

# Hvordan præsenteres data?

## Torben Stener Nielsen

En kvantitativ undersøgelse resulterer i en masse informationer. En spørgeskemaundersøgelse med 1800 besvarede skemaer med hver

6 spørgsmål giver ca. 10.000 enkeltoplysninger. For at kunne komme videre i arbejdet er det derfor nødvendigt at få ordnet disse rå data, så de fremstår på en overskuelig måde.

Figur 3.1 En tabels opbygning

		<b>Tabelhoved</b>			
		1.kolonne	2.kolonne	3.kolonne	4.kolonne
<i>Overskrift</i>		<b>Intelligens og social baggrund</b>			
		<b>Restgruppeprocenter i 38-års alderen</b>			
		Point i intelligens test	0-30	31-40	41- Alle
<b>Forspalte</b>		<b>Oprindelsesklasse</b>			
	<i>1.række</i>	Professionerne	20	-	8 7
	<i>2.række</i>	Funktionærer	23	16	7 13
	<i>3.række</i>	Selvstændige	40	15	6 18
	<i>4.række</i>	Faglærte arbejdere	41	14	14 20
	<i>5.række</i>	Ikke-faglærte arbejdere	54	29	15 34
	<i>6.række</i>	Uden for erhverv	63	32	4 33
	<i>7.række</i>	Alle	44	19	9 21
	<i>Note</i>	NB: Tabellen er baseret på den generation, der i 1992 var 38 år. Da de i 1968 gik i 7. klasse blev deres intelligens testet ved hjælp af en verbaltest. Deres oprindelsesklasse er bestemt ved forsørgers erhverv, da de gik i 7. klasse. Restgruppen er her dem, der i 38-års alderen ikke har en erhvervs-kompetancegivende uddannelse.			
	<i>Kilde</i>	Erik Jørgen Hansen: <i>En generation blev voksen. SFI 1995.</i>			

## Tabeller ordner data

En god tabel er bygget op som vist i figur 3. 1. Den består af et antal rækker (linier) i forspalten og kolonner (søjler), anbragt i hovedet. Række nr. 4 er fordelingen af de faglærte arbejdere, kolonne nr. 2 er fordelingen af points i en intelligens test. Række nr. 4 og kolonne ny. 2 (markeret) viser altså, at 14% af børn af faglærte arbejdere i 38-års alderen, som opnåede mellem 31 og 40 points i denne intelligens test, ikke havde fået en kompetancegivende uddannelse.

De fleste tabeller er *todimensionale*, dvs. at man krydser 2 variable med hinanden (f.eks. indkomst og socialgruppe), men der er ikke noget til hinder for at udvide med flere oplysninger. Det kan gøres ved at underindele tabellens kolonner og rækker. Det er klart, at jo mere det sker, jo mere vil det gå tid over overskueligheden. Tabel 3.6 er et eksempel på en ret sindrig tabel med mange variable. Den er kompliceret at læse, men giver til gengæld mulighed for flere interessante konklusioner.

Tabel 3.6

**Social generationsmobilitet fører til politisk generationalmobilitet.  
Pct. vandret.**

Faders stilling/ Egen stilling	Mobili- tet	Samme parti	Parti								Antal (=100%)	
			V	K	CD	R	Frp.	SF	S	Små parti		
<i>Landmand</i>												
Landmand	Stabil	68	70	4	1	8	8			2	7	219
Middelklasse	Mobil	38	25	17	3	13	7	3	22		9	276
Arbejder	Mobil	29	19	4	4	11	7	5	46		5	316
<i>Middelklasse</i>												
Middelklasse	Stabil	36	9	23	7	14	6	5	26	10		547
Arbejder	Mobil	53	7	6	2	5	3	9	63	5		198
<i>Arbejder</i>												
Middelklasse	Mobil	45	7	13	4	9	8	8	4	8		419
Arbejder	Stabil	69	2	2	1	5	4	9	70	7		619

Kilde: Torben Worre: *Dansk vælgeradfærd, 1987***Huskeliste, når du læser en tabel****Hvad viser tabellen noget om?**

Læs overskriften omhyggeligt. Se hvilke variable, der indgår i hoved og forspalte. Se, om der er ekstra oplysninger i fodnoter eller anmærkninger. Find måleenheden. Undersøg, om det er absolutte eller relative tal.

**Hvordan er tabellen konstrueret?**

Skal der læses lodret eller vandret? Er der foretaget sammenlægning af tallene?

**Læs tabellen!**

Læs systematisk. Søg ikke hist og pist. Er der tal, der er mere centrale end andre? Kan der spores en tendens eller en udvikling i tid? Sammenhold forskellige søjler eller linjer, f.eks. en procentfordeling i en socialgruppe med fordelingen i hele befolkningen.

**Hvad viste tabellen?**

Tabeller "taler" ikke for sig selv. Formulér en konklusion, der giver et dækkende svar på de spørgsmål, du var interesseret i at tabellen skulle besvare.

**Vurdér tabellen**

Hvor stammer oplysningerne fra? Er der tale om en totaltælling eller en stikprøve-undersøgelse? Hvor stor er stikprøven. Sammenhold relative tal med absolutte tal: er procentgrundlaget rimeligt, hvor stor er den statistiske usikkerhed?

**Huskeliste, når du laver en tabel****Overskrift**

Hvad omhandler tabellen? Overskriften må gerne være så kort som mulig, men samtidig dækkende. Den må gerne samtidig fange interessen, evt. ved at stille spørgsmål og konkludere. Se f.eks. tabel 3.6.

**Tidsangivelse**

Hvornår er oplysningerne fra? Angiver tabellen en bestemt tidsperiode?

**Dataangivelse**

Hvad er optalt? Personer eller kroner? Årets priser eller faste priser?

**Variable**

Hoved- og forspalte angiver efter hvilke egenskaber (variable), oplysningerne i tabellen er fordelt. Husk også her at angive måleenheden (år, kr., socialgruppe)

**Noter og anmærkninger**

Måske skal der tages forbehold overfor nogle af tallene, f.eks. fordi procentgrundlaget er meget lille, tallene er foreløbige, opgørelsesmetoderne er ændrede eller lignende.

**Kildeangivelse**

Hvor stammer tabellens oplysninger fra?

# Figurer giver overblik

Ved hjælp af tabeller får vi styr på vores data, således at de gøres egnet til at arbejde videre med. Ofte er de overskuelige nok i sig selv og velegnede til at trække oplysninger ud af. Men ved meget omfattende tabeller, som vi f.eks. møder dem i publikationerne fra Danmarks Statistik og i Samfundsstatistik, vinder materialet ved at blive yderligere bearbejdet og visualiseret i form af en figur. Det er det, der er sket i figuren "Eksporten og importen fordelt på lande" (Se Samfundsstatistik i afsnittet udenrigsøkonomi.), hvor et ret overvældende talmateriale først er blevet omregnet til procentfordelinger, som så sidenhen er "kogt ned" til 2 cirkeldiagrammer. Herved fremtræder materialet langt mere tilgængeligt og pædagogisk, men "prisen" er samtidig, at der smides noget væk. Nemlig alle de mellem-liggende år og udviklingen i selve eksporten: er den steget eller faldet? Vurder i hvert enkelt tilfælde, om det giver nogen mening, når du bruger grafik. Hvad vindes, hvad mistes?

## Kurver

I Samfundsstatistik i afsnittet "Priser og forbrug" er der en kurve, der viser udviklingen i inflationen i Danmark. Oplysningerne om de årlige inflationstal er prikket ind i et *diagram*, hvor y-aksen angiver procent og x-aksen tiden. Herefter er de enkelte punkter forbundet med hinanden ved rette linjer. Vi kalder det en *knækket* kurve. Figuren viser ikke mere og andet, end man kunne se i en tabel, men vi ser dog fortiden nogle store svingninger også to tydelige langtidstendenser, *trend'er*, én før og én efter 1973. Ved aflæsning af en figur, hvor y-aksen angiver en relativ stigning skal man være meget omhyggelig. Fra 1980 til 1993 falder den årlige inflation fra ca. 10% til knap 2%. Men priserne falder ikke, for så skulle kurven jo skære nullinjen. Derimod falder *prisstigningstakten*. Og den falder ikke med 8%, men med 8 procentpoints (jfr. tidligere, s. 60).

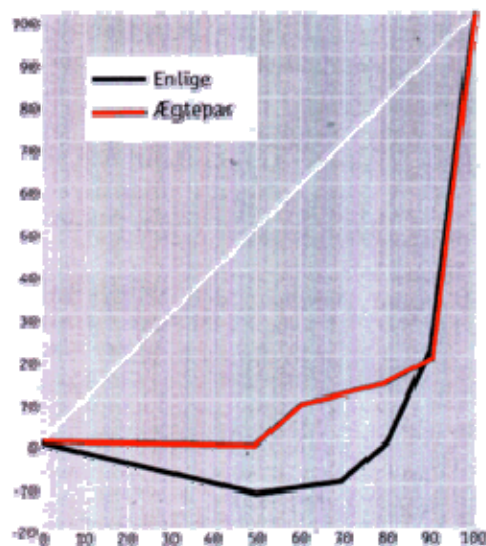
Også figuren "Reallønsindeks" (Se Samfundsstatistik i afsnittet "Indkomst og formue") er baseret på relative tal. Lønudviklingen for de 3 stillingsgrupper er omregnet til indeks med samme basisår.

Figuren kaldes et *viftediagram*. Vi kan her se, hvordan *forholdet* mellem grupperne har udviklet sig - ikke noget om, hvem der tjener

mest. Som i ovenstående figur har diagrammer ofte flere kurver indtegnet, hvor man i aflæsningen kan skyde sig direkte ud på y-aksen. Det kan man ikke gøre i den næste figur, vi skal se på, figuren over bruttofaktoringkomsten fordelt på primære, sekundære og tertiære erhverv (Se Samfundsstatistik i afsnittet "Erhvervsstrukturen". Her har man "*stabet*" de enkelte kurver ovenpå hianden. Tallet for sekundære erhverv i 1930 var altså ikke ca. 50%. De 50% er i stedet den andel, de varemestillende erhverv (primære + sekundære) tilsammen udgjorde af bruttofaktoringkomsten.

Et *Lorenzdiagram* viser en fordeling, men skiller sig ud på den måde, at der ligger angives procent på både x- og y-aksen. Figur 3.2 viser, at 10% af familierne ejer 80% af den samlede formue i samfundet. Jo længere væk fra 45grader linjen den afsatte kurve befinder sig, jo mere skæv/ulig er fordelingen.

Figur 3.2 Lorenzdiagram.  
Formuefordelingen 1992



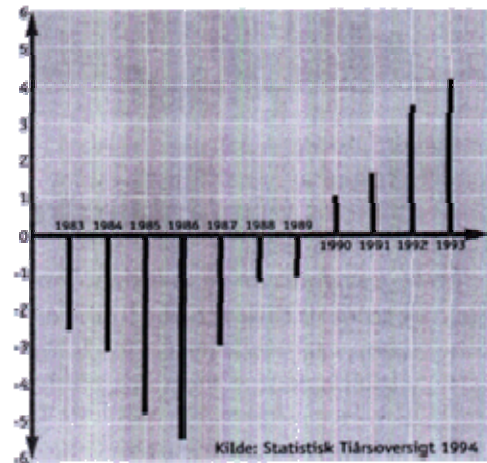
## Pinde, søjler, bjælker og cirkler

Et **pindediagram** som figur 3.3 minder på mange måder om en knækkurve, men i stedet

for at forbinde punkterne angives afstanden til nullinjen ved en lodret linie. Der er dog flere fordele ved at erstatte pindene med **søjler**. De markerer sig tydeligere grafisk, navnlig ved sammenligninger. (Se f.eks. søjlediagrammet Arbejdsløshedsprocenter i EU i afsnittet "Arbejdsmarkedet" i Samfundsstatistik).

Søjler egner sig desuden fint til at vise fordelinger. Både når det gælder faktiske og relative tal.

*Figur 3.3 Pindediagram. Betalingsbalancens løbende poster i % af BNP*



**Bjælke­diagrammer** er søjlediagrammer, der er vendt 90 grader. I figuren "Ændringer i beskæftigelsen fordelt på erhverv" (Se Samfundsstatistik i afsnittet "Erhvervsstruktur") har den vandrette akse både en positiv og negativ skala.

**Cirkeldiagrammer** bruges udelukkende til fordelinger. Her svarer 1% til 3,6 grader. (Find eksempler i Samfundsstatistik).

## Logaritmisk skala

Får man til opgave at lave en kurve over udviklingen i Danmarks eksport i udvalgte år fra 1975 til 1991, kan den, som i figur 3.4 løses ved at prikke tallene for de enkelte varekategorier ind i et koordinatsystem og derefter forbinde punkterne til knækkede kurver. Husk, at afstanden mellem de enkelte år skal angives korrekt på tidsaksen. En aflæsning af figuren viser, at *eksport i alt* samt *industri* har den største vækst, hvad de jo også har, målt i absolutte tal. P.g.a. den store forskel i tallene for de forskellige varegrupper er det derimod umuligt at få et præcist billede af udviklingen på de nederste kurver.