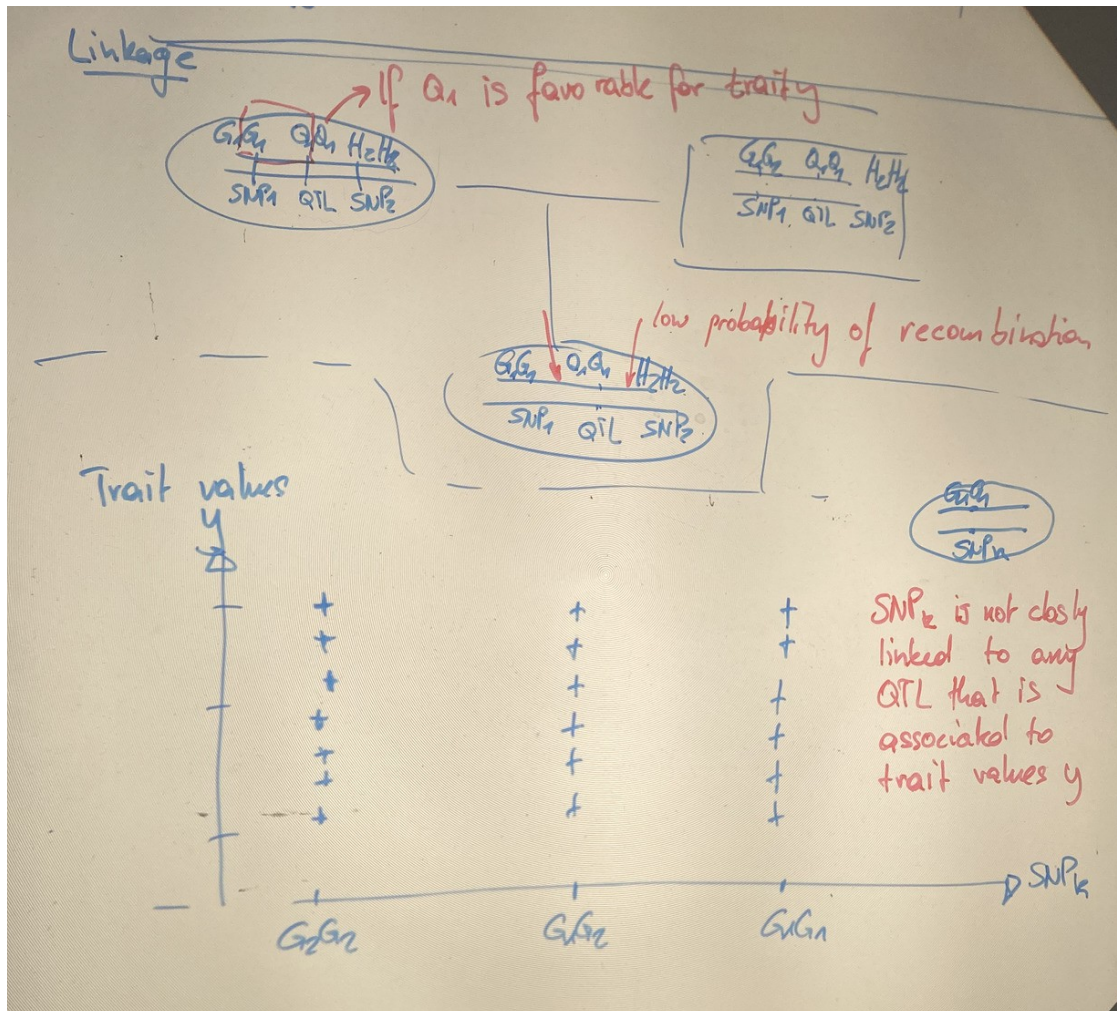
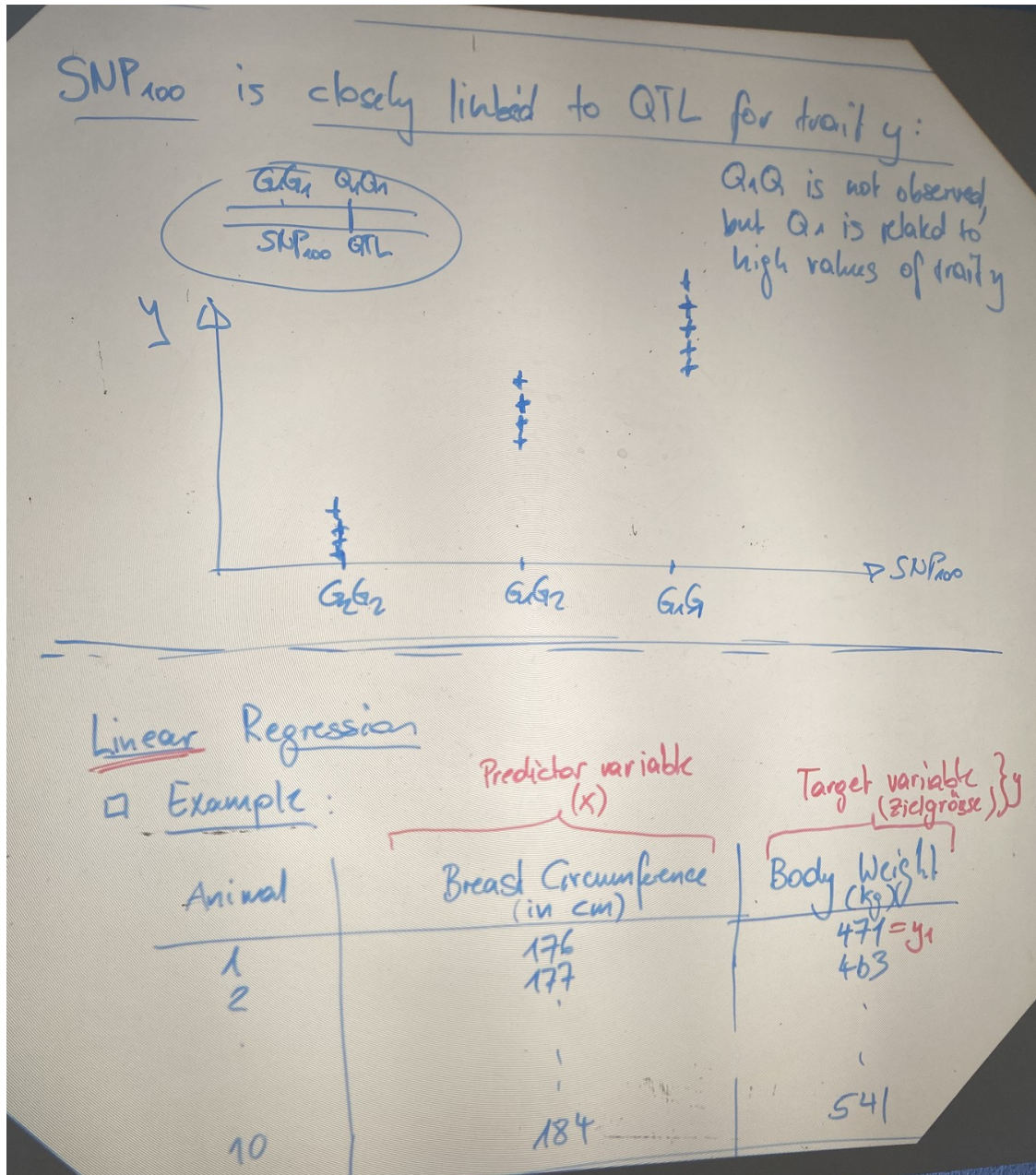


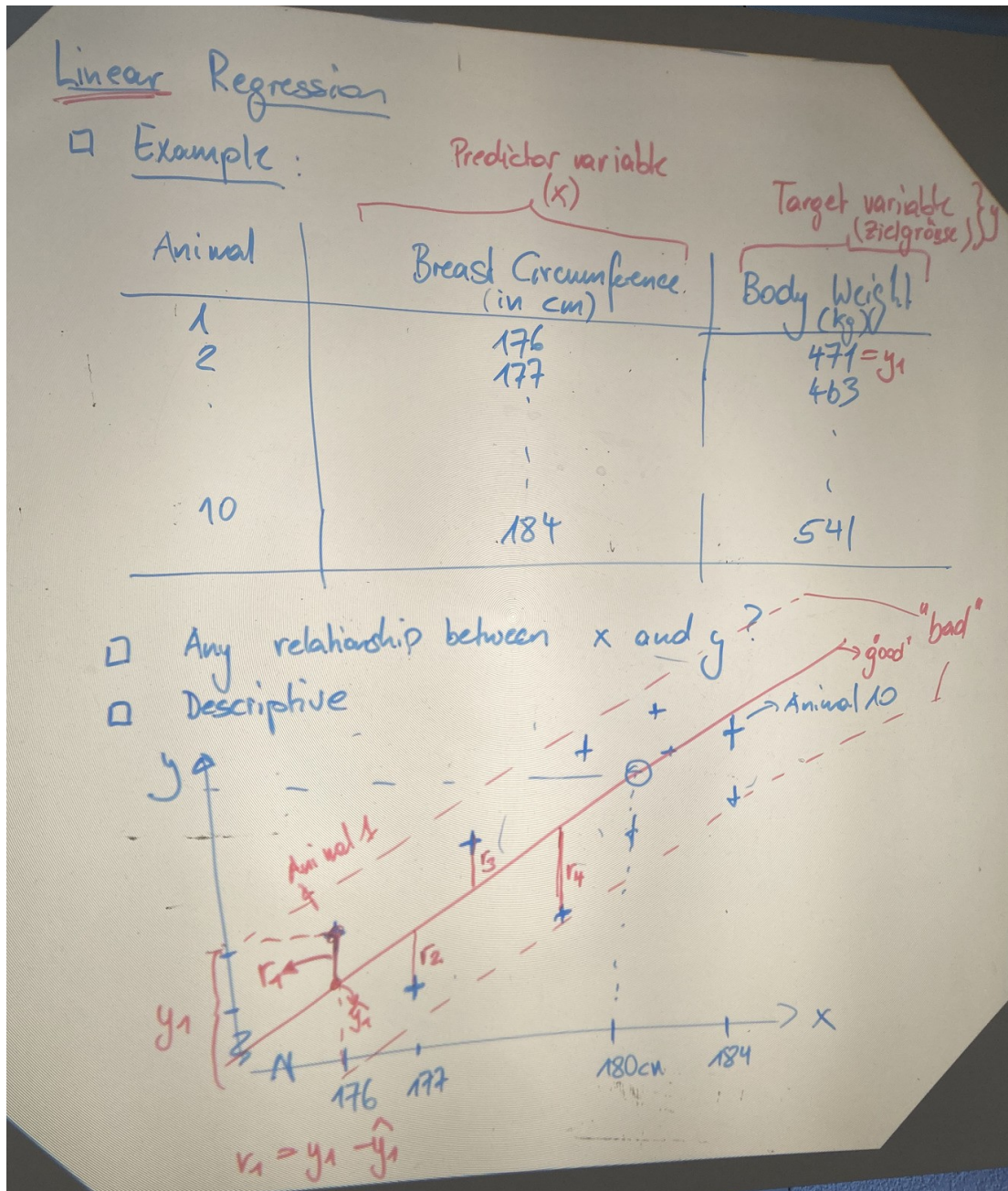
OHP Picture 1



OHP Picture 2



OHP Picture 3

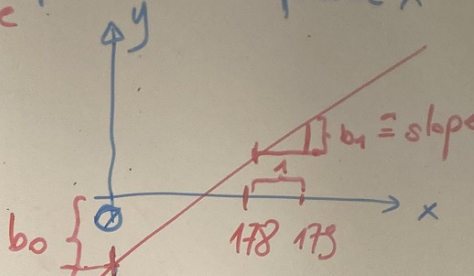


How to find the red-line (Regression line)

□ Red-line gives the expected body weight ($E(y)$) based on a given value of breast circumference x

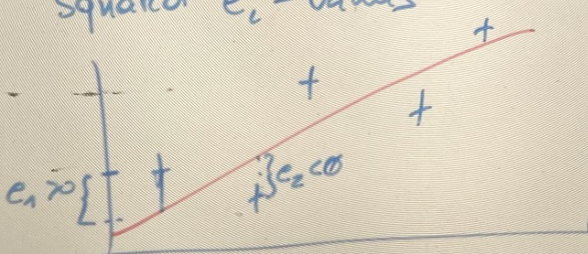
$\hat{y} = E(y) = \underbrace{b_0}_{\text{intercept}} + \underbrace{b_1 \cdot x}_{\text{slope}}$

Every observation
 $y_i = E(y) + e_i$
 $= \underbrace{b_0}_{\text{unknown}} + \underbrace{b_1}_{\text{unknown}} \cdot \underbrace{x_i}_{\text{observed, known}} + e_i$



□ Regression Line (defined by b_0 and b_1) is determined by minimizing the sq sum of the squared e_i -values

Use: $L = e_1^2 + e_2^2 + \dots + e_n^2$
 $= \sum_{i=1}^n e_i^2$



⇒ Goal: Determine b_0 and b_1 such that L is minimal

Use $y_i = b_0 + b_1 x_i + e_i$

OHP Picture 5

Minimize L : Use $y_i = b_0 + b_1 x_i + e_i$

$$L = \sum_{i=1}^N e_i^2 = \sum_{i=1}^N [y_i - b_0 - b_1 x_i]^2$$

$$\frac{\partial L}{\partial b_0} = \sum_{i=1}^N 2 \cdot [y_i - b_0 - b_1 x_i] \cdot (1)$$

$$= -2 \sum_{i=1}^N [y_i - b_0 - b_1 x_i] = -2 \left[\sum_{i=1}^N (y_i) - N b_0 - b_1 \sum_{i=1}^N x_i \right]$$

$$\frac{\partial L}{\partial b_1} = \sum_{i=1}^N -2 [y_i - b_0 - b_1 x_i] x_i$$

$$= -2 \left[\sum_{i=1}^N (x_i y_i) - \sum_{i=1}^N (b_0 x_i) - \sum_{i=1}^N (b_1 x_i^2) \right]$$

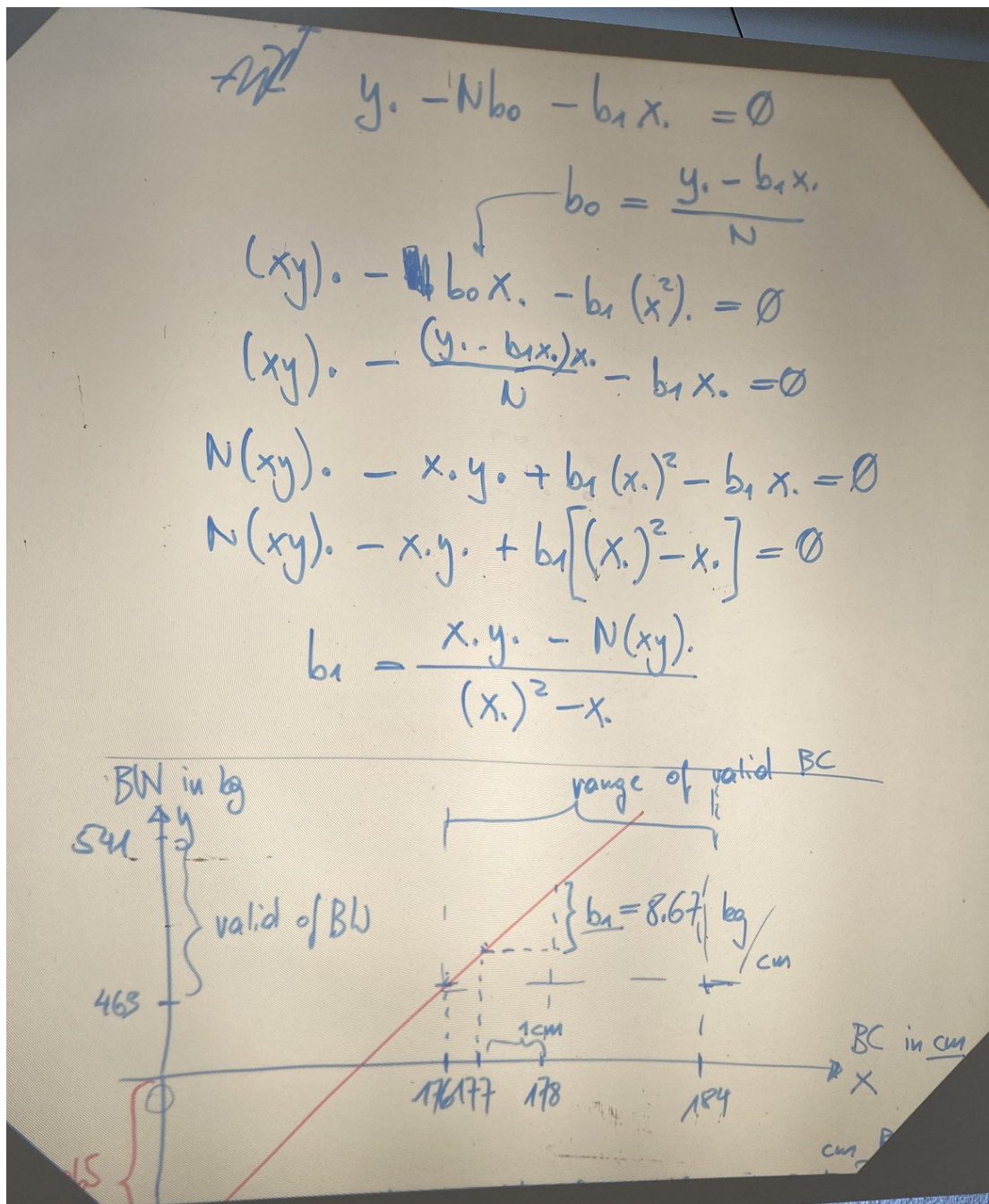
$$= -2 \left[\sum_{i=1}^N (x_i y_i) - b_0 \sum_{i=1}^N (x_i) - b_1 \sum_{i=1}^N (x_i^2) \right]$$

Determine b_0 and b_1 where $\frac{\partial L}{\partial b_0} = 0$ and $\frac{\partial L}{\partial b_1} = 0$

$$\frac{\partial L}{\partial b_0} = -2 \left[\sum y_i - N b_0 - b_1 \sum x_i \right] = 0$$

$$\frac{\partial L}{\partial b_1} = -2 \left[\sum (x_i y_i) - b_0 \sum x_i - b_1 \sum (x_i^2) \right] = 0$$

OHP Picture 6



OHP Picture 7

