II. Linear Correlation Coefficient r

A <u>correlation</u> exists between two variables when the values of one or more associated with the values the others.

A <u>linear correlation</u> exists between any two variables when there is a pattern that can be approximated by a straight line.
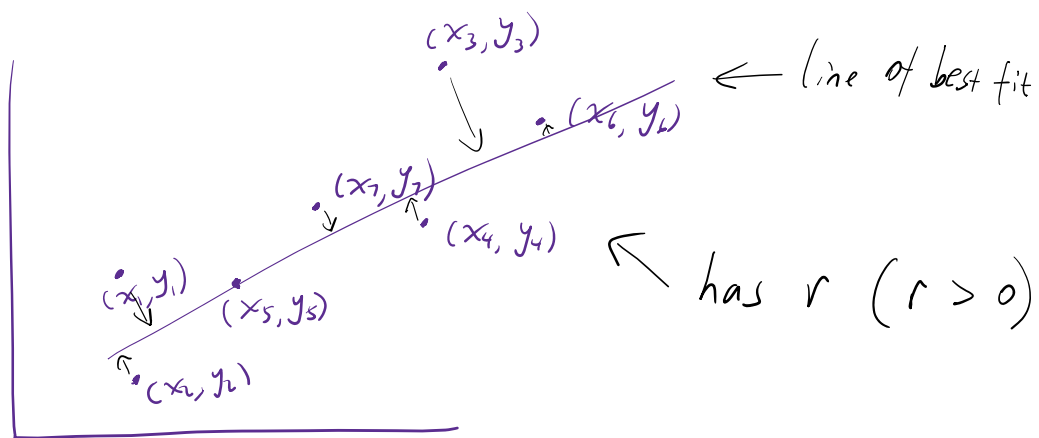
A <u>linear correlation coefficient</u>, r, ← special slope is a number that measures how well the <u>pairs of sample data values</u> would fit in a straight line.

Property :

- $-1 \leq r \leq 1$

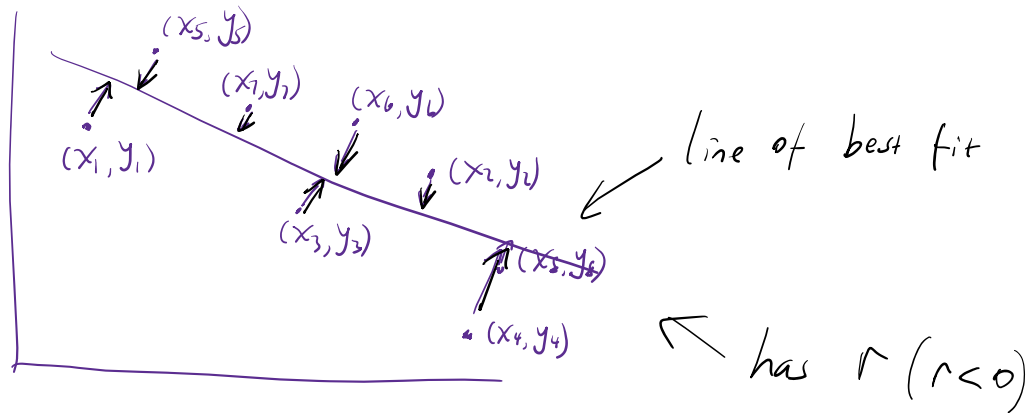- It represent the "strength" of linearization

- Affected by outlier(s).

eg (x_1, y_1)
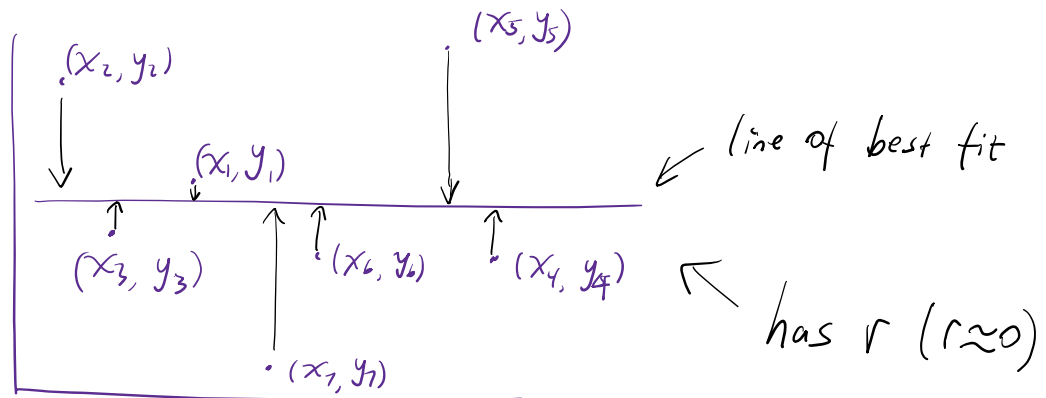(x_2, y_2)
$\vdots$ $\Longrightarrow$
(x_n, y_n)

(x_3, y_3)

← line of best fit

(x_6, y_6)

(x_7, y_7)

(x_4, y_4)

(x_1, y_1)

(x_5, y_5)

(x_2, y_2)

↖ has r (r > 0)

eg (x_1, y_1)
    (x_2, y_2)
    $\vdots$ $\Rightarrow$

(x_5, y_5)
(x_7, y_7)
(x_6, y_6)
(x_1, y_1)
(x_2, y_2)
(x_3, y_3)
(x_5, y_6)
(x_4, y_4)

    (x_n, y_n)

← line of best fit

← has r (r < 0)

eg (x_1, y_1)
    (x_2, y_2)
    $\vdots$ $\Rightarrow$

(x_5, y_5)
(x_2, y_2)
(x_1, y_1)
(x_3, y_3)
(x_6, y_6)
(x_4, y_4)
(x_7, y_7)

    (x_n, y_n)

← line of best fit

← has r (r ≈ 0)

Overall,
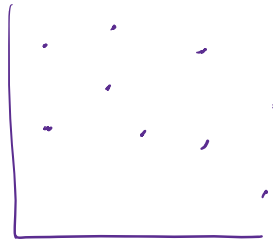
Positive correlation
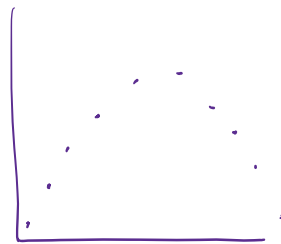
$r \approx 0.62$

negative correlation

$r = -0.57$

non linear relation

$|r| > 1$

not for this class

no relation

$r = 0$

Formula $$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}}$$

Eg Find the linear corrlelation coefficient for the pair of datas:
(1, 12), (2, 6), (3, 5), (4, 2), (6, 3)

S: Eg. Find the linear correlation coefficient.
(X , y)

X	Y	XY	X²	Y²
1	12	12	1	144
2	6	12	4	36
3	5	15	9	25
4	2	8	16	4
6	3	18	36	9
Total: 16	28	65	66	218

← Σ

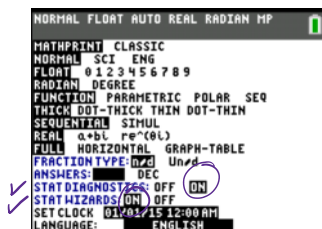n is 5, since 5 points

$$r = \frac{5 \cdot 65 - (16)(28)}{\sqrt{5(66)-(16)^2}\sqrt{5(218)-(28)^2}} \approx \boxed{-0.82}$$
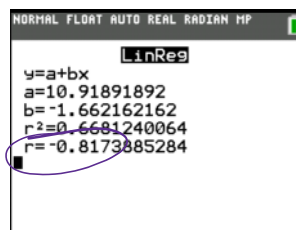
TI-84: stat → CALC → 8: LinReg $(a+bx)$

with condition

L_1 for x's
L_2 for y's

$x_1 \ y_1 \quad x_2 \ y_2 \quad x_3 \ y_3 \quad x_4 \ y_4 \quad x_5 \ y_5$
(1, 12), (2, 6), (3, 5), (4, 2), (6, 3)


```
LinReg(a+bx)
Xlist:L1
Ylist:L2
FreqList:
Store RegEQ:
```

→

```
LinReg
y=a+bx
a=10.91891892
b=-1.662162162
r²=0.6681240064
r=-0.8173885284
```

Eg Find the linear corrlelation coefficient r:
(3, 5), (6, 6), (2, 4), (8, 15), (10, 25)

S:
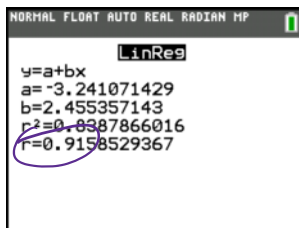
III. Scatterplot

It is a correlation that refers to the existance of a relationship between two variables.

TI-84 : 2nd → statplot → "on with ⌐⋅⋅ → graph → Zoom → 9: ZoomStat

Eg Construct a scatterplot from the following:
(3, 5), (6, 6), (2, 4), (8, 15), (10, 25)

S: