla, jonka jälkeen muuttuja voidaan siirtää yhdistetyn tiedoston muuttujajoukkoon. Jos toisesta aineistosta puuttuu tilastoyksiköitä, valitse muuttuja(t), jonka avulla aineistot voidaan yhdistää. Aineistojen yhdistämisessä apuna käytettävän avainmuuttujan (key variable) perusteella lisättävät muuttujat osoitetaan oikeille tilastoyksiköille. Avainmuuttujan on oltava kummassakin aineistossa, ja yhdistettävien aineistojen on oltava järjestetyt (kts. luku 3.4.) ko. muuttujan suhteen. Tällöin Match cases on key variables in sorted file -vaihtoehdoista valitaan vaihtoehto Both files provide cases.

3. HAVAINTOAINEISTON MUOKKAAMINEN

3.1. Muuttujan arvot lausekkeiden avulla

Jos uuden muuttujan arvot tai jo olemassa olevan muuttujan uudet arvot halutaan määrätä lausekkeella, jossa esiintyy muuttujia, vakioita, funktioita ja laskutoimituksia, valitaan **Transform**-valikosta **Compute Variable**, jolloin avautuu uusi määrittelyikkuna (kuvio 21).



Kuvio 21. Compute Variable -määrittelyikkuna, muodostetaan *lasnaolopros*-muuttuja käyttämällä *RND*-funktiota, muuttujia *osallist* ja *opetus* sekä aritmeettisia laskutoimituksia

Määrittelyikkunan **Target Variable** -ruutuun kirjoitetaan tulosmuuttujan nimi (uusi tai entinen) eli sen muuttujan nimi (esimerkkikuviossa: *lasnaolopros*), johon uusien muuttuja-arvojen halutaan tallentuvan. Muuttujan tyyppi ja nimike voidaan (ja joskus on pakkokin määrätä) näpäyttämällä optionäppäintä **Type&Label**. Muuttujan nimike (**Label**) voidaan kirjoittaa itse tai se voi olla lauseke, jolla uusi muuttuja muodostetaan (**Use expression as label**). Tätä ominaisuutta kannattaa hyödyntää dokumentoinnin takia.

Numeric Expression -ruutuun muodostetaan lauseke kaksoisnäpäyttämällä muuttujat ja funktiot listoistaan sekä valitsemalla esimerkiksi näytön laskinnäppäimistä sopivat merkit ja/tai kirjoittamalla (esimerkkikuviossa: *RND(100*osallis/opetus*). Funktiot on ryhmitelty (**Function group**), esimerkkikuviossa valittu ryhmä *Arithmetic*. Kun valitaan ryhmä, niin siihen kuuluvat funktiot näkyvät kohdassa **Functions and Special Variables**. Kustakin funktiosta saa lyhyen esittelyn näpäyttämällä

funktion nimeä, esimerkkikuviossa valittuna funktiosta *Rnd* versio 1: Rnd(1). Funktion kuvauksessa kerrotaan mm. montako pakollista argumenttia funktion käyttämiseksi tarvitaan (esimerkkikuviossa: 1 argumentti pakollinen) ja minkä tyyppinen tulos on (esim. Numeric/String).

Jos arvot lasketaan vain osalle tilastoyksiköistä, esitetään käytettävä ehto **If**-option avulla (kts. luku 3.5.). Funktioita, muuttujia ja operaattoreita voidaan käyttää myös ehtolausekkeissa. Kun lauseke on muodostettu, saadaan muuttuja-arvot muodostettua näpäyttämällä **OK**-painiketta. Uusi muuttuja arvoineen näkyy aineistoikkunan viimeisenä sarakkeena. Jos korjataan olemassa olevan muuttujan arvoja, varmistaa SPSS vielä ennen muutosta, halutaanko arvot todella vaihtaa (**Change existing variable?**). Aineistoikkunasta kannattaa heti muodostamisen jälkeen tarkistaa, että uudet muuttuja-arvot ovat järkeviä.

Näytön laskinnäppäimet ovat

| + | vhteenlasku | - | vähennyslasku |
|----|------------------------------|----|-------------------------------------|
| * | kertolasku | / | jakolasku |
| ** | potenssiin korotus | 0 | laskutoimitusten järjestys suluilla |
| < | pienempi kuin | > | suurempi kuin |
| <= | pienempi tai yhtä suuri kuin | >= | suurempi tai yhtä suuri kuin |
| = | yhtä suuri kuin | ~= | erisuuri kuin |
| & | ja | | tai |
| ~ | ei | • | desimaalierotin (piste) |
| | | | |

Tilastollisia funktioita (Function group: Statistical) ovat mm.

| , |
|---|
| variaatiokerroin (ainakin kaksi numeerista argumenttia) |
| maksimi (kahdesta tai useammasta arvosta) |
| keskiarvo ainakin kaksi numeerista argumenttia) |
| minimi (kahdesta tai useammasta arvosta) |
| keskihajonta (ainakin kaksi numeerista argumenttia) |
| summa (ainakin kaksi numeerista argumenttia) |
| varianssi (ainakin kaksi numeerista argumenttia) |
| |

Kukin tilastollinen funktio vaatii siis vähintään kaksi argumenttia, jotka voivat olla numeerisia muuttujia, vakioita tai lausekkeita. **Laskutoimitukset suoritetaan tilastoyksiköittäin** kuten muitakin funktioita käytettäessä. Funktioiden argumentit erotetaan toisistaan pilkulla. Siksi desimaalierotin lausekkeita kirjoitettaessa on aina piste (.).

Huom! Tilastolliset funktiot antavat tilastoyksikölle jonkin numeerisen tuloksen, jos edes yhdellä argumentilla on arvo. Aritmeettisia laskutoimituksia suoritettaessa tilastoyksikkö saa tulosarvon vain silloin, jos kaikilla argumenteilla on arvo. Jos esimerkiksi halutaan laskea tilastoyksiköittäin yhteen muuttujien x1, x2 ja x3 arvot, niin yhteenlasku x1+x2+x3 ja funktio SUM(x1, x2, x3) voivat tuottaa eri tuloksen.

Aritmeettisia funktioita (Function group:Arithmetic) ovat mm.

| ABS(numexpr) | itseisarvo |
|----------------------------|---|
| COS(radians) | kosini, argumentti radiaaneina |
| EXP(numexpr) | e:n potenssi |
| LG10(numexpr) | 10-kantainen logaritmi |
| LN(numexpr) | luonnollinen logaritmi |
| MOD(numexpr, modulus) | jakojäännös |
| RND | pyöristäminen |
| RND (numexpr) | pyöristäminen kokonaisluvuksi; esim. RND(23.54) = 24 |
| RND (numexpr, mult) | pyöristäminen haluttuun tarkkuuteen; esim. $RND(23.54, 0.1) = 23.5$ |
| SIN(radians) | sini, argumentti radiaaneina |

| SQRT(numexpr) | neliöjuuri |
|----------------|---|
| TRUNC(numexpr) | katkaisu kokonaisluvuksi; esim. TRUNC(23.54) = 23 |

Satunnaisluku- tai jakaumafunktioita löytyy funktioryhmistä CDF & Noncentral CDF, Inverse DF, PDF & Noncentral P, Significance ja Randon Numbers. Esim. funktiolla RV.NORMAL(keskiarvo, keskihajonta) saadaan satunnaislukuja normaalijakaumasta, jonka keskiarvo ja keskihajonta annetaan argumentteina.

Ajan (kalenteritiedon ja kellonaikatiedon) käsittelyyn liittyviä funktioita löytyy funktioryhmistä **Current Date/Time, Date Arithmetic, Date Creation, Date Extraction, Time Duration Creation** ja **Time Duration Extractions**. Esim. kahden päivämäärän välinen päivien lukumäärä voidaan laskea funktiolla DATEDIFF(pvm2, pvm1, "days").

Huom! Päivämäärien ja kellonaikojen tarkasteluun on olemassa myös "ajantarkasteluvelho" (**Transform**-valikon **Date and Time Wizard**), jonka avulla on helpohko ohjatusti käsitellä päivämääräja kellonaikatietoja.

If-painiketta painamalla avautuu lisämääritysikkuna. Jos ikkunasta valitaan Include all cases, niin edellisessä ikkunassa määritelty muunnos koskee kaikkia aineiston tilastoyksiköitä. Jos valitaan kohta Include if case satisfies condition:, niin edellisen ikkunan muunnos koskee niitä tilastoyksiköitä, jotka täyttävät sen ehdon, jonka käyttäjä muodostaa tähän lisämääritysikkunaan.

3.2. Muuttujan arvojen vaihtaminen

Muuttujan luokkien yhdisteleminen, jatkuvan muuttujan luokittelu ja muut arvojen vaihtamiseen liittyvät proseduurit löytyvät **Transform**-valikosta.

Transform-valikon valinnan **Recode** avulla voidaan muodostaa itse uudet luokat tai muuttuja-arvot. Uudelleenkoodaus voidaan toteuttaa joko muuttujan sisältöä vaihtamalla (valinta **Into Same Variables**), jolloin muuttujan nimi ja sijainti aineistossa säilyvät ennallaan, tai muodostamalla olemassa olevan muuttujan avulla uusi muuttuja (valinta **Into Different Variables**).

Kun alivalikosta valitaan Into Different Variables, avautuu uusi määrittelyikkuna (kuvio 22).

| a nimi | Numeric <u>V</u> ariable -> Output Variable: tentti> tenttil | Output Variable |
|--|---|----------------------------|
| Pääaine [paaaine] Vuosikurssi [kurs | | tenttil Label: |
| Ikä vuosina [ika] Sukupuoli [sukup | * | Tenttiin lukemiseen käytet |
| onko avo- tal avio Onko ansiotyöss Ansiotyön määrä | | |
| Opintojen edisty | | |
| Osallistuminen o | Old and New Values | |
| Edellisen mat/til | [[f] (optional case selection condition) | |

Kuvio 22. Recode into Different Variables-määrittelyikkuna

Luokittelu voidaan määritellä koskemaan vain osaa havainnoista. Osajoukko määritellään **If...**-option avulla avautuvassa ikkunassa (kts. luku 3.5.). Muuttuja(t) valitaan listasta ja siirretään käsiteltävien muuttujien ruutuun (**Input/Numeric/String Variable** -> **Output Variable**), esimerkkikuviossa valittu muuttuja *tentti*. Samantyyppiset, samalla tavalla luokiteltavat muuttujat voidaan valita samaan ruutuun, jolloin määrittelyt voidaan tehdä kaikille samanaikaisesti. Muodostuvan uuden muuttujan nimi kirjoitetaan **Output Variable Name** –ruutuun (esimerkissä: *tenttil*), sen nimike haluttaessa **Label**-ruutuun (esimerkissä: *Tenttiin lukemiseen kä...*) a sen jälkeen näpäytetään **Change**-painiketta.

Old and New Values -painikkeella saadaan auki määrittelyikkuna (kuvio 23) arvojen muutosten esittämiseksi.



Kuvio 23. Arvojen määrittelyikkuna

| (Old Value) määritetään vasemmalla puolella vaihtoehdoista: |
|---|
| yksittäinen muuttuja-arvo, arvo kirjoitetaan ruutuun |
| SPSS:n määrittämä puuttuva tieto |
| SPSS:n määrittämä tai itse määritelty puuttuva tieto |
| väli (esitetään alaraja ja yläraja) |
| väli (esitetään yläraja), muuttujan pienimmät arvot |
| väli (esitetään alaraja), muuttujan suurimmat arvot |
| kaikki muut arvot |
| |

Vanhalle merkkijonomuuttujalle voi määritellä yksittäiset arvot, mutta välejä ei voi määritellä.

| Uuden muuttujan arvoina ikk | tunan oikealla puolella (New Value) käytetään |
|-----------------------------|--|
| Value | yksittäinen muuttuja-arvo, arvo kirjoitetaan ruutuun |
| System-missing | SPSS:n määrittämä puuttuva tieto |
| Copy old value(s) | kopioidaan vanhat arvot |

Muutos tehdään seuraavasti: valitaan sopiva Old Value -kohta, kirjoitetaan ruutuun arvo/arvoväli, valitaan sopiva New Value -kohta, kirjoitetaan ruutuun arvo ja näpäytetään Add-painiketta, jolloin

listaan **Old** --> **New** tulee tieto ensimmäisestä muutoksesta jne. Jokaisen arvomäärityksen jälkeen näpäytetään **Add**-painiketta. Määritys voidaan poistaa (**Remove**) tai sitä voidaan muokata (**Change**). Kuvion 23 esimerkissä: *tentti*-muuttujan puuttuvat tiedot ovat uudessa *tenttil*-muuttujassa puuttuvia, *tentti*-muuttujan arvo 0 kopioidaan arvoksi myös uudella *tenttil*-muuttujalla, , *tentti*-muuttujan arvoväli 1 – 9 kirjautuu uuteen *tenttil*-muuttujan arvona 1, *tentti*-muuttujan arvoväli 10 - 19 kirjautuu uutena arvona 2 ja kaikki muut *tentti*-muuttujan arvot kirjautuvat uutena arvona 3. Lopuksi näpätään **Continue** ja vielä päämäärittelyikkunasta **OK**.

Jos uusi muuttuja halutaan merkkijonomuuttujaksi, valitaan kohta **Output variables are strings**. Ko. valinta ei ole useinkaan järkevä tilastolliselta kannalta. Jos numeromerkit halutaan muuntaa reaaliluvuiksi, valitaan **Convert numeric strings to numbers**.

Muunnokset voidaan kirjata mielivaltaisessa järjestyksessä, mutta ne toteutetaan aina seuraavasti:

- 1° yksittäiset arvot
- 2° puuttuvat tiedot
- 3° arvovälit
- 4° muut arvot.

Kun uusi muuttuja on muodostettu, kannattaa aineistoikkunasta heti tarkistaa, että muuttujaarvot ovat järkeviä.

Myös **Transform**-valikon valinnan **Visual Binning** avulla voidaan numeerisen järjestysasteikollisen tai määrällisen muuttujan arvot ryhmitellä uuden muuttujan luokiksi. Ensimmäisessä määrittelyikkunassa valitaan ne muuttujat, joita halutaan käsitellä. Sen jälkeen painetaan **Continue**-painiketta, jonka jälkeen avautuu uusi määrittelyikkuna (kuvio 24). Tämä tehtävä käy automaattisesti läpi kaikki havainnot valituista muuttujista ja esittää aktiiviseksi valitun muuttujan pienimmän ja suurimman arvon sekä frekvenssijakauman kuviona, mikä helpottaa luokituksen tekemistä.

Kohtaan **Binned Variable** kirjoitetaan sen muuttujan nimi, johon luokitustiedon halutaan tallentuvan kokonaislukuina 1, 2, jne. (esimerkissä *Tenttiluokitus*). Luokkien rajat saa muodostaa mielivaltaisesti, kun **Value**-sarakkeelle kirjoitetaan kunkin muodostuvan luokan yläraja. Kohdasta **Upper Endpoints** valitaan se, kuuluvatko kirjoitetut ylärajat luokkiin mukaan (**Included**) vai ei (**Excluded**) (esimerkissä kirjoitettu itse arvot 10 ja 20 ja tehty valinta Excluded). **Make Cutpoints** -valintaa taas kannattaa käyttää silloin, kun luokituksen haluaa tehdä tasavälisesti (valinta **Equal Width Intervals**), tai kun haluaa jakaa muuttujan arvot samankokoisiin ryhmiin (**Equal Percentiles Based on Scanned Cases**), tai kun luokkarajat perustuvat keskiarvoon ja keskihajontaan (**Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases**).

Make Labels -valinnalla SPSS nimeää luokat automaattisesti (esimerkissä käytetty tätä valintaa ja siten Label-kohdassa näkyvät: <10, 10 - 19 ja 20+). **Copy Bins** –valinnalla voidaan luokitustapa kopioida toisesta muuttujasta (**From Another Variable**) tai toisille muuttujille (**To Other Variables**). **Reverse Scale** -valinnalla luokkien numerointi tapahtuu käänteisessä järjestyksessä (siis viimeinen luokka saa koodiksi numeron 1, toiseksi viimeinen luokka numeron 2 jne.).



Kuvio 24. Visual Bander - määrittelyikkuna

3.3. Tilastoyksiköiden järjestäminen

Jotkin havaintoaineiston muokkaamiseen liittyvät tehtävät vaativat, että tilastoyksiköt on järjestetty tietyn muuttujan arvojen perusteella nousevaan tai laskevaan järjestykseen. Järjestäminen aloitetaan valikon **Data** valinnalla **Sort Cases**, jolloin avautuvassa määrittelyikkunassa (kuvio 25) ruutuun **Sort by** valintaan ne muuttujat, joiden perusteella järjestäminen tehdään. Ensimmäisenä esitetty muuttuja (esimerkissä *ika*) on ensisijainen järjestäjä ja toisena esitetty (esimerkissä *nimi*) on toissijainen järjestäjä jne. Kohdassa **Sort Order** valitaan, asetetaanko muuttujan arvot nousevaan (Ascending) vai laskevaan (**Descending**) järjestykseen. Valinta tehdään jokaiselle muuttujalle erikseen, ja tehty valinta näkyy muuttujan nimen perässä. Esimerkin valinnoilla lajittelu tapahtuu niin, että nuorimmat sijoittuvat aineiston alkupäähän ja vanhimmat loppupäähän, ja jos samanikäisiä henkilöitä on useita, heidät järjestetään nimen perusteella aakkosjärjestykseen.

Tämän järjestetyn aineiston voi tallettaa valinnalla Save file with sorted data.



Kuvio 25. Havaintojen järjestämisen määrittelyikkuna

3.4. Tilastoyksiköiden painottaminen

Tilastoanalyyseissä tilastoyksiköitä kohdellaan samanarvoisina, jolloin jokaisella tilastoyksiköllä on sama painoarvo esim. keskiarvoa laskettaessa. Joskus tilastoyksikkö (esim. yritys) kuitenkin edustaa useata havaintoa (esim. yrityksen työntekijöitä). Laskettaessa sellaisia tilastollisia tunnuslukuja, joissa halutaan ottaa huomioon se, että tilastoyksiköt eivät olekaan samanarvoisia, voidaan ottaa käyttöön painomuuttuja

(esim. yrityksen työntekijöiden lukumäärä) **Data**-valikon valinnalla **Weight Cases** (tai työkalu). Kohtaan **Weight cases by Frequency Variable** valitaan se muuttuja, johon painottaminen perustuu. Aineistoikkunan oikeassa alakulmassa tilarivillä lukee **Weight On** silloin, kun painotus on voimassa. Painotus poistetaan käytöstä **Weight Cases** –proseduurin valinnalla **Do not weight cases**.

3.5. Osajoukon valitseminen

Joskus tilastoanalyyseissä halutaan tarkastella vaan osaa tilastoyksiköistä. Joissakin tilastoproseduureissa on mahdollista tehdä osajoukkovalinta silloin, kun valinta perustuu vain yhteen muuttujaan ja siihen liittyvään ehtoon (esimerkiksi valitaan sukupuolen perusteella vain naiset). Useimmiten aineiston osajoukko (esimerkiksi satunnaisotos tai tietyn ehdon täyttävät tilastoyksiköt) joudutaan kuitenkin muodostamaan ennen varsinaisen tilastoproseduurin käyttöä **Data**-valikon valinnan **Select Cases** (tai

työkalu 💻) avulla.

Aukeavan määrittelyikkunan (kuvio 26) Select -kohdan toiminnot ovat

All cases

kaikki tilastoyksiköt ovat mukana seuraavissa tehtävissä. Osajoukkovalinta voidaan kumota tällä valinnalla.

If condition is satisfied

valitaan tietyn ehdon täyttävät tilastoyksiköt. Valinnan jälkeen painetaan If -painiketta, jolloin avautuvaan määrittelyruutuun (kuvio 27) kirjoitetaan sopiva ehto. Aineistoon muodostuu muuttuja

filter_\$, jonka arvo on 1, kun tilastoyksikkö on valittu ja 0, kun ei ole valittu. Muodostuvan muuttujan arvoja voi tarvittaessa muokata.

Random sample of cases

tilastoyksiköistä poimitaan satunnaisotos. **Sample**-painikkeella avautuvassa ikkunassa voidaan valita otoskoko prosenttiosuutena (**Approximately** % of all cases) tai lukumääränä (Exactly ... cases from the first ... cases) aineiston alkuosan tilastoyksiköistä. Aineistoikkunaan muodostuu nytkin *filter_\$* -muuttuja.

Based on time or case range

poimitaan tietyn välin tilastoyksiköt. **Range**-painikkeella avautuvassa ruudussa määrätään ensimmäisen (**First**) ja viimeisen (**Last**) valittavan tilastoyksikön numero.

Use filter variable:

valitaan suodinmuuttuja. Jos valitun muuttujan arvo on 0, tilastoyksikkö ei tule mukaan seuraaviin tehtäviin.

| O Delete unselected cases | Tiedekunta [tdk] Pääaine [paaaine] Vuosikurssi [kurssi] Ikä vuosina [ika] Sukupuoli [sukupuol] Onko avo- tai avioliit Onko ansiotyössä [Ansiotyön määrä vii Opintojen edistymin Ohjatun opetuksen Osallistuminen opet Yksinumeroinen sat Edellisen mat/til kur Matemaattiset taidot vuosi | Select |
|---------------------------|--|--------|
|---------------------------|--|--------|

Kuvio 26. Select Cases -määrittelyikkuna

Valitsematta jääneet tilastoyksiköt merkitään käsittelystä poistetuiksi valintapainikkeella **Filter out unselected cases**. Aineistoikkunassa tilastoyksikön järjestysnumeron päälle tulee vino viiva osoittamaan, että ko. tilastoyksikkö ei ole seuraavissa tehtävissä mukana. Aineistoikkunan oikeassa alakulmassa tilarivillä lukee **Filter On** tai **Use On** silloin, kun vain osa tilastoyksiköistä on valittu.

Valitsematta jääneet tilastoyksiköt voidaan myös tuhot aineistoikkunasta **Delete unselected cases** - valinnalla. Tällöin täytyy kuitenkin olla varovainen, ettei vahingossa korvaa alkuperäistä aineistotiedostoa vajaalla aineistolla. Valitut tilastoyksiköt voidaan myös kopioida uuteen

havaintoaineistoon **Copy selected cases to a new dataset** –valinnalla, jolloin uudelle aineistolle annetaan nimi kohdassa **Dataset name:**

Kuvion 27 esimerkissä on käytetty kahteen muuttujaan liittyvää ehtoa, jonka mukaan valituksi tulevat ne tilastoyksiköt, joilla *sukupuol*-muuttujan arvo on 2 ja *tdk*-muuttujan arvo on 1.

| Tiedekunta (tdk) | | 01 - 2 a | | | | | |
|--|---|----------|----|---|---|---|--------------------------------------|
| Pääaine (paaaine) | | | | | | | Function group: |
| kä vuosina (ika) Sukupuoli (sukupuol) | + | < | > | 7 | 8 | 9 | All Arithmetic |
| Onko avo- tai avioliit Onko ansiotyössä [| - | <= | >= | 4 | 5 | 6 | CDF & Noncentral CDF Conversion |
| Ansiotyön määrä vii Opintojen edistymin | | = | ~= | 1 | 2 | 3 | Current Date/Time Date Arithmetic |
| Ohjatun opetuksen Osallistuminen opet | 1 | & | | | 0 | | Date Creation |

Kuvio 27. Tietyn ehdon täyttävien tilastoyksiköiden määrittelyikkuna (ikkunan yläosa)

3.6. Tilastoyksikköryhmien muodostaminen

Muutamat SPSS-proseduurit voidaan toteuttaa havaintoaineiston eri tilastoyksikköryhmille, jotka muodostetaan yhden tai useamman muuttujan arvojen perusteella. Esimerkiksi ristiintaulukoita käsiteltäessä tämä onnistuu suoraan **Crosstabs**-proseduurin omilla valinnoilla. Toiset proseduurit taas vaativat, että ryhmät muodostetaan ennen varsinaisen tilastoproseduurin käyttöä.

Jos kaikkia muodostuvia ryhmiä on tarkoitus käsitellä seuraavissa tehtävissä erikseen, voi ryhmittelyn tehdä **Data**-valikon valinnalla **Split File** (tai työkalu).

Avautuvassa ikkunassa (kuvio 28) olevat valinnat ovat

Analyze all cases, do not create groups

seuraavissa tehtävissä on mukana kaikki tilastoyksiköt, ryhmiä ei muodosteta

Compare groups

seuraavat tehtävät suoritetaan erikseen jokaiselle muodostettavalle ryhmälle; eri ryhmien tulokset yritetään esittää kuitenkin niin, että ryhmien vertailu on helppoa

Organize output by groups

seuraavat tehtävät suoritetaan erikseen jokaiselle muodostettavalle ryhmälle; eri ryhmien tulokset esitetään erikseen

Groups Based on:

kohtaan valitaan muuttujat (korkeintaan 8 kpl), joiden perusteella ryhmät muodostetaan. Ensimmäiseksi valittu muuttuja on ensisijainen ryhmittelijä, toisen muuttujan ryhmät muodostetaan ensimmäisen muuttujan ryhmien sisällä jne. Esimerkissä jaetaan tilastoyksiköt ryhmiin ensisijaisesti *Tiedekunnan* ja sitten *Sukupuolen* mukaan.

Sort the file by grouping variables

aineisto lajitellaan valittujen muuttujien mukaan järjestykseen

File is already sorted

jos aineisto on jo ennalta lajiteltu muuttujien mukaan, ilmoitetaan se tällä valinnalla

| Animi Pääaine (paaaine) Vuosikurssi (kurs Vikä vuosina (ika) | Analyze all cases, do not create groups Compare groups Organize output by groups |
|--|--|
| Onko avo- tai avio Onko ansiotyöss Ansiotyön määrä Opintoien edisty | Groups Based on: |
| Ohjatun opetuks Osallistuminen o | Sort the file by grouping variables File is already sorted |
| Current Status: Analysis by | groups is off. |

Kuvio 28. Ryhmien muodostamisen määrittelyikkuna

Kun ryhmiinjako on voimassa, lukee aineistoikkunan oikeassa alareunassa tilarivillä **Split File On**. Aineistoikkunan havaintojen järjestys voi muuttua alkuperäisestä järjestyksestä.

3.7. Muuttujajoukon muodostaminen

Havaintoaineistoa analysoitaessa esim. mitta-asteikoiltaan samanlaisille muuttujille tehdään samanlaisia tehtäviä. SPSS:ssä voidaan muodostaa muuttujajoukkoja, joista voidaan valita yksi tai useampi aktiiviseksi. Tällöin tehtävämäärityksissä esiintyvä muuttujaluettelo sisältää vain aktiivisen muuttujajoukon muuttujat.

Muuttujajoukko muodostetaan valikon Utilities valinnalla Define Sets. Avautuvassa määrittelyikkunassa ruutuun Set Name kirjoitetaan muodostettavan muuttujajoukon nimi. Sen alapuolella näkyy käyttäjän muodostamat muut muuttujajoukot. Joukon muuttujat valitaan ruutuun Variables in Set.

Muuttujajoukko otetaan käyttöön Utilities-valikon valinnalla Use Sets (tai työkalu), jolloin avautuvasta ikkunasta tehdään joukon valinta. Muuttujajoukoista oletusarvoisesti ovat käytössä ALLVARIABLES (kaikki muuttujat) ja NEWVARIABLES (ne muuttujat, jotka on lisätty aineistoon sen avaamisen jälkeen).

3.8. Aineiston rakenteen vaihtaminen

Käytettävän tilastomenetelmän tai väärän tallentamistavan takia voi tulla tarve muuttaa aineiston rakennetta esim. transponoimalla havaintoaineisto (=vaihdetaan rivit sarakkeiksi) tai jakamalla yksi muuttuja useaksi muuttujaksi.

Aineiston muokkausvelho otetaan käyttöön valikosta **Data** valinnalla **Restructure**. (Pelkän transponoinnin voi suorittaa **Data**-valikon **Transpose**-valinnallakin.) Aukeavassa määrittelyikkunassa on kolme vaihtoehtoa, joista ensimmäinen yhdistää useita muuttujia yhden muuttujan useaksi havainnoksi (**Restructure selected variables into cases**), toinen jakaa usean havaintorivin tietyn muuttujan arvot yhden havaintorivin (=tilastoyksikön) useaksi muuttujaksi (**Restructure selected cases into variables**) ja kolmas vaihtoehto muuttaa kaikki sarakkeet riveiksi ja rivit sarakkeiksi (**Transpose all data**).

Seuraamalla muokkausvelhon ohjeita aineiston muokkaus onnistuu melko helposti.

3.9. Yhteenvetoaineiston muodostaminen

Aineistosta voidaan tehdä valittujen muuttujien mukaan ryhmiteltyjä yhteenvetoaineistoja valikon **Data** valinnalla **Aggregate**. Avautuvassa määrittelyikkunassa ruutuun **Break Variable**(s) valitaan se muuttuja (tai useita), jonka arvojen perusteella muodostetaan ryhmät ja ruutuun **Aggregate Variables**(s) se muuttuja (tai useita), jonka yhteenvetotietoja halutaan määrittää eri ryhmille. Oletuksena proseduuri antaa muodostuvalle yhteenvetomuuttujalle nimeksi muuttujan nimen lisättynä alaviivalla ja järjestysnumerolla ja laskentafunktioksi keskiarvon (**Mean**). Yhteenvetomuuttujan nimen voi vaihtaa kohdassa **Name & Label** ja laskentafunktion voi valita kohdasta **Function**. Valinnalla **Number of cases** lisää aineistoon muuttujan, joka kertoo kuhunkin ryhmään kuuluvien havaintojen lukumäärän.

Valinnalla Add aggregated variables to active dataset, lisätään yhteenvetomuuttujat aineistoikkunan havaintoaineistoon. Valinnalla Create a new dataset containing only the aggregated variables aineistoikkunaan muodostuu havaintoaineisto vain yhteenvetomuuttujista. Valinnalla Write a new data file containing only the aggregated variables talletetaan yhteenvetoaineiston muuttujat tiedostoksi, jonka nimi oletusarvoisesti on aggr.sav. Nimen voi vaihtaa kohdassa File.

4. GRAAFISET ESITYKSET

4.1. Kuviotyypit

SPSS:n päävalikon kohdasta Graphs löytyvät kuviotyypit ja niiden ryhmittely.

Ensimmäinen vaihtoehto **Chart Builder** avaa ikkunan, jonka avulla saadaan muodostettua tilastokuvioita valmiin mallin perusteella tai elementti elementiltä. Vaihtoehto **Graphboard Template Chooser** avaa ikkunan, jossa muuttujavalintojen jälkeen SPSS ehdottaa sopivaa kuvatyyppiä. **Legacy Dialogs** - vaihtoehdon avulla voidaan tilastokuviota muodostaa perinteisellä eli aikaisempien SPSS-ohjelmaversioiden tavalla.

Tilastokuviovalikoimat ovat melkein samat kaikissa kolmessa vaihtoehdossa:

| Bar | pylväskuvio (pystypylväät, vaakapylväät, summapylväät, pylväsryhmät) |
|-------------|--|
| Dot | piste- tai viivakuvio |
| Line | viivakuvio |
| Ribbon | kolmiulotteinen viivakuvio |
| Drop-Line | piste- tai viivakuvio |
| Area | alue- eli pinta-alakuvio |
| Pie | ympyräkuvio eli piirakka |
| Boxplot | laatikko-viikset-kuvio |
| Error Bar | virhejanakuvio |
| Histogram | frekvenssihistogrammi |
| Scatterplot | pisteparvikuvio |
| | |

4.2. Esimerkkikuvioita Chart Builder –vaihtoehdon avulla

Kun tilastokuvion laatiminen aloitetaan **Chart Builder**in avulla, tulee esiin huomautus siitä, että muuttujien mitta-asteikoiden pitää olla kunnossa, jotta **Chart Builder**illä saadaan muodostettua järkevä kuvio. Jos muuttujien mitta-asteikot ovat kunnossa, voi kuvion laatimista jatkaa näpäyttämällä OKpainiketta. Jos muuttujien määrittelyt kaipaavat tarkennuksia, pääsee niitä tekemään näpäyttämällä **Define Variable Properties** –painiketta. Joskus kesken aineiston käsittelyn täytyy kuitenkin mittaasteikkotietoa tilapäisesti vaihtaa. Esimerkiksi, jos 5-portaisen Likertin asteikollisen muuttujan jakaumasta haluaa saada arvojen lukumääriä kuvaavan pylväskuvion, pitää asteikkona olla **Ordinal**, mutta jos halutaan useasta Likertin asteikollisesta muuttujasta esittää keskiarvot viivakuviona, pitääkin asteikkona olla **Scale**.

Chart Builder -ikkunan (kuvio 29) **Gallery**-välilehdellä valitaan muodostettavan tilastokuvion perustyyppi. Kuviovalikoiman viereen tulee kuvakkeet erilaisista vaihtoehdoista, jotka pohjautuvat valittuun perustyyppikuvioon. Haluttu kuviovaihtoehto laahataan esikatseluruutuun (**Chart Preview**) ja tarkasteltavat muuttujat laahataan kuvion akselialueelle ja muille alueille muuttujalistasta. Silloin, kun jokin kuvioalue sisältää sinistä tekstiä, on kyseiseen kohtaan valittava muuttuja. Jos kuvioalue sisältää mustaa tekstiä, on kohtaan jo tullut itse valittua jokin muuttuja tai tunnusluku, tai kuvioproseduurissa on oletus siitä, mitä kyseinen alue kuviossa esittää. (Tilastokuvion voi tehdä myös **Basic Elements** – välilehden avulla, jolloin kuvion osat laahataan yksi kerrallaan esikatseluruutuun. Tämä tapa on aika vaativa aloittelevalle tilastotieteilijälle.)

Jos kuvioon halutaan lisätä uusia muuttujia (esimerkiksi ryhmittelemään tilastoyksiköitä), voidaan **Groups/Point ID** –välilehdeltä valita yksi tai useampi vaihtoehto, jolloin esikatseluruutuun tulee uusia alueita, johon laahataan tarvittavat muuttujat muuttujalistasta. **Title/Footnotes** –välilehdellä voidaan kuvioon lisätä otsikkotekstit ja alaviitteet.



Kuvio 29. Chart Builder – määrittelyikkuna

Element Properties –ikkunassa, voidaan vaihtaa kuvion osasten määrittelyjä. Esimerkiksi pylväskuvioissa voidaan vaihtaa pylväiden muoto. Ikkunassa voidaan myös valita oletusarvoisesti esitettävän tilastollisen tunnusluvun (**Statistic**) tilalle jokin toinen tunnusluku. Frekvenssihistogrammikuviossa voidaan valita mm. joko pylväiden määrä tai luokkavälin pituus. **Options**-välilehdellä voidaan määritellä mm. puuttuvien arvojen käsittelyä ja kuvan kokoa.

Pylväskuvio diskreetin muuttujan frekvenssijakaumasta

1. Valitaan Graphs - Chart Builder. Jos ikkunassa on aikaisempien kuvioiden määrittelyjä, ne voi poistaa Reset-painikkeella.

2. Valitaan **Gallery**-valikoimasta vaihtoehto **Bar** ja laahataan yksinkertaisen pylväskuvion (**Simple Bar**) kuva esikatseluruutuun (**Chart preview**).

3. Tarkasteltava muuttuja laahataan esikatseluruudun X-Axis? -kohtaan ja näpäytetään painiketta OK.

Pylväskuvio kahden diskreetin muuttujan yhteisjakaumasta

1. Valitaan Graphs - Chart Builder.

2. Valitaan **Gallery**-valikoimasta vaihtoehto **Bar** ja laahataan pylväsryhmäkuvion (**Clustered Bar**) (tai summapylväskuvion (**Stacked Bar**)) kuva esikatseluruutuun.

3. Laahataan ensimmäinen tarkasteltavista muuttujista esikatselunruudun X-Axis? -kohtaan.

4. Laahataan toinen tarkasteltavista muuttujista esikatseluruudun kohtaan Cluster on X: set color (tai Stack: set color) ja näpäytetään painiketta OK.

Usean muuttujan ryhmäkeskiarvojen esittäminen viivakuviona

1. Valitaan Graphs - Chart Builder.

2. Valitaan **Gallery**-valikoimasta vaihtoehto **Line** ja laahataan usean viivan viivakuvion (**Multiple Line**) kuva esikatseluruutuun.

3. Laahataan ryhmittelevä muuttuja esikatselunruudun X-Axis? -kohtaan.

4. Näpäytetään painiketta **Element Properties** ja valitaan **Edit Properties of** -kohdasta **Line1** ja valitaan **Statistic**-alasvetovalikosta **Mean**.

5. Laahataan ne muuttujat, joista keskiarvot halutaan esittää esikatseluruudun **Mean**-kohtaan ja näpäytetään **OK** ja vielä **OK**.

Laatikko-viikset -kuvio

1. Valitaan Graphs - Chart Builder.

2. Valitaan Gallery-valikoimasta vaihtoehto Boxplot ja laahataan yksiulotteisen laatikko-viikset -kuvion

(1-D Boxplot) kuva esikatseluruutuun.

3. Laahataan tarkasteltava muuttuja esikatselunruudun X-Axis? -kohtaan ja näpäytetään OK.

Frekvenssihistogrammi

1. Valitaan Graphs - Chart Builder.

2. Valitaan **Gallery**-valikoimasta vaihtoehto **Histogram** ja laahataan yksinkertaisen frekvenssihistogrammin (**Simple Histogram**) kuva esikatseluruutuun.

3. Laahataan tarkasteltava muuttuja esikatselunruudun X-Axis? -kohtaan. Kuvioon voi lisätä myös normaalijakauman tiheysfunktion kuvaajan Element Properties -valinnalla Display normal curve. Näpäytetään lopuksi OK.

Pisteparvikuvio

1. Valitaan Graphs - Chart Builder.

2. Valitaan **Gallery**-valikoimasta vaihtoehto **Scatter/Dot** ja laahataan yksinkertaisen pisteparven (**Simple Scatter**) kuva esikatseluruutuun.

3. Laahataan ensimmäinen tarkasteltavista muuttujista esikatselunruudun X-Axis? -kohtaan ja toinen Y-Axis? -kohtaan ja näpäytetään OK.

4.3. Esimerkkikuvioita Legacy Dialogs –vaihtoehdon avulla

Pylväskuvio diskreetin muuttujan frekvenssijakaumasta

1. Valitaan Graphs - Legacy Dialogs - Bar

2. Avautuvasta määrittelyikkunasta (kuvio 30) valitaan Simple ja Data In Chart Are Summaries for groups of cases ja lopuksi näpäytetään Define-painiketta.

| 🔚 Bar Cha | arts 🧶 | x | | | | | |
|-----------|-----------------------------------|---|--|--|--|--|--|
| | Simple | | | | | | |
| | Clustered | | | | | | |
| | Stacked | | | | | | |
| Data in | Data in Chart Are | | | | | | |
| Sum | Summaries for groups of cases | | | | | | |
| O Sum | O Summaries of separate variables | | | | | | |
| O Valu | es of individual cases | | | | | | |
| De | fine Cancel Help | | | | | | |

Kuvio 30. Bar Charts -määrittelyikkuna

<u>HUOM.</u> Kuvion 30 määrittelyikkunassa oleva **Data In Chart Are -**ruutu esiintyy myös monissa muussa kuviomäärittelyikkunassa. Ruudusta valitaan esitettävän aineiston muoto:

Summaries for groups of cases

Kuviossa esitetään tilastoyksikkö*ryhmittäin* (esim. miehet ja naiset) laskettuja tietoja, esim. havaintojen lukumäärät ja määrällisen muuttujan keskiarvot lasketaan ryhmille erikseen **Summaries of separate variables**

Kuviossa esitetään useasta eri muuttujasta laskettuja tietoja (esim. keskiarvot)

Values of individual cases

Kuviossa esitetään yksittäisten tilastoyksiköiden tietoja (esim. muuttuja-arvot)

3. Avautuvassa määrittelyikkunassa (kuvio 31) ruutuun **Category Axis** valitaan se muuttuja (**Matemaatt...**), jonka ryhmiä pylväät edustavat ja näpäytetään painiketta **OK**, jolloin tulosikkunaan saadaan pylväskuvio (kuvio 32). Muita valintoja ei välttämättä tarvitse tehdä.



Kuvio 31. Yksinkertaisen pylväskuvion määrittelyikkuna

<u>HUOM.</u> Kuvion 31 määrittelyikkunassa voidaan pylvään korkeudeksi valita kohdassa **Bars Represent** muutakin kuin vain havaintomäärä. Vaihtoehtoja ovat

| | vunitomaara. Vunitoentoja ovat |
|-----------------|---|
| N of cases | pylvään pituus vastaa ryhmän havaintojen lukumäärää eli frekvenssiä |
| % of cases | pylvään pituus vastaa ryhmän havaintojen prosenttiosuutta koko havaintojoukosta |
| Cum. N | pylvään pituus vastaa ryhmän summafrekvenssiä |
| Cum. % | pylvään pituus vastaa ryhmän prosentuaalista summafrekvenssiä |
| Other statistic | (e.g., mean): Variable-kohtaan valitaan muuttuja, jonka arvoista lasketaan Change |
| | Statistic -kohdassa valittava tunnusluku jokaiselle ryhmälle erikseen ja pylvään |
| | pituus vastaa ko. tunnusluvun arvoa ko. ryhmässä. Valittavia tunnuslukuja ovat |
| | mm. keskiarvo, mediaani, moodi, keskihajonta ja varianssi sekä fraktiilit. |
| | |



Kuvio 32. Yksinkertainen pylväskuvio

<u>HUOM.</u> Kuviotyyppien valintaikkunoissa (kuvio 31) on yhteisiä painikkeita/valintoja, olkoon valittu kuviotyyppi mikä tahansa.

Panel by Rows/Columns -valinnoilla saadaan muodostettua useita kuvioita erikseen rinnakkain/allekkain sellaisille ryhmille, jotka muodostuvat **Rows/Columns** -kohtiin valittujen muuttujien arvojen perusteella

Template-ruudussa voidaan valita aikaisemmin muodostettu, tiedostoon talletettu mallikuvio, jonka muotomäärittelyjä käytetään nyt tehtävässä kuviossa.

Titles-valinnalla saadaan muodostettua otsikot ja alaviitteet (, jotka voi lisätä kuvioon myös muokkausvaiheessa, mutta se on hankalahkoa).

Options-painikkeen avulla voidaan määrittää puuttuvan tiedon käsittelytapa

Missing Values -kohdassa vaihtoehtoja ovat yleensä

Exclude cases listwise, jos tilastoyksiköllä on puuttuva arvo yhdessäkin tehtävään valitussa muuttujassa, ei tilastoyksikköä oteta tehtävää suoritettaessa huomioon lainkaan

Exclude cases variable by variable, jos tilastoyksiköllä on puuttuva arvo jossakin valituista muuttujista, otetaan tilastoyksikkö mukaan niihin tehtävän kohtiin, jossa puuttuva-arvoista muuttujaa ei tarkastella

Display groups defined by missing values, jolla voidaan valita, esitetäänkö puuttuvien havaintojen ryhmät kuviossa.

Pylväskuvio kahden diskreetin muuttujan yhteisjakaumasta

1. Valitaan Graphs - Legacy Dialogs - Bar.

2. Avautuvasta määrittelyikkunasta (kuvio 30) valitaan **Clustered** ja **Data In Chart Are Summaries for groups of cases** ja lopuksi näpäytetään **Define**-painiketta.

3. Avautuvassa määrittelyikkunassa (kuvio 33) valitaan ruutuun **Define Clusters by** se muuttuja, joka jakaa havaintoaineiston ryhmiin (*Sukupuoli*) ja ruutuun **Category Axis** se muuttuja (*Matemaatti...*), jota halutaan vertailla ryhmien (miesten ja naisten) kesken. Näpäytetään painiketta **OK**, jolloin tulosikkunaan saadaan pylväsryhmäkuvio (kuvio 34).

| | Bars Represent | | Titles |
|------------------------|----------------------------|-------------------|------------|
| 🔏 nimi | N of cases | ◎ % of cases | Options |
| Pääaine [naaaine] | © <u>C</u> um. N | © Cu <u>m</u> . % | (<u>-</u> |
| Vuosikurssi [kurssi] | Other statistic (e.g., mea | in) | |
| 🔗 lkä vuosina [ika] | Variable | | |
| Onko avo- tai avioliit | | | |
| Ansiotvön määrä vii | | | |
| Opintojen edistymin | Change S | Statistic | |
| 🔗 Ohjatun opetuksen | | | |
| 🔗 Osallistuminen opet | Category Axis: | | |
| Fenttiin lukuaika ten | Matemaattiset | taidot [mattaito] | |
| Yksinumeroinen sat | Define Clusters by: | | |
| Vuosi | 🎽 🔒 Sukupuoli [suk | supuol] | |
| 10031 | Panel by | | |
| | | | |

Kuvio 33. Pylväsryhmien määrittelyikkunan yläosa



Kuvio 34. Pylväsryhmäkuvio

Usean muuttujan ryhmäkeskiarvojen esittäminen viivakuviona

1. Valitaan Graphs - Legacy Dialogs – Line.

2. Avautuvasta määrittelyikkunasta (kuvio 35) valitaan Multiple ja Data In Chart Are Summaries of separate variables ja lopuksi näpäytetään Define-painiketta.

| Line Charts |
|---------------------------------|
| Simple |
| Multiple |
| Drop-line |
| Data in Chart Are |
| O Summaries for groups of cases |
| Summaries of separate variables |
| O Values of individual cases |
| Define Cancel Help |

Kuvio 35. Viivakuvioiden määrittelyikkuna

3. Avautuvassa määrittelyikkunassa (kuvio 36) valitaan ruutuun **Category Axis** se muuttuja, joka jakaa havaintoaineiston ryhmiin (*Pääaine*) ja ruutuun **Lines Represent** ne muuttujat (*Tenttiin..., Osallist..., Ohjatun..., Ansiot...*), joiden haluttu tunnusluku (**MEAN**) valitaan **Change Summary** -painikkeen avulla. Näpäytetään painiketta **OK**, jolloin tulosikkunaan saadaan viivakuvio (kuvio 37).



Kuvio 36. Viivakuvion määrittelyikkunan yläosa

40



Kuvio 37. Viivakuvio

Laatikko-viikset -kuvio

1. Valitaan Graphs - Legacy Dialogs - Boxplot.

2. Avautuvasta määrittelyikkunasta (kuvio 38) valitaan Simple ja Data In Chart Are Summaries of separate variables ja lopuksi näpäytetään Define-painiketta.

| Boxplo | t | |
|------------------|------------------------------|--|
| ₿₿®₿ | Simple | |
| ₿ [₿] € | Clustered | |
| Data in | Chart Are | |
| Sum | maries of separate variables | |
| | | |

Kuvio 38. Boxplot -määrittelyikkuna

3. Avautuvassa määrittelyikkunassa (kuvio 39) ruutuun **Boxes Represent** valitaan se muuttuja (*Ansiot*...), josta kuvio halutaan ja näpäytetään painiketta **OK**. Tulosikkunaan saadaan kuvio (kuvio 40), jossa laatikon pohja kuvaa alakvartiilia ja kansi yläkvartiilia. Laatikon keskellä oleva vaakaviiva kuvaa mediaania. Laatikon kannesta lähtee jana ylöspäin yleensä suurimpaan muuttuja-arvoon asti ja pohjasta alaspäin pienimpään arvoon. Jos tarkasteltavassa muuttujassa on kuitenkin sellainen arvo, joka on laatikon kannesta tai pohjasta 1.5 – 3 kvartiilivälin pituuden etäisyydellä, esitetään havainto poikkeushavaintona (merkintänä O ja vieressä tilastoyksikön numero/tai sellainen muuttuja-arvo, joka tilastoyksiköllä on Label Cases by: -ruutuun valitulla muuttujalla) ja jos havainto on vieläkin kauempana, on sen merkintänä *.



Kuvio 39. Yksinkertaisen laatikko-viikset-kuvion määrittelyikkunan yläosa



Kuvio 40. Laatikko-viikset-kuvio

Frekvenssihistogrammi

1. Valitaan Graphs - Legacy Dialogs – Histogram.

2. Avautuvassa määrittelyikkunassa (kuvio 41) ruutuun **Variable** valitaan se muuttuja (*Ikä...*), josta frekvenssihistogrammi halutaan. Kuvioon voi lisätä myös normaalijakauman tiheysfunktion kuvaajan valinnalla **Display normal curve**. Näpäytetään painiketta **OK**, jolloin tulosikkunaan saadaan histogrammi (kuvio 42). SPSS luokittelee havainnot tasavälisesti luokkiin. Frekvenssihistogrammikuviossa on vaaka-akselilla näkyvissä muuttuja-arvojen viitekohtia ja pystyakseli

kuvaa luokan havaintomäärää. Kuvan vieressä on tiedot muuttujan keskiarvosta, keskihajonnasta sekä havaintomäärästä. Kuviossa näkyvä käyrä kuvaa sellaisen normaalijakauman tiheysfunktiota, jonka keskiarvo ja keskihajonta ovat samat kuin tarkasteltavalla muuttujalla.

| | Variable: | Titles |
|----------------------|---------------------------|--------|
| 💑 nimi | 🛀 🔣 🚺 🖉 Ikä vuosina (ika) | |
| 💑 Tiedekunta [tdk] | Display normal curve | |
| 💑 Pääaine [paaaine] | | |
| Vuosikurssi (kurs | Panel by | |
| 💑 Sukupuoli [sukup | Rows: | |
| 💑 Onko avo- tai avio | | |

Kuvio 41. Histogram-määrittelyikkunan yläosa



Kuvio 42. Frekvenssihistogrammi

Pisteparvikuvio

1. Valitaan Graphs – Legacy Dialogs - Scatter/Dot.

2. Avautuvasta määrittelyikkunasta (kuvio 43) valitaan Simple Scatter ja näpäytetään Define-painiketta.



Kuvio 43. Scatterplot-määrittelyikkuna

3. Avautuvassa määrittelyikkunassa (kuvio 44) valitaan ruutuun **Y** Axis pystyakselille tuleva muuttuja (*Ansiot...*), **X** Axis vaaka-akselille tuleva muuttuja (*Ikä*). Lisäksi ruutuun **Set Markers by** voidaan haluttaessa valita muuttuja (*Sukupuoli*), jonka eri ryhmät merkitään esim. eri väreillä ja ruutuun **Label Cases by** voidaan valita muuttuja, jonka perusteella tilastoyksiköt nimetään, ja nimet tulevat näkyviin kuvioon. Näpäytetään painiketta **OK**, jolloin tulosikkunaan saadaan pisteparvikuvio (kuvio 45).



Kuvio 44. Yksinkertaisen pisteparvikuvion määrittelyikkunan yläosa



Kuvio 45. Pisteparvikuvio

4.4. Kuvion muokkaaminen ja mallikuvio

Kaikki kuviot esitetään tulosikkunassa (**SPSS Viewer**). Kuviota pääsee muokkaamaan näpäyttämällä sitä kahdesti, jolloin avautuu kuvionmuokkausikkuna **Chart Editor** (kuvio 46).

Kuvion muokkauksen perusidea on se, että muokattavaa kohtaa näpäytetään kerran tai kahdesti, jonka jälkeen useimmiten avautuu ominaisuusikkuna (**Properties**). Tämä ikkuna koostuu useasta välilehdestä, joista voi valita ne toimenpiteet, mitä kuviolle halutaan tehdä. Esimerkiksi, jos edellisen kuvion vaaka-akselin nimikohtaa näpäytetään kerran, tulee nimen ympärille kehys. Kun perään tehdään toinen näpäytys pienellä viiveellä, pääsee vaaka-akselin nimeä muokkaamaan. Jos nimikohtaa kaksoisnäpäytetään nopeasti, avautuu ominaisuusikkuna, jossa voi valita esim. mihin kohtaan vaaka-akselin nimi sijoitetaan. Ominaisuusikkunoita saa auki myös esimerkiksi **Edit**-valikon avulla.



Kuvio 46. Kuvionmuokkausikkuna ja sen valikot

File-valikon valinnat ovat

Save Chart Template, jolla kuvio talletetaan mallikuvioksi myöhempää käyttöä varten Apply Chart Template, jolla otetaan käyttöön mallikuvio Export Chart XML, jolla kuvio talletetaan kuvaksi Close, jolla kuvanmuokkausikkuna suljetaan

Help-valikon Topics-valinnalla saa apua kuvionmuokkausikkunan välineiden käyttöön

View-valikon valinnat ovat

Status Bar, jolla valitaan tilarivi näkymään ikkunan alareunaan

Edit Toolbar, jolla valitaan näkymään ne työkalut, jotka koskevat muokattavan akselin valintaa **Options Toolbar**, jolla valitaan näkymään ne työkalut, jotka liittyvät **Options**-valikon valintoihin

Element Toolbar, jolla valitaan näkymään ne työkalut, jotka liittyvät **Elements**-valikon valintoihin

Format Toolbar, jolla valitaan näkymään ne työkalut, jotka koskevat mm. kuvion tekstimuokkauksia

Large Buttons, jolla valitaan työkalupainikkeiden koko

Edit-valikon valintoja ovat mm.

Properties (myös työkalu is tai kuvion kaksoisnäpäytys), jolla voidaan muuttaa kuvion koko halutuksi; **Maintain aspect ratio** –valinta pitää kuvion leveys- ja korkeussuhteen vakiona, jos muuttaa vain toista leveys/korkeusvalinnoista

Select Chart, jolla valitaan kuvio, jotta koko kuviolle voidaan tehdä muokkaustoimenpide)

Select X Axis (X tai vaaka-akselin kaksoisnäpäytys), jolla avautuu määrittelyikkuna, jonka välilehtien avulla voidaan muokata kuvion vaaka-akselin ominaisuuksia

Select Y Axis (**Y** tai vaaka-akselin kaksoisnäpäytys), jolla avautuu määrittelyikkuna, jonka välilehtien avulla voidaan muokata kuvion pystyakselin ominaisuuksia

Select Z Axis), jolla avautuu määrittelyikkuna, jonka välilehtien avulla voidaan muokata kuvion kolmannen akselin ominaisuuksia

Options-valikon valintoja ovat mm.

X-Axis/ Y-Axis Reference Line ([], [-]) akselin suuntaisia viiteviivoja

Title (, kuvion otsikko

Annotation (上) huomautustekstejä, jotka ankkuroidaan tiettyyn kuvion kohtaan

Text Box (Ш) tekstilaatikkoja haluttuun kohtaan kuviota

Show/Hide Derived Axis (U), jonka avulla saadaan näkyviin tai poistettua y-akselista muunnettu akseli

Show/Hide Legend (), jonka avulla saadaan näkyviin tai poistettua esim. selitykset ryhmien merkintätavoista

Transpose Chart (**b**), jonka avulla vaihdetaan kuvion akselit (vaaka-akselista tulee pystyakseli)

Elements-valikon valintoja ovat mm.

Data Label Mode (), jonka avulla kuviosta voidaan valita ne yksittäiset osiot (esim. tilastoyksiköt), joiden nimet/nimikkeet halutaan lisätä/poistaa

Show/Hide Data Labels (L.), jonka avulla saadaan näkyviin tai poistettua esim. pylväiden pituusnumeroinnit tai esim. tilastoyksiköiden numerot/nimet

Add Markers (^{LC}), jonka avulla saadaan näkyviin viivakuvioiden havaintopisteet merkkien avulla

Fit Line at Total () sovitetaan erilaisia soviteviivoja (esim. regressiosuora)

Fit Line at Subgroups () sovitetaan ryhmille erilaisia soviteviivoja (esim. regressiosuora)

Interpolation Line () sovitetaan erilaisia arviointiviivoja

Show Distribution Curve () sovitetaan kuvion päälle tietyn todennäköisyysjakauman tiheysfunktion kuvaaja

Explode/Return Slice (¹), jonka avulla piirakkakuviosta voidaan irrottaa tai palauttaa sektori

Jos aineiston muuttujista aiotaan tehdä useita samantyyppisiä kuvioita, voidaan muodostaa yksi peruskuvio malliksi, ja tähän kuvioon liittyviä määrityksiä voidaan soveltaa muiden kuvioiden tuottamisessa. Mallikuvio muodostetaan ja muokataan aivan normaalisti kuvionmuokkausikkunassa, jonka jälkeen se talletetaan sopivalla nimellä tiedostoksi. Kun mallikuvion määrityksiä halutaan käyttää, valitaan uuden kuvion määrittelyikkunassa (esim. kuvio 31) mallikuviotiedosto **Template**-ruudun kohdassa **Use chart specifications from: File**.