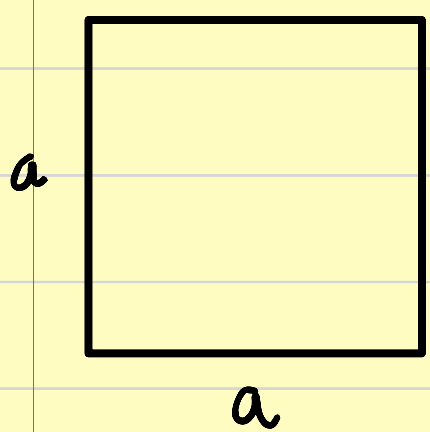
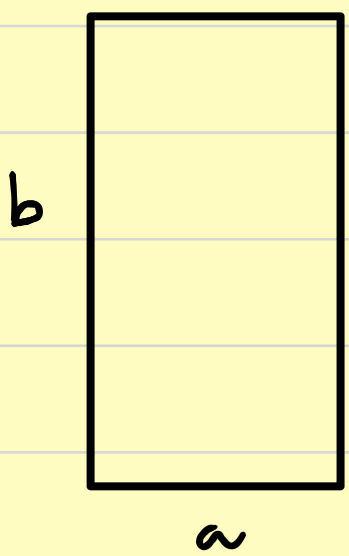
20240320_Data_Analytics_BU6

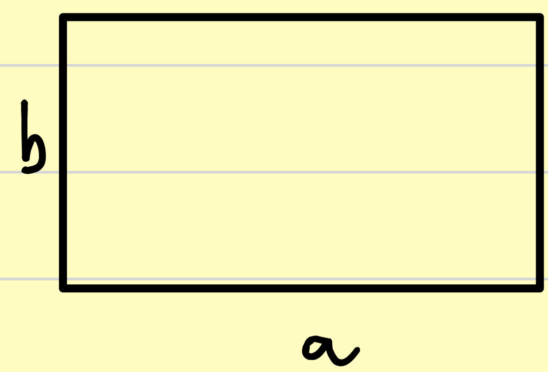
## Lineare Prognose . LINEARE REGRESSION



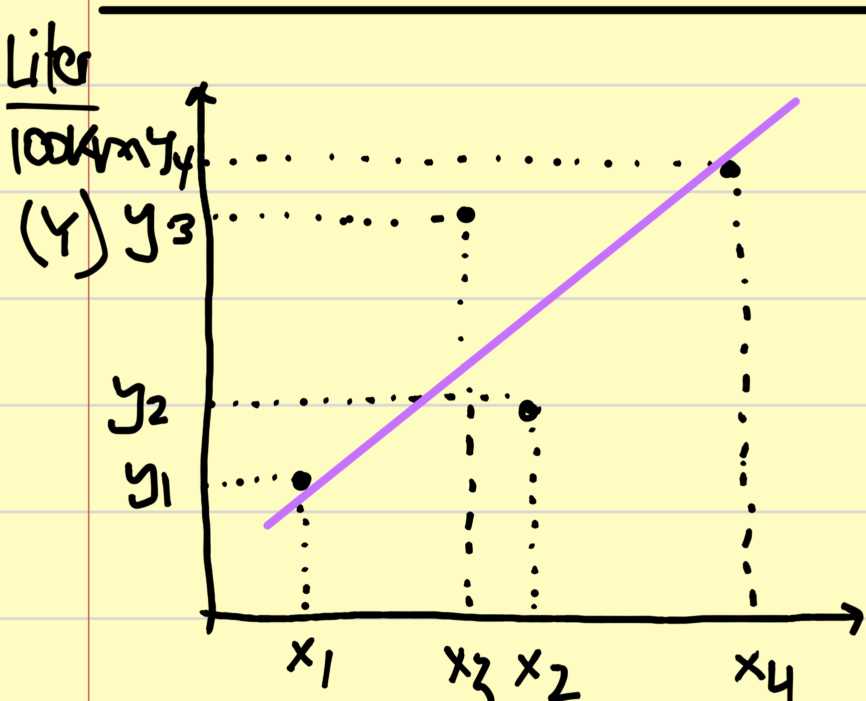

Fläche $= a \cdot a = a^2$




Fläche $= a \cdot b$




Fläche $= a \cdot b$




Gewicht FZG (x)

Wir streben die Ermittlung einer Linie $y = b_0 + b_1 x$, dessen Abstand zu den Punkten des Datasets minimal ist.

Somit sind wir in der Lage, gegeben x (Gewicht), den Verbrauch (y) zu prognostizieren (und umgekehrt)

# SCHRITTE DER LINEAREN PROGNOSE (REGRESSION)

## 1. Datenermittlung.

| | x | Y |
|---|---|---|
| 1 | 3 | 6'5 |
| 2 | 4 | 8'5 |
| 3 | 6 | 13 |
| 4 | 3 | 3'5 |

## 2. Mittelwert der Variabeln (Schwerpunkt des Datasets)

Die lineare regressionslinie geht den Mittelwert: $\bar{y} = b_0 + b_1 \bar{x}$

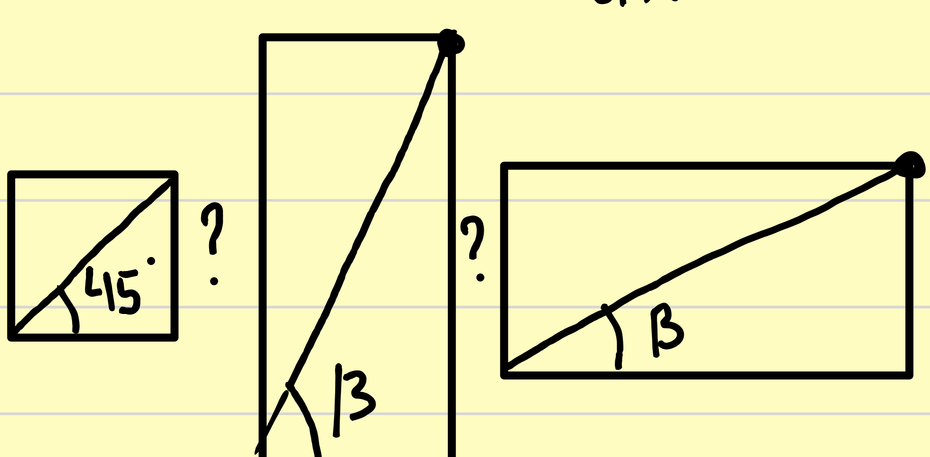
$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{4}(3+4+6+3) = 4$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i = \frac{1}{4}(6'5+8'5+13+3'5) = $$
$$= 7'875$$

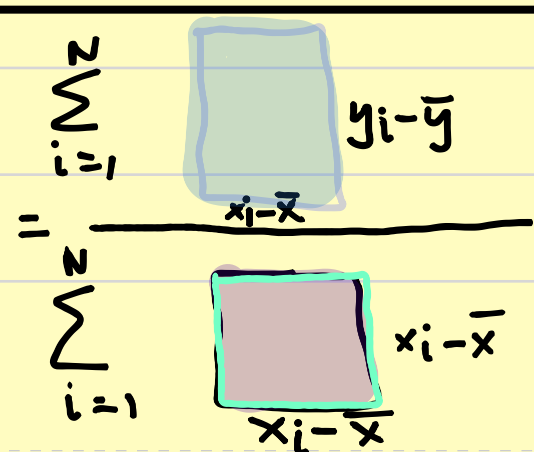
## 3. NEIGUNG DER LINEAREN REGRESSIONSLINIE

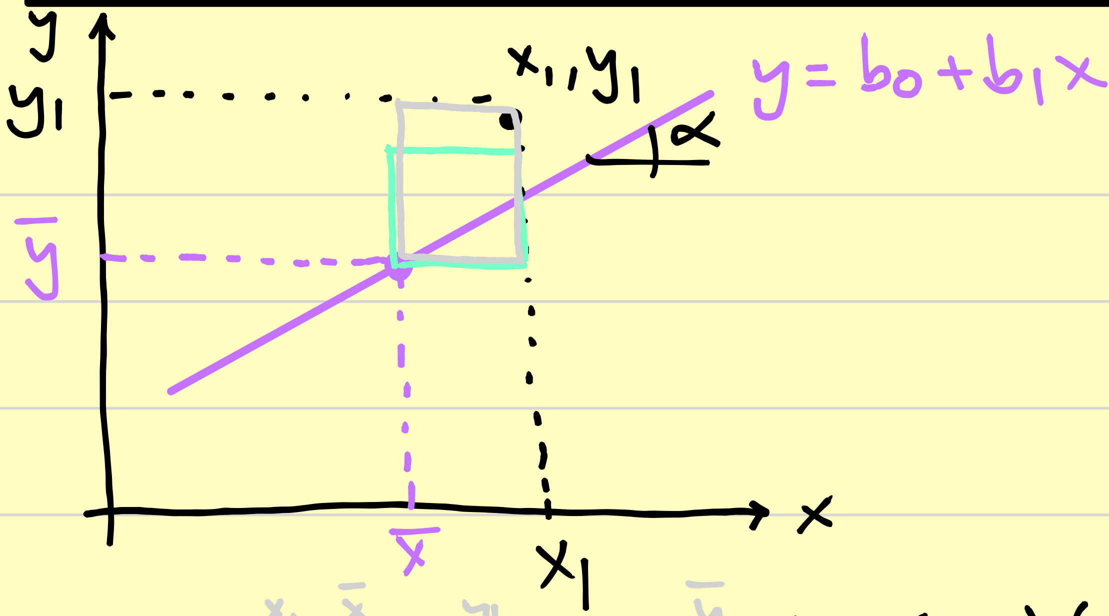

$$y = b_0 + b_1 x$$

$$b_1 = tg\alpha = \frac{dy}{dx}$$




$$b_1 = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{N}(x_i-\bar{x})^2} =$$






$$b_1 = \frac{(3-4)(6'5-7'875)+(4-4)(8'5-7'875)+(6-4)(13-7'875)+(3-4)(3'5-7'875)}{(3-4)^2+(4-4)^2+(6-4)^2+(3-4)^2}$$

wir gehen davon aus $b_1$ ist bekannt

$$y = b_0 + b_1 x$$

4. ERMITTLUNG VON $b_0$:

$$\bar{y} = b_0 + b_1 \bar{x} \longrightarrow \boxed{b_0 = \bar{y} - b_1 \bar{x}}$$

$$b_0 = 7'875 - b_1 \cdot 4$$

✓