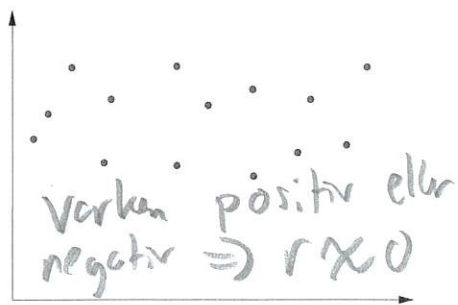


FACIT

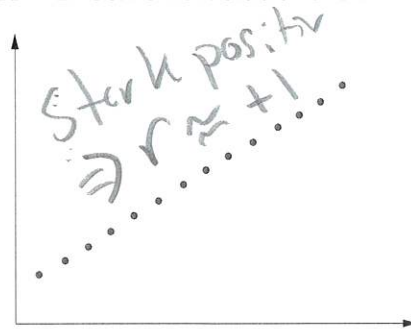
Korrelationskoefficient och linjärregression

Del 1 – Utan digitala hjälpmedel

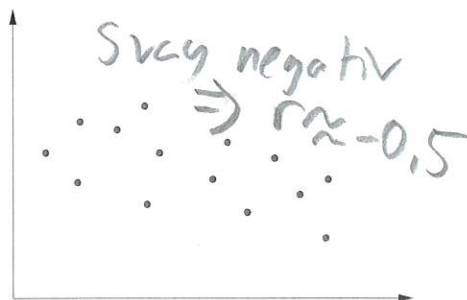
1. Nedan visas fyra spridningsdiagram, A – D och 6 korrelationskoefficienter, r_1 till r_6



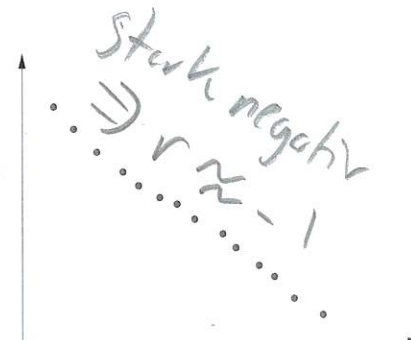
A



B



C



D

$$r_1 = 0,9998$$

$$r_2 = -0,5446$$

$$r_3 = -0,9971$$

$$r_4 = 0,4412$$

$$r_5 = -0,8204$$

$$r_6 = 0,0131$$

Para ihop korrelationskoefficienterna med diagrammen.

(2/0/0)

Observera att det blir två korrelationskoefficienter över.

Storleken på r avgörs av hur spridda prickarna är: Utspridda \Rightarrow litet r , inte utspridda $\Rightarrow r \approx \pm 1$

A: $r \approx 0$ || B: $r \approx 1$ || C: $r \approx -0,5$ || D: $r \approx -1$
 $r_6 = 0,0131$ || $r_1 = 0,9998$ || $r_2 = -0,5446$ || $r_3 = -0,9971$

2. Ztatte Zteak läser i en mattebok att för korrelationskoefficienten, r alltid gäller att $-1 \leq r \leq 1$.

Förklara för Ztatte vad detta innebär.

(1/1/0)

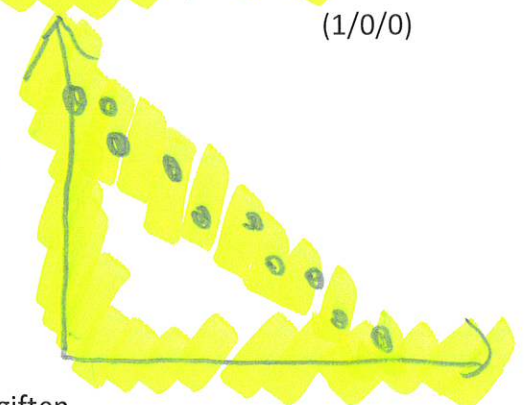
r är som en slags procentsiffra som visar hur väl punkternas följer en rät linje. Därför är r mellan 0% och 100%. r -värdet visar även om linjen lutar uppåt eller nedåt, därav + eller -

3. Rita ett spridningsdiagram med minst 10 prickar där korrelationskoefficienten, $r < -0,9$.

exempelvis

(1/0/0)

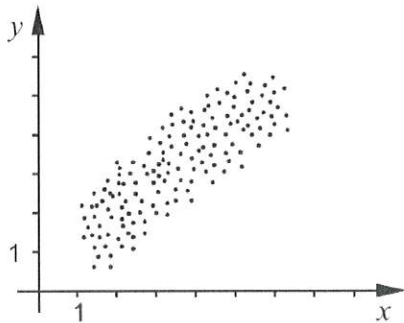
$r < -0,9 \Rightarrow$
stark negativ korrelation
 \Rightarrow inte så spridda punkter
med negativ lutning



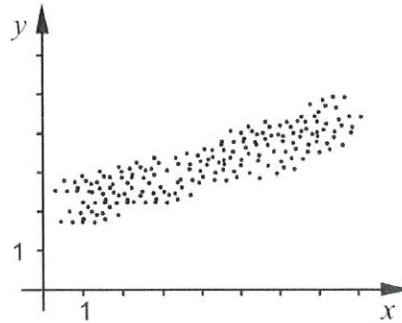
4. Uppgiften nedan är ifrån ett gammalt nationellt prov. Lös uppgiften.

Nedan visas fyra spridningsdiagram A-D.

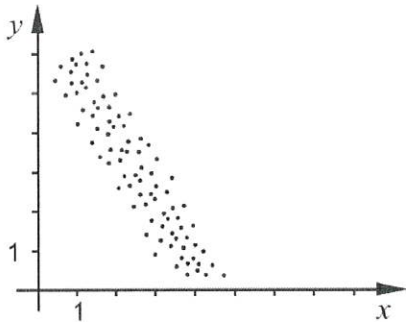
A.



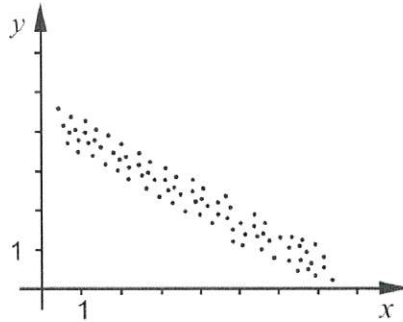
B.



C.



D.

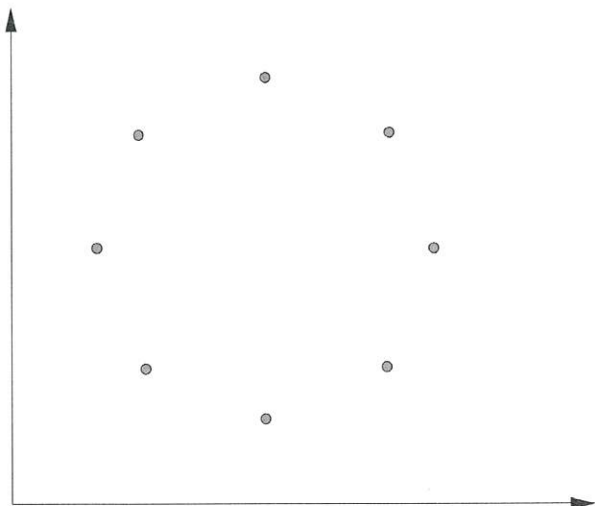


- a) Vilket/vilka av diagrammen A-D visar en negativ korrelation? Motivera. (1/0/0)
b) Vilket av diagrammen A-D visar starkast korrelation mellan variablerna x och y? Motivera. (0/1/0)

a) Negativ korrelation \Rightarrow lutar nedåt
C och D

b) Starkast korrelation \Rightarrow minst utspridda punkter
D

5. Figuren visar ett ovanligt spridningsdiagram med 8 punkter placerade symmetriskt i en cirkel.



Mattias undrar vad det är för korrelationskoefficient för dessa punkter och tänker följande:

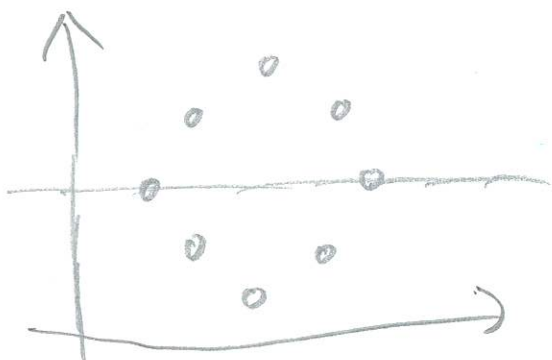
Det är ju ett jättetydligt samband! Alltså borde r vara ett högt tal. Frågan är bara om r är $+1$ eller -1 ...?

Hjälp Mattias att bestämma vad som gäller för korrelationskoefficienten för dessa punkter.

(0/2/0)

Korrelationskoefficienten r visar hur väl punkterna följer linjära samband.

Dessa punkter är symmetriskt placerade kring en linje med lutningen noll:



\Rightarrow

r är mycket närmare noll än någon av ytterligheterna $+1$ el. -1

Del 2 – Med digitala hjälpmedel

- D1. Johanna går en kurs i keramik, och har gjort en fin vas i lera. Denna ska brännas i en ugn. Ugnen värms upp med vasen i. Vid upphettningen höjs temperaturen långsamt.

Johanna mäter temperaturen på ugnen vid några tillfällen, vilket visas i tabellen nedan.

	A	B	C	D	E
Tid (timmar)	0,4	1,25	2,05	2,7	3,5
Temperatur (Celsius)	49	86	142	179	226

← x
← y

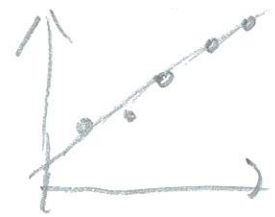
- a) Använd tabellen för att ta fram ett linjärt samband mellan temperaturen och tiden där y är ugnens temperatur i Celsius och x är tiden i timmar räknat från att ugnen sattes i gång.

(2/0/0)

1) Lägg in datat som punkter: $(0,4, 49), (1,25, 86), \dots$

2) "Regression Lin(A, B, C, D, E)"

$$\Rightarrow y = 58,36x + 20,86$$



- b) Enligt sambandet i a), hur lång tid tar det att värma ugnen från 320°C till 450°C ?

(0/1/0)

$$320^\circ\text{C} \Rightarrow y = 320 \Rightarrow \text{Skärning (f, 320)} \Rightarrow x \approx 5,13$$

$$450^\circ\text{C} \Rightarrow y = 450 \Rightarrow \text{Skärning (f, 450)} \Rightarrow x \approx 7,36$$

$$\text{Tid mellan} \Rightarrow 7,36 - 5,13 = 2,23 \text{ h}$$

- D2. Forskare har undersökt antalet hur antal fågelarter på ett visst berg varierar med höjden över havet.

Deras mätvärden visas i tabellen till höger.

- a) Ta fram *korrelationskoefficienten* för mätvärdena. (1/0/0)

	Höjd (m)	Antal fågelarter
A	1250	78
B	1380	76
C	1510	68
D	1500	62
E	1560	60
F	1740	46
G	2080	39
H	2100	36

- b) Tolka svaret i a)

$$r \approx -0,97$$

(0/1/0)

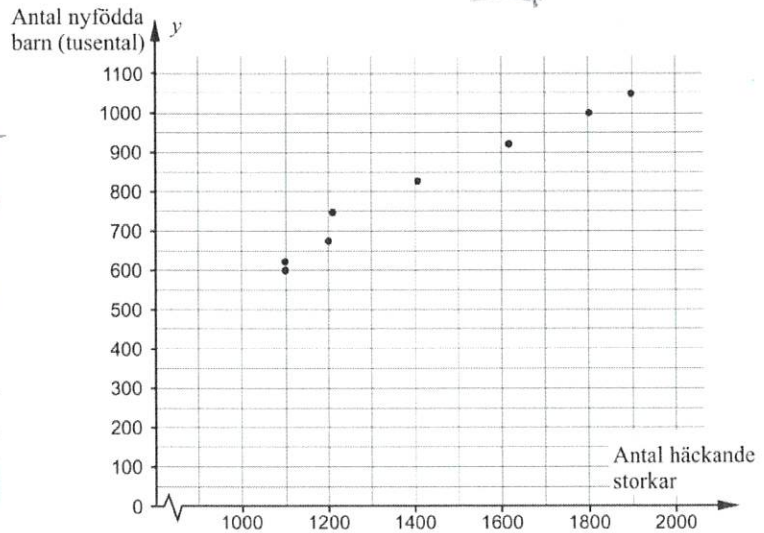
Det finns en stark korrelation mellan antalet fågelarter och höjden över havet, så att en ökning av höjden innebär en minskning av antalet fågelarter. (Troligtvis finns även en biologisk kausalitet.)

D3. Uppgiften nedan är ifrån ett gammalt nationellt prov. Lös uppgiften.

Nedanstående tabell och diagram visar antal häckande storkar respektive antal nyfödda barn i Västtyskland mellan åren 1965 och 1978.



År	Antal häckande storkar	Antal nyfödda barn (tusental)
1965	1900	1050
1966	1800	1000
1968	1610	920
1970	1405	825
1972	1208	750
1974	1200	675
1976	1100	620
1978	1100	600



- a) Bestäm ett linjärt samband mellan antal nyfödda barn i tusental, y , och antal häckande storkar, x .

(2/0/0)

"Linjärt samband" = Rät linje

* Lägg in datat som punkter: (1900, 1050), (1800, 1000)...

* "Regression Lin(A, B, C, D, E, F, G, H)"

$$\Rightarrow y = 0,53x + 49,88$$

- b) Korrelationen mellan x och y är 0,99. Simon drar slutsatsen att det finns ett starkt orsakssamband mellan antal nyfödda barn i tusental och antal häckande storkar i Västtyskland.

Har Simon rätt? Motivera ditt svar.

(0/1/0)

Simon har tyvärr fel. Att $r \approx 0,99$ visar på en stark korrelation (dvs storkarna och barnen kom samtidigt), men ingenting om VARFÖR. Troligtvis är det bara slump som gör att det händer samtidigt, men oavsett säger inte r något om orsak.

D4. Mattias har köpt kebabrullar ifrån samma pizzeria under lång tid.

Han har noterat att priset har ökat genom åren.



Tre av värdena visas i tabellen nedan.

	A	B	C
År	2016	2018	2020
Pris (kr)	65	79	90

x
y

a) Utgå från tabellen och ta fram ett linjärt samband. Använd detta samband för att bestämma vad priset borde vara år 2030.

(1/1/0)

Skriv in punkterna: $(2016, 65), (2018, 79), (2020, 90)$
 "Regression Lin(A, B, C)" $\Rightarrow y = 6,25x - 12534,5$
 År 2030 $\Rightarrow x = 2030 \Rightarrow y = 153 \text{ kr}$

b) Under denna tid har Mattias noterat att kebabrullarnas längd minskat enligt tabellen:

	D	E	F	G	H
År	2016	2017	2018	2019	2020
Längd (cm)	39,4	39,2	38,6	38,3	37,6

x
y

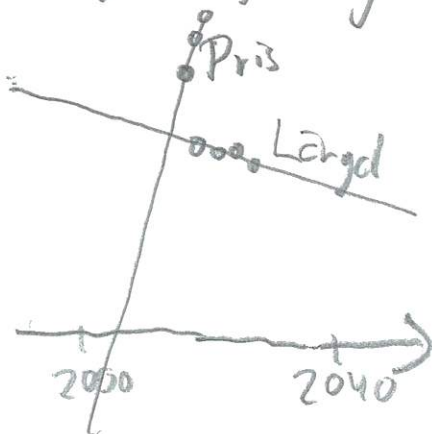
Antag att även längden kan skrivas som ett linjärt samband.

Hur många gånger har priset per cm ökat från år 2016 till det år då kebabrullen är 20 cm lång?

(0/2/1)

Lägg in punkterna: $(2016, 39.4), (2017, 39.2), \dots$
 "Regression Lin(D, E, F, G, H)" $\Rightarrow L(x) = -0,45x + 946,72$

Enligt a) gäller att $\text{Pris}(x) = 6,25x - 12534,5$



År då längden är 20 cm:
 Skärning $(L, 20) \Rightarrow 2059,38$

Pris / cm	
År 2016	År 2059
$\frac{65}{39,4} = 1,65$	$\frac{336,6}{20} = 16,83$

\Rightarrow Ökning: $\frac{16,83}{1,65} \approx 10,2 \text{ ggr}$