

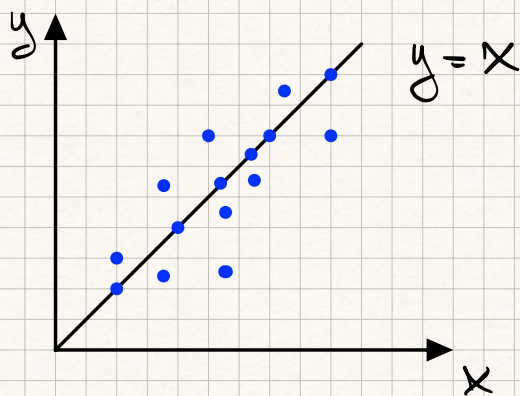
# REGRESSIONE

Metodo parametrico per ridurre il volume dei dati. L'idea è quella di trovare un modello matematico che fitte i dati, e salvare solo quello (Al più salvare solo eventuali outlier).

Per modellare un modello matematico appropriato definiamo:

- 1) VARIABILI DIPENDENTI :  $y$
- 2) VARIABILI INDIPENDENTI :  $x_1 \dots x_n$

Bisogna poi trovare i parametri che meglio adattano il modello ai dati



## REGRESSIONS LINEARS

I dati seguono un andamento lineare che si adatta a una retta

$$y = w_1 x + w_0$$

Abbiamo due coefficienti di regressione,  $w_0$  e  $w_1$ .

Possiamo stimare i coefficienti tramite il metodo dei minimi quadrati, conoscendo i punti  $y_i$  e  $x_i \in D$

$$w_1 = \frac{\sum_{i=1}^{|D|} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{|D|} (x_i - \bar{x})^2}$$

$$w_0 = \bar{y} - w_1 \bar{x}$$

## REGRESSIONS LINEARS MULTIPLA

$$y = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_n x_n$$