

Introduktion till SPSS

Start

Kör gärna ”tutorial” för att bekanta dig med programmet!

3½ fönster i SPSS

Data- och variabelfönster

Datafönster

Visar datafilen.

Endast en datafil kan vara öppen i taget.

Härifrån gör man

- Sorteringar
- Selektioner och filter
- Inställningar
- Diverse kommandon från menyer

Variabelfönster (öppnas genom att klicka på flik eller dubbelklicka på variabelnamn)

Variabelnamn: 1-8 tecken, inga mellanslag.

Variable label: klartextetikett för variabelns hela namn (max 255 tecken).

Type: typ av data, t ex numeriskt, datum/tid, text (max 255 tecken).

Format: antal tecken, inkl minus, decimaltecken och decimaler.

Value labels: Värdeetiketter ersätter kodvärde med klartextetikett, men tillåter fortsatt beräkning med numeriskt värde.

Output-fönster

Visar resultat av kommandon, t ex tabeller och grafer och andra beräkningar.

Visar varningar, kommandon etc (om inställningarna anger det).

Kan redigeras **mycket**.

Kan exporteras till Word, html mm.

Flera outputfönster kan vara öppna samtidigt.

Syntax-fönster

Visar programmeringsspråk.

Kommandon kan klistras in från dialogrutor, från outputfönster (om man visar kommandon), journalfilen, eller från en annan fil.

Sparas som sps-fil, men är egentligen en textfil, som kan redigeras i ordbehandlare.

Varje kommando avslutas med en punkt, men kan löpa över flera rader (orsakar 98% av alla syntaxfel).

OBS! Endast punkt används som decimaltecken i syntaxfönster! (vid beräkningar med formler)

Kommentarer kan läggas in; börja raden med stjärna (*) och avsluta med punkt.

Man kan köra ett enskilt kommando, ett urval kommandon eller hela fönstrets innehåll på en gång.

Flera syntaxfönster kan vara öppna samtidigt.

Inställningar

Inställningar av programmet gör under **Edit|Options**

Viktiga inställningar

- Placering av logfil, och ”append file”
- Decimaltecken = komma
- Val av ”labels” och/eller ”name” i listor och output
- Kommandon ska visas i output

Log-filen är en textfil, även om den har ändelsen ”.jnl”. Därför kan den öppnas i en vanlig ordbehandlare.

Log-filen infogar datum och klockslag varje gång SPSS startas, och visar alla kommandon som körs, liksom varningar och felmeddelanden. På detta sätt kan man alltid gå tillbaka och kolla hur man gjorde den gången det fungerade, eller vad som gick fel.

Om man väljer alternativet ”Append file” lägger SPSS till de nya kommandona i slutet av den gamla logfilen, annars skrivs den över varje gång man startar om SPSS.

Kommandon

Kommandon körs från menyer och dialogrutor, eller syntaxfönster.

Kommando-syntax visas i logfil och output (om den funktionen aktiverad).

I dialogrutor klistrar knappen ”Paste” in kommandot med alla gjorda inställningar i syntaxfönstret.

Dialogrutornas kommandon klarar det mesta, men alla kommandon och inställningar finns inte med. Titta i hjälpen för respektive kommando.

Inklistring från dialogrutor ger bra grund för att kunna göra ändringar i syntax, t ex variabelnamn eller -antal, eller köra långa sekvenser av kommandon.

Menykommandon, exempel

File

Open
Open Database | New query

Edit

Vanliga klipp och klistra-kommandon
Options

View

Value Labels

Data

Split File (separate analys för varje värde för vald variabel)
Select cases (analys endast för de poster som uppfyller urvalskriterier)
Aggregate
Merge

Transform

Compute
Recode

Analyze

Statistisk analys av olika slag, t ex

- Tabeller
- t-test
- Regression
- Icke-parametriska test

Graphs

Mer specialiserade grafer än snabbgrafer i analyser

Import av data från andra program

Exempel: Import från Access

Välj kommandot **File | Open Database | New query**

Välj "MS Access" i listan

Leta upp "C:\Databasen.mdb" (eller var nu databasen ligger)

Välj tabell som ska importeras genom att dra den till det högra fönstret

Stega igenom importfunktionen (använd bara standardalternativ)

Data visas nu i Datafönstret, men

- Variabelnamnen är konstiga
- Kategorivariabler innehåller bara siffror (t ex om Sex är kodat 0=kvinnor och 1=män finns bara 1:or och 0:or)
- Långa tal eller datum visas bara som "spjälstaket"

Gå till Variabel-fönstret och

- Ändra variabelnamn till bättre alternativ (max 8 tecken)
- Ange Variable labels (klartextnamn för varje variabel, max 255 tecken) om så önskas
- Ändra format på variablerna så de kan visa all information
- Ange värdeetiketter (Value labels) för varje värde för kategorivariabler

Visserligen kan man använda å, ä, ö och tecken som &, % och komma och punkt, men undvik helst dessa tecken i labels och variabelnamn. Det fungerar så länge man håller sig till SPSS, men om man behöver spara data i ett annat filformat, t ex till en kollega som använder ett annat program, ställer dessa tecken ofta till stora problem.

Slå på "Labels" i Data-fönstret

Spara som filen som en spss-datafil, sav-fil

I Outputfönstret har följande hänt:

- Utförda kommandon visas under rubriken "Log"
- Dock ej format-ändringar och angivna etiketter...
- Varningar för eventuella fel vid import och/eller konvertering till SPSS-format skrivs ut.

Upprepa ovanstående för varje tabell du vill importera. Eller gör en syntaxfil...

På motsvarande sätt kan man importera data från andra databaser, enskilda blad i en Excel-fil, tabbade textfiler etc.

Beräkna nya variabler

Nya variabler kan skapas på flera sätt. Gör hellre en ny variabel för resultatet än att ändra värden i rådata, om du inte är 200% säker på att resultatet blir det du önskar. Det är lättare att ta bort en överflödigt variabel i efterhand än att försöka återskapa värden man skrivit över...

Kommandot Transform | Compute

Värden (=resultat av en beräkning) sparas i datafilen, inte som formler som i Excel. Ändras utgångsvärdena måste man alltså göra en ny beräkning för att uppdatera en resultatvariabel.

Kan användas för att beräkna t ex ålder, index för flera variabler tillsammans mm.

Kommandot Transform | Recode

Används för att t ex göra en kontinuerlig variabel till grupper, t ex ålder till åldersgrupp, eller slå samman flera kategorier till ett färre antal, tex ålder i 10-årsintervall till yngre resp. äldre än 50 år.

Ändra datafiler

Spara först datafilen under nytt namn med "Save as"-kommandot, för att inte riskera att skada ursprungsfilen.

Merge

Kommandot "Merge" slår samman data från två eller flera filer till en gemensam fil. Filerna måste vara sorterade på samma sätt, t ex efter stigande löpnr. Merge-kommandot kan utföra flera olika slags sammanslagningar.

1. Sammanslagning av filer som innehåller olika variabler för matchande poster

Man kan både slå samman filer som matchar varandra exakt (med en post för varje nyckelvariabel i var fil, t ex en fil med persondata, och en fil med lab-resultat från ett blodprov för varje person), och filer som inte matchar (endera saknas några matchande poster i en fil som finns i den andra, eller så har man en fil med en post per nyckelvariabel och en fil med flera poster per nyckel, t ex upprepade labresultat för samma patient).

Exempel:

I forskningsstudier gör man ofta flera sorters undersökningar på samma patient, t ex blodtrycksmätningar, en enkät och en lista på använda mediciner. Alla mätningar registreras tillsammans med ett löpnummer (t ex personnr) som identifierar patienter, men de olika undersökningarna lagras i olika tabeller. Alla dessa data behöver slås samman till gemensamma filer för att kunna analysera samband mellan dem. T ex kan man behöva uppgifter om ålder och kön från enkäten tillsammans med blodtrycksvärden. För detta ändamål använder man Merge-kommandot

Arbetsgång:

- Öppna enkätfilen
- Sortera efter löpnr.
- Spara filen
- Öppna, sortera och spara blodtrycksfilen på samma sätt
- Välj kommandot "Merge", "Add variables"
- Välj enkätfilen som andra tabell
- Välj löpnr som "Key variabel"
- Välj Enkätfilen som "keyed table" (dvs den fil som bara har en post med respektive löpnr).
- Kör!

Den nya filen innehåller alla blodtrycksdata som tidigare, men efter dessa variabler syns nu alla variabler från enkäten. Om patienten har flera registrerade blodtryck upprepas enkätsvaren för varje blodtrycksmätning.

Spara resultatfilen under lämpligt namn.

2. Sammanslagning av filer som innehåller samma variabler men omatchade poster

Om data samlas in i flera omgångar eller på olika platser får man ofta filer med samma slags data, men för olika personer. Dessa data kan också slås samman med kommandot "Merge", men man använder då "Append data"-versionen.

Aggregate

Kommandot kan användas för att sammanfatta data för grupper av poster, t ex antal poster, min/max för en variabel, värdet för första/sista post eller summan av en variabel för alla poster med samma löpnr.

Exempel: Beräkna data om blodtryck

- Öppna blodtrycksfilen
- Välj kommandot "Aggregate"
- Ange att följande nya värden ska beräknas:
 - o Break = löpnr
 - o Antal = Antal registrerade blodtryck för varje patient
 - o BTmin = lägsta registrerade tryck
 - o BTmax = högsta registrerade BT
 - o BTmean = medel av blodtryck
- Ange att resultatet ska vara den nya aktuella filen
- Kör kommandot och studera den nya resultatfilen.

Den aggregerade filen kan sedan slås samman med annan fil, t ex personuppgifter, med hjälp av kommandot "Merge" med löpnr som nyckelvariabel.

De första analyserna

För att kontrollera datakvalitet körs bas-statistik

- Frekvens/antal
- Antal Missing values, både helt saknade värden och Missing value-koder
- Min/max för varje variabel
- Mean, Median, SD, SEM, Skewness, Kurtosis
- Korstabeller för att kontrollera grupperingar och kodningar, och fördelning mellan två grupperingar t ex kön vs åldersgrupp.
- Enkla grafer, t ex stapel- eller punktdiagram visar om det finns oväntade mönster eller extremvärden

Andra användbara kommandon

Select

Efter att man angivit villkoren för "Select" körs efterföljande kommandon enbart för poster som uppfyller urvalskriterierna.

Kan slås av och på enligt önskemål. Bortvalda poster finns kvar, dvs blir ej raderade.

Om man vill ha en fil med endast de selekterade posterna kan man kryssa för detta alternativ i dialogrutan för "Select"-kommandot, men gör gärna en vanlig selektion först, för att kolla att urvalet blir korrekt, och spara filen under nytt namn INNAN du kör Select delete.

Group by

Kör följande analyser för varje värde av den grupperande variabeln, t ex för män resp kvinnor, eller flera åldersgrupper