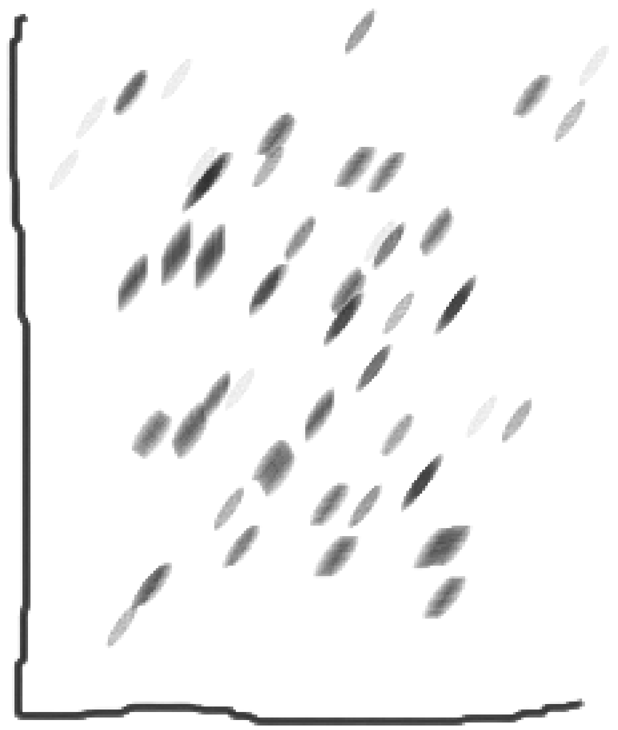
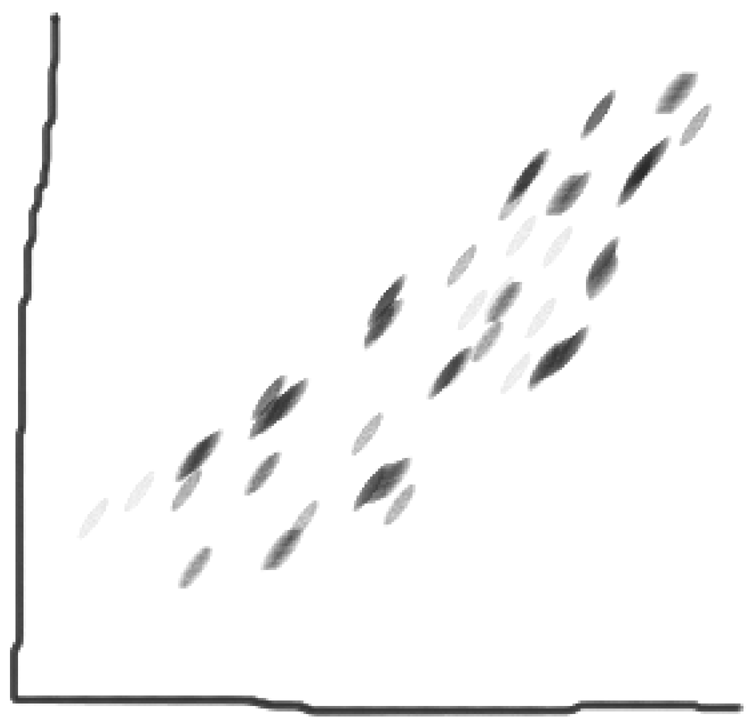
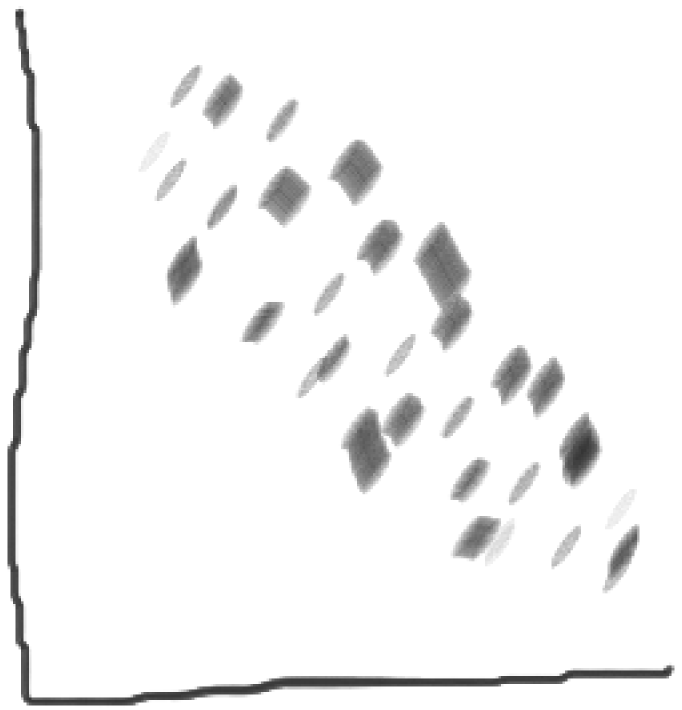
Correlation



R near 0



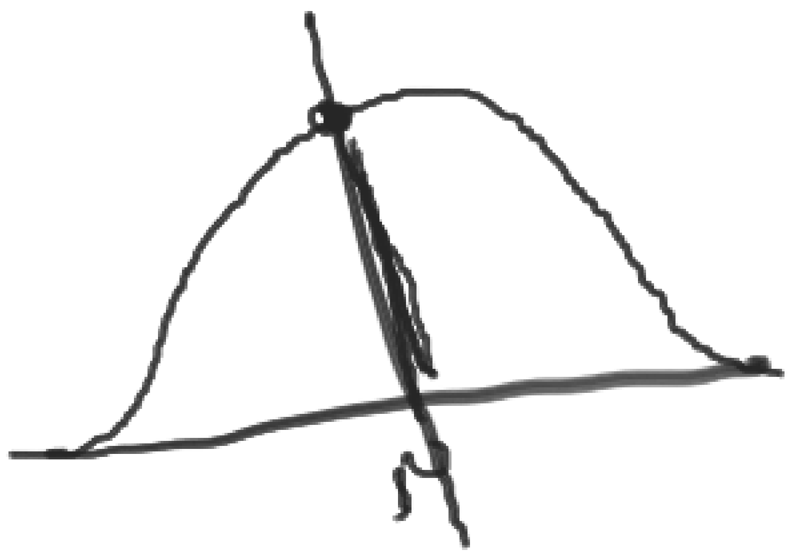
+ve
R near 1



-ve
R near -1

Variance

$$V = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{200}{10} = 20$$



when sample are less

$$\frac{\quad}{n-1}$$

$$= \frac{20}{5-1} = \frac{20}{4} = 4+1 = 5$$

$+2 \quad -2 \quad -3$

$+5$

$= 1$

$2^2 \quad -2^2 \quad -3^2$

4^2

$= 20$

$x_i =$

$\bar{x} =$

$n =$

when 1 var is there

$(x_i - \bar{x})^2$

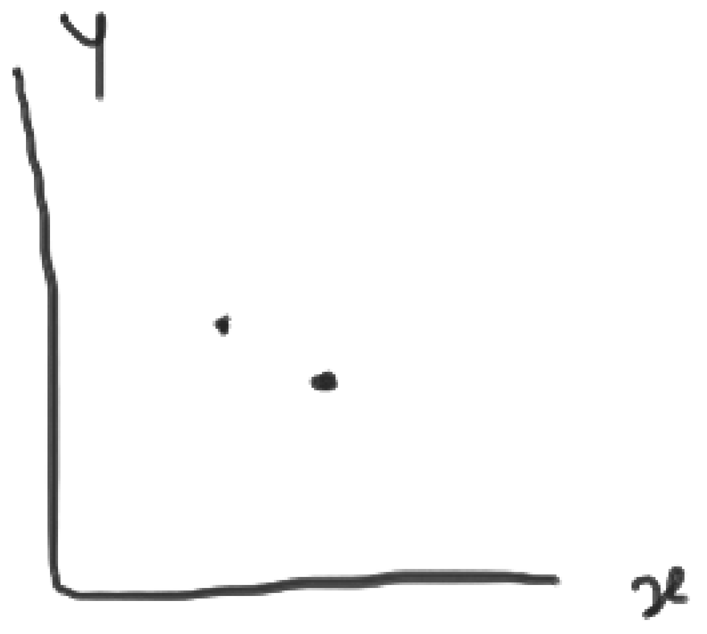
cv = central value

x dual var

158
178
172

Covariance

$$\text{Covariance} = \frac{(x_i - \bar{x}) \times (y_i - \bar{y})}{n}$$



v 1
v 2

$\text{cov} = 0$

$\text{corr} = 0$

$$\text{Var} = \frac{\Sigma (x_i - \bar{x})^2}{n}$$

$$(\sigma) \ \text{std} = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}} \quad - \sqrt{\text{var}}$$

Pearson's Correlation Coefficient

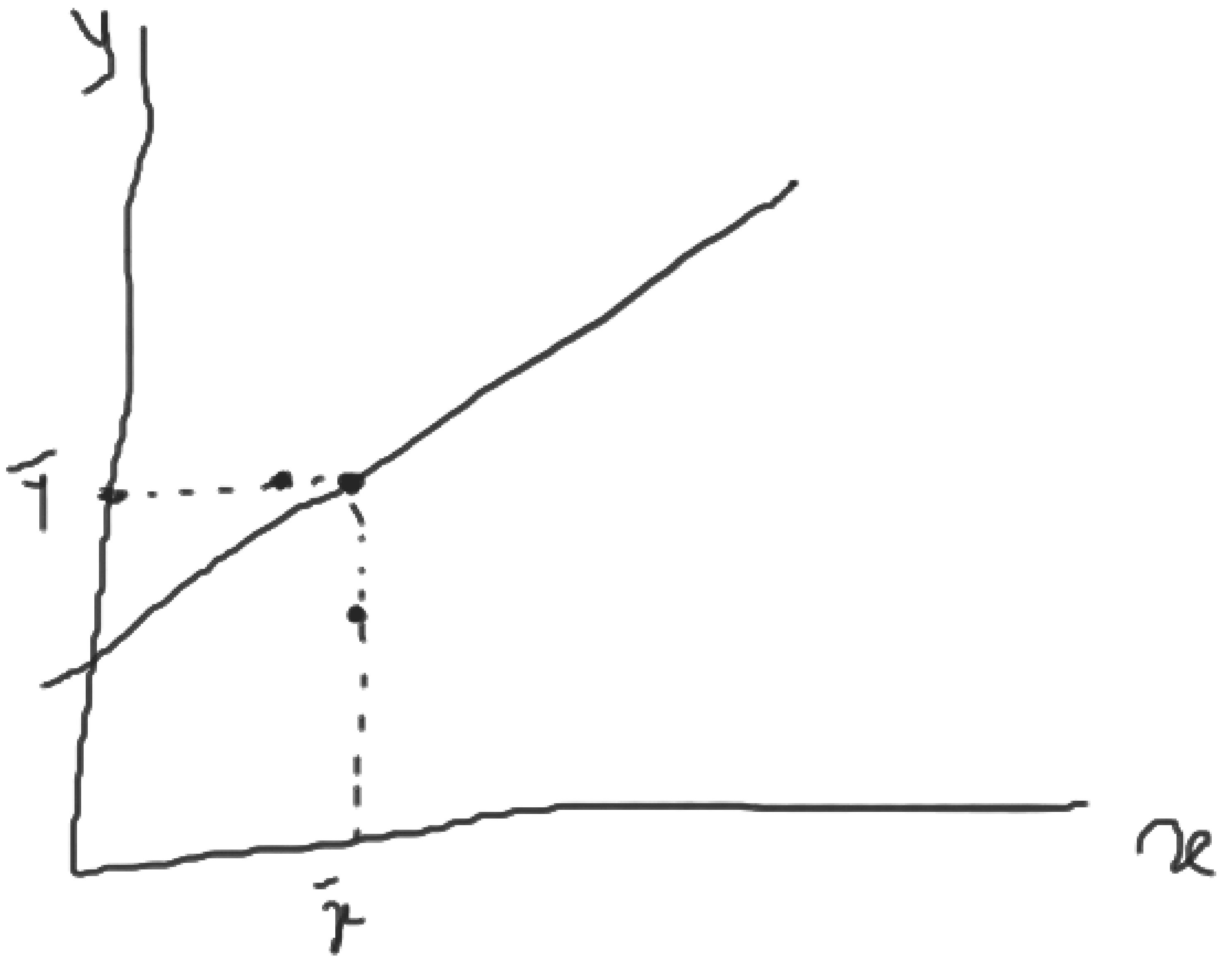
$$\frac{\frac{1}{\cancel{x}}}{\frac{3}{\cancel{x}}}$$

$$R_{(x,y)} = \frac{\Sigma (x_i - \bar{x}) \times (y_i - \bar{y})}{\sqrt{\Sigma (x_i - \bar{x})^2} \times \sqrt{\Sigma (y_i - \bar{y})^2}} \rightarrow \text{covariance}$$

$$- \text{std}(x), \text{std}(y)$$

$$r_{(x,y)}$$

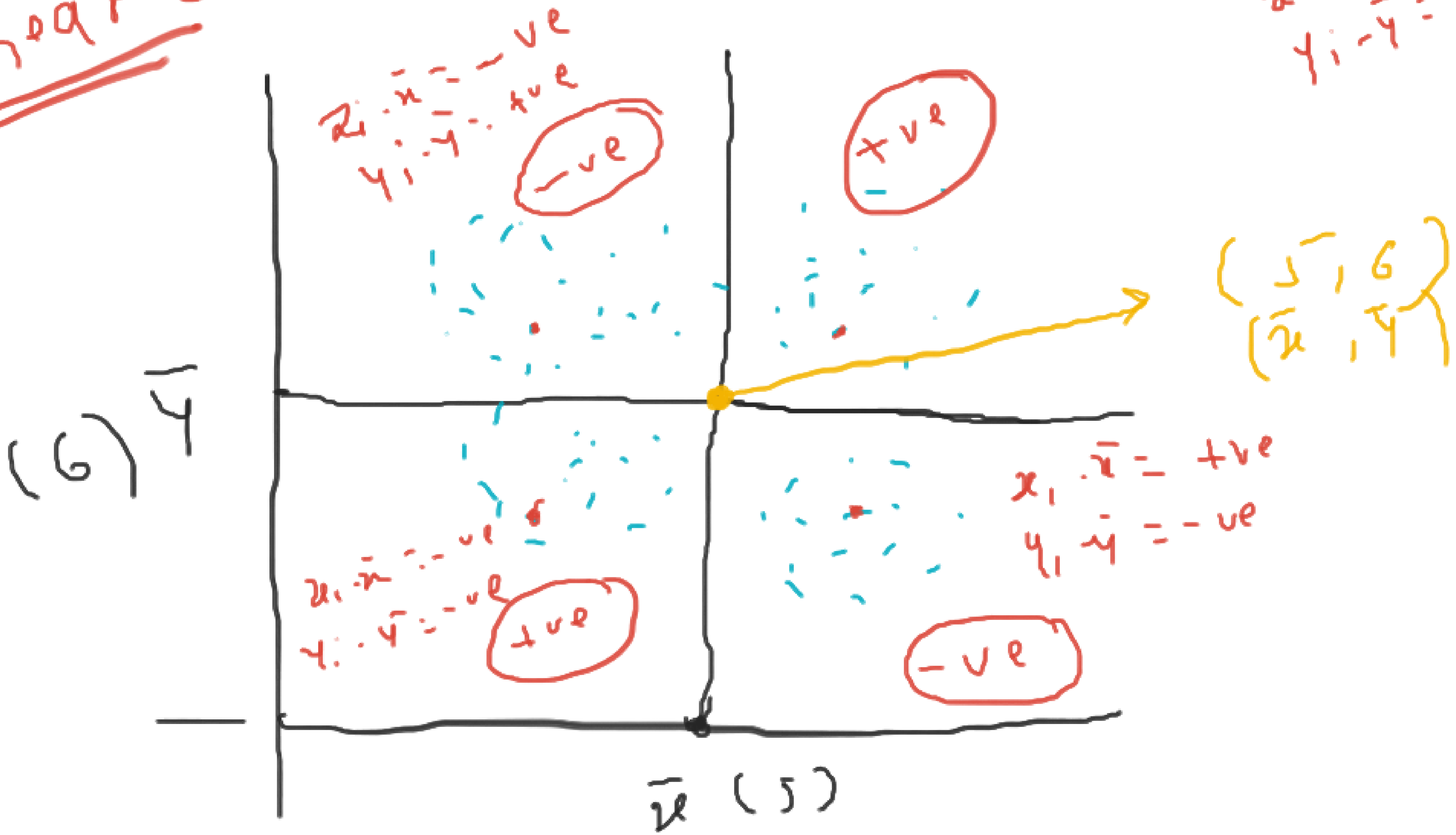
$$r_{(x,y)} = \frac{\text{covariance}}{\sigma_x \sigma_y}$$


y
$\bar{y}$
$\bar{x}$
x


$\bar{y}$
$\bar{x}$

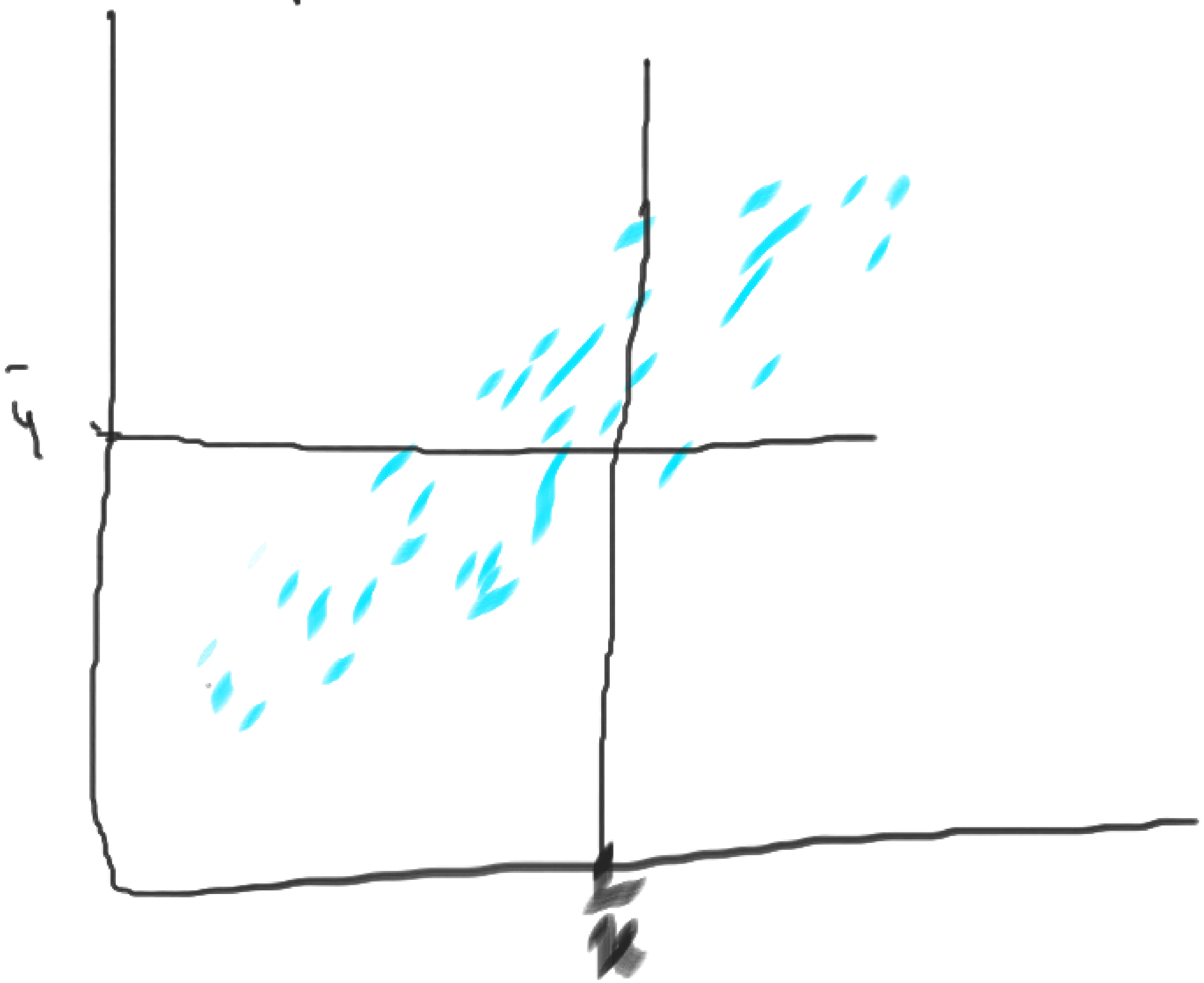
R = near 0

$x_i - \bar{x} = 8 - 5 = +3$
$y_i - \bar{y} = 8 - 6 = +2$



x_i
x_n
28
-
28
-

R near 1



20

24

- 66

1

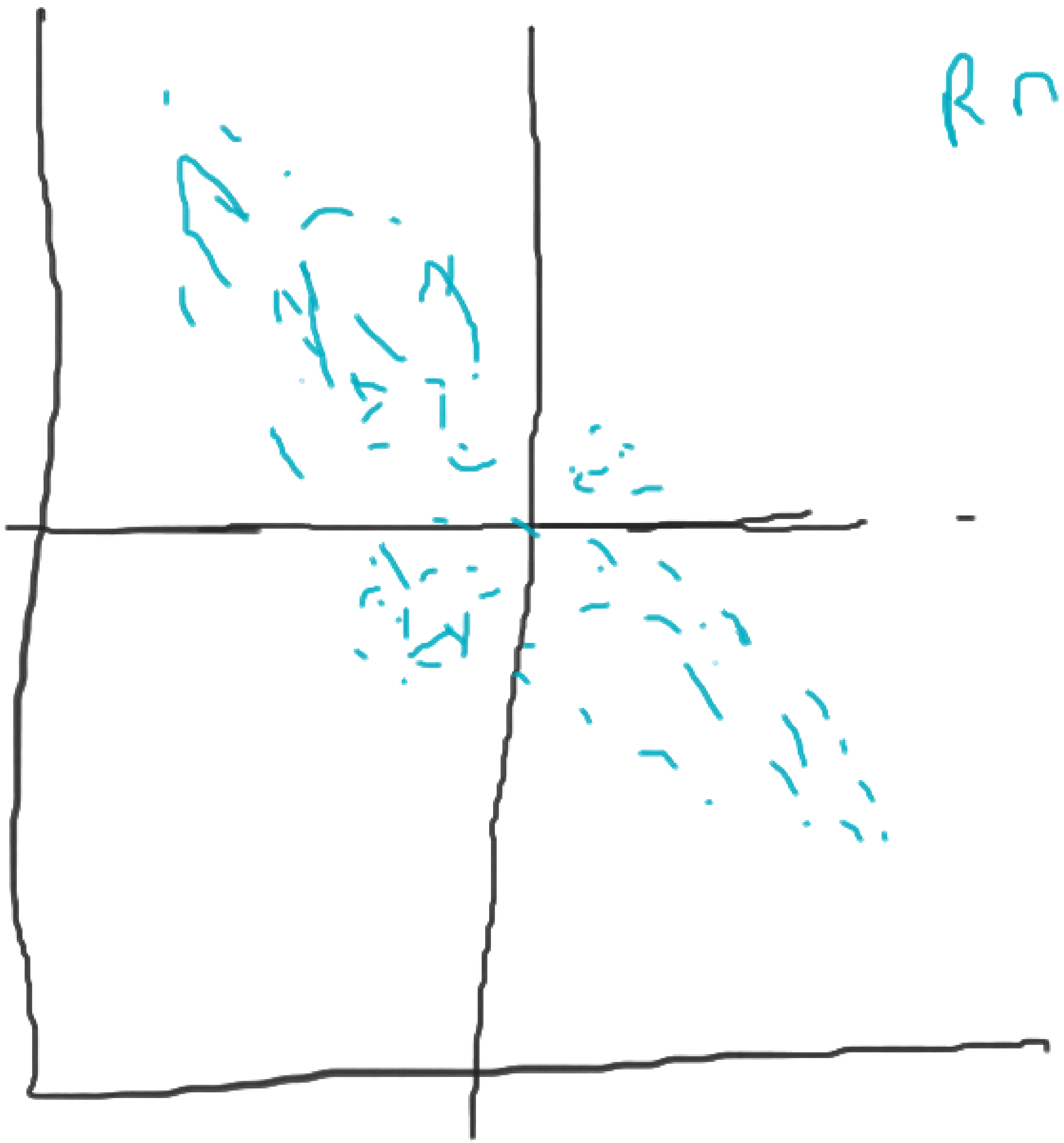
-

37

=

$\frac{37}{51}$

-0.7



Rneqr — 1

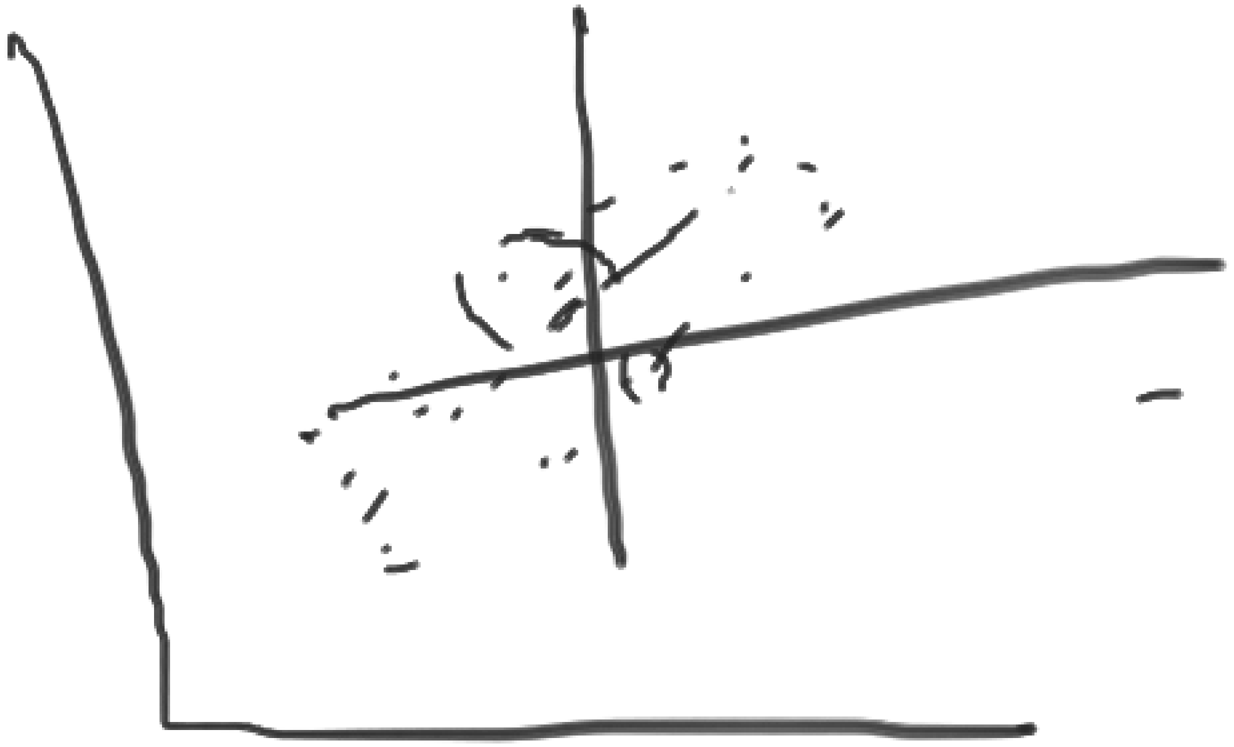
$+5$
$+10$
-68
69
$-$

117

132

-0.9

Beacuse of Random sampling while computing correlation there could be statistical fluke.



statistical fluke

~~0.5 - 0.7~~

0.3

or 0.9