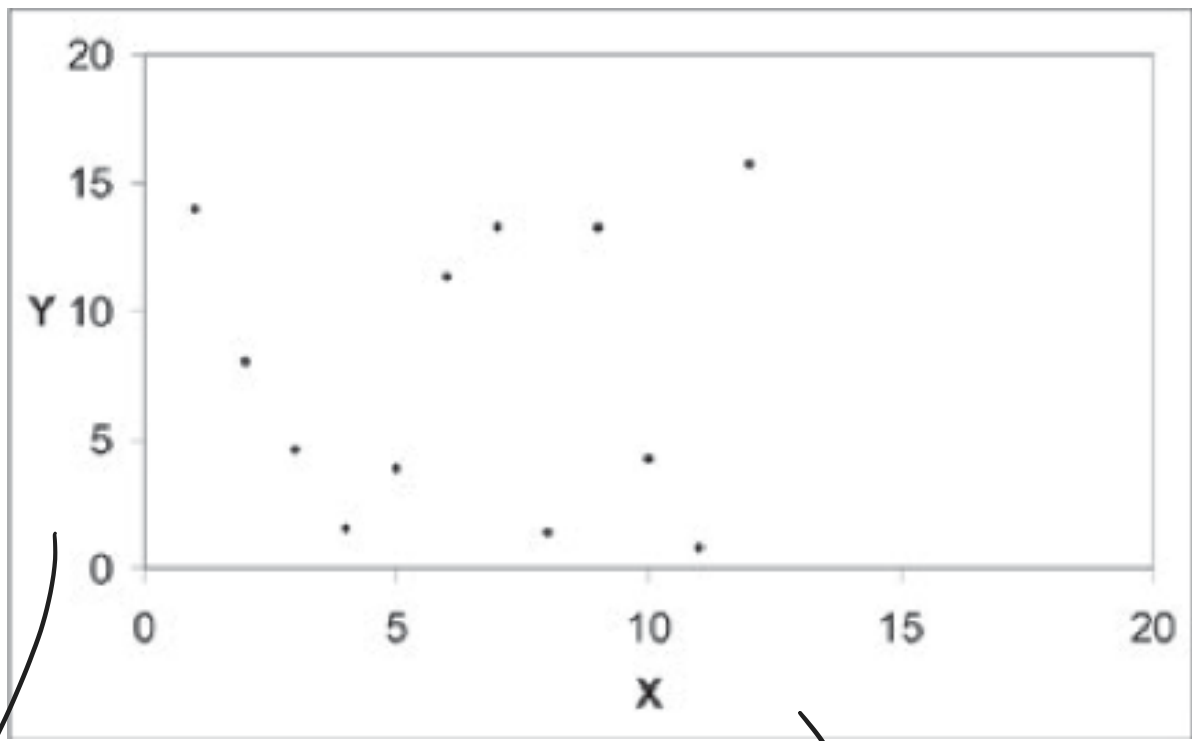
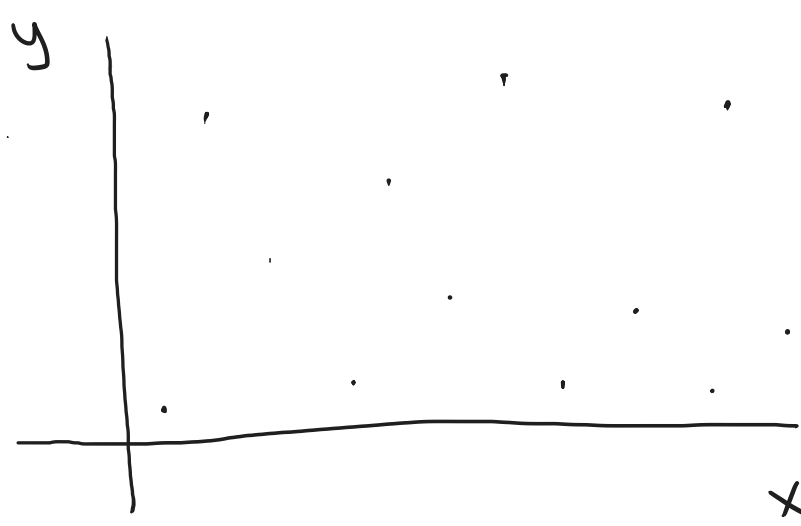
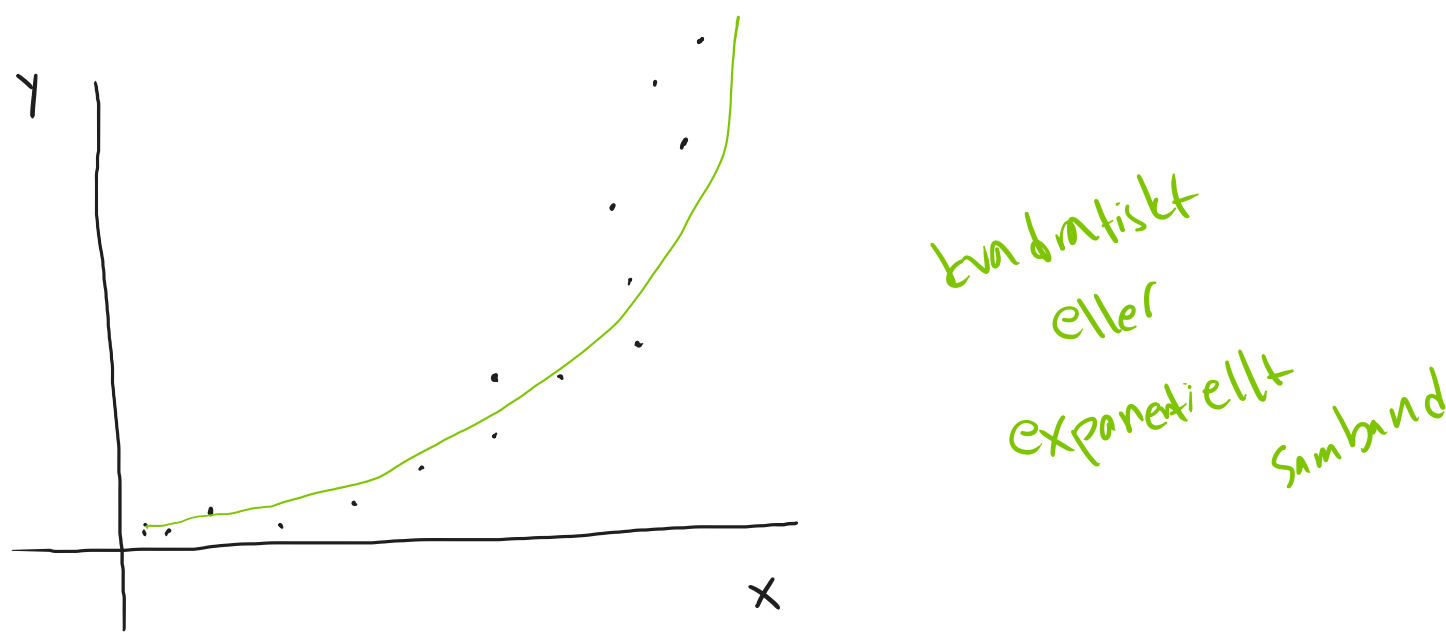
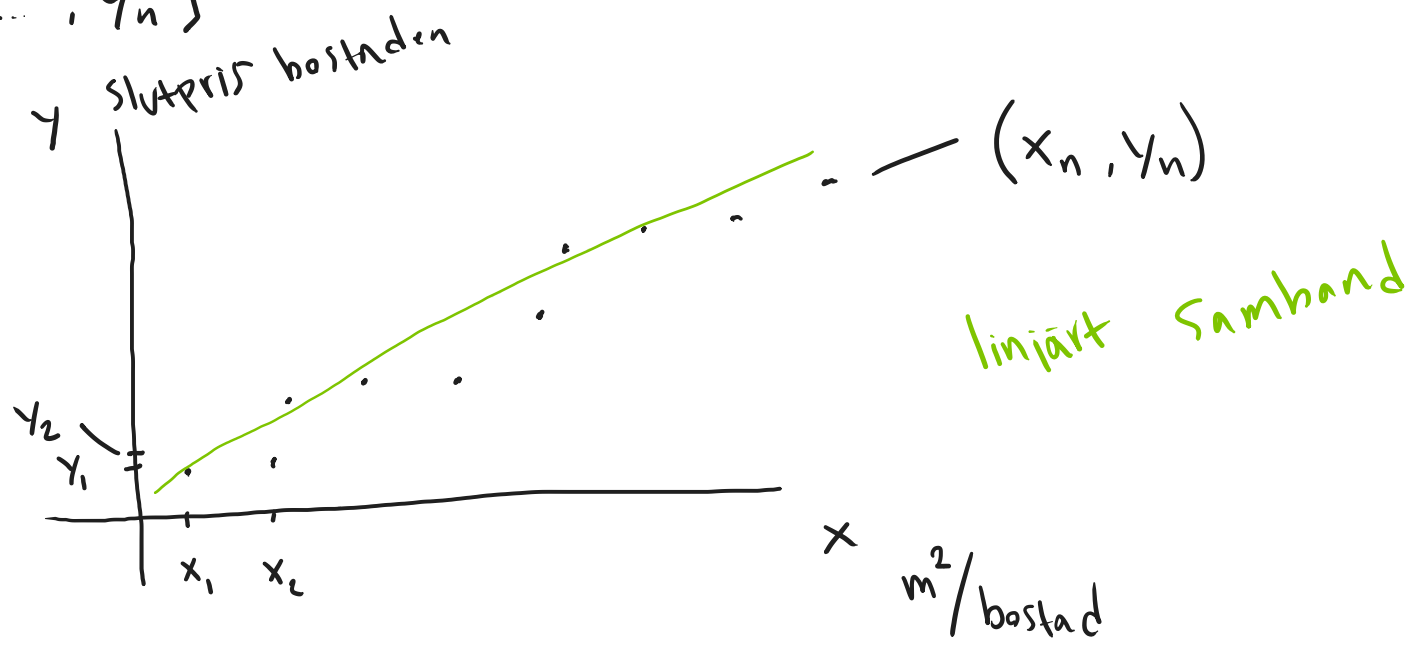


Korrelation

$$x = [x_1, x_2, x_3, \dots, x_n]$$

$$y = [y_1, y_2, y_3, \dots, y_n]$$

plt.plot(x,y)



Finns inget samband

antall byggnadsstreck i
Malmö köper
längden på vår väst

Korrelationskoefficienten

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}}$$

där

$$SS_x = \sum_{i=1}^n (x_i - \bar{x})^2$$

$$SS_y = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$SS_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

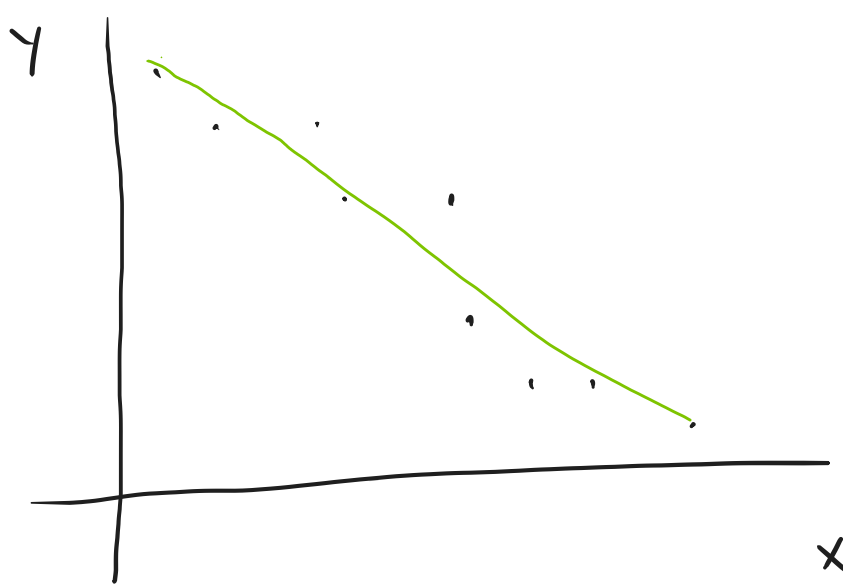
Summan av kvadrat differenser

$$\rightarrow -1 \leq r \leq 1$$

$$r = 1 \rightarrow \text{Starkt positivt samband / korrelation}$$

$$r = 0 \rightarrow \text{Inget samband / korrelation}$$

$$r = -1 \rightarrow \text{Starkt negativt samband / korrelation}$$



$$r > 0 \rightarrow \text{positivt samband / korrelation}$$

$$r < 0 \rightarrow \text{negativt samband / korrelation}$$

Uppgift

Bygg en function i Python som givet
två listor x,y beräknar ^{och returnerar} korrelationskoefficienten r

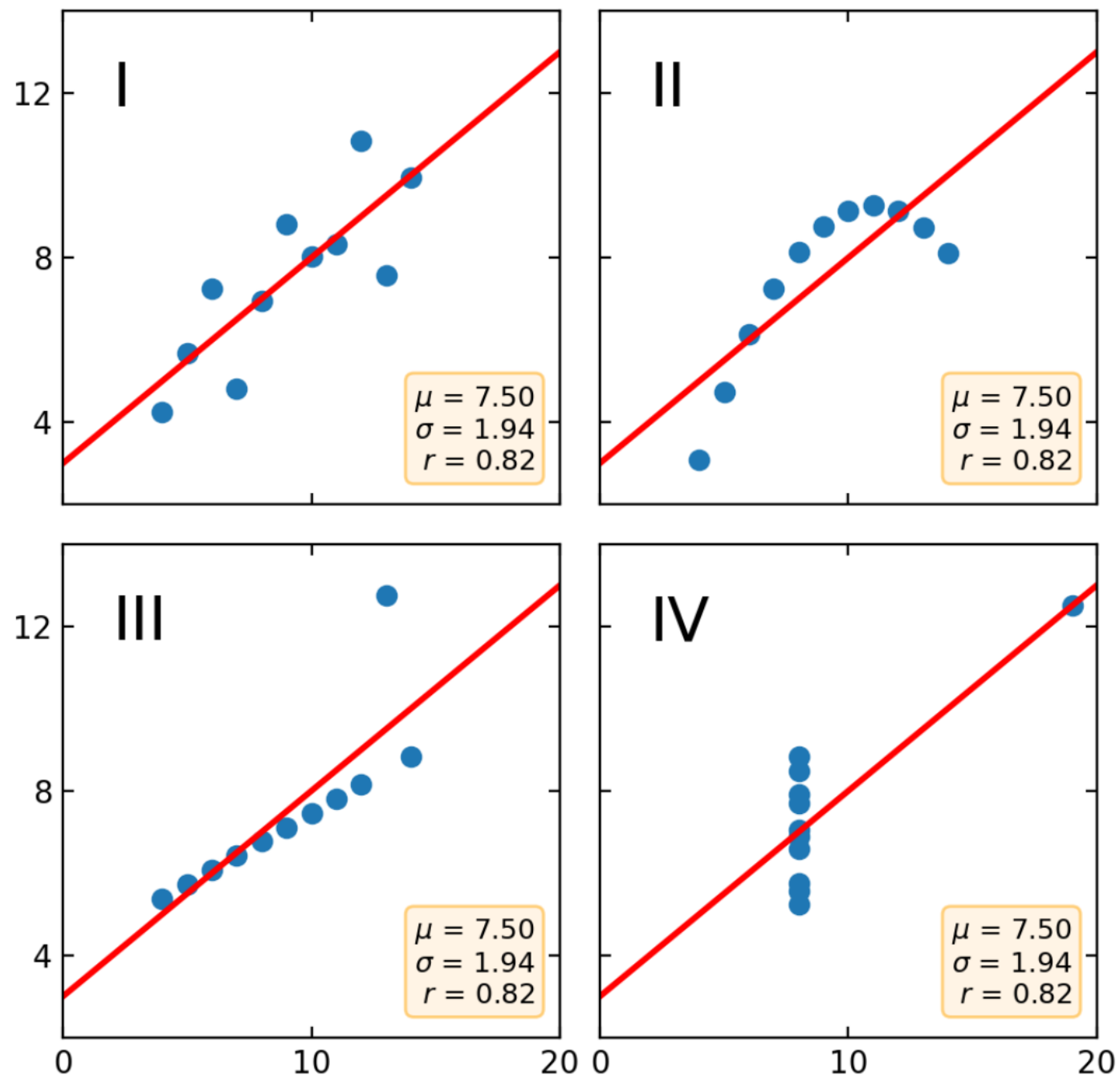
$$SS_x = \sum_{i=1}^n (x_i - \bar{x})^2$$

$$SS_y = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$SS_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}}$$

plotta även x och y



ORSS

Var försiktig
med er tolkning
av korrelationskoefficienten!