ØF-notat nr. 17/2007

Det lille kvantitative metodeheftet

av

Vegard Johansen

Østlandsforskning

Østlandsforskning er et forskningsinstitutt som ble etablert i 1984 med fylkeskommunene og høgskolestyrene/de regionale høgskolesentra i fylkene Oppland, Hedmark og Buskerud som stiftere i samarbeid med Kommunaldepartementet.

Østlandsforskning er lokalisert i høgskolemiljøet på Lillehammer og har i tillegg kontorer i Hamar. Instituttet driver anvendt, tverrfaglig og problemorientert forskning og utvikling.

Østlandsforskning er orientert mot en bred og sammensatt gruppe brukere. Den faglige virksomheten er konsentrert om to områder:

> Næringsliv og regional utvikling Velferd, organisasjon og kommunikasjon

Østlandsforsknings viktigste oppdragsgivere er departement, fylkeskommuner, kommuner, statlige etater, råd og utvalg, Norges forskningsråd, næringslivet og bransjeorganisasjoner.

Østlandsforskning har samarbeidsavtaler med Høgskolen i Lillehammer, Høgskolen i Hedmark og Norsk institutt for naturforskning. Denne kunnskapsressursen utnyttes til beste for alle parter. ØF -notat nr. 17/2007

Det lille kvantitative metodeheftet

av

Vegard Johansen



Tittel:	Det lille kvantitative metodeheftet
Forfatter:	Vegard Johansen
ØF-notat nr.:	17/2007
ISSN nr.:	0808-4653
Prosjektnummer:	10123
Prosjektnavn:	NFR Basisbevilgning
Oppdragsgiver:	Grunnbevilgning fra NFR
Prosjektleder:	Vegard Johansen
Referat:	Dette notatet er basert på undervisningsnotater i forbindelse med undervisning i samfunnsvitenskapelig metode ved Høgskolen i Lille- hammer og Norges Teknisk - Naturvitenskapelige Universitet. I notatet gjennomgås noen grunnleggende måter å analysere kvantitative data på ved statistikkprogrammet SPSS.
	Notatet er beregnet for studenter, forskere og andre som har liten eller ingen erfaring med bruk av kvantitativ metode. I heftet redegjøres det kortfattet for forskningsprosessen og deretter fokuseres det på kvantita- tiv metode. Veiledningen er laget til versjon 14.0 av SPSS.
Emneord:	Forskningsprosess, statistiske begreper, survey, SPSS, beskrivende analyse, univariat analyse, bivariat analyse, estimering, feilmargin, hypotesetesting, kjikvadrattest, t-test, korrelasjon, multivariat analyse, omkoding, konstruksjon av sammensatt variabel
Dato:	Oktober 2007
Antall sider:	49
Pris:	Kr 120,-
Utgiver:	Østlandsforskning Serviceboks 2626 Lillehammer
	Telefon 61 26 57 00 Telefax 61 25 41 65 e-mail: post@ostforsk.no http://www.ostforsk.no
	Dette eksemplar er tremstilt etter KOPINOR, Stenergate 1 0050 Oslo 1. Ytterligere ek- semplarfremstilling uten avtale og strid med åndsverkloven er straffbart og kan medføre erstatningsansvar.

Forord

Dette notatet er basert på mine undervisningsnotater i forbindelse med undervisning i samfunnsvitenskapelig metode ved Høgskolen i Lillehammer og Norges Teknisk - Naturvitenskapelige Universitet. På bakgrunn av hyggelige tilbakemeldinger fra studenter og kolleger ved Østlandsforskning fikk jeg lyst til å gjøre mine notater tilgjengelig for enda flere.

Hensikten med dette notatet er å gjøre det praktiske analysearbeidet lettere for studenter og andre forskere som har liten eller ingen erfaring med bruk av kvantitativ metode. I notatet gjennomgås noen grunnleggende måter å analysere kvantitative data på ved statistikkprogrammet SPSS.

I heftet benyttes datasett fra tre prosjekter gjennomført av Østlandsforskning.

Evaluering av Jysla Løye i TastaSkolen. Et prosjekt om fysisk aktivitet i skolen Prosjektleder: Hans Olav Bråtå Oppdragsgiver: Stavanger kommune

Mot mer makt for kvinner i lokalpolitikken Prosjektleder: Ingrid Guldvik Oppdragsgiver: KS – Kommunesektorens interesse- og arbeidsgiverorganisasjon

Ung i Gjøvik Prosjektleder: Vegard Johansen Oppdragsgiver: Gjøvik kommune

Lillehammer, oktober 2007

Torhild Andersen forskningsleder

Vegard Johansen prosjektleder

Innhold

F	orord.		5
In	nhold	l	7
Fi	gurer	og tabeller	8
1	Ser	ntrale begreper	9
	1.1	Forskningsprosessen	9
	1.2	Innsamling av "survey-data"1	0
	1.3	Populasjon, sannsynlighetsutvelging og utvalg1	2
	1.4	Enhet, variabel, verdi, målenivå og årsakssammenhenger1	2
2	Int	roduksjon til SPSS 1	5
	2.1	Meny og ikoner i SPSS 1	5
	2.2	Registrering av svar i SPSS1	6
3	Bes	skrivende analyse	1
	3.1	Univariat analyse	1
	3.2	Bivariat analyse	4
4	Est	imering3	1
	4.1	Estimering av populasjonsgjennomsnitt	2
	4.2	Estimering av feilmargin	4
5	Hy	potesetesting	5
	5.1	Kjikvadrattesten	5
	5.2	t-test for to uavhengige utvalg	8
	5.3	Signifikanstest av korrelasjon4	1
	5.4	Vurdering av slutningsstatistikk	2

Figurer og tabeller

Figur 1. Forskningsprosessens faser (kvantitativ tilnærming)	10
Tabell 1. Huskeliste når vi utvikler spørreskjema.	11
Figur 2. Populasjon - nettoutvalg.	12
Figur 3. Fire målenivåer	13
Figur 4. Årsaksmodell	13
Figur 5. Meny og alternativer	15
Figur 6. Ikoner i SPSS	16
Figur 7. Datamatrisen	16
Figur 8. Variabelmatrisen	17
Figur 9. Resultatvinduet	17
Figur 10. Punching av variabelinformasjon	17
Tabell 2. Punching av variabelinformasjon	18
Figur 11. Punching av svar	18
Tabell 3. Kodebok eller allerede utfylte verdier i spørreskjema	19
Figur 12. Gjennomføring av univariat analyse av kategorisk variabel	21
Figur 13. Resultater av univariat analyse av kategorisk variabel	22
Tabell 4. Presentasjon av variabelen Ordførers kjønn.	22
Figur 14. Univariat analyse av kategorisk variabel	23
Figur 15. Resultater (output) av univariat analyse av kontinuerlig variabel	24
Tabell 5. Presentasjon av variabelen Kvinneandel i kommunestyret	24
Figur 16. Gjennomføring av bivariat analyse. Avhengig variabel "medlem idrettslag"	. 25
Figur 17. Resultater av bivariat analyse. Avhengig variabel "medlem idrettslag"	26
Tabell 6. Presentasjon av sammenhengen mellom kjønn og medlemskap i idrettslag	3.26
Figur 18. Gjennomføring av bivariat gjennomsnittsanalyse av variabel "skoletrivsel".	27
Figur 19. Resultater av gjennomsnittsanalyse av variabel "skoletrivsel"	28
Figur 20. Gjennomføring av Pearson korrelasjonsanalyse	29
Figur 21. Resultater av Pearson korrelasjonsanalyse	29
Figur 22. Slutte om fordelingen i populasjonen basert på utvalgsfordelingen	31
Figur 23. Illustrasjon av konfidensintervall.	32
Tabell 7. Tradisjonell beregning av konfidensintervall	32
Figur 24. Gjennomføre beregning av 95 prosent konfidensintervall	33
Figur 25. Tolkning av resultater for beregning av 95 prosent konfidensintervall	33
Tabell 8. Huskeregler angående variasjoner i feilmargin	34
Figur 26. Gjennomføring av moderne utregning av kjikvadrat	37
Figur 27. Resultater (output) av kjikvadrat-test	38
Figur 28. Gjennomføring av moderne utregning av t-test for to uavhengige utvalg	40
Figur 29. Resultater moderne utregning av t-test	40
Figur 30. Mulige feilslutninger ved slutningsstatistikk.	42

1 Sentrale begreper

I kapittel 1 gjennomgås følgende:

- Forskningsprosessen
- Tips for innsamling av survey-data
- Sentrale begreper ved gjennomføring av kvantitativ analyse

1.1 Forskningsprosessen

Forskningsprosessen kan deles inn i fire faser (se illustrasjon på neste side).

- **Forberedelse:** Her utvikles/formuleres forskningsspørsmålet ut fra et gitt tema, og så finner man ut hvordan man best kan besvare dette spørsmålet (valg av tilnærming).
- **Datainnsamling:** Finne relevante og pålitelige data som gjenspeiler virkeligheten som undersøkes, enten ved å samle inn egne data via survey eller å bruke tilgjengelige data.
- Analyse: Data bearbeides ved beskrivende statistikk og/eller slutningsstatistikk.
- **Rapportering:** Kan foregå både muntlig og skriftlig.





1.2 Innsamling av "survey-data"

En survey er en systematisk og strukturert utspørring av et, ofte stort, utvalg personer om et hvilket som helst tema. En survey-undersøkelse har fire faser:

Forberedelse

- a. Valg av tema for undersøkelsen og utarbeidelse av spørreskjema
- b. Prestudie
- c. Hvis populasjonen er stor trekkes et bruttoutvalg

Datainnsamling

- a. Valg av innsamlingsteknikk; telefon, post, besøkintervju eller nettbasert
- b. De som svarer på undersøkelsen utgjør nettoutvalget
- c. Enhetenes svar "punches"/registreres i statistikkprogram

Dataanalyse

- a. Beskrivende statistikk analyser bestemmes av variablenes målenivå
- b. Generaliserende statistikk analyser bestemmes av variablenes målenivå

Rapportering

- a. Skriftlig
- b. Muntlig

Spørreskjemaer er prekodede skjema med faste spørsmål og (ofte) faste svaralternativer. Denne standardiseringen innebærer at vi kan samle inn mye data på kort tid, undersøke likheter og variasjoner i respondenters svar, kartlegge fenomener, generalisere resultater og undersøke sammenhenger mellom fenomener.

Tabell 1. Huskeliste når vi utvikler spørreskjema.

Struktur	 Introdusere undersøkelsen på første side ved å fortelle om hensikten med undersøkelsen, hvem som gjennomfører den, eventuell godkjennelse av NSD eller etiske komiteer, svarfrist og kontaktopplysninger Spørreskjemaet må ikke bli for langt, da kan det tenkes at vi bare får med de respondentene som er mest interessert i temaet Rekkefølgen på spørsmålene kan pirre respondentens interesse, tema legges tydelig opp, og sensitive spørsmål tas inn underveis
Spørsmål	 Forholde oss til tidligere forskning for å velge ut relevante spørsmål Gi svar på forskningsspørsmålene Entydige, presist og enkelt formulert Ikke ledende
Svaralternativer	 Gjensidig utelukkende og mest mulig uttømmende Både positive og negative svarmuligheter Alternativet "vet ikke" ved vanskelige spørsmål
Prestudie	- Teste skjemaet på noen utvalgte respondenter
Punching	- Utforme skjemaet slik at det er lett å punche

1.3 Populasjon, sannsynlighetsutvelging og utvalg

Populasjon, sannsynlighetsutvalg og nettoutvalg er tre sentrale begreper.

Figur 2. Populasjon - nettoutvalg.



Bruttoutvalg – Nettoutvalg = *Frafall*.

Fremgangsmåten ved en spørreundersøkelse kan være slik:

- 1. Populasjonen defineres
- 2. Fra populasjonen trekkes et bruttoutvalg ved sannsynlighetsutvelging.
- 3. Innsamlingsmetode velges og innsamlingen gjennomføres
- 4. De som svarer på undersøkelsen utgjør nettoutvalget
- 5. Vi undersøker om nettoutvalget gjenspeiler populasjonen
 - a. *Svarprosent (netto/brutto):* Avhenger ofte av flere faktorer som innsamlingsmetode, typen populasjon, tema for undersøkelsen m.m.
 - b. Frafallsanalyse: Reduserer usikkerheten ved frafallet og innebærer å sammenligne utvalgsfordelingen med fordelingen i populasjonen på sentrale områder.

1.4 Enhet, variabel, verdi, målenivå og årsakssammenhenger

Enheter: Forskningsspørsmålene angir hvem vi skal vite noe om og i kvantitative undersøkelser kaller vi de som undersøkes for enheter. Enhetene er oftest mennesker, men kan også være dyr, ting, hendelser, land m.m. Hvis datainnsamlingen foregår ved hjelp av spørreskjema kaller vi enhetene for respondenter.

¹ Tre metoder er. 1. Enkel tilfeldig trekking (ETT): Fra en populasjon på N enheter trekkes et utvalg på n enheter tilfeldig, der hver enhet har lik sannsynlighet for å komme med i utvalget. 2. Stratifisert utvelging (to trinn): Populasjonen deles inn i grupper (strata) før utvalget trekkes ved ETT, og hensikten er å sikre at utvalget er representativt på stratifiseringsvariablene. 3. Klyngeutvelging (to eller flere trinn): Først trekkes det ut et utvalg av klynger, og deretter trekkes det utvalg av enheter fra hver klynge ved ETT.

Variabel: Fenomener ved enhetene beskrives ved at vi operasjonaliserer variabler. Operasjonalisering vil si å knytte teoretiske begreper til empiriske indikatorer; vi gjør generelle fenomener konkrete slik at de kan måles eller klassifiseres. Denne prosessen resulterer i at vi beskriver fenomenet med en eller flere variabler. Hvis enheten er mennesker er kjønn, utdanning, yrke og inntekt eksempler på typiske variabler.

Verdier: Variabler varierer med ulike verdier/kategorier som kan skilles logisk. Hvis vår variabel er kjønn kan vi skille mellom verdiene mann og kvinne.

Målenivå: Variabler med ulike egenskaper deles inn i et hierarki med fire målenivåer.

Målenivå	Karakteristikk	Eksempel
Nominalnivå	Gjensidig utelukkende kategorier	Kjønn
Ordinalnivå	Gjensidig utelukkende kategorier Kategoriene kan rangeres	20 0 Ikke Litt Noe Interessert interessert interessert <i>Politisk interesse</i>
Intervallnivå	Gjensidig utelukkende kategorier Kategoriene kan rangeres Lik avstand mellom kategoriene	20 - 10 - 0
Farholdstallsnivå	Gjensidig utelukkende kategorier Kategoriene kan rangeres Lik avstand mellom kategoriene Absolutt nullpunkt	100000 40000 ? Inritekt

Figur 3. Fire målenivåer.

Å**rsakssammenhenger:** Noen ganger ønsker vi å undersøke sammenhenger mellom variabler, og ofte er det en retning på disse sammenhengene. D.v.s. at ett fenomen forklares ved ett eller flere andre fenomener. Vi snakker ikke om lovmessighet, men ulik grad av sannsynlighet for at dersom X inntreffer så inntreffer også Y.

Figur 4. Årsaksmodell



Når vi ønsker å "forklare" variasjonen i et fenomen er det en rekke krav som bør innfris:

- X skal komme før Y
- Samvariasjon mellom X og Y
- Kontroll for andre variabler (X_1, X_2, X_n)

2 Introduksjon til SPSS

I kapittel 2 gjennomgås følgende:

- Meny og ikoner i SPSS
- Registrering av svar i SPSS

2.1 Meny og ikoner i SPSS

Øverst i SPSS-vinduet er det en menylinje med 10 forskjellige alternativer:

Figur 5	5. Meny	og alterr	nativer
---------	---------	-----------	---------

🖬 *Untitled1 [DataSet0] - SPSS Data Editor											
File	Edit	View	Data	Transform	Analyze	Graphs	Utilities	Window	Help		

Filehar alternativer for å lagre, skrive ut resultater og hente inn nye filerEdit gir deg muligheten til å kopiere (f.eks. kopiere tall eller datamatrise)View gir deg mulighet til å tilpasse utseendet av SPSSDatalar deg endre datamatrisen systematiskTransformhar kommandoer for å endre verdier og variablerAnalyzeer menyen for å kjøre statistisk analyse av datamaterialetGraphsgir tilgang til grafer og grafisk analyseUtilitiesgir informasjon fra datamatrisen og SPSSWindowgir muligheten til å holde orden på de ulike vinduene i SPSSHelp gir tilgang til hjelpemenyen i SPSS

i igui ei merier						
Open file	Goto	Case Ins	ert Variable			
Print	Undo	Find	Weigł	t Cases		
🔁 *Untitled1	[DataSet0] - S	PSS Data Edi	tor			
Fie Edit View	Data Transfor	m Analyze G	iraphs Utilities	Window	Help	
🖕 📮 🗐 🛯	i 🔸 🔶 🖥	= 🛛 🗚 י	作曲 🛱 🖗	b 🛒 🖻	<u>i</u> @+	Use Sets
Save file	Redo	Inse	rt Cases S	elect Cas	es Va	lue Labels
Dialo	g recall 🛛 🛛 🛛 🛛	ariables	Split	file		

Figur 6. Ikoner i SPSS

<u>Open file:</u> Åpner filer avhengig av hvilket vindu som er aktivt <u>Save file:</u> Lagrer filer avhengig av hvilket vindu som er aktivt <u>Print:</u> Skriver ut vinduet som er aktivt (brukes oftest på resultater) <u>Dialog recall:</u> De siste "dialogboksene" huskes av SPSS og kan fås tilbake <u>Undo:</u> Her kan du angre på operasjoner en har utført, såfremt de ikke er lagret <u>Redo:</u> Her kan du gjøre om igjen operasjoner som du har angret på <u>Goto Case:</u> Her kommer det opp en boks og du kan velge hvilken enhet du vil se på <u>Variables:</u> Her er en boks der du kan finne informasjon om samtlige variabler i datafilen <u>Find:</u> Her kan du søke etter verdier i de ulike ved å plassere musa i kolonnen du ser på <u>Insert Cases:</u> Ved å klikke på en bestemt enhet i datavinduet setter du inn en ny enhet <u>Insert Variables:</u> I datavinduet kan innholdet vises som eller verditeksten som er definert

2.2 Registrering av svar i SPSS

I SPSS er det tre vinduer vi kan møte²:

Fig	Figur 7. Datamatrisen																		
ι	Intitl	ed 1 (I	DataS	et0] -	SPS:	S Da	ita Ed	litor											
File	Edit	Edit View Data Transfori			form	Analyze Graphs Utilities Window Help						Dobbelklikk og du kommer til variabelmatrisen				nmer			
Þ	≥ 🖬 🏯 📴 🔶 🐜				*	62	🖗 👫 🏦 🏦 🖽 🖽 🐺 🎯												
1:																			
		Vi	ar	1	/ar		٧a	ď		var			var		Y	ar		var	
	1			<u> </u>		_			_										_
	2		~			+			+					_					+-
4	G	nta Vie	N.W	/ariabl	e Viev	NJ			1										

Data view er utformet som et ruteark med vertikale kolonner og horisontale rader:

- a. Hver kolonne representerer en variabel
- b. Hver rad representerer en enhet

² I tillegg har vi Syntax-vinduet der man må skrive inn kommandoer. Dette vinduet har sine fordeler når man skal ha store, presise operasjoner, men bruk av syntax-vinduet skal vi ikke si mye mer om.

Figur 8. Variabelmatrisen

File Edit	View Data	Transform Anal	yze Graph	ns Utilities Win	dow Help								
ا 🖨	● 🛯 🚔 🖶 🖗 梯 挿 莆 🖩 亟 再 ≫ 🔕												
	Name	Туре	Width	Decimals	Label	Values	Missing	Columns	Align	Measure			
Variabel 1				1									
Variabel 2													
3													
4													
5													
)ata View 🔍	ariable View		· · · · · · · · · · · · · · · · · · ·		<							
Dobbelkli			Antall tec	an La	ngt navn på var	iabel				Målenivå			
log du kommer :	til på vari	abel Variabe	eltype) 🖟	Antall desimal	er	Navn på tallve	rdier						
datamatri	isen .		=										

Variable View gir oss opplysninger om variablene som er med i datasettet, herunder:

- a. informasjon om variablenes verdier
- b. informasjon om variablenes målenivå

Figu	igur 9. Resultatvinduet													
ie c	Cutput1 - SPSS Viewer													
File	Edit	View	Data	Trans	form	Insert	Forma	at Ai	nalyze	Graphs	Utilities	Window	Help	
⊳	8	ê R		 ;	•	1	= [?	0	垕	•				
+	٠	+ -			7									
	•	Output												
	Kata viha	alog ov ar kjør	verko t	mma	ndo	ene		Res	ultaten	e av vår	e bereg	ninger		

Output gir oss resultatene av de beregningene SPSS har utført. Outputvinduet kommer automatisk opp når vi kjører en beregning.

I SPSS bruker vi Variable view og Data view når vi skal registrere svar.

Variable view brukes til å registrere informasjon om variablene våre.

🖬 *Untitled1 [DataSet0] - SP	SS Data Editor	,						
File Edit View Data Iransform	Analyze Graphs L	<u>I</u> tilities <u>W</u> indow	Help					
Name Type	Width Decimals	Hva er din høve	Label	Values 1. Grunnskole	Missing	Columns 8	Align Right	Measure
Variable Type	Width: 8 al Places: 0	Cancel	(alue Labels Value Labels Value: Label: Add 1 = "Grunnskole" 2 = "Videregåenc 3 = "Bachelorgrad" 4 = "Mastergrad"	' de skole"	Canc		1.03.	✔ Scale ✔ Ordinal ✔ Nominal

Figur 10. Punching av variabelinformasjon

Hva?	Symbol	Forklaring
Name:	Name Dobbelklikk	Kort navn på variabel i ett ord
Туре:	Type Numeric Trykk her	 Avgjøre hvilken type data vi skal registrere Numeric = talldata String = tekstdata Width: Hvor mange tegn vi skal ha plass til Decimal Places: Antall desimaler
Label:	Label Dobbelklikk	Registrere spørsmålet vi har stilt
Values:	Values Value Labels Value: 2 Label: Kvinne Add 1,00 = "Mann" Oppsamlingsboks	Om vi har talldata kobler vi verdinavn til tall gjennom boksen Value Labels. Det er tre rom: Value: Her skriver vi inn verdien Value Label: Her skriver vi inn det verdinavnet vi vil ha på verdien Oppsamlingsboks: Når vi har fylt ut Value og Label trykker vi på Add, og det nye verdinavnet legges inn i oppsamlingsboksen Eksempel: I variabelen Kjønn er verdi 1 = Mann og verdi 2 = Kvinne. Vi setter inn 2 i Value, Kvinne i Label, og trykker Add
Measure	Measure Klikk på pilen Scale Ordinal Nominal	Klikk på pilen for Measure og vi får da tre alterna- tiver til målenivå: <i>Scale:</i> Om variabel er på forholdstallsnivå <i>Ordinal:</i> Om variabel er på ordinalnivå <i>Nominal:</i> Om variabel er på nominalnivå

Tabell 2. Punching av variabelinformasjon

Data view brukes til å registrere respondentenes svar.





I figuren viser vi noen eksempler ut fra tre variabler: Utdannelse, Kjønn og Alder. For å være effektive puncher vi verdiene til hver enkelt enhet (vertikalt). Når vi puncher benyttes enten en *kodebok* eller at man allerede har skrevet inn *verdier i spørreskjemaet*. Se eksempler nedenfor.

	ing to relater reporteengema
Kodebok	Avkrysset spørreskjema
Utdannelse	Hva er din høyeste fullførte utdannelse?
1 = Grunnskole	1 Grunnskole
2 = Videregående skole	🔀 Videregående skole
3 = Bachelorgrad	3 Bachelorgrad
4 = Mastergrad	4 Mastergrad

Tabell 3. Kodebok eller allerede utfylte verdier i spørreskjema

3 Beskrivende analyse

Vi ser i kapittel 3 på fordelingsanalyse for en variabel (univariat) og to variabler (bivariat).

3.1 Univariat analyse

Variablenes målenivå bestemmer hvordan vi analyserer dem. Kategoriske variabler, d.v.s. variabler på nominal- eller ordinalnivå, analyseres ut fra prosentuerte tabeller eller grafer. Gjennomføringen illustreres i figuren under (eks. kjønn på ordfører):



Figur 12. Gjennomføring av univariat analyse av kategorisk variabel

Den andre delen handler om å tolke Output-feltet - resultatene som vises i figuren på neste side.

Figur 13. Resultater av univariat analyse av kategorisk variabel Statistics Ordførerkjønn

Ν	Valid Missing	433 C	} <mark>- Antall</mark>) - Antall	<u>svar</u> "hull" i data	amatrisen
			Ordføre	erkjønn	
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mann	358	82,7	82,7	82,7
	Kvinne	75	17,3	17,3	100,0
	Total	433	100,0	100,0	
		Antallet	Prosent	Prosent valide svar	Kummulativ prosent

Frequency = Faktisk observerte antallet enheter med ulike verdier

Percent = Frekvensen delt på det totale antallet enheter

Valid Percent = Frekvensen delt på enhetene med oppgitt verdi

Cumulative Percent = Summen av prosentandelene fra den første verdien til den aktuelle verdien en

Dersom vi skal bruke resultatene som vises i Output-vinduet i en rapport bør vi redigere den først. Det vi kan gjøre er å kopiere resultatene fra SPSS til andre program (eks. Word/Excel) og så lage grafer/tabeller. I figuren under presenteres resultatene fra vår analyse av variabelen *Ordførerkjønn*.



 Tabell 4. Presentasjon av variabelen Ordførers kjønn.

Kort konklusjon

Nesten 5 av 6 ordførere er menn, mens 1 av 6 ordførere er kvinner.

Når vi har kontinuerlige variabler, d.v.s. variabler på forholdstalls-, intervall- eller ordinalnivå med mange verdier, blir det for mange verdier til å få en god oversikt over fordelingen ved å se på hver enkelt verdis prosentandel. Vi benytter statistiske mål på fordelingens tyngdepunkt (gjennomsnitt, median modus) og spredning (variasjonsbredde, kvartiler, standardavvik). I vårt eksempel brukes en variabel som viser til andel kvinner i kommunestyret:

*ммм	l utvidet un	dersøkelse	: [DataSet1] ·	- SPSS Dat	a Editor						
Edit	View Data	Transform	Analyze Grap	hs Utilities	Window	Help					
. 🔡	🖹 🖬 🖷	🔶 🐜	Reports	•	k 📼	RA					
12			Descriptive !	Statistics 🔸	Frequ	iencies	🔲 Freq	uencies			
Komi	mnr		Tables	•	Descr	iptives				V-2-11-(-)	
	Kommnr	Na	Compare Me	eans 🕨	Exploi	re <u>1</u>	🛷 Kon	nmunenr (Komn 📥	ŕ	vanable(s):	
1	101	Halden	General Line	ar Model 🕨	Cross	tabs	🛃 🛃	/n 🚬 🖳			nstyre Paste
2	104	Moss	Mixed Model	IS P	Ratio		💑 Pro:	sentkykomm	\sim		
3	105	Sarpsborg	Pegression		4	45 "Sarpsbor	🍓 Sva	artkommune			<u>R</u> eset
4	106	Fredriksta	Loglinear		4	42 0	🛷 Sva	arkomm2	<u> </u>		Cancel
5	111	Hvaler	Classify	•	2	24 0	💞 Рук 🔗 Раб	.e alkning			
6	118	Aremark	Data Reduct	ion 🕨	3	35 0		ukning 11			Help
7	119	Marker	Scale	•	ę	56 "Marker"	→ v:		L		
8	121	Rømskog	Nonparamet	ric Tests 🕩		31 0	🔽 Displa	au frequencu tables			
9	122	Trøgstad	Time Series	•		36 0		by mequency (abies			
10	123	Spydeberg	Survival	•	4	48 "Spydebe			Statistics.	Charts	Format
11	124	Askim	Multiple Res	ponse 🕨	3	31 0					
12	Frequencie	s: Statisti	cs				3				
13	, in the particular							Frequencies:	Charts		
14	Percentile	Values		Central 1	fendency	Continue		Ch. J. T.			1
15	🔽 <u>Q</u> uartile	s		Mear	n	Conned		Chart Type-	· · ·	Continue	
16	Cut noir	nts for: 10		Medi	an	Cancer		O None		Cancel	
17			equal groups		-	Help		O <u>B</u> ar charts			
18	Percent	tile(s):		Mode Mode	9			O <u>P</u> ie charts		нер	
19	A	dd		<u>S</u> um				<u> H</u> istograms			
20	<u>C</u> ha	ange						✓ With no	rmal curve		
21	Rer	nove		Values		enide sints					
22				- vajues	are group	millipoints		- Chart Values-			
23	Dispersion			Distributi	on			0.5		Descenteres	
24	Std. der	viation	Minimum	Ske <u>v</u>	vness					rercentages	
25	<u>V</u> arianc	e 🗌	Ma <u>x</u> imum	Kurte	sis						J.
26	✓ Range		S. <u>E</u> . mean								

Figur 14. Univariat analyse av kategorisk variabel

a. Analyze + Descriptive Statistics + Frequencies

b. Vi velger variabelen "Kvinnandelkommstyre" og trykker på piltegnet

c. Vi merker ikke av for Display frequency tables

d. Vi velger statistiske mål: Trykk på *Statistics* + *Quartiles* + *Mean* + *Median* + *Mode* + *Std*. *Deviation* + *Range* + *Continue*

e. Vi viser fordelingen i et histogram: *Charts* + *Histograms* + *With normal curve* + *Continue* f. Trykk *OK*

Den andre delen handler om å tolke Output-feltet, d.v.s. resultatene som vises på neste side

Figur 15. F	Resultater	(output) a	v univariat	analyse av	, kontinuerlig	variabel
Kvinnandelkom	mstvre					

N	Valid	433 Antall svar
	Missing	0–
Mean		34,94 🕂 Gjennomsnitt: Summen av alle verdiene fordelingen dividert med antall enheter
Median		35,00 🕂 Median: Verdien som splitter en ordnet fordeling i to like store mengder av enheter
Mode		33 🕂 Modus: Den hyppigst forekommende verdi eller verdiklasse i en fordeling
Std. Deviation		7,824 Variasjonsbredde: Differansen mellom den høyeste og laveste verdien i en fordeling
Range		42 - Standardavvik: Det typiske avviket fra gjennomsnittet: Beregnes som kvadratroten til
Percentiles	25	29,00 summen av de kvadrerte avvik fra gjennomsnittet for alle enheter i utvalget
	50	35,00 Kvartiler. Variabelverdier som splitter en rangert fordeling i fire like store deler enheter
	75	41,00-

Dersom vi ønsker å bruke tabellen/figuren som vises i Output-vinduet bør vi redigere den først. Vi kan presentere resultatene av variabelen *Kvinneandel i kommunestyret* på ulike måter.



 Tabell 5. Presentasjon av variabelen Kvinneandel i kommunestyret.

Kort konklusjon

Variasjonen mellom kommunene er stor – 42 prosentpoeng – men halvparten av kommunene har mellom 29 og 41 prosent kvinnelige representanter. Det typiske avviket fra gjennomsnittet er lite, under 8 prosentpoeng. Sentraltendensmålene ligger svært nært hverandre.

3.2 Bivariat analyse

Valg av bivariat analyse avhenger av variablenes målenivå. Når begge variabler er på nominalnivå eller ordinalnivå med få verdier benyttes prosentuerte krysstabeller/grafer. Eks. Medlemskap i idrettslag (ja/nei) i forhold til kjønn (mann/kvinne).



Figur 16. Gjennomføring av bivariat analyse. Avhengig variabel "medlem idrettslag"

Den andre delen handler om å tolke Output-feltet (neste side).

- Den øverste tabellen forteller hvor mange enheter som har valide verdier, det vil si at de inkluderes i analysen, og hvor mange som ikke har valide svar.
- Den neste tabellen tolkes vertikalt: 59 prosent av guttene og 48 prosent av jentene er medlem av idrettslag, totalt er 52 prosent medlem i idrettslag.

		Cases										
			Valid				sing		Total		al	
		N		Perce	ent	N		Percen	t	N		Percent
Er du medlem i idrettslag * Kjønn		12	288	99,2%			10	,8	%	129	98	100,0%
											Ϊ	\sim
ir	Ant: iklud	allet dert i	som anal	er γsen	A eri	ntallet s nkluder	om ik tian:	ke alγsen		Ar andeien	talle The	et og ele utvalget
Andelen av utvalget som er inkludert i analysen Er du medlem i idrettslag * Kjønn Crosstabulation								en]]]				
						Ŕ	jønn)	r -			L a	5
						Gutt	Τ	Jente		Total	i ja	ת
Ær du medlem 👌 🛚	lei	С	ount			227		387		614	/ P	
Nidrettslag		%	with	in Kjønr	1 I	41,2%		52,5%		47,7%	<u> </u>	
	а	С	ount			324		350		674	j.	5
		%	with	in Kjønr	n	58,8%		47,5%		52,3%	l an	5
Total		С	ount			551		737		1288		:
		%	with	in Kjønr	1	100,0%		100,0%		100,0%		
Avhengig variabel											_	

Figur 17. Resultater av bivariat analyse. Avhengig variabel "medlem idrettslag" Case Processing Summary

Dersom vi ønsker å bruke tabellen som vises i Output-vinduet bør vi redigere den først. Vi kan presentere resultatene av den bivariate analysen på ulike måter.



Tabell 6. Presentasjon av sammenhengen mellom kjønn og medlemskap i idrettslag.

Kort konklusjon: Gutter i utvalget er oftere medlem av idrettslag komparert med jentene.

Andre ganger er avhengig variabel på forholdstalls-, intervall- eller ordinalnivå med mange verdier, mens den andre variabelen er på ordinalnivå med få verdier eller nominalnivå. Ved slike tilfeller kan vi sammenligne statistiske mål som gjennomsnitt, median og standardavvik. Analysen går ut på å studere forskjeller mellom verdier på uavhengig variabel, og gjennomføres slik:





Den andre delen handler om å tolke Output-feltet (neste side).

Case Processing Summary												
		Cases										
		Included				Excluded					Total	
		N		Percent		N		Pe	Percent		N	Percent
Hvor godt triv på skolen? *	es du ' Skole	115	51	88,7%		147	11,3%		1298	100,0%		
Antallet som er inkludert i analysen er inkludert i analysen andelen i hele utvalget									et og ele utvalget			
	Andelen av utvalget som er inkludert i analysen											
Hvor godt triv	es du på	skolen?	R	Report								
Ungdomsski	ole	Mean	ו ו	Std. D	eviatio	on	N					
Ungdomsski	ole	8,	,18	1,8		82		587				
Videregåend	e skole	7,		1,7	18		564					
Total		7,	,96	-	1,8	17	1	151				
Uavhengig	Uavhengig variabel Standardawik											
	Gj	ennomsr	nittsta	Illene	Ant	allet n	espon	dente	er			

Figur 19. Resultater av gjennomsnittsanalyse av variabel "skoletrivsel" o n : o

Den øverste tabellen forteller hvor mange enheter som har valide verdier, det vil si at de inkluderes i analysen, og hvor mange som ikke har valide svar. Tabellen under er den vi konkluderer ut fra:

- a. Gjennomsnittet for elever på ungdomsskolen i utvalget er 8,2 mens gjennomsnittet er 7,7 for elever på videregående skole
- b. Standardavvikene viser videre at svarene til elevene på ungdomsskolen er noe mer spredt enn svarene til elevene på videregående skole
- c. Den siste delen N viser til antallet respondenter

Når begge variabler er på forholdstalls-, intervall- eller ordinalnivå med mange verdier, kan vi foreta Pearson korrelasjonsanalyse.



R = 1 R = 0R = -1

Korrelasjonstesten gjennomføres slik. I eksempelet på neste side ser vi på andelen kvinner i kommunestyret og andelen kvinner i formannskapet.

	gui 20. C	ajerinion.	noring av i cars	01	попсазјонз	anaryse	
TΜ	MM utvidet	undersøkel:	se [DataSet1] - SPSS D)ata	a Editor		
dit	View Data	Transform	Analyze Graphs Utiliti	es	Window Help		
3	🔒 🖬 🖢	•	Reports Descriptive Statistics	*		Bivariate Correlations	×
ntgr	uppeleder	9	Tables	۲		Kommunenr [Komm]	
	Kommnr	Na	Compare Means Coperal Lipper Model	ł	innandelk Svartko	Svarkomm2	ns <u>P</u> aste
1	627	Røvken	Mixed Models	÷	mmstyre une 13 Ravker	Befolkning	<u>R</u> eset
2	2012	Alta	Correlate	•	Bivariate	I I I I I I I I I I I I I I I I I I I	Cancel
3	1001	Kristiansa	Regression Loglinear	+	Partial Distances		
4	1253	Osterøy	Classify	+		✓ KignnΩPU2	
5	1241	Fusa	Data Reduction	۲	32 "Fusa"		
6	236	Nes	Scale	۲	40 "Nesi.		
7	1124	Sola	Nonparametric Tests	۲	29 "Sola "		
8	211	Vestby	Time Series	×	34 "Vestb	Test of Significance	
9	1201	Bergen	Survival	1	40 "Berge	Two-tailed	o.c. 1
10	1627	Bjugn	Multiple Response	•	31 "Bjugn'	Elag significant correlations	Uptions

Figur 20, Giennomføring av Pearson korrelasionsanalyse

a. Analyze + Correlate + Bivariate

b. Vi merker av for "Kvinnandelkommstyre" og "Prosentkvinnerformannskap" og trykker på pilen.

- c. Vi haker av for Pearson som korrelasjonskoeffisient
- d. Trykk OK

Den andre delen handler om å tolke Output-feltet.

	Correlations	Kvinnandel kommstyre	Prosentkvinne rformannskap	Korrelasjonskoeffisienten r
Kvinnandelkommstyre	Pearson Correlation	1	,332-	
	Sig. (2-tailed)		,000-	Signifikansnivå
	N	433	330~	
Prosentkvinnerforman	Pearson Correlation	,332	1	Antall enheter
nskap	Sig. (2-tailed)	,000		
	N	330	330	

I den første ruten og siste ruten ser vi korrelasjonen for hver variabel satt opp mot seg selv, og korrelasjonen er naturligvis 1. De rutene som er interessante er når variablene krysser hverandre. Vi ser i rute 2 (og 3) at Pearson Correlation er 0,33, signifikansnivået er lavere enn 0,000 og antallet enheter (N) er lik 330. Vi konkluderer med at korrelasjonen mellom andel kvinner i kommunestyret og andel kvinner i formannskapet er svak og positiv (+ 0,33).

Boks 1. Andre korrelasjonsmål

Det finnes også en rekke andre korrelasjonsmål som måler styrken i sammenhengen mellom to variabler. Når en avgjør hva slags korrelasjonsmål en ønsker å bruke må en se på målenivået og det er den variabelen med det laveste målenivået som bestemmer hva slags mål vi kan bruke. Vi skal ikke se eksplisitt på disse målene, men gi en kort oversikt over dem.

- a. Nominalnivå: Phi og Cramer's V
- b. Ordinalnivå: Gamma, Kendalls's tau-b og Kendall's tau-c
- c. Kontinuerlig variabel: Pearson r (som vi har gjennomgått)



4 Estimering

Når vi bruker et utvalg av populasjonen som grunnlag for våre analyser er det knyttet en viss usikkerhet til resultatene når vi generaliserer fra utvalget til populasjonen. Vi kjenner jo ikke fordelingen i populasjonen, men vi kan generalisere resultater fra utvalget til populasjonen.



Figur 22. Slutte om fordelingen i populasjonen basert på utvalgsfordelingen

Generalisering

- a. Estimere: Anslå en parameter i populasjonen ved hjelp av en estimator i utvalget.
- b. Hypotesetesting: Undersøke om forskjeller mellom utvalg kan generaliseres til å gjelde for populasjoner.

4.1 Estimering av populasjonsgjennomsnitt

Når vi har kontinuerlige variabler (forholdstall, intervall og ordinalnivå med mange verdier) beregnes utvalgsgjennomsnittet og standardavvik. Vi kan bruke variabelens utvalgsgjennomsnitt og standardavvik til å estimere populasjonsgjennomsnittet. Tre forhold spiller inn på hvor presist dette estimatet blir:

- a. Utvalgets størrelse: Flere enheter gir sikrere estimat
- b. Spredningen: Dess større standardavvik, dess større estimatområde
- c. Valg av sikkerhetsintervall: Dess sikrere vi vil være, dess større estimatområde

I slutningsstatistikk er det en konvensjon å benytte seg av en sikkerhetsmargin på 95 prosent når vi beregner populasjonsgjennomsnitt. Det er da 2,5 prosent sannsynlighet for at populasjonsgjennomsnittet er større estimatet og 2,5 prosent sannsynlighet for at populasjonsgjennomsnittet er mindre. Sikkerhetsmarginen kalles konfidensintervall.

Ved 1,96 standardfeil vil 95 prosent av de mulige utvalgsgjennomsnittene vi kunne trukket, befinne seg. Det er m.a.o. 95 prosent sannsynlig at vi har rett når vi sier at populasjonsgjennomsnittet befinner seg innenfor et område på +/- 1,96 standardfeil.





Vi skal nå vise tradisjonell og moderne utregning av konfidensintervall.

Tabell 7. Tradisjonell beregning av konfidensintervall

Beregne standardfeilen (Samplingsfordelingens estimerte standardavvik)	Standardfeil til gjennomsnittet = <mark>Standarda∨vik</mark> √Antall enheter							
Eks: Snitthøyden for jenter i e	t utvalg på 200 er 168 cm. Standardavvik på 10 cm.							
1SF = <u>Standarda∨vik</u> = - √Antall enheter -	$1SF = \frac{\text{Standardavvik}}{\sqrt{\text{Antall enheter}}} = \frac{10}{\sqrt{200}} = 0,7$							
95 % sikkerhetsmargine	en = 0,7 * (+/- 1,96 standardfeil) = +/- 1,4 cm							
Konklusjon: Det er 95 prosent blant jenter i populasjonen er i	sikkert når vi konkluderer med at gjennomsnittshøyden intervallet <u>166,6 til 169,4 cm.</u>							

Vi kan også beregne et 95 prosent konfidensintervall i SPSS. I eksempelet vårt har et utvalg av elever i en norsk storby krysset av for hvilke aktiviteter som de har hatt i gymtimene. I alt kan man krysse av for 10 aktiviteter. Vi skal beregne gjennomsnittet for populasjonen av elever på ungdomsskolene i storbyen i forhold til antall aktiviteter.



Figur 24. Gjennomføre beregning av 95 prosent konfidensintervall

d. Trykk OK

Den andre delen handler om å tolke output-feltet.

Figur 25. Tolkning av resultater for beregning av 95 prosent konfidensintervall



I første tabell er det to viktige resultater: Gjennomsnittet = 3,97 og Standardavvik = 2,58. I andre tabell ses konfidensintervallet: Gjennomsnittet i populasjonen kan bli så lavt som 3,75 og så høyt som 4,19.

Konklusjonen: Med 95 prosent sannsynlighet har populasjonen av elevene i storbyen gjennomført mellom 3,75 og 4,2 av de 10 utvalgte aktiviteter.

4.2 Estimering av feilmargin

Ved variabler på nominal eller ordinalnivå med få verdier estimeres konfidensintervall for prosentfordelinger i utvalg. Vi foretar kun tradisjonell beregning av feilmarginer.

Eks: I en landsdekkende undersøkelse, med et utvalg på 700, sier 9 prosent av ungdom mellom 13 og 15 år at de mobbes. Hva er feilmarginen med 95 prosent sikkerhet? Først beregnes standardfeilen til prosentandelen.

$$SF(p) = -\sqrt{\frac{p(100-p)}{n}} \qquad SF(p) = -\sqrt{\frac{p(100-p)}{n}} = \sqrt{\frac{9(100-9)}{700}} = \sqrt{\frac{1,17}{1,17}} = 1.08$$

Så beregnes feilmarginen, og den er gitt ved: +/- 1.96 SF * 1.08 = 2.1 prosent.

Feilmarginen blir mindre med økende	Antall		Potensiell feilmargin ved 95 % sikkerhet			
	500 respondenter	+/- 4,4				
utvalgsstørrelse	1000 respondenter				+/- 3,1	
	10000 respondenter				+/- 1	
Feilmarginen			Feilmar	gin ved 95 % s	ikkerhet	
resultatet	Antall	1	0 prosent	50 prosent	90 prosent	
	500 respondenter		+/- 2,6	+/- 4,4	+/- 2,6	
	1000 respondenter		+/- 1,9	+/- 3,1	+/- 1,9	
Feilmarginen		F	otensiell fei	Imargin ved 95	i % sikkerhet	
størrelsen på	Antall	20	000 i pop	10000 i pop	Uendelig pop	
populasjonen	500 respondenter		+/- 3,8	+/- 4,2	+/- 4,4	
	1000 respondenter		+/- 2,2	+/- 2,9	+/- 3,1	

 Tabell 8. Huskeregler angående variasjoner i feilmargin

5 Hypotesetesting

Når vi har påpekt en sammenheng mellom to variabler, er det interessant å undersøke om denne sammenhengen er signifikant eller ikke. Forskjeller mellom utvalg må være av en viss størrelse for at vi skal konkludere med at det også er forskjell mellom de respektive populasjonene.

Prinsippet bak hypotesetesting er å formulere en nullhypotese (H_0) som sier at det ikke er differanse mellom populasjonene og en alternativ hypotese (H_1) om at det er forskjell mellom populasjonene. Så beregner vi, ved bruk av en testobservator, om det er sannsynlig at H_0 bør beholdes eller forkastes.

5.1 Kjikvadrattesten

Når avhengig variabel er på nominalnivå, eller på ordinalnivå med få verdier, benyttes kjikvadrattesten for å teste om resultater i utvalget kan generaliseres til populasjonen.

Tradisjonell utregning

Det antas å være en sammenheng mellom kjønn og medlemskap i idrettslag.

- 1. Vi formulerer hypoteser:
 - H₁ = Det er en sammenheng mellom kjønn og medlemskap i idrettslag
 - H_0 = Det er ikke en sammenheng mellom kjønn og medlemskap i idrettslag
- 2. Vi bestemmer oss for å benytte oss kjikvadrattesten.
- 3. Vi velger et signifikansnivå på 0.05, d.v.s. 5 prosent usikkerhet ved resultat
- 4. Beregne tabellens kjikvadratverdi

0	Observert fordeling (O)							Estimert fordeling (E)				
Med	llem i	Kjø	inn	Total					Kjø	٥nn	Total	
idrett	slag	Gutt	Jente			Med	dlem i lo	drettslag	Gutt	Jente		
	Ja	324	350	674	\rightarrow		Ja				674	
	Nei	168	281	449			Nei				449	
Т	otal	492	631	1123		Tota	1		492	631	1123	
								/		\sim		. (
Rute	0	E	O-E	(O-E) ²	(O-E)	² /E	/	É E :	= <u>n (</u>	(oloni	ne) x i	<u>п (гекке)</u>
1	324	295	29	841	2,8	5	1	Bereani	nder	n (ne	le tab	ellen)
2	350	379	-29	841	2,2	2		Gutt med	llem	= (6	674 x 49	2)/1123 = 295,3
3	168	197	-29	841	4,2	7		Jente me	edlem medlen	= (6 (7	674 x 63	1)/1123 = 378,7 2)/1123 = 196 7
4	281	252	-29	841	3,3	4		Jente ikk	e medle	em = (4	49 x 63	(1)/1123 = 252,3
Sum	1123	1123	0		$X^2 = 12$	2,68						

5. Beregning av kritisk verdi (signifikansnivået (0.05) og antallet frihetsgrader (1).³

Antall frihetsgrader	5 prosent	1 prosent	
1	3,84	6,64	Beholde Ho Kritick Forkaste Ho
2	5,99	9,21	verdi

6. Sammenligning: Tabellens kjikvadratverdi (12,7) er høyere enn kritisk verdi (3,84).

Konklusjon: Med 95 prosent sikkerhet kan vi slutte at vi også vil finne kjønnsforskjeller når det gjelder medlemskap i idrettslag i populasjonene av gutter og jenter.

Moderne utregning

Vi antar at det er en sammenheng mellom kjønn og medlemskap i dans-/teatergruppe.

- 1. Vi formulerer hypoteser:
 - H_1 = Det er en sammenheng mellom kjønn og medlemskap i dans-/teatergruppe.
 - H_0 = Det er ikke en sammenheng mellom kjønn og medlemskap i d-/t-gruppe.
- 2. Vi bestemmer oss for kjikvadrattesten og velger et signifikansnivå på 0.05.
- 3. Finne utvalgsobservatoren

 $^{^{3}}$ Formel: (m-1)x(n-1), der m er antallet verdier på den første variabelen (kjønn) og n er antallet verdier på den andre variabelen (medlemskap i idrettslag).

*Ung	i Gjøvik [DataSet1] - S	PSS Data Editor			
Edit	View Data Transform	Analyze Graphs Utilities W	indow Help	Constale	
: Sko 1 2 3 4	Image: Second system Image: Se	Reports Image: Compare Means Tables Image: Compare Means General Linear Model Image: Compare Means Mixed Models Image: Compare Means Correlate Image: Compare Means Regression Image: Compare Means	Frequencies Descriptives Explore Crosstabs Ratio 2 1	Crossrabs Alder [Alder] Hvor bor du? [Bos] Hvor du? [Bos] Hvor du? Hvor bor du samn Hvor dan opplever Hvor ofte har det h Hvor mance av sk	Rgw(s): OK
5	1	Classify 1	6	🖉 Jeg synes at jeg jo 🛛 🖓 Layı	er 1 of 1
6	2	Data Reduction 🔹 🕨 1	6	De voksne hjemme Pre	vious <u>N</u> ext
7	2	Scale	1	W Hvordan røler du d	
0 9	1	Time Series	6	n 🖉 Har du mobbet/pla	
10	2	Survival 1	1	🖉 Har du opplevd å t 💽 👘	
11	1	Multiple Response 1	2	Display clustered <u>b</u> ar charts	
12	2	1 1	1	Suppress tables	
13	Crosstabs: Statistics				
14 16 17 18 20 21 22 23 24 25 26 27 26 27	Chi-square Nominal Contingency coefficien Phi and Cramér's V Lambda Uncertainty coefficient Nominal by Interval Eta Cochran's and Mantel-H Test common odds ratio	t Correlations Continual Continual Somers' d Kendall's tau-b Kappa Risk McNemar aenszel statistics o equals: 1	Continue Cancel Help	Crosstabs: Counts Deserve Expecte Bow Columni I otal Noninteger O Round	tistics Cells Eormat Cell Disclay Image: Continue d Continue d Cancel d Image: Cancel Image: Cancel Image: Cancel d Image: Cancel Image: Cancel Image: Cancel Imag
				O Trunca O No adj	ite cell counts O Truncate case weights ustments

Figur 26. Gjennomføring av moderne utregning av kjikvadrat

a. Analyze + Descriptive statistics + Crosstabs.

b. Vi velger avhengig variabel "Medlemdans" og trykker piltegnet for Rows.

c. Vi velger uavhengig variabel "Kjønn" og trykker piltegnet for Columns.

d. Vi bestemmer "celleinnholdet": Trykk *Cells* og velg *Observed og Expected* under *Counts* (observerte og estimerte frekvenser) og *Column*.

e. Vi velger kjikvadrattesten: Trykk på Statistics og merk av for Chi-square.

Den andre delen handler om å tolke Output-feltet, d.v.s. resultatene på neste side.

Meddans2 * Kjønn Crosstabulation								
	Kjø	nn						
	Gutt	Jente						
Meddans2 Nei Count	529	645						

Figur 27. Resultater (output) av kjikvadrat-test

			Guii	Jente	TOLAT
Meddans2	Nei	Count	529	645	1174
		Expected Count	502,2	671,8	1174,0
	Ja	Count	22	92	114
		Expected Count	48,8	65,2	114,0
Total		Count	551	737	1288
		Expected Count	551,0	737,0	1288,0

Count = O (Observert fordeling) Expected count = E (Estimert fordeling)

						Asymp. Si	g.	Exact Sig.	Exact Sig.
_		[Va <u>lue</u>	df		(2-s <u>ideď</u>)	(2-sided)	(1-sided)
1	Pearson Chi-Square	I	28,17 <u>1</u> 5	₽ <	Ð		002	λ	
	Continuity Correction a	<u> </u>	27,129		1	,0	þo		
	Likelihood Ratio		30,725		1	,0	þo		
	Fisher's Exact Test							,000	,000
	Linear-by-Linear Association		28,149		1	0,	po		
	N of Valid Cases		1288						
	a. Computed only fo	ra:	2x2 table						
	b. O cells (,0%) have	ected co	unt less t	han	5. The minir	hum	expected cou	int is	

Chi-Square Tests

Antall frihetsgrader 48,77. L

> p-verdien Kjikvadratverdien

Value: Den faktiske kjikvadratverdien = 28,2

DF: Antallet frihetsgrader = 1

Asymp. Sig.: p-verdien som SPSS har regnet ut for tabellen er lavere enn 0,000

Konklusjon: Med 95 prosent sikkerhet kan vi slutte at vi vil finne forskjeller i forhold til medlemskap i dans-/teatergruppe i populasjonene av gutter og jenter

5.2 t-test for to uavhengige utvalg

Når vi skal undersøke forskjeller mellom grupper og avhengig variabel er kontinuerlig, kan vi benytte oss av t-testen. Hensikten med t-tester er å undersøke om eventuelle forskjeller mellom grupper i utvalget er så store at vi også kan hevde at disse gruppene har ulike gjennomsnitt i populasjonen. Spørsmålet vi da må stille i en t-test er om denne forskjellen skyldes tilfeldige

avvik mellom utvalgene, eller om forskjellen er så stor at vi også må anta at det er forskjeller mellom populasjonene av de ulike gruppene.

t-fordelingen er en symmetrisk statistisk sannsynlighetsfordeling som er litt flatere enn normalfordelingen i små utvalg og i store utvalg identisk med normalfordelingen.t-testen finnes i en rekke varianter for ett og to utvalg: I toutvalgstester testes forskjellen mellom to gjennomsnitt, og vi kan både se på avhengige (før-etter test av det samme utvalget) og uavhengige utvalg.

Vi kan se på et eksempel av t-test for to uavhengige utvalg.

Tradisjonell utregning

Vi antar at det er en sammenheng mellom kjønn og tid som brukes til trening.

- 1. Vi formulerer følgende hypoteser:
 - H_1 = Det er en sammenheng mellom kjønn og tid som brukes til trening.
 - H_0 = Det er ikke en sammenheng mellom kjønn og tid som brukes til trening.

2. Avhengig variabel er på forholdstallsnivå og er operasjonalisert som antall minutter man trener hver dag. Vi kan da benytte oss a t-test for uavhengige utvalg.

3. Vi bestemmer oss for et signifikansnivå på 0.05.

4. Finne utvalgsobservatoren

En standardfeil for differansen i gjennomsnittlig treningstid er:



Når vi har mange enheter (over 120) benytter vi oss +/- 1,96 SF for finne kritisk verdi. Utregning: +/- 1,96 * 3,84 minutter = 7.5 min

5. Sammenligning: Kritisk verdi er 7,5 minutter, mens gjennomsnittlig forskjell i utvalget av kvinner og menns daglige trening er 14,5 min.

6. Konklusjon: Forskjellen mellom kvinner og menns treningsmengde er større enn kritisk verdi, og vi kan på denne bakgrunn forkaste H_0 med 95 prosent sikkerhet.

Moderne utregning

Vi gjennomfører den samme beregningen i SPSS.

- a. Analyze + Compare Means + Independent Samples t-test
- b. Valg av avhengig variabel: "Trening" og trykk pilen
- c. Valg av uavhengig variabel: "Kjønn" og trykk pilen
- d. Deretter defineres gruppene: 1 = mann & 2 = kvinne
- e. Vi velger konfidensintervall: Trykk Options og merk av for 95 % konfidensintervall



Figur 28. Gjennomføring av moderne utregning av t-test for to uavhengige utvalg Edit View Data Transform Analyze Graphs Utilities Window Help

Den andre delen handler om å tolke Output-feltet, d.v.s. resultatene i figuren under

Fiaur 29.	Resultater	moderne	utreanina	av t-test
	riocanator	1110 0 0 1 1 10	auoginig	a, 11001

Antall Giennomsritt Standardi								lfeilen	
Group Statistics <u>Standardavvik</u>									
		,						Std. Error	7
	Kjønn2		N	Mear	۱	Std. Devi	ation	Mean	
Trening	1,00		64	47,26	56	24,08	335	3,01042	1
	2,00		71	32,74	85	20,02	338	2,37634	

Gjennomsnitt for menn er 47 min Gjennomsnitt for kvinner er 33 min Standardavviket for menn er 24 min Standardavviket for kvinner er 20 min



Levene's Test viser om det er sannsynlig at den avhengige variabelen har lik varians i de to populasjonene av menn og kvinner. Sannsynligheten er så liten (sign<0,05) at det er tryggest å anta at variansen er ulik. Vi bruker t-testen som ikke forutsetter lik varians.

Vi får følgende resultater:

- t: t-verdien er 3,8
- df: Antallet frihetsgrader er 123
- p-verdien er lavere enn 0,000
- Gjennomsnittsforskjellen er 14,5 minutter

Konklusjon: Med 95 prosent sikkerhet slutter vi at vi vil finne kjønnsforskjeller i forhold til antall minutter trening i populasjonene av menn og kvinner



- a. t-testen for parede utvalg brukes hvis vi ønsker å sammenligne to variabler for det samme utvalget. Vi tester H_0 om at gjennomsnittsverdien for disse forskjellene er null
- b. En-veis-Anova brukes når vi skal teste forskjeller mellom 3 eller flere ulike grupper. Vi beregner gjennomsnitt for hver enkelt gruppe, og beregner også spredningen innad i gruppen og mellom gruppene

5.3 Signifikanstest av korrelasjon

Når begge variabler er på forholdstall, intervall eller ordinalnivå med mange verdier kan vi foreta en signifikanstest ved Pearson korrelasjonsanalyse. Vi tester om korrelasjonen er så sterk at den sannsynliggjør at vi vil finne en korrelasjon mellom de samme variablene i populasjonen.

1. Formulere hypoteser:

- H_1 = Korrelasjonen mellom kvinneandel i kommunestyret og kvinneandel i formannskapet er signifikant
- H_0 = Korrelasjonen mellom kvinneandel i kommunestyret og kvinneandel i formannskapet er ikke signifikant

2. Resultater (gjennomføring i SPSS se kap. 3.2) Correlations

		Kvinnandel kommstyre	Prosentkvinne rformannskap	Korrelasjonskoeffisienten r
Kvinnandelkommstyre	Pearson Correlation	1	,332	
	Sig. (2-tailed)		,000-	- Signifikansnivä
	N	433	330-	
Prosentkvinnerforman	Pearson Correlation	,332	1	Antall enheter
nskap	Sig. (2-tailed)	,000		
	N	330	330	

3. Bestemmer oss for et signifikansnivå på 0.05

Tradisjonell utregning

Standardfeilen for korrelasjonen beregnes slik: SF (r) = $\sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-(0,33^*0,33)}{330-2}} = 0,05$ n = antallet enheter

Nullhypotesens gyldighetsområde er gitt ved: 0,052 * 1,96 = 0,101. Korrelasjonen 0,33 er større enn gyldighetsområdet til H₀. Vi konkluderer med at H₀ forkastes.

Moderne beregning

Sig. (2-tailed) verdien er lavere enn 0,000 (se figuren over). Vi kan forkaste H₀ med 95 prosent sikkerhet.

5.4 Vurdering av slutningsstatistikk

Ved hypotesetesting er det nullhypotesen vi tester. Testingen går ut på å beregne sannsynligheten for å få det observerte resultat dersom nullhypotesen er sann; d.v.s. sannsynligheten for å forkaste en riktig nullhypotese. Det betyr at slutningsstatistikk innebærer muligheter for å gjøre feil:

	H₀ er sann	H₀ er falsk				
H₀ forkastes	Feil (type I) –	Korrekt				
H₀ beholdes	Korrekt /	Feil (type II) –				
Feil av type I vil vær nullhypotese: Vi kan for å forkaste H₀ ved signifikansnivå (eks.	e å forkaste en sann redusere muligheten å bruke et strengere 0,01)	il av type II vil være å beholde Ilhypotesen selv om den er gal. Vi kan e muligheten for å forkaste H₀ ved å nke et mindre strengt signifikansnivå :s. 0,1)				
Ergo: Reduserer vi faren for å begå type I- feil øker vi faren for å begå feil av type II						

Figur 30. Mulige feilslutninger ved slutningsstatistikk.

Boks 3. Omkoding av variabel og konstruksjon av sammensatte variabler

Hvis variabler har mange verdier og vi ønsker å benytte dem i krysstabeller, må vi kode variablene om til færre kategorier.

Vi omkoder variabelen "Klassetrinn" fra seks til to verdier:

Før: 1 = 8. klasse, 2 = 9. klasse, 3 = 10. klasse, 4 = 1VG, 5 = 2VG og 6 = 3VG

Etter: 1 = ungdomsskole (verdiene 1, 2 og 3) og 2 = videregående (verdiene 4, 5 og 6)

Konstruksjon av den nye variabelen							
Edit View Data	Transform Analyze Graphs U	tilities Window Help	Pacada inte Diffora	nt Variabler			
🔚 🖨 📴 🧄	Compute	he RA			<u></u>		
: Trivselskole Kjønn 1 2 2 2 3 2 4 2 5 1	Recode Visual Bander Count Rank Cases Automatic Recode Date/Time Create Time Series Replace Missing Values	Into Same Variables Into Different Variables Trivselskole Trivselskole 2 5 1 5 1 5 1 4 2 1 4 3	Er du gutt eller jen Hvor gammel er dt Hvirken skole går Hvordan trives du Hvor mange dage Fotball (Fotball) Händbal/Aaskett Volleyball/Sandvol	Numenc Vanable > Uutput Vanable: Klassetrinn -> Klassetrinn2	Cutput Variable Name: Klassetrinn2 Labet Klassetrinn2 Change		
6 2 7 1 8 2 9 Recode into Diffe	Random Number Generators Run Pending Transforms 13 1 erent Variables: Old and Net	1 5 3 1 4 2 1 5 4 W Values Versjo	Findrett [Findrett] Dansing/teater/tu Gâtt på tur [Tur] Svømming [Svøm 1 (få verdier)]	Old and New Values If (optional case selection condition)	te Reset Cancel Help		
(• Value:		Value: 2		Recode into Different Variables: Old and N	ew Values		
5		O System-missing		Old Value Versjon 2 (mange verdier)	New Value		
O System-missing		Copy old value(s)		O Value:	O Value: 2		
O System- or user-	missing	Old> New:			O System-missing		
C Hange:		Add 1 -> 1 2 -> 1 2 -> 1 3 -> 1 4 -> 2 Remove		System-missing Range:	Copy old value(s) Old -> New: Lowest thru 3 -> 1		
Range, LOWES	ST through value: rough HIGHEST:	Output variables are strings	Width: 8	Range, LOWEST through value	Change Remove		
All other values		Continue C	ancel Help	Range, value through HIGHEST:	Output variables are strings Convert numeric strings to		

- Trykk på *Transform* + *Recode* + *Into Different Variables* (lager ny variabel)
- Velger variabelen "Klassetrinn" og trykker på piltegnet variabelen havner i oppsamlingsboksen som *Numeric Variable*
 - Vi skriver navn for den nye variabelen: Name = "Klassetrinn2" og Label = "Klassetrinn2"
 - Vi trykker Change og da navn i oppsamlingsboks som Output Variable
- Så trykker vi oss inn på Old and New Values og får opp et nytt vindu
 - Til venstre skal vi skrive de gamle verdiene (*Old Value*) og til høyre skriver vi de nye verdiene (*New Value*)
 - Ved få verdier bruker vi tekstfeltet for de gamle verdiene:
 - 1 = 1 + Add, 2 = 1 + Add og 3 = 1 + Add
 - 4 = 2 + Add, 5 = 2 + Add og 6 = 2 + Add
 - Ved mange verdier bruker vi Range-funksjonen: Merk av *Range*, *LOWEST through value* 3 = 1 + Add Merk av *Range value through*, *HIGHEST* 4 = 2 + Add
- Trykk Continue
- Kommer tilbake til *Recode into Different Variables* og trykker *OK*

Det nes	te skrittet e	r teksti	ng av ny	e verdier	
Nam	e Type	Width	Decimals	Label Values Missing	Columns
ເບງົາທະຣະເກກູ	j Numeric	U	U U	Jey name Weber tekste	D
27 Konsent	rasj Numeric	8	0	Jeg syne value Labels	
28 Kjenneri	man Numeric	8	0	Jeg kjenr Value Labels	
9 Gøymed	lfys Numeric	8	0	Jeg syne Value: 2	
30 JYorg	Numeric	8	0	Elevene I Label Midage Sandal	Cancel
31 Fysak	Numeric	8	0	Hvor mar	Help
32 Fotball2	Numeric	8	0	Fotball Add 1 = "Ungdomsskole"	
33 Ball1ver.	2 Numeric	8	0	Håndball. Change	
34 Ball2ver	2 Numeric	8	0	Volleyba	
5 Friidrett	2 Numeric	8	0	Friidrett	
6 Dansing	2 Numeric	8	0	Dansing/	
37 Tur2	Numeric	8	0	Gått på t	
8 Svømmi	ng2 Numeric	8	0	Svømmir	
9 Vintersp	ort Numeric	8	0	Vinteridre	
O Sykling	Numeric	8	0	Sykling	
1 NaturSk	og2 Numeric	8	0	Vært i na	
2 hvormar	gea Numeric	8	0	Hvor mange av de 10 nevnte aktivitetene nar elevene givone	22
3 Klasseti	inn Numeric	8	0	Klassetrinn2 None None	8

Data View Variable View

- Vi går til Variable View og finner den nye omkodede variabelen "Klassetrinn2".
- Vi går til *Values*, velger vinduet *Value Labels* og foretar registrering (se kap 2.2)

Vi har vist omkoding. Andre ganger ønsker vi å slå sammen variabler for å gjennomføre analyser. Vi skal lage variabelen "Hvormangeaktiviteter" som vi så på tidligere. Vårt utgangspunkt er 10 aktiviteter der elevene har svart Ja = 1 og Nei = 0.

Konstruksjon av den nye variabelen

Compute. Compute. Compute Variable 19: Trivselskole Recode Ysual Bander Count 20: 33: 22 Rank Cases Automatic Recode Foball 2 + Ball ver2 + Fridrett 2 + Dansing2 + Tur2 + Syking2 + NatuSkog2 3: 22 Pate/Time Create Time Series Function group: 4: 4: 22 Create Time Series Function group: 6: 22 Pate/Time Create Time Series 7: 1 Random Number Generators Pate/Time Date/Time Date Arithmetic Date Creation Date Arithmetic Date Creation Date Stration Date Stratinana Stration Date Stratina Stration Date Strat	File Edit View Dat	a Transform Analyze Graphs Utilit	ies Window Help
Kjørn Count Rank Cases 1 2 2 3 2 4 2 5 1 Rank Cases Pate/Time Create Time Series Replace Missing Values Random Number Generators Potball (Fotball2) 7 1 8 2 13 1 10 1 11 1 12 1 13 1 12 1 13 1 16 2 13 1 16 2 17 1 18 1 19 1 19 1 19 1 13 1	😕 📮 🚔 📴 19 : Trivselskole	Compute Recode Visual Bander	Compute Variable Ariable Ariget Variable Fotball2 + Ball ver2 + Ball2ver2 + Friidrett2 + Dansing2 + Tur2 +
3 2 Date/Time 4 2 Create Time Series 6 2 Replace Missing Values Random Number Generators Potal [Fotball?] 4 1 1 7 1 Run Pending Transforms 8 2 13 10 1 13 11 1 1 12 1 13 16 2 13 16 2 13 17 1 13 18 1 13 19 1 13	Kjønn 1 2	Count 2 Rank Cases 2 Automatic Recode	Type & Label Svæmming2 + Vintersport2 + Sykling2 + NaturSkog2 Jeg synes det er gi Function group:
7 1 Run Pending Transforms 8 2 13 1 9 2 13 1 10 1 13 1 11 1 1 1 12 1 13 1 13 1 12 2 14 1 13 1 15 2 13 1 16 2 13 1 17 1 13 1 18 1 13 1 19 1 13 1 19 1 13 1	3 4 5 6	2 Date/Time 2 Create Time Series 1 Replace Missing Values 2 Random Number Generators	✓ Elevene har fått væ ✓ (> ?) 89 All ✓ Hvor mange timer i - <=>= 456 CDF & Noncentral CDF ✓ Fotball [Fotball2] = == 123 Conversion ✓ Håndball/basketba = == 123 Current Date/Time
10 1 13 1 11 1 1 1 12 1 13 1 13 1 12 2 14 1 13 1 15 2 13 1 16 2 13 1 17 1 13 1 18 1 13 1 19 1 13 1	7 8 9	Run Pending Transforms 2 13 1 2 13 1	Volleyball/sandvoll / & O, Date Arithmetic Date Arithmetic Date Cartain Date Cartain Date Extraction Date Extract
13 1 12 2 14 1 13 1 15 2 13 1 16 2 13 1 17 1 13 1 18 1 13 1 19 1 13 1	10 11 12	1 13 1 1 . 1 1 13 1 1 4 12 2	Svarming [Svarm Vinteridrett (Ski, sk Sykling [Sykling2]
10 2 13 1 17 1 13 1 18 1 13 1 19 1 13 1	13 14 15 16	1 12 2 1 13 1 2 13 1 2 13 1	Vent Hadderrakt Vent Hadderrakt Vent Hadderrakt Klassetrinn2 [Klass
UK Paste Heset Cancel Heip	17 17 18 19	1 13 1 1 13 1 1 13 1 1 13 1	UK Paste Reset Cancel Help

Vi velger Transform + Compute

- Under Target Variable skriver vi navn på variabelen "Hvormangeaktiviteter"
- Vi merker av for *Fotball2* trykker på piltegnet og *Fotball 2* kommer i boksen *Numeric Expression*. Vi trykker på tegnet for "+" og tar for oss neste variabel.
- Når vi har lagt inn alle variablene trykker vi på *OK*

Vi går til Variable View og finner den nye sammenslåtte variabelen nederst

Boks 4. Multivariat analyse

Det er sjelden nok at vi bare analyserer sammenhenger mellom to variabler. Vi skal nå redegjøre for hvordan vi gjør analyser der vi kontrollerer for flere uavhengige variabler.

- I forhold til den beskrivende analysen skal vi se på trivariat analyse
- I forhold til slutningsstatistikken skal vi se på multivariat regresjon

Trivariat analyse

I trivariat analyse kontrolleres effekten for uavhengig variabel med en annen uavhengig variabel. Eks: Avhengig variabel = Dans med to verdier (1 = artig og 2 = Kjedelig)

Kjønn og D	ans		Mestring og	g Dans		Kjønn og Mestri	ng	
	Gutt	Jente		Lav	Høy		Gutt	Jente
Artig	58	84	Artig	58	78	Lav mestring	44	20
Kjedelig	42	16	Kjedelig	42	22	Høy mestring	56	80
Sum	100	100	Sum	100	100	Sum	100	100
Ν	106	110	Ν	69	147	Ν	106	110

- Vi har funnet en sammenheng mellom Kjønn og Dans.
- Vi har funnet en sammenheng mellom opplevelse av Mestring og Dans.
- I tillegg har vi funnet en sammenheng mellom Kjønn og Mestring

Kan den positive sammenhengen mellom mestring og holdning til dans egentlig skyldes at kjønn påvirker begge to? Vi gjennomfører analysen i SPSS.

- Analyze + Descriptive Statistics + Crosstabs
- Vi legger Dans som avhengig variabel (*Row*), vi legger Mestringsdans som uavhengig variabel 1 (*Column*) og Kjønn er uavhengig variabel 2 (*Layer 1 of 1*)

File Edic view Data Transform	Analyze Graphs Oulicles W			
🗁 📑 🚔 📴 🤝 🥐 🦢	Reports			
19 · Triveelekole	Tables	Prequencies	Crosstabs	×
TO . HIVSEISKOIE		Descriptives		
Kjønn Alder	Compare Means	Explore	MUTS nigt det ve 🔨	HOWISE OK
94 2 1	General Linear Model	Crosstabs	V Utenom skoletiden	Dans [Dans]
95 . 1 1	Mixed Models	Ratio	A Sterion stoletiden	Paste
96 \ 1 1	Correlate •	5	V leg kjenner områd	Reset
97 1 1	Regression	5	Vi vår klasse har vi	Column(s):
	Loglinear 🕨 🕨		Det er mue bråk i k	Mestringdans Mestrine Cancel
90 \ 1	Classify 🕨 🕨	3	leg her god methos	Help
99 1 1	Data Reduction 🕨 🕨	5	log or godt formaw	
100 1 1	Scale 🕨	4	Jeg er god form	LayerTorT
101 2 1	Nonparametric Tests 🔹 🕨	4	A leg ber fått til de fl	Previous
102 2 1	Time Series	5		
103 1 1	Survival 🕨 🕨	4	A Jeg synes det er re	Kjønn [Kjønn2]
104 2 1	Multiple Response 🕨 🕨	3	V Jey Kjermer mange	
104 2 1	Missing Value Analysis	5	V Jeg synes det er g	
	Complex Samples 🕨 🕨	5	Display clustered bar chart	2
106 2 1	4 Z J	- 4		-
107 2 1	4 2 3	4	Suppress tables	
108 2 1	4 2 3	4		
109 1 1	4 2 3	3	Exact	Statistics Cells
			,	

V	Vi får ut følgende resultat i Output-vinduet: Dans * Mestringdans * Kjønn Crosstabulation								
				Mestr	ingdans				
				Lav grad	Høy grad av				
	Kjønn			mestring	mestring	Total			
	Gutt	Dans	Artig	46,8%	67,8%	58,5%			
			Kjedelig	53,2%	32,2%	41,5%			
		Total		100,0%	100,0%	100,0%			
	Jente	Dans	Artig	81,8%	84,1%	83,6%			
			Kjedelig	18,2%	15,9%	16,4%			
		Total		100,0%	100,0%	100,0%			

Vi konsentrerer oss i det videre arbeidet om en av verdiene i den avhengige variabelen - når avhengig variabel er dikotom (2 verdier) spiller det ingen rolle hvilken verdi vi velger for analysen – og velger ut verdien "Artig". Vi får følgende tabell:

Kjønn	Jente		Gutt	
Mestring	Lav	Høy	Lav	Høy
Dans er "Ar-	82	84	47 ₁	68/
tig"				

Tabellen omskrives og vi beregner effekter

	10 1		/
	Jente	Gutt	Delsammenhenger
Høy mestring	84	68	16
Lav mestring	82	47 🕈	35
Delsammenhenger	2	19	

a. Gjennomsnittlig effekt av "Kjønn" kontrollert for "Mestring" = (16 + 35)/2 = 25,5 %.

b. Gjennomsnittlig effekt av "Mestring" kontrollert for "Kjønn" = (2 + 19)/2 = 10,5 %.

c. Delsammenhengene er ulike

- Det betyr at effekten av "Kjønn" varierer etter hvilken verdi av "Mestring" vi holder konstant: Effekten av "Kjønn" kontrollert for "Mestring" er større for gutter enn jenter.
- Det betyr at effekten av "Mestring" varierer etter hvilken verdi "Kjønn" vi holder konstant: Effekten av "Mestring" kontrollert for "Kjønn" er større for "Lav" enn "Høy".

Regresjonsanalyse

Regresjonsanalysen er basert på tanken om at sammenhengen mellom en avhengig variabel Y og et sett uavhengige variabler (X), kan framstilles i form av en lineær funksjon.



I funksjonen finner vi følgende:

- Variasjonen i avhengig variabel (Y) forklares av uavhengige variabler $(X_1...X_n)$.
- Regresjonskonstanten (b₀) er den predikerte verdi av Y når x-variablene er 0.
- Regresjonskoffesienten (b₁) viser endringen i Y når X endres med en måleenhet, kontrollert for de andre variablene i modellen.
- Residualene (e) er forskjellen mellom observerte og predikerte verdier av Y.

Vi skal gjennomføre en lineær multivariat regresjonsanalyse i SPSS. Vår avhengige variabel er en sumskårevariabel, sammenslått variabel, av ulike påstander som går på trivsel i gymtimene. Avhengig variabel "gymtrivsel" har laveste verdi 1 (stor grad av mistrivsel) og høyeste verdi 10 (stor grad av trivsel). Vi lager følgende modell:

				Njønn:	Fysisk form:		
Kiønn	7~			1 = Gutt	1 = I god form		
, goint]	~		0 = Jente	0 = I dărlig form	1	
Fysisk 1	form			Elevano de d	dasis a.		
.,			ivsel i gymtimene	1 – Eleven	rkning: o har fått væro mov	d å hostommo sumaktivit	otor
Elevme	dvirkning		55	0 = Eleven	e har ikko fåttvære	a bestemme gymakuvit med å bestemme gyma	eter ktivitatar
Licome	<u>avning</u>			0 - Lieven		e med a bestemme gyma	Kuvitetei
Mestrino	a			Mestring		Lærers stil	
	<u></u>			1 = Høy gra	d av mestring	1 = lkke-autoritær	
Lærers	stil			0 = Lav gra	d av mestring	0 = Autoritær	
x 7· ·	C /	1					
vi gjen	nomføre	er analysen					
•	Analyze	e + Regression	n + Linear				
	A 1		7 4 1 17 1	1 7		• 11	
•	Avheng	gig variabel "C	symtrivsel" legg	ges under L	Dependent Vari	iable	
•	Uavher	igige variablei	r legges i bokser	n for <i>Indep</i>	endent		
-	T1.1.		66	1			
•	тукк (
File Edit Vie	ew Data Transf	orm Analyze Graphs Util	ilities Window Help				
🗁 🔒 🚔	📴 🦘 🚸	Reports		💻 Linea	r Regression		×
🗁 🔚 🔒 217 : Kjønn2	2	Reports Descriptive Statistics Tables	s 🕴 🐼 🔕	Linear	Regression	Bependent:	
🗁 📙 🏝 217 : Kjønn2	E 🕈 🕈 🏕 2 Kienn 🛛 Ali	Reports Descriptive Statistics Tables Compare Means	s	Er du	r Regression	Pependent:	
🗁 틙 📇 217 : Kjønn2	₽ 🗗 🔶 2 Kjønn Al	Reports Descriptive Statistics Tables General Linear Mode	s Fivselskole JYLITS	iofte F(Hvik	r Regression	Dependent:	OK Paste
Image: Participation of the second	E 🎝 👼 2 Kjønn Ale 1	Reports Descriptive Statistics Tables Compare Means General Linear Mode Mixed Models	s Fivselskole JYLITS	iofte Freduction Hvilk	r Regression gutt eller jent ▲ gammel er du et klassetrinn en skole går c ■	Bependent: Gymtrivsel [Gymtrivsel] f 1 Next	OK Paste Reset
Image: Constraint of the second se	2 Kjønn Alı 1 2	Reports Descriptive Statistics Tables Compare Means General Linear Mode Mixed Models Correlate Correlate	s Trivselskole JYLITS	tofte Fr du 4 Vor Hvoik	r Regression gutt eller jent ▲ gammel er du et klassetrinn en skole går d dan trives du	Gymtrivsel [Gymtrivsel]	OK Paste Reset
Image: Constraint of the second se	E + + 2 Kjønn Al 1 2 2 2	Reports Descriptive Statistics Tables General Linear Models Correlate Regression Regression	s investigation	iofte Fr du	r Regression gutt eller jent ▲ gammel er du et klassetrinn en skole går (= dan trives du mange dager	Gymtrivsel [Gymtrivsel]	OK Paste Reset Cancel
Image: Constraint of the second se	Image: Constraint of the second sec	Reports Descriptive Statistics Tables Compare Means General Linear Model Mixed Models Correlate Regression LogInear Collever	s i ivselskole JYLITS 4 Linear Curve Estimation	iofte F(+ With A - With	Regression gutt eller jent gammel er d. et klassetrinn en skole går (dan trives du mange dager all (Fotball)	Gotterm Electron	OK Paste Reset Cancel Help
▷	Image: Non-State Image: Non-State Kjønn Alu 1 2 2 2 2 2 2 2 1 1 2 2 1 1 2 1	Reports Descriptive Statistics Tables General Linear Mode Mixed Models Compare Means General Linear Mode Loginear Loginear Classify Data Reduction	s i i i i i i i i i i i i i i i i i i i	iofte Fr + Vik 4 - Vik - Vik	Regression gutt eller jent gammel er du et klassetinn en skole går (dan trives du mange dager all [Fotball] bibal/baskebb		OK Paste Reset Cancel Help
 κjønn2 217 : Kjønn2 958 959 960 961 962 963 	Image: Second	Reports Descriptive Statistic: Tables General Linear Mode 1 Compare Means General Linear Mode 1 Correlate 1 Loglinear 1 Classify 1 Data Reduction 1 Scale	s vivselskole JYLITS	Linear	Regression gutt eller jent A gammel er du et klassetinn en skole går (dan tirves du mange dager all [Fotball] iball/basketba yball/sandvoll	Dependent: Comparison of the second	OK Paste Reset Cancel Help
Image: Product of the system Image: Product of the system 217 : Kjørn2 958 959 960 961 962 963 964	Image: symbol with the symbol withe symbol with the symbol with the symbol with the sym	Reports Descriptive Statistic: Tables General Linear Model Compare Means General Linear Model Correlate Cassify Data Reduction Scale Nonoparametric Tests:	s ivselskole JYLITS 4 Curve Estimation Binary Logistic Multinomial Logistic S	Linear	Regression gutt eller jent gammel er d. et klassetrinn en skole går c dan trives du mange dager il (Fotball) tball/basketb: yball/sandvoll ett (Fridrett) ing/teater/tur	Dependent: Cymtrivsel [Gymtrivse] (1 Next Jedependent(s): Godtom [Godform] Bestemmeaktiviteter [// Method: Enter	OK Paste Reset Cancel Help
 Image: Image: Im	Image: style="text-align: center;">Image: style="text-align: center;"/>Image: style="text-align: center;"////////////////////////////////////	Reports Descriptive Statistics Tables Compare Means General Linear Models Ocrrelate Nixed Models Correlate Loginear Classify Data Reduction Scale Nonparametric Tests Nonparametric Tests	s i ivselskole JYLITS d ivselskole JYLITS d curve Estimation Binary Logistic Multinomial Logistic ordinal Probit	Linear iofte Fe Hvor 4 2 - - - - - - - - - - - - -	Regression gutt eller jent gammel er d. et klassetrinn en skole går (dan trives du mange dager il (Fotball) ball/baskebz yball/sandvoll att [Fiidrett] ing/teater/tur jå tur [Tur]	Selection Variable	OK Paste Reset Cancel Help
 	Image: style="text-align: center;">Image: style="text-align: center;"/>Image: style="text-align: center;"////////////////////////////////////	Reports Descriptive Statistics Tables Compare Means General Linear Model Mixed Models Correlate Correlate Regression LogInear Classify Data Reduction Scale Nonparametric Tests: Time Series Survival	s i i i i i i i i i i i i i i i i i i i	Linear Softe Fr 4 2 4 2 4 2 4 2 4 4 2 4 4 4 4 4 4 4 4 4 4 4 4 4	Regression gutt eller jent A gammel er d. et klassetrinn en skole går (dan trives du mange dager all (Fotball) blal/basketb: bjal/basketb: på tur [Tur] ming (Svømr	Gymtrivsel [Gymtrivsel] Mext Metpendent(s): Godform [Godform] Bestemeaktiviteter [YYM Method: Enter Selection Variable:	OK Paste Reset Cancel Help
217 : Kjønn2 958 960 961 962 963 964 965 966 966 966	Image: style="text-align: center;">Image: style="text-align: center;"/>Image: style="text-align: center;"////////////////////////////////////	Reports Descriptive Statistics Tables General Linear Mode Mixed Models Compare Means General Linear Mode LogInear LogInear Classify Data Reduction Scale Nonparametric Tests Tomical Series Survival Multiple Response	s i i i i i i i i i i i i i i i i i i i	iofte Fr Fr Hvik 4 2 Fridu Fri	Regression gutt eller jent A gammel er du et klassetinn en skole går (dan trives du mange dager all [Fotball] biball/basketbiz yball/sandvoll et [Fiidrett] ing/teater/tur på tur [Tur] ming [Svammr ridrett (Ski, sk	Gependent: Gyntrivsel (Gyntrivsel) Next Udependent(s): Gutt (Gutt) Godform (Godform) Method: Enter Selection Variable: Pude	OK Paste Reset Cancel Help
Image: Product of the system Image: Product of the system <t< td=""><td>Image: Second second</td><td>Reports Descriptive Statistics Tables Compare Means General Linear Mode Mixed Models Correlate Reports Loglinear Classify Data Reduction Scale Nonparametric Tests: Time Series Survival Multiple Response Missing Value Analys</td><td>s introduction in the second s</td><td>es</td><td>Regression gutt eller jent A gammel er du et klassetrinn en skole går (dan tives du mange dager all [Fotball] iball/basketbi yball/sandvoll tt [Friidrett] ing/teater/tur på tur [Tur] ming [Svømr riidrett [Ski, sk ng [Sykling]</td><td>Vependent: Case Labels:</td><td>OK Paste Reset Cancel Help</td></t<>	Image: Second	Reports Descriptive Statistics Tables Compare Means General Linear Mode Mixed Models Correlate Reports Loglinear Classify Data Reduction Scale Nonparametric Tests: Time Series Survival Multiple Response Missing Value Analys	s introduction in the second s	es	Regression gutt eller jent A gammel er du et klassetrinn en skole går (dan tives du mange dager all [Fotball] iball/basketbi yball/sandvoll tt [Friidrett] ing/teater/tur på tur [Tur] ming [Svømr riidrett [Ski, sk ng [Sykling]	Vependent: Case Labels:	OK Paste Reset Cancel Help
Image: Product of the system Image: Product of the system 217 : Kjønn2 958 959 960 961 962 963 964 965 966 967 968 969	Image: Second	Reports Descriptive Statistic: Tables General Linear Mode Mixed Models Compare Means General Linear Mode 1 Correlate Cassify Data Reduction Scale 1 Nonparametric Tests 1 Survival Missing Value Analys Complex Samples	s ivselskole JYLITS ivselskole JYLITS 4 4 5 6 6 7 7 7 7 7 8 7 8 7 8 7 8 7 9 7 9 7 9 7 9	Linear iofte Fr & Fr d Hvik 4 2 Volle Fotb	Regression gutt eller jent gammel er d. et klassetinn en skole går d dan tives du mange dager all (Fotball) tball/basketba yball/sandvoll ett (Friidrett) ing/teater/tur på tur [Tur] ming (Syamm ing (Syafing) i naturen/skc	Dependent: Capendent: Capendent(s): Cadom (Godform) Restemmeaktiviteter (FY) Method: Enter Selection Variable: Case Labels:	OK Paste Reset Cancel Help
 ▶ ■ ▶ ■ ■ ■ ■ 968 ■ 969 960 961 962 963 964 965 966 967 968 969 969 970 	Image: Second	Reports Descriptive Statistics Tables Compare Means General Linear Model Mixed Models Correlate Repression Loglinear Classify Data Reduction Scale Scale Nonparametric Tests Time Series Multiple Response Mixed Models Correlate	s i i i i i i i i i i i i i i i i i i i	es	Regression gutt eller jent gammel er di et klassetrinn en skole går (dan trives du mange dager all (Fotball) tball/basketb: yball/b	Selection Variable:	OK Paste Reset Cancel Help
 217 : Kjønn2 958 969 960 961 962 964 966 967 968 969 969 961 962 963 964 963 	Image: Second	Reports Descriptive Statistics Tables Compare Means General Linear Mode Mixed Models 1 Regression LogInear Classify 1 Data Reduction Scale Scale 1 Nonparametric Tests 1 Nonparametric Tests 1 Multiple Response 1 Scale 1 Multiple Response 1 Complex Samples 10 1	s inviselskole JYLITS rivselskole JYLITS 4 4 Curve Estimation Binary Logistic Ordinal Probit Nonlinear Weight Estimation 2-Stage Least Squar Optimal Scaling 6 5 6 5	Linear softe Fr 4 2 4 2 4 2 4 2 4 2 4 4 2 4 4 4 4 4 4 4 4 4 4 4 4 4	Regression gutt eller jent A gammel er du et klassetrinn en skole går d dan trives du mange dager all (Fotball) Iball/baskebt yball/sandvoll ett (Friidrett) ing/keater/tur på tur (Tur) ming (Svømr ridrett (Ski, sk g (Sykling) i naturen/skc I'S gjør det ve om skolekiden	Vependent: Gymtrivsel [Gymtrivsel] Next Vependent(s): Godform [Godform] Bestemeaktiviteter [VVV Method: Enter Selection Variable: Case Labels: WLS Weight:	OK Paste Reset Cancel Help
Image: Product of the system Image: Product of the system 217 : Kijønn2 958 959 960 961 961 962 963 966 966 966 966 968 968 969 970 971 972 972	Image: Second	Reports Descriptive Statistics Tables General Linear Mode Mixed Models Compare Means General Linear Mode LogInear Classify Data Reduction Scale Nonparametric Tests: Time Series Survival Multiple Response Multiple Response Complex Samples 10 1 10 1	s investigation in the second	es	Regression gutt eller jent A gammel er du et klassetinn en skole går (dan tirves du mange dager all [Fotball] iball/basketbi yball/sandvoll ett [Firidrett] ing/teater/tur på tur [Tur] ming [Svæmr ridrett (Ski, sk ng (Sykling) i naturen/skc S gjør det ve om skoletiden S har gitt ma V	Vependent: Gymtrivsel (Gymtrivsel)	OK Paste Reset Cancel Help
Image: Product of the system Image: Product of the system 217 : Kjørm2 958 959 960 961 961 962 963 966 966 966 967 968 969 970 971 971 973 973	Image: Second	Reports Descriptive Statistics Tables Compare Means General Linear Mode Mixed Models Correlate Repression LogInear Classify Data Reduction Scale Nonparametric Tests Survival Multiple Response Missing Value Analys Complex Samples 10 1 10 1 10 1 15 6	s introduction in the second s	es	Regression gutt eller jent ▲ gammel er du et klassetinn en skole går (dan tives du mange dager all [Fotball] iball/basketbi yball/sandvoll ett [Friidrett] ing/teater/tur på tur [Tur] ming [Svømr ridrett [Ski, sk ng [Sykling] i naturer/skc S gjør det ve om skoletiden S har gitt ma ♥	Vependent: Gyntrivsel [Gymtrivsel]	OK Paste Reset Cancel Help
Image: Product of the system Image: Product of the system 917 : Kjørnn2 968 969 960 960 961 963 962 963 964 966 965 966 966 967 968 969 970 971 972 973 974 973	Image: Second	Reports Descriptive Statistics Tables Compare Means General Linear Mode 1 Correlate Regression 1 Cassify Data Reduction Scale 1 Scale 1 Survival Multiple Response Missing Value Analys Complex Samples 10 1 10 1 15 6	s introduction in the second s	es	Regression gutt eller jent ▲ gammel er du et klassetinn en skole går (dan tirves du mange dager all [Fotball] iball/basketbz yball/sandvoll tt [Friidjett] ing/teater/tur på tur [Tur] ming [Syking] i naturen/skc IS giar det veo miskoletiden IS har gitt ma ♥ Statistics	Vependent: Guntrivsel [Gymtrivsel] I I I I I I I I I I I I I I I I I I I	OK Paste Reset Cancel Help

får ut	følgende resultate	er som sk	al tolkes:				
	Model	Summary	~	Den iv	sterte kvadr	erte korrela	sionskoeffisienten
Mode 1 a.	el R R Squa ,432ª ,1 Predictors: (Constant), Bestemmeaktiviteter, G	re Adjust R Squ 87 Ikke-autoritæ utt, Mestreta	ed Std. E are the Es 179 er, Godform, aktiviteter	rror of timate 1,735			
	\frown)VÆ			_	
<u>Model</u> 1	Regression 350,07	f s df	Mean S	quare F 0,015 23,27	Sig.) ^a	
	Total 1872,48	1 51 39 51		3,009			
b. De Kv va	pendent Variable: Gymt vadratsummene viser iriansen i avhengig va	nisel her til hvor riabel: 350/	mye regresj 1872 = 0,187	inetsgråder son jonen vår forkla ′ (som er lik R si	n ner er anta rer av den tø quare i tabe	otale Ilen over)	igige variabler
	Regres	onskoeffisie	entene				
	,				t-verdien	signifikan	snivăet
	,	c	Coefficients		t-verdien	signifikan	snivăet
		C Unstanc Coeffi	Coefficients ^a dardized cients	Standardized Coefficients	t-verdien	signifikan	snivăet
Model	(Constant)	Unstand Coeffi B	dardized cients Std. Error	Standardized Coefficients Beta	t-verdien	Sig.	snivăet
Model 1	(Constant) Gutt	Unstanc Coeffi JB 4,660 392	Coefficients dardized cients Std. Error ,166 157	Standardized Coefficients Beta	t-verdien	Sig. ,000	snivăet
Model 1	(Constant) Gutt Godform	Unstand Coeffi B 4,660 ,392 460	Coefficients dardized cients Std. Error ,166 ,157 .164	Standardized Coefficients Beta ,103 118	t-verdien t 28,124 2,495 2,811	Sig. ,000 ,013 005	snivăet
Model 1	(Constant) Gutt Godform Bestemmeaktiviteter	Unstanc Coeffi B 4,660 ,392 ,460 .707	Coefficients [®] dardized cients Std. Error ,166 ,157 ,164 ,161	Standardized Coefficients Beta ,103 ,118 .177	t-verdien t 28,124 2,495 2,811 4,398	Sig. ,000 ,013 ,005 ,000	snivăet
Model 1	(Constant) Gutt Godform Bestemmeaktiviteter Mestretaktiviteter	C Unstanc Coeffi JB 4,660 ,392 ,460 ,707 ,656	Coefficients [®] Jardized cients Std. Error ,166 ,157 ,164 ,161 ,101	Standardized Coefficients Beta ,103 ,118 ,177 ,276	t-verdien t 28,124 2,495 2,811 4,398 6,506	Sig. ,000 ,013 ,005 ,000 ,000	snivăet

a. Dependent Variable: Gymtrivsel

Adjusted R^2 : Forteller oss at de uavhengige variablene forklarer rundt 18 prosent av variasjonen i avhengig variabel (Y). Det betyr m.a.o. at andre faktorer enn de vi har inkludert i analysen forklarer 82 prosent av variasjonen i Y.

Regresjonskonstanten (b_0) er gitt ved verdi 4,7: Det vil si at en elev som er jente, i dårlig form, som ikke får være med å bestemme gymaktiviteter, som har lav mestring av aktivitetene og som har en autoritær lærer har en skåre på 4,7 trivselspoeng i gym

Regresjonskoeffisienter forteller oss endringer når vi går fra verdi 0 til 1. Det betyr:

- at gutter trives bedre i gymtimene enn jenter
- at de som er i god form trives bedre enn de som er i dårlig form
- at de som får være med å bestemme aktiviteter trives bedre enn de som ikke får det
- at de som mestrer aktiviteter trives bedre enn de som ikke mestrer aktivitetene
- at de som har en ikke-autoritær lærer trives bedre enn de som har en autoritær

Signifikansverdiene er under signifikansnivået 0.05: Kjønn, fysisk form, medvirkning, mestring og lærerstil har signifikant betydning for trivsel i gymtimene

Boks 5. Oversikt univariat og bivariat analyse
Univariat analyse
Nominal/ Ordinal m/få verdier Slutningsstatistikk Beregning av feilmargin Gjernomsnitt Median
Ordinal m/mange verdier Intervall Beskrivende analyse Modus Spredningsmål (variasjonen i fordelingen) Variasjonsbredde Slutningsstatistikk Estimere et gjennomsnitt i populasjonen
Bivariat analyse
Variabel 1 — Variabel 2 Nominalnivå — Nominalnivå Nominalnivå — Ordinalnivå (få verdier) Ordinalnivå (få verdier) – Ordinalnivå (få verdier)
Nominalnivå Ordinalnivå (mange verdier) Nominalnivå Intervallnivå/forholdstallsnivå Ordinalnivå (få verdier) Ordinalnivå (mange verdier) Ordinalnivå (få verdier) Ordinalnivå (få verdier) Ordinalnivå (få verdier) Intervallnivå/forholdstallsnivå
Ordimalnivå (mange verdier) — Ordinalnivå (mange verdier) Ordimalnivå (mange verdier) — Intervallnivå/forholdstallsnivå Intervallnivå/forholdstallsnivå - Intervallnivå/forholdstallsnivå Slutningsstatistikk — Signifikanstest av korrelasjon

Det lille kvantitative metodeheftet

Dette notatet er basert på mine undervisningsnotater i forbindelse med undervisning i samfunnsvitenskapelig metode ved Høgskolen i Lille-hammer og Norges Teknisk - Naturvitenskapelige Universitet. I notatet gjennomgås noen grunnleggende måter å analysere kvantitative data på ved statistikkprogrammet SPSS.

Notatet er beregnet for studenter, forskere og andre som har liten eller ingen erfaring med bruk av kvantitativ metode. I heftet redegjøres det kortfattet for forskningsprosessen og deretter fokuseres det på kvantitativ metode. Veiledningen er laget til versjon 14.0 av SPSS.

Notat nr.: 17/2007 ISSN nr: 0808-4653