

# ALT OM STATISTIK OG DIAGRAMMER

## TÆNK SELV!

Statistik er noget af det sværeste i stoffet, fordi man ud fra statistik og diagrammer skal sige noget selvstændigt. Selve de ting man skal kunne er dog ikke svære og det meste kan gøres med hjælp af regneark, Geogebra eller WordMat.

## SÅDAN ARBEJDER DU MED STATISTIK

Start med at finde nogle **Statistiske Deskriptorer**,

find sammenhænge og sig noget om deskriptorerne,

Lav et passende diagram, find sammenhænge og sig noget om dem.

## GRUNDLÆGGENDE STATISTISKE DESKRIPTORER

En samling data hedder et **datasæt** eller et **observationsæt**. Følgende tal (deskriptorer) kan være en god ide at finde for et observationsæt, men det kommer an på hvad det handler om. I det følgende betyder (...) at man skal markere de celler der skal behandles på dette sted. Der står først hvad du skal skrive på dansk Excel, der efter hvad du skal skrive på engelsk Excel.

**Gennemsnit eller middelværdi:** Læg alle tal sammen og del med antallet, eller brug =MIDDEL(...) eller =AVERAGE(...).

**Mindsteværdi:** Den mindste værdi i sættet. =MIN(...) (eller KVARTIL(...;0) eller QUARTILE(...;0))

**Størsteværdi:** Den største værdi i sættet. =MAKS(...) eller =MAX(...)(eller KVARTIL(...;4) eller QUARTILE(...;4))

**Nedre kvartil:** Den største af de 25 % mindste data (den midterste af den nedre halvdel). =KVARTIL(...;1), eller QUARTILE(...;1)

**Median:** Den største af de 50 % mindste data. (den midterste). =MEDIAN(...) eller =KVARTIL(...;2), eller =QUARTILE(...;2).

**Øvre kvartil:** Den største af de 75 % mindste data (den midterste af den øvre halvdel). =KVARTIL(...;3), eller QUARTILE(...;3)

**Variationsbredde:** Forskellen på størsteværdi og mindsteværdi. =MAKS(...) - MIN(...)

**Typetal:** Det tal der forekommer flest gange. Hvis flere tal forekommer flest gange er de alle typetal. =HYPPIGST.FLERE(...) eller =MODE.MULT(...). Vær opmærksom på at dette giver flere tal som så skal sorteres, f.eks. med MIN(..). *TIP!* Det er ofte en fordel at lave en hyppighedstabel som beskrevet nedenfor.

**Hyppighed(-stabel):** Hvor mange gange et tal forekommer i datasættet. Tælles for hvert tal, gerne i en tabel, altså en hyppighedstabel. Få f.eks. excel til at sortere data først. Du kan bruge =TÆL.HVIS(...;tal) til at tælle hvert enkelt tal (hedder COUNTIF på engelsk)

**Frekvens(tabel):** Hvor mange procent af gangene forekommer et tal. (hyppighed/antal). Skal ganges med 100 hvis du laver det i hånden, ellers har excel en tast der kan få data vist som procent...

## DIAGRAMMER TIL VISNING AF DATASÆT

Ofte bliver data mere forståelige når de sættes ind i et diagram. Men hvor skal du bruge hvad? BRUG PINDEDIAGRAM HVIS DU ER I TVIVL, MEN TÆNK DIG OM!

**Histogram** viser en fordeling af de enkelte datas antal. Laves ud fra en hyppigheds- eller frekvenstabel. Se hvordan nedenfor.

**Boksplot** bruges til at vise sammenligninger mellem datasæt. Se hvordan nedenfor.

**Pindediagram** eller **søjlediagram** til at vise niveauer og niveauforskelle.

**Kurvediagram** til at vise hvordan noget udvikler sig.

**Cirkeldiagram** til at vise hvordan noget er fordelt ud af et hele. Fordeling. Det er MEGET sjældent cirkeldiagram du skal bruge.

Lad være med at få noget til at se fancy ud! Det forvirrer ofte mere end godt er.

Brug fanen Anbefalede diagrammer. Den virker mega godt. Er der noget galt med akserne eller labels, må du forsøge dig med forskellige ting med at højreklikke på diagrammet og trykke "vælg data" eller lignende.



## DIAGRAMMER TIL SAMMENLIGNING AF DATASÆT

Hvis du skal sammenligne datasæt, skal de behandles på en eller anden måde. Boksplot er skabt til dette, men en anden mulighed er at fremstille et histogram, en frekvens- eller hyppighedstabel, eller et boksplot.

## STATISTISK FREMSKRIVNING AF DATA

Nogle gange vil man bruge målinger til at forudsige fremtiden. Her gælder det om at finde en form for vækst som man kan bruge til at *modellere* hvad der vil ske. Til at finde en vækstmodel bruges regression (se ALT om *Find den rette funktion eller regression*).

## SÅDAN LAVER DU ET BOKSPLOT

Åben Geogebra!

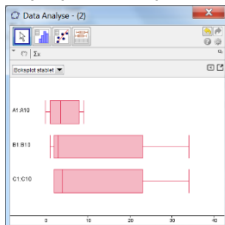
1. Tryk vis regneark (CTRL+SHIFT+S)
2. Paste data ind i regnearket. Hvis der er flere datasæt skal de stå ved siden af hinanden.
3. Marker alt det data du vil analysere (lave boksplot på).

	A	B	C
1	9	23	23
2	1	34	34
3	1	4	4
4	3	23	23
5	4	2	2
6	7	3	3
7	9	3	4
8	9	3	7
9		1	4
10		2	2

4. Tryk på den knap der hedder "Flervariabelanalyse" (eller "enkeltvariabelanalyse" hvis du kun har et datasæt). Der fremkommer nu et nyt, mindre vindue.

	A1:A10	B1:B10	C1:C10
0	23	23	
1	34	34	
1	4	4	
3	23	23	
4	2	2	
7	2	2	
8	3	4	
8	3	7	

5. Tryk på "Analyser".



6. Højreklik på dette billede og tryk på "kopier til tegneblok".

-Nu kan du ændre navnene på dine boksplot (og andre forhold), ved at højreklikke på dem.

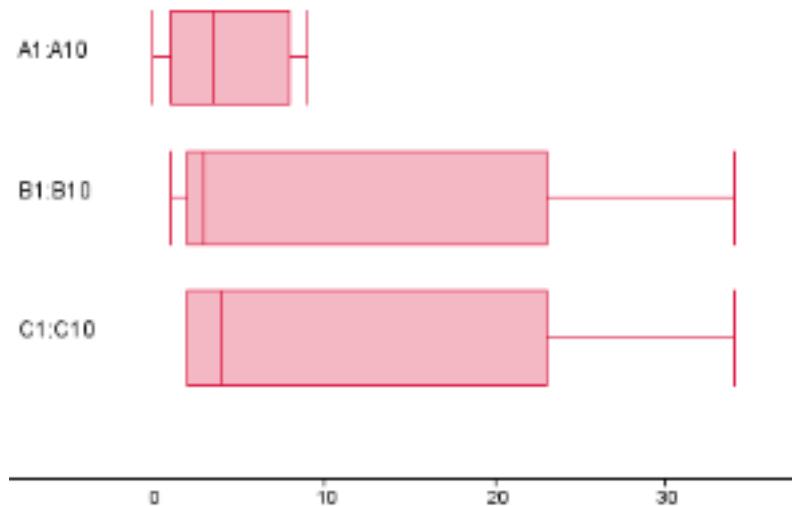
## SÅDAN LÆSER DU ET BOKSPLOT

Boksplot bruges til at sammenligne datasæt, eventuelt af forskellig størrelse.

Boksplot viser mindsteværdi, nedre kvartil, median, øvre kvartil og størsteværdi i et samlet billede.

Husk at mellem hver streg er der repræsenteret lige meget data fra sættet. En smal boks har derfor mange ens data og en lang boks har data spredt meget ud.

*Eksempel:*



Sæt A har data i et meget smalt område. Der er mange ens værdier i toppen og bunden af dets datasæt.

Sæt B har data mere spredt, men rigtig mange ens værdier i de data med lavest værdi.

Sæt C har også meget spredte data, men 50% af data for C ligger mellem 2? og 4? (anslået).

## SÅDAN LAVER DU ET HISTOGRAM I EXCEL (OG GEOGEBRA)

1. Start med at lave en hyppighedstabel, altså hvor meget af datasættet der ligger på en bestemt værdi, eller interval af værdier, hvis du grupperer dine data.
2. Gør det samme for hvert datasæt, hvis du vil sammenligne.
3. Giv hver kolonne en fornuftig titel.
4. Marker dine data og titler og indsæt et pindediagram.

Hvis hellere vil gøre det i geogebra, skal du paste data ind i geogebbras regneark, markere dem og vælge enkeltvariabelanalyse. Her er der en mulighed for at lave et histogram.