Lineær regression i GeoGebra

*Lineær regression* består i at lade GeoGebra finde den bedst mulige linje gennem en samling punkter fra et datasæt. Linjen kaldes *regressionslinjen* hørende til datasættet. Du skal senere lære andre typer af regression.

**Sådan gør du:**

1. Få vist [regnearket](http://ggbkursus.dk/geogebra/1-kom-godt-i-gang/regnearket-geogebra/) under menuen ‘Vis’ og derefter Regneark (*Tastaturgenvej CTRL + SHIFT + S*).
2. Indtast din tabel i de 2 første kolonner kaldet A og B - I den første række kan du skrive navnene på dine variable (her ”x-værdi” og ”y-værdi”)



1. Marker alle tallene.
2. Tryk på **den lille** **trekant** ved -knappen. I den fremkomne menu vælger du ‘*Regressionsanalyse’ .* Derefter kommer “Datakilde” frem som vist nedenunder.

3. Tryk på Analyser.

 Nu ser du følgende vindue, hvor punkterne er blevet plottet i et koordinatsystem:


1. Vælg ‘lineær’ under *regressionsmodel. (Hvis du tænker, at det er en lineær udvikling. Ellers afprøv polynomiel, eksponentiel etc.)
*
2. Nu har GeoGebra indtegnet den bedste rette linje gennem punkterne. Nederst i billedet kan du aflæse ligningen for denne rette linje (regressionsligningen). I dette tilfælde y=2,2x+1,5 (husk at bruge det danske komma og ikke punktum, når du skal fortælle, hvad dit facit bliver!)

**Kopiering af graf over på tegneblokken**

1. Du skal nu vælge at overføre grafen til tegneblokken (så du senere kan kopiere grafen og sætte over i et word-dokument til din aflevering). Det kan du gøre ved at højreklikke på grafen og derefter vælge “Kopier til tegneblok”.

2. Du skulle gerne få følgende billede



Bemærk at det kan være nødvendigt at forstørre/formindske tegneblokken, så vi kan se akserne.

Læg også mærke til at vores punkter nu er blevet gemt i en liste, som GeoGebra her har kaldt liste1 (ses i Algebra-vinduet).

Regressionsligningen er også blevet gemt som en funktion, i dette tilfælde under navnet g(x) (ses ligeledes i Algebra-vinduet). Vi kan altså udregne funktionsværdier direkte i inputfeltet ved at skrive for eksempel g(1), hvis vi vil udregne værdien af g i x=1.

**Residualdiagram:**

Til at afgøre hvor godt vores modelværdier passer med vores måledata, kan vi lave et *residualplot*. Et residualplot giver et grafisk billede af afvigelserne mellem måledata og værdierne ifølge modellen.

Bemærk, at GeoGebra faktisk laver residualplottet samtidig med den lineære regression. Når du er i det lille vindue kaldet *Data Analyse* (der hvor du valgte at lave den lineære regression), skal du i stedet for *Punktplot* vælge *Residualdiagram*:



**Forklaringsgraden** $r^{2}$**:** Når vi laver regression, skal vi også finde den såkaldte forklaringsgrad, $r^{2}$. Hvis denne værdi er tæt på 1, er sammenhængen klart lineær. Er værdien tæt på 0, er sammenhængen ikke lineær. Forklaringsgraden siger altså noget om hvor god en model vi har fået lavet. Før du finder $r^{2}$, så tjek at GeoGebra er sat til at lave afrundinger med mere end 2 decimaler (Gøres under Indstillinger 🡪Afrunding 🡪vælg antal decimaler - for eksempel 4)

1. For at finde forklaringsgraden skal du skrive følgende kommando i inputfeltet:

Rkvadreret[ <Liste med Punkter>, <Funktion> ]

<Liste med Punkter>: Her skriver du navnet på listen over vores punkter, i dette tilfælde liste1.

<Funktion>: Her skriver du navnet på vores regressionsligning, i dette tilfælde g.

Når du trykker Enter, vises værdien af forklaringsgraden i Algebra-vinduet (under numerisk). I denne øvelse giver den 0,9308, dvs. ca. 0,93.

Ekstra hjælp: På følgende hjemmeside finder du en video, der viser punkt 1-8 udført i praksis og ”live” (videoen er placeret et stykke nede på siden - efter det sidste punkt): <http://ggbkursus.dk/geogebra/3-kan-det-meste/sadan-laves-en-regressionsanalyse/>