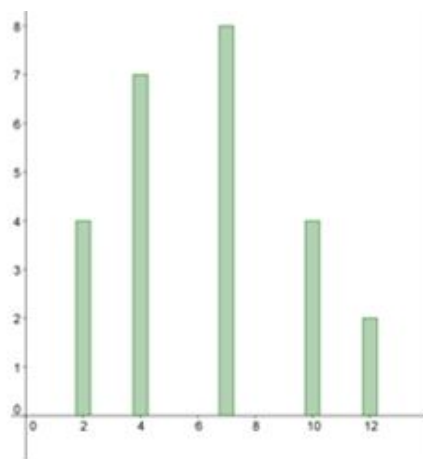


<b>STATISTIK - Fase 3</b> <b>Kritisk vurdere undersøgelser og data</b>			
<i>Vurdering fra 1 til 5 (hvor 5 er højst)</i>			
Læringsmål	Selv	Lærer	Beviser og forslag til forbedring
1. Jeg kan vise, fordele og ulemper ved grafiske fremstillinger og hvordan data kan manipuleres i grafiske fremstillinger.			
2. Jeg kan vurdere usikkerheden i en statistisk undersøgelse på baggrund af datasættets størrelse ( <i>mundtligt – "store tals lov"</i> ).			
3. Jeg kan forklare, hvad et " <b>repræsentativt udvalg</b> " er i tilknytning til en statistisk undersøgelse, og hvorfor repræsentativitet kan have stor betydning ( <i>f.eks. bias</i> ).			
4. Jeg kan bruge et digitalt værktøj til <b>regressionsanalyse</b> .			
5. Jeg kan løse forskellige opgaver vedrørende statistik.			
6. Jeg kender til begreberne nederst.			
<b>Begreber/noter:</b> Stor tals lov, bias, repræsentativt udvalg, tendenser, spredning, variation			



Figur 1. Pindediagram

Observationssæt

7	10	4	4	4
7	2	4	7	7
4	7	2	12	7
2	7	10	4	10
10	12	4	7	2

Middeltal	6,2
Median	7
Variationsbredde	10
Typetal	7

Figur 2. Lille tabel

Karakter	Hyppighed	Frekvens	Summeret frekvens
-3	0	0,00	0,00
00	0	0,00	0,00
02	4	0,16	0,16
4	7	0,28	0,44
7	8	0,32	0,76
10	4	0,16	0,92
12	2	0,08	1,00

Figur 3. Stor tabel



Figur 4. Boksplot

1. Ovenfor ser du fire forskellige repræsentationer af en undersøgelse af karakterer i matematik i 9.b. Du skal vurdere fordelene og ulemper ved de fire repræsentationsformer.
2. Du skal lave en undersøgelse, hvor du sammenligner karakterer i to kommuner. Du må vælge to repræsentationer - begrund dit valg.

### Opgave 1 : Lav et manipuleret diagram

år	besøgstal
2007	12000
2008	12500
2009	12750
2010	12745
2011	12700
2012	12750
2013	13000
2014	13100
2015	12400

Siden 2007 er besøgstallet til Palles Gavebod steget og steget grundet god kampagne fra deres reklame bureau. Men i 2015 går der noget galt.

Sidemakker A: Du er ejer af reklamebureauet, og vil gerne sørge for at blive ved med at lave reklame for "Palles Gavebod" - Lav et diagram og en tekst til diagrammet, der kan overtale Palle til at beholde dem til fortsat at lave reklame for Palles Gavebod.

Sidemakker B: Du er ejer af et andet reklamebureau, og du vil gerne overtage Palles Gavebod som kunde. Lav et diagram og en tekst, der gør at du får kontrakten for den fortsatte reklame.

- Slå sten, saks papir om hvem der er sidemakker A og hvem der er sidemakker B*
- Lav diagram og tekst i forhold til om du er A eller B*
- Vis det til din lærer.*

## Opgave 1 : Vurdering af grafik

- Find fejlen
- Begrund hvordan fejlen er fremkommen?
- Lav et stolpediagram der er rigtig



Kilde: <http://politik.tv2.dk/2015-05-07-her-gaar-det-helt-galt-for-news-jeg-koeber-nye-briller>

## Opgave 2 : Samme statistik:

- Begrund hvorfor de 2 grafer er ens
- Hvilket diagram tror du Socialdemokraterne har brugt i deres annonce?

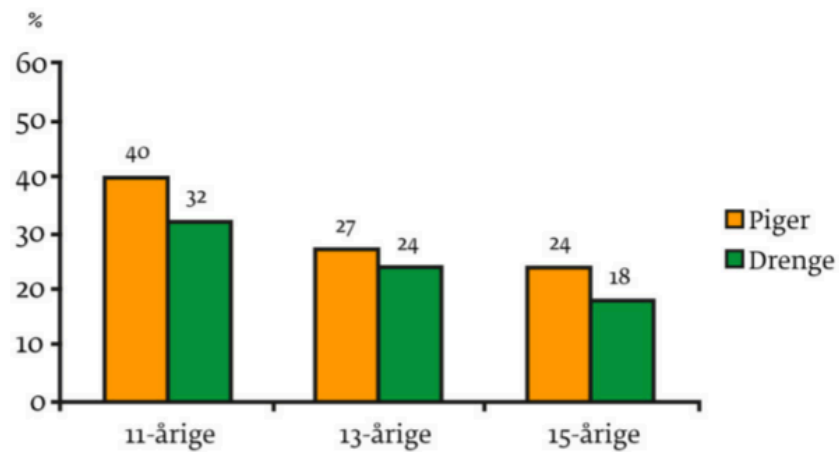


Kilde: <http://www.karenmelchior.eu/ventetid-va-operationer-manipulation-med-grafer/>

### 3.1 Statistik

---

I 2010 gennemførte Statens Institut for Folkesundhed en undersøgelse blandt ca. 4500 repræsentativt udvalgte 11-årige, 13-årige og 15-årige i Danmark. Diagrammet herunder viser børnenes svar på et af undersøgelsens spørgsmål:



Kilde: Statens Institut for Folkesundhed.

Din opgave er at fremstille et eller flere nye stolpediagrammer, der viser de samme svar på måder, som giver læseren et andet indtryk end diagrammet herover. Du skal beskrive, hvordan du har fremstillet diagrammerne, og hvilket indtryk du synes de giver.

### Store tals lov

Forklar med dine egne ord hvad de store tals lov er.

### Hvad er statistisk usikkerhed?

"Hvor stor er usikkerheden på undersøgelsen?" kan man ofte høre nogen spørge. Men usikkerheden er ikke knyttet til undersøgelsen, men til de konklusioner, vi ønsker at uddrage af undersøgelsen ... og hvad værre er, kan den være vidt forskellig fra spørgsmål til spørgsmål og fra konklusion til konklusion.

F.eks. Nogle medarbejdere har svaret på spørgeskemaet i en organisation på udsagnet: *"Min nærmeste leder er alt i alt en god leder"*. I gennemsnit har det givet værdien 3,73. Hvis gennemsnittet er regnet rigtigt ud, er der ikke nogen usikkerhed på dette udsagn. Det er simpelt hen en konstatering af, hvordan krydserne er sat på spørgeskemaet. Men det er jo sjældent den konklusion, vi er ude efter.

Alle ansatte i organisationen giver i gennemsnit værdien 3,73 på udsagnet: *"Min nærmeste leder er alt i alt en god leder"*.

Hvis samtlige ansatte i organisationen har svaret på spørgeskemaet passer udsagnet med "alle", men hvis ikke alle ansatte i organisationen har svaret på spørgeskemaet – så er der en usikkerhed forbundet med at generalisere besvarelsenerne fra de, som har besvaret skemaet, til at gælde for hele organisationen.

Selvom de fleste statistiske undersøgelser gennemføres som *"totalundersøgelser"*, hvor mange inviteres til at deltage, er det alligevel uhyre sjældent, at svarprocenten ender med at blive 100. Spørgsmålet er således hvordan man forholder sig til den usikkerhed, der er forbundet med, at man ikke har svar fra alle, men ønsker at udtale sig om alle?

### Et eksempel

Lad os sige, at vi har en organisation med 1.000 medarbejdere, hvoraf 800 har besvaret spørgsmålet. Med hvilken sikkerhed kan vi på baggrund af de 800 besvarelser udtale os om samtlige 1.000 medarbejdere?

For overhovedet at kunne sige noget fornuftigt om den statistiske usikkerhed må vi lave den forudsætning, at de 200, der ikke har svaret, ikke er fundamentalt anderledes, end de 800 der har. Hvis vi på forhånd har en mistanke om, at de 200 der ikke svarede er væsentlig mere kritiske, væsentlig mere positive eller væsentlig mere ligeglade end de 800, der svarede, har vi et skævt estimat og dermed større problemer end statistisk usikkerhed.

Selv i mindre organisationer vil den statistiske usikkerhed således være til at overskue, så længe man blot har en høj svarprocent. Dvs. det er her "store tals lov" kommer i spil i forhold til at få mere sikkerhed i en undersøgelse.

## 1. Hvad er regression?

Regressionsanalyse handler kort sagt om, at vi skal lave en analyse af, hvordan en udvikling bedst kan beskrives.

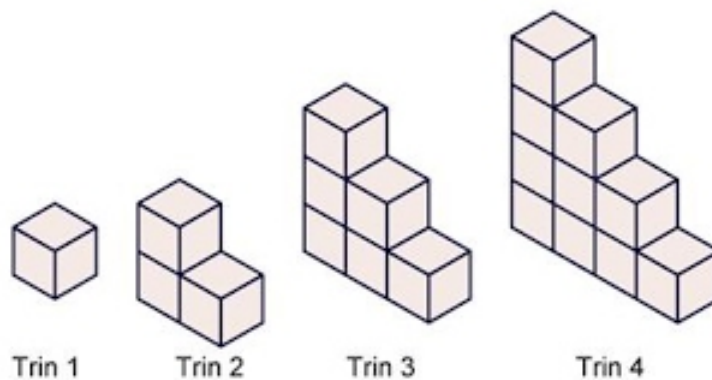
Udviklingen kan beskrives vha. en formel, ligning og/eller en forskrift. Vi får dermed et bud på, hvordan udviklingen "i gennemsnit" udvikler sig.

Nogle gange udvikler ting sig lineært (vokser med samme værdi hele tiden) og andre gange ikke-lineært.

Den linje der bedst passer til punkterne (dataet) hedder "**tendenslinjen**"

## 2. Du skal lave en regressionsanalyse i Geogebra.

Udfyld tabellen der viser antal Centicubes fra Trin 1 til 8.



Trin	1	2	3	4	5	6	7	8
Antal								

- I Geogebra skal du åbne for regneark funktionen. Indtast dine resultater fra tabellen.
- Marker alle tal i regnearket og derefter vælge "To variabel regressionsanalyse"



- Derefter kommer der et "regressionsanalyse" vindue frem.
- Hvis du tænker at det ligner en lineær udvikling skal du vælge "lineær".  
Hvis du tænker at det ligner en ikke-lineær udvikling kan du vælge "potens".  
Hvis du tænker at det ligner en ikke-lineær udvikling kan du vælge "eksponentiel".  
Hvis du tænker at det ligner en ikke-lineær udvikling kan du vælge "polynomial".  
Osv.

7. Nu kan du aflæse forskriften/formlen/ligningen for udviklingen.  
Hældningstallet for fortæller hvilket tal udviklingen udvikler sig med.
8. Kopier grafen ind til tegnefladen og derefter kan du give grafen titler på både x og y-aksen og til hele grafen. Husk linjens forskrift på tegningen.



I skal nu lave jeres egen undersøgelse, hvor I bruger statistik, som I selv indsamler. I kan arbejde sammen i grupper. Arbejdet kan godt strække sig over flere timer.

- 1 Begynd med at diskutere, hvad I gerne vil undersøge. I kan søge inspiration med de links, der findes på Kolorits hjemmeside.  
Skriv det eller de spørgsmål ned, som I gerne vil have svar på.
- 2 Diskuter, hvordan I bedst kan indsamle jeres data. Skal I foretage en meningsmåling – eller en anden form for stikprøveundersøgelse?  
Skal I fremstille et spørgeskema – og hvem skal svare på det?  
Kan I finde jeres data på internettet, i bøger eller i andre medier?  
Skal I bruge et simuleringsprogram til at simulere et eksperiment?
- 3 Gennemfør jeres undersøgelse.
- 4 Analyser jeres data.  
Hvilke tabeller eller diagrammer kan hjælpe med at skabe overblik over materialet?  
Hvilke tal fortæller noget vigtigt om jeres data?
- 5 Diskuter, hvad I har fået at vide ved at gennemføre undersøgelsen.  
Har I fået svar på de spørgsmål, som I stillede fra begyndelsen?  
Var der noget, der overraskede jer?
- 6 Afslut arbejdet med at præsentere jeres resultater for hinanden.  
Hvordan bliver det lettest for andre at få indblik i jeres undersøgelse og resultater?





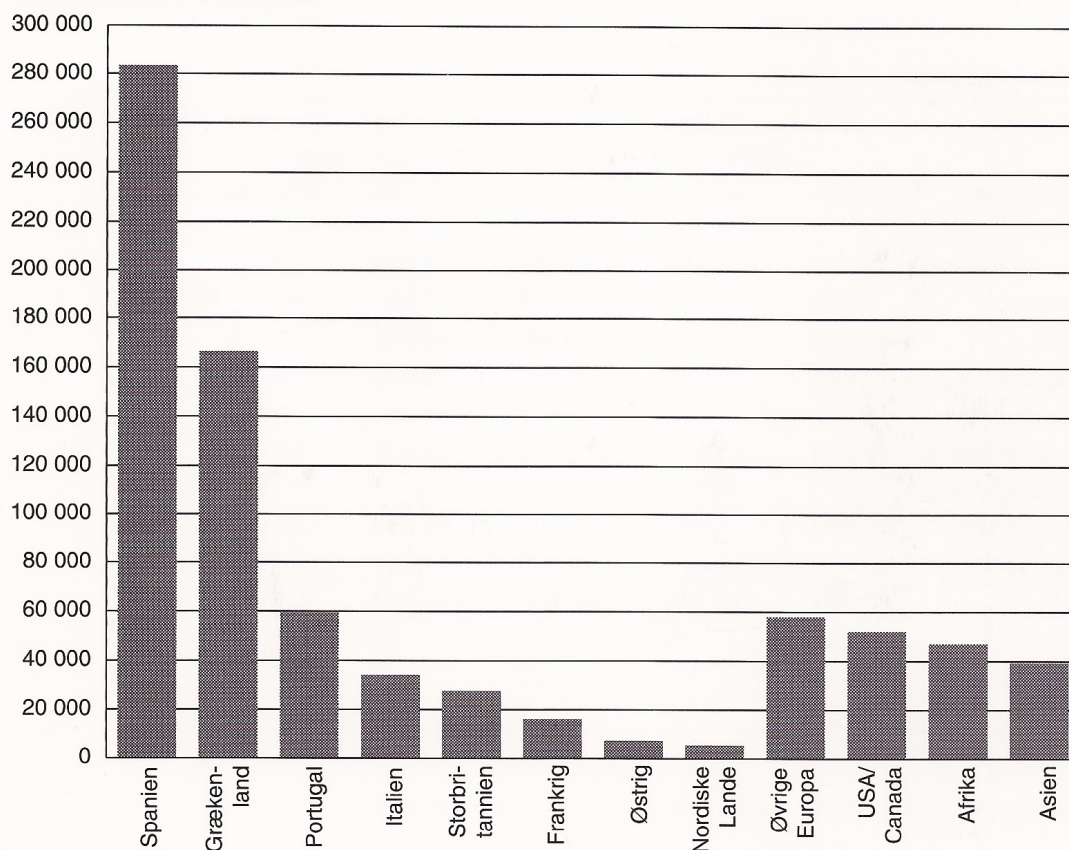
19 Tallene viser hvor mange registrerede ledige, der har været i procent af arbejdsstyrken.

1982	9,8 %
1983	10,5 %
1984	10,1 %
1985	9,1 %
1986	7,9 %
1987	7,9 %
1988	8,7 %
1989	9,5 %
1990	9,7 %
1991	10,6 %
1992	11,4 %

a Illustrer tallene ved hjælp af en kurvetegning i et koordinatsystem, hvor årene afsættes på 1. akse og procenterne på 2. akse med enheden 1 cm = 1%.

c Kommentér forskellen.

## 20 Danskernes feriemål



I 1990 rejste 650.000 på charterferie. I diagrammet ses de mest populære rejsemål. Kilde: Stat.årbog 91.

a Aflæs i diagrammet, hvor mange der tog til Spanien på charterferie.

b Hvor mange procent er det af samtlige charterturister?

c Hvor mange danskere var på charterferie i hele Europa?

d Hvor mange procent af disse var i Spanien?

e Lav et cirkeldiagram, der viser fordelingen på verdensdele.