

PSYC 3101: ØVING I FAKTORANALYSE

I forrige SPSS-øving undersøkte vi reliabiliteten til en skala bestående av åtte holdningsspørsmål. Basert på de analysene vi foretok, endte vi opp med at disse seks av de åtte spørsmålene kunne utgjøre en reliabel skala:

	Helt uenig						Helt enig	
SP1. Ofte er det slik at jo bedre kjøreferdigheter en sjåfør har, jo fortere kan sjåføren kjøre	1	2	3	4	5	6	7	
SP2. Det er helt i orden å kjøre over fartsgrensen hvis trafikforholdene er bra	1	2	3	4	5	6	7	
SP3. Jeg vil heller gå 5 km enn å sitte på med en som råkjører	1	2	3	4	5	6	7	
SP4. Jeg sitter på med en uforsiktig sjåfør hvis det ikke er andre måter å komme seg hjem på	1	2	3	4	5	6	7	
SP6. Jeg sitter på med venner som kjører litt rått så lenge jeg stoler på dem	1	2	3	4	5	6	7	
SP8. Noen ganger må man tøye reglene for å komme seg fram i trafikken	1	2	3	4	5	6	7	

Cronbach's alfakoeffisient for disse seks spørsmålene ble forrige gang beregnet til å være .817, noe som indikerer tilfredsstillende reliabilitet. Man kan dermed bli fristet til å lage en samlet skåre for alle disse seks spørsmålene, dvs. at de seks spørsmålene betraktes som indikatorer på den samme underliggende latente variabelen.

Nå kan det imidlertid være slik at det er flere latente variabler som er årsak til samvariasjonen (korrelasjonene) mellom de seks spørsmålene. Hensikten med dagens øving er å undersøke om dette er tilfelle. For å undersøke dette benytter vi oss av faktoranalyse.

Oppgave 1. Kikk på spørsmålene og korrelasjonene mellom disse

Faktoranalyse tar som kjent utgangspunkt i korrelasjonene mellom de observerte variablene. I dette tilfellet dreier det seg om korrelasjonene mellom de seks holdningsspørsmålene. Vi tar en kikk på korrelasjonsmatrisen før vi foretar faktoranalysen. Først åpner dere datafilen ”**Datafil - øving faktoranalyse.sav**” som ligger på Fronter. Korrelasjonsmatrisen får du frem på følgende måte:

1. Velg **Analyze** på menylinjen og klikk deretter på **Correlate** og velg **Bivariate**.
2. Klikk alle seks variablene over i boksen *Variables*: og klikk deretter på **OK**-knappen.

A. Du får nå frem korrelasjonene mellom alle variablene i SPSS Output vinduet. Kikk på matrisen. Er enkelte grupper av variablene høyere korrelert med hverandre enn med andre variabler? I tilfelle ja, hvilke variabler tenderer til å være høyt korrelerte med hverandre? (*Tips til løsningsforslag: se siste side*).

B. Hvis de seks spørsmålene ikke var korrelerte med hverandre i det hele tatt ($r = 0.00$), hadde det da vært noen vist i å foreta faktoranalyse?

C. Synes du det er teoretisk begrunnet å utføre en faktoranalyse basert på de seks spørsmålene (les: synes du det er rimelig å betrakte de seks spørsmålene som indikatorer på en eller flere underliggende faktorer)?

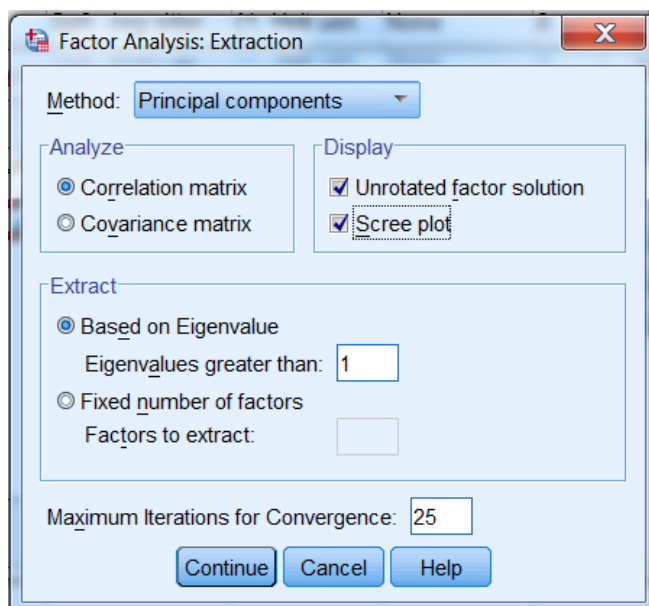
Oppgave 2. Faktoranalyse

Vi ønsker nå å undersøke hvorvidt korrelasjonene mellom de seks holdningsspørsmålene kan forklares ut i fra færre faktorer (eller dimensjoner), og hvilke spørsmål som evt. grupperer seg sammen (har høye faktorladninger fra samme faktor). Vi har ingen antagelse om den underliggende faktorstrukturen på forhånd. Hva kalles denne typen faktoranalyse?

	Sett ett kryss:
Konfirmatorisk	
Eksplorerende	

Slik foretar du faktoranalysen:

1. Velg **Dimension Reduction** under **Analyze** på menylinjen
2. Velg **Factor**
3. Klikk så de variablene som skal inngå i faktoranalysen over i boksen *Variables*:
4. Klikk deretter på boksen **Extraction**



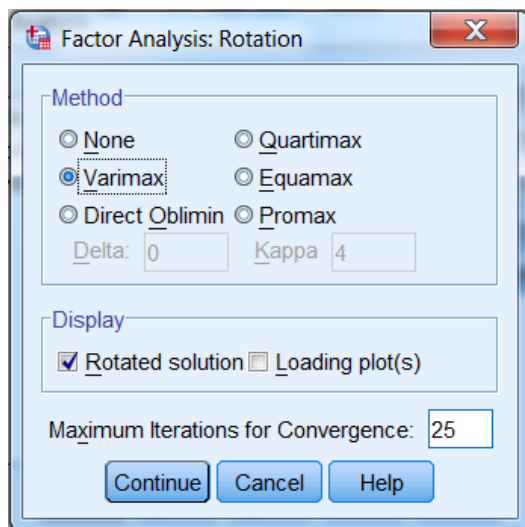
Extraction betyr utrekning, og vi blir bedt om å spesifisere hvilken metode vi skal bruke når vi trekker ut faktorene. I boksen for *Method*: står det principal components, og det er denne vi skal velge.

Under *Extract*: har vi to valgmuligheter: "Eigenvalues over" eller "Number of Factors". SPSS foreslår automatisk Eigenvalues over 1, og det er dette vi skal velge.

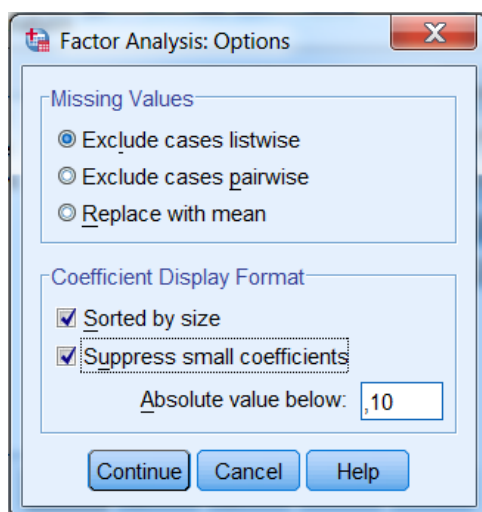
Velg også "**Scree Plot**"

Klikk deretter på **Continue**

5. Klikk deretter på boksen der det står **Rotation**. Her velger vi om vi skal rotere og i så fall hvilken type rotering av faktorene vi skal ha. Vi vil ha ortogonale faktorer i første omgang, og velger **Varimax**. Klikk deretter på **Continue**.



6. Klikk på boksen **Options** og kryss av for "Sorted by size" og "Suppress values less than". (Dette vil gjøre det lettere å tolke resultatutskriften). Klikk deretter på **Continue**.



7. Klikk til slutt på **OK**-knappen, og SPSS utfører faktoranalysen.

Før du kikker på resultatutskriften i SPSS Output, skal du svare på noen kontrollspørsmål, slik at du er sikker på at du forstår hva du har bedt SPSS om å gjøre.

- a. Under Extraction valgte du Eigenvalues over 1 som uttrekningskriterie. Dette betyr:

- ☐ SPSS trekker kun ut en faktor
- ☐ Faktorene som trekkes ut er ukorrelerte med hverandre
- ☐ SPSS trekker kun ut faktorer som har en Eigenvalue på 1 eller mer før rotering

- b. Under Extraction valgte vi også "Scree plot". Hva er et Scree plot og hva brukes det til?

c. Under Rotation valgte vi Varimax, noe som gir ortogonale faktorer. Ortogonale faktorer betyr at:

- ☐ Faktorene er korrelerte med hverandre
☐ Faktorene er ukorrelerte med hverandre

3. Studer utskriften og tolk resultatene

I SPSS Output vinduet har du nå flere tabeller. Den første heter Communalities, og viser kommunaliteten til de observerte variablene. "Extraction" betyr kommunaliteten som er beregnet ut i fra faktorladningene som vi kommer frem til.

Den neste tabellen, **Total Variance Explained**, inneholder mye informasjon.

Dette viser Eigenvalue og andel forklart varians **før** uttrekning. Legg merke til at det må like mange faktorer som variabler til for å forklare all variasjon i variablene.

Dette viser Eigenvalue og andel forklart varians **etter** uttrekning. Legg merke til at kun faktorer med Eigenvalue over 1 er tatt med.

Dette viser Eigenvalue og andel forklart varians **etter rotering**. Legg merke til at Eigenvaluene og prosent forklart varians er forandret for hver faktor. *Men total forklart varians er den samme*

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,177	52,949	52,949	3,177	52,949	52,949	2,416	40,271	40,271
2	1,259	20,976	73,925	1,259	20,976	73,925	2,019	33,654	73,925
3	,591	9,847	83,772						
4	,415	6,910	90,682						
5	,318	5,303	95,985						
6	,241	4,015	100,000						

Extraction Method: Principal Component Analysis.

Noen kontrollspørsmål (les ut av tabellen Total Explained Variance):

A. Hvor mange faktorer har SPSS trukket ut etter uttrekning og rotering?

Antall faktorer: _____

B. Hvorfor tror du at akkurat dette antallet faktorer er trukket ut?

C. Hva er Eigenvalue og andel forklart varians til den enkelte faktor **etter rotering**? Skriv opp dette for hver faktor.

Faktor: Eigenvalue (etter rotering): Andel forklart varians:

D. Hvor mye total forklart varians har faktormodellen?

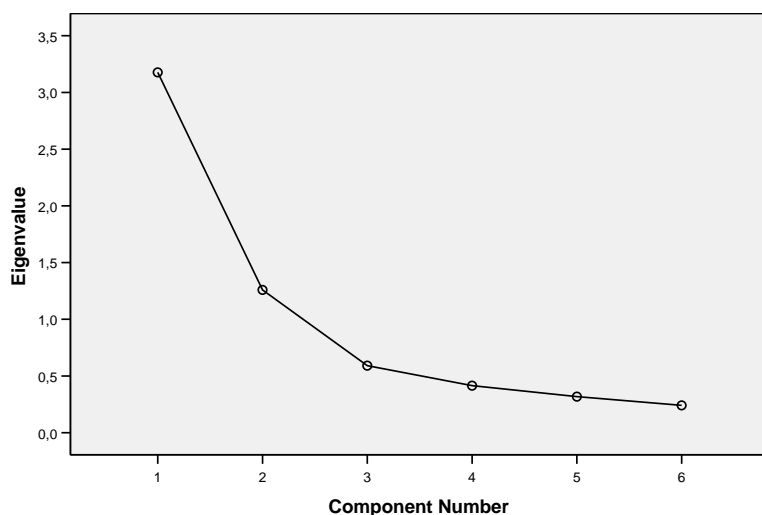
Total forklart varians: _____%

E. Hvorfor endres faktorenes Eigenvalue seg etter rotering?

- ☐ Fordi faktorladningene endrer seg når vi roterer faktorene
- ☐ Fordi total forklart varians endrer seg
- ☐ Fordi hver enkelt variabels kommunalitet endrer seg etter rotering

4. Scree plot

Har du fulgt instruksjonene riktig, skal du få et "Scree plot" som ser slik ut:



Vurder hvor grafen flater ut. Kan det være aktuelt å velge flere eller færre faktorer enn det som var tilfelle når vi brukte Kaisers kriterium?

5. Rotering

Neste tabell heter **Component matrix**, og viser faktorladningene til de uttrukne faktorene **før rotering**. Dette er egentlig ikke så interessant for oss, siden vi har bedt om akserotasjon (det ba vi om når vi valgte varimax under Rotation).

Den nest siste tabellen heter **Rotated Component Matrix**, og viser faktorladningene **etter at faktorene er roterte**. Ut i fra denne tabellen får vi de endelige faktorladningene. Ut fra denne tabellen kan man lettere se hvilke spørsmål som grupperer seg sammen, dvs. har høye faktorladninger på samme faktor. I tabellen er spørsmålene sortert etter styrke på faktorladning til den enkelte faktor (dette fikk vi gjennom å be om *Sorted by size* når vi bestilte faktoranalysen i SPSS).

A. Finn ut i fra denne tabellen hvilke spørsmål som grupperer seg på samme faktor:

B. Vil du si at resultatene etter rotering gir en "Enkel struktur?" Begrunn ditt svar:

C. Et av spørsmålene har negative faktorladninger (etter rotering). Hva tror du er årsaken til det?

6. Tolkning av faktorene

Siste del av en faktoranalyse er av ren teoretisk art, der vi tolker faktorene. Vi undersøker da den praktiske betydning av faktorene, dvs om de spørsmål som grupperer seg sammen har noe felles bortsett fra det å ha høye faktorladninger på samme faktor.

Vi kan også gi de to faktorene navn ut i fra de spørsmål som inngår i faktorene. Kikk også nøye på spørsmålene som grupperer seg sammen – og navngi faktorene ut i fra dette:

Navn Faktor 1: _____

Navn Faktor 2: _____

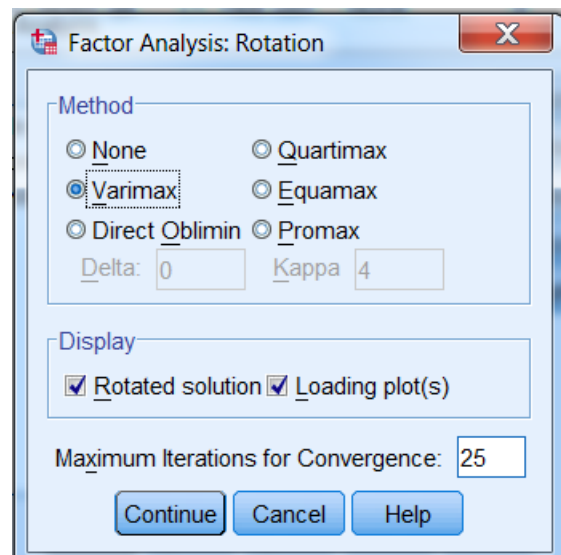
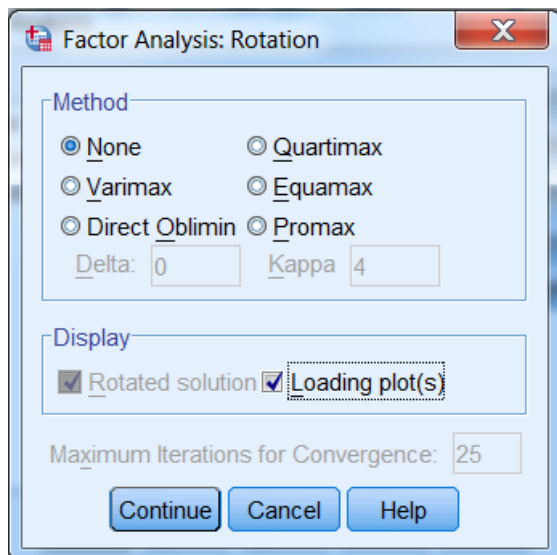
7. Loading plots (Vektordiagram)

Du kan også be om å få presentert variablene i et vektordiagram som tar utgangspunkt i faktorladningene. Dette bestilles gjennom å krysse av for "Loading plots" under **Rotation**.

NB! Så lenge du velger å rotere faktorene, vises kun den roterte løsningen i et slikt vektordiagram. For å se vektordiagrammet *før* rotering av faktorene, må du velge "None" under **Rotation**, dvs. en faktoranalyse uten noen rotering av faktorene.

Med andre ord, forta faktoranalysen på nytt to ganger. Denne gangen velger du å ikke rotere faktorene, men bestiller "Loading plot", den andre gangen velger du å rotere og bestiller samtidig også "Loading plot".

(Fortsettelse på neste side...)



Sammenlign ”Loading plot” før og etter rotering – hva har skjedd?

8 Lag skårer på hver faktor (hver person i datafila får en skåre på hver faktor)

Det er som kjent to hovedmåter å lage skårer på faktorene på; faktorskårer (vektede skårer) og faktor-baserte skårer (uvektede skårer).

Faktorskårer lager du gjennom å ”bestille” disse når du utfører faktoranalysen. Gjør det samme som du gjorde i oppgave 2 på nytt, men under **Scores..** klikker du på avkrysningsboksen for ”Save as variables”. Klikk deretter på **Continue** og så **OK**.

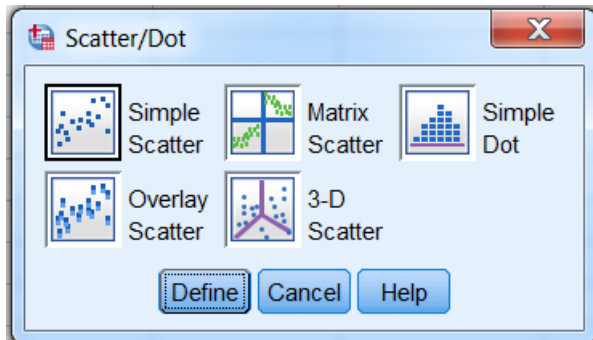
Du vil nå få to nye variabler i **datafila** – en skåre for hver faktor. Hva tror du en høy skåre på henholdsvis faktor 1 og faktor 2 betyr?

Hvis du ønsker å lage uvektede faktor-baserte skårer, så gjør du det gjennom å benytte **Compute** kommandoen som ligger **Transform** hovedmeny. Her kan du lage en *sumskåre* for hver faktor, basert på skårer på de variablene som inngår i den enkelte faktor. Vi har ikke skrevet opp fremgangsmåten her, men bruk den kunnskapen du tilegnet deg under forrige øving til å gjøre dette. Husk at ett av spørsmålene må rekodes før man lager en sumskåre (finn ut hvilket!)

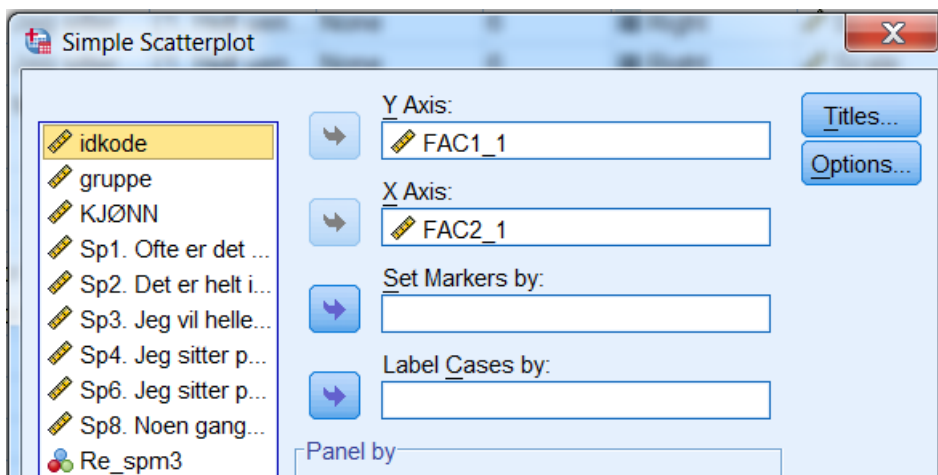
Etter at du både har laget faktorskårer og faktor-baserte skårer, kan du korrelere disse med hverandre. Blir korrelasjonen mellom faktorskårer og faktorbaserte skårer høy eller lav?

9. Presenter skårene på de to faktorene i et spredningsdiagram.

Du kan også studere hvordan hver person skårer på de to faktorene gjennom å lage et spredningsdiagram. Klikk på **Graphs, Legacy Dialoges** velg deretter **Scatter/dot** og velg ”Simple scatter”



Klikk på **Define** og klikk den ene faktorskåren over i **Y-axis** og den andre over i **X-axis**. Deretter **OK**



Studer spredningsdiagrammet (scatter plot). Hva sier dette deg?

10. Tilleggsoppgave: Oblik rotering

Den roterte faktormatrisen tyder på at de to faktorene har ladninger til alle de seks spørsmålene, selv om faktorladningene ikke er så veldig sterke i alle tilfellene. Imidlertid kan dette tyde på at vi bør velge en oblik faktorløsning - prøv ut faktoranalysen med oblike faktorer. Dette gjør du gjennom å velge *Direct Oblimin* under **Rotation**. Resten av faktoranalysen gjennomføres som før.

Utskriften du nå får skiller seg litt fra den du får når du velger en ortogonal løsning. Du får ikke en factor matrix, men to: en mønstermatrise (*Pattern Matrix*) og en strukturmatrikse (*Structure Matrix*). I tillegg vises korrelasjonen mellom de to faktorene nederst i utskriften i tabellen *Component Correlation Matrix*.

Når du skal tolke resultatene, er det mest oversiktlig å se på mønstermatrisen (pattern matrix) og deretter på korrelasjonen mellom faktorene (component correlation matrix).

Pattern Matrix^a

	Component	
	1	2
SP6_t1 Jeg sitter på med venner som kjører litt røtt så lenge jeg stoler på dem	,910	
SP4_t1 Jeg sitter på med en uforsiktig sjåfør hvis det ikke er andre måter å komme seg hjem på	,905	
SP3_t1 Jeg vil heller gå 5 km enn å sitte på med en som råkjører	-,873	
SP1_t1 Ofte er det slik at jo bedre kjøreferdigheter en bilfører har, jo forttere kan han eller hun kjøre		,823
SP2_t1 Det er helt i orden å kjøre over fartsgrensen hvis trafikkforholdene er bra	,103	,813
SP8_t1 Noen ganger må man tøy e reglene for å komme seg fram i trafikken		,804

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 4 iterations.

Component Correlation Matrix

Component	1	2
1	1,000	,412
2	,412	1,000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

Sammenlign resultatene ved oblik rotering med ortogonal rotering. Grupperer de observerte spørsmålene seg annerledes enn etter ortogonal rotering? Hvor høy er korrelasjonen mellom de to faktorene?

Tips til løsning av oppgave 1A.

For hver rad i korrelasjonsmatrisen, identifiser hvilket *variabelpar* som korrelerer *høyest* med hverandre. Noter ned *nummeret* på dette variabelparet (f.eks.: 1, 2). Hvis et variabelpar allerede er notert, noter ned hvilket variabelpar som har *nest høyest* korrelasjon på den aktuelle raden. Gjenta prosessen for hver rad.

Du skal nå ha notert ned seks variabelpar. - Ser du et mønster i hvordan variablene grupperer seg? Grupper variablene som korrelerer med hverandre.