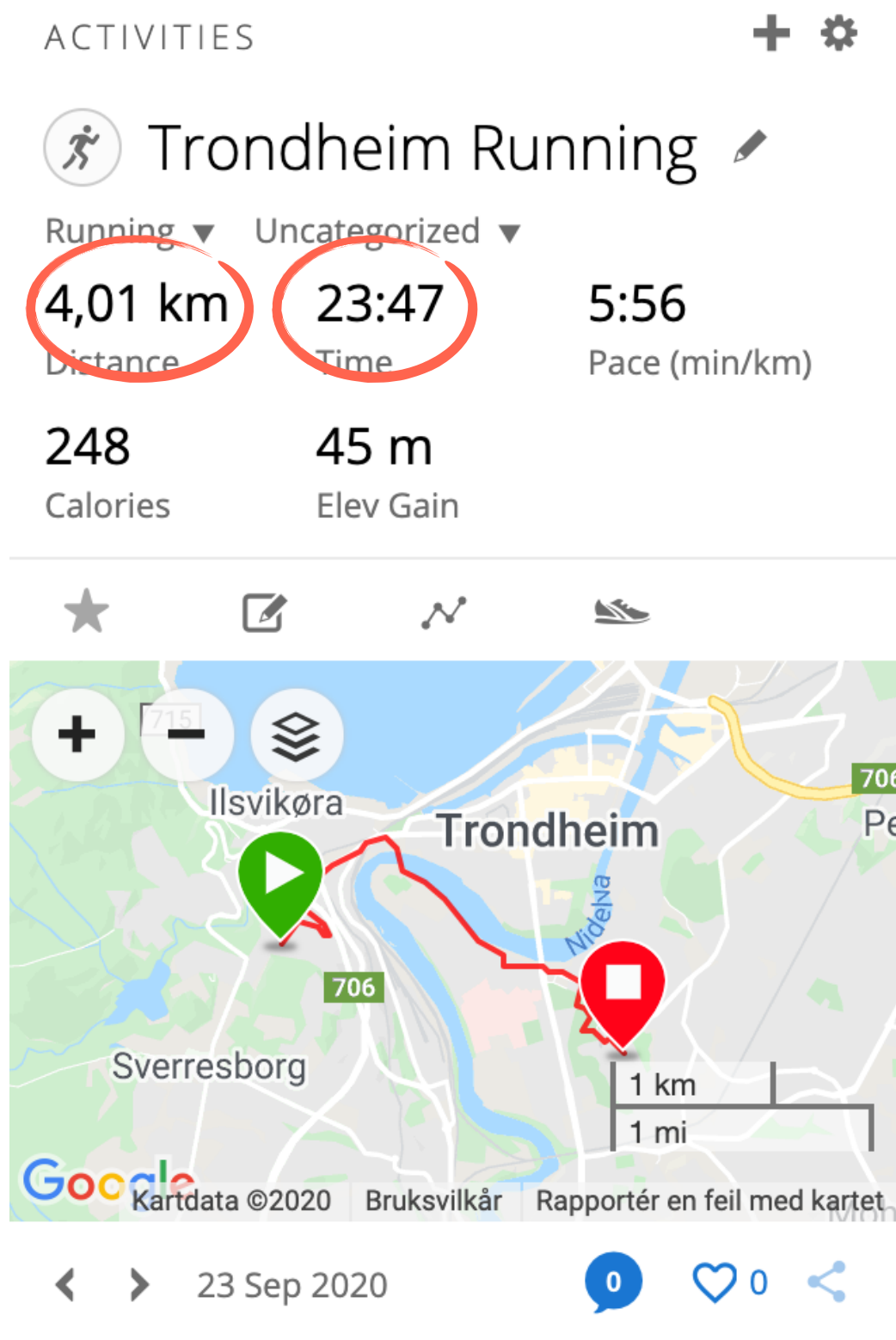
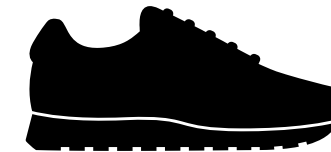


Lineær regresjon

Del 1

Thea Bjørnland
Institutt for matematiske fag
NTNU

Eksempel: Treningsdata



Hvis jeg skal løpe 5 km i morgen,
hvor langt tid bør jeg regne med å bruke?

X : distanse løpt (km)

$$E(X) = \mu_x \quad SD(X) = \sigma_x$$

Y : tid brukt (minutter)

$$E(Y) = \mu_y \quad SD(Y) = \sigma_y$$

Er det en lineær sammenheng mellom X og Y ?

$$Cov(X, Y) = E \left((X - \mu_x)(Y - \mu_y) \right)$$

$$\rho = \frac{Cov(X, Y)}{\sigma_x \sigma_y} \quad -1 \leq \rho \leq 1$$

Eksempel: Treningsdata

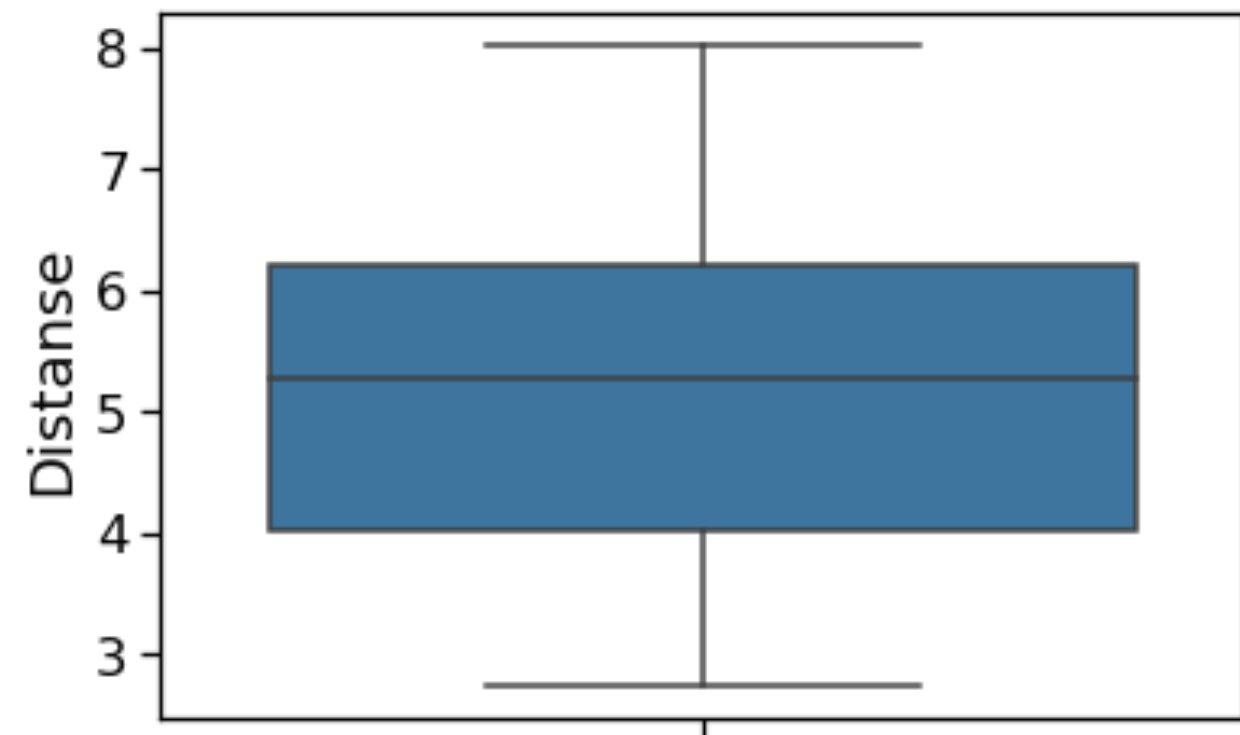
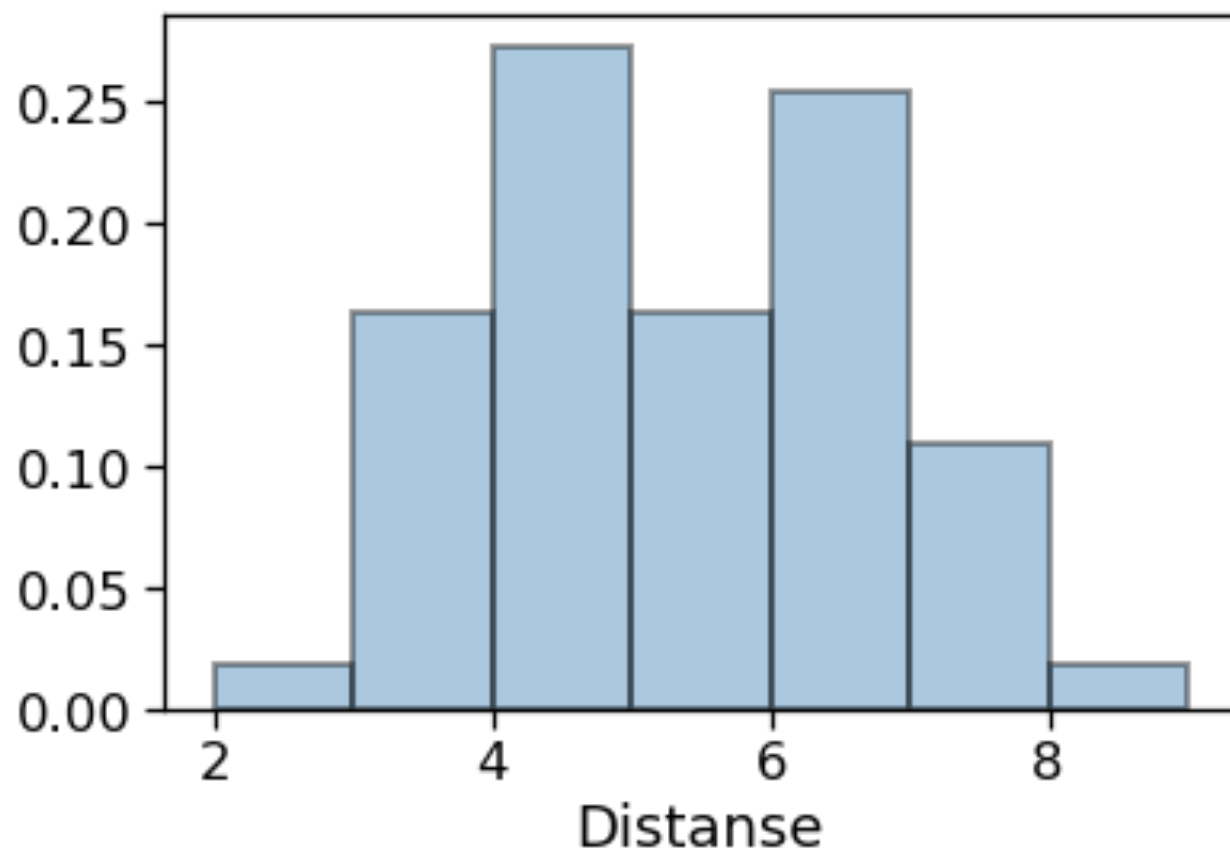


Data: 55 løpeturer lastet ned fra connect.garmin.com

Juni - Oktober 2020

Parvise observasjoner (x_1, y_1) , (x_2, y_2) , ..., (x_{55}, y_{55}) **Anta: uavhengige!**

Distance (km) $\curvearrowright$ $\curvearrowleft$ Tid (min)



$\bar{x} \approx 5.2$ km $s_x \approx 1.4$ km

Eksempel: Treningsdata

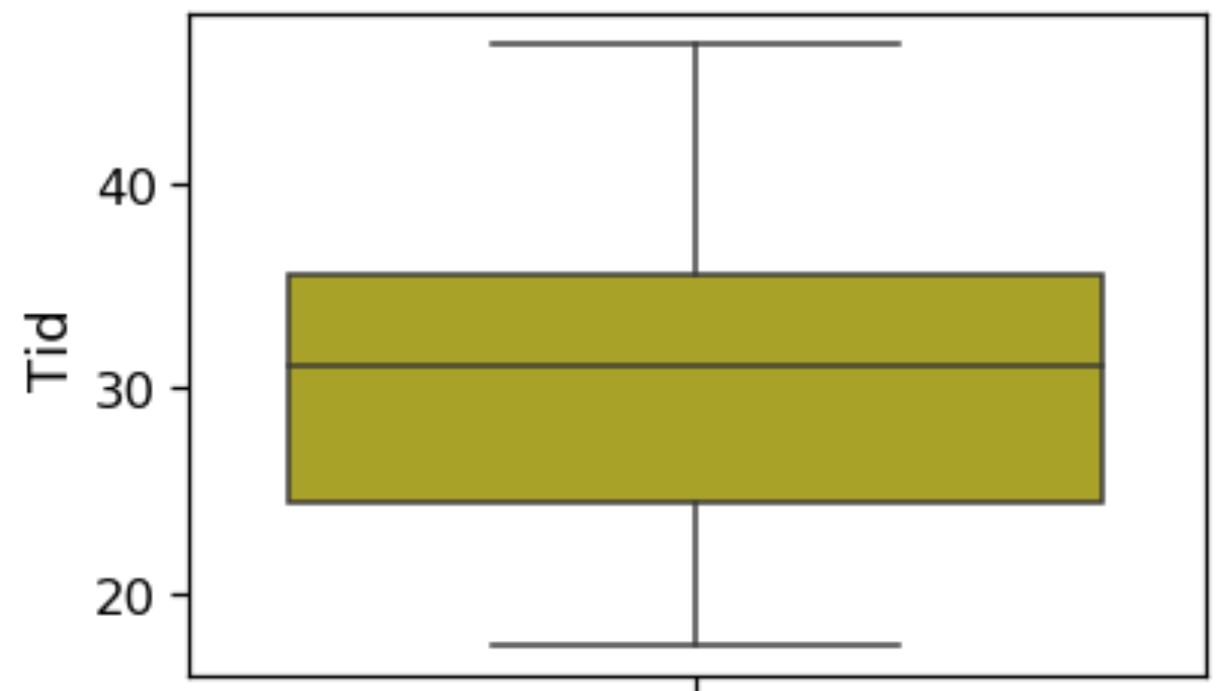
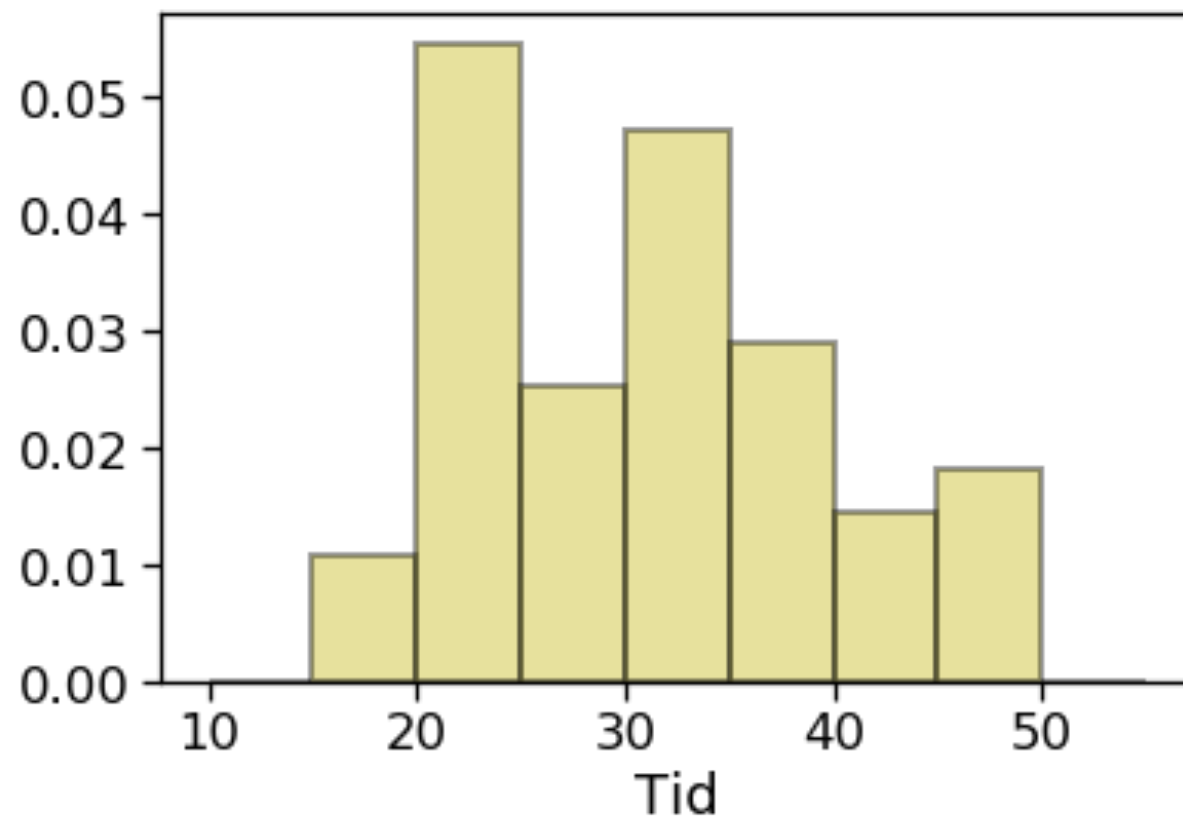


Data: 55 løpeturer lastet ned fra connect.garmin.com

Juni - Oktober 2020

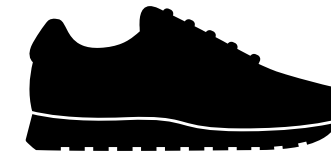
Parvise observasjoner $(x_1, y_1), (x_2, y_2), \dots, (x_{55}, y_{55})$ **Anta: uavhengige!**

Distanse (km) $\curvearrowright$ $\curvearrowleft$ Tid (min)



$$\bar{y} \approx 31.1 \text{ min} \quad s_y \approx 7.9 \text{ min}$$

Eksempel: Treningsdata

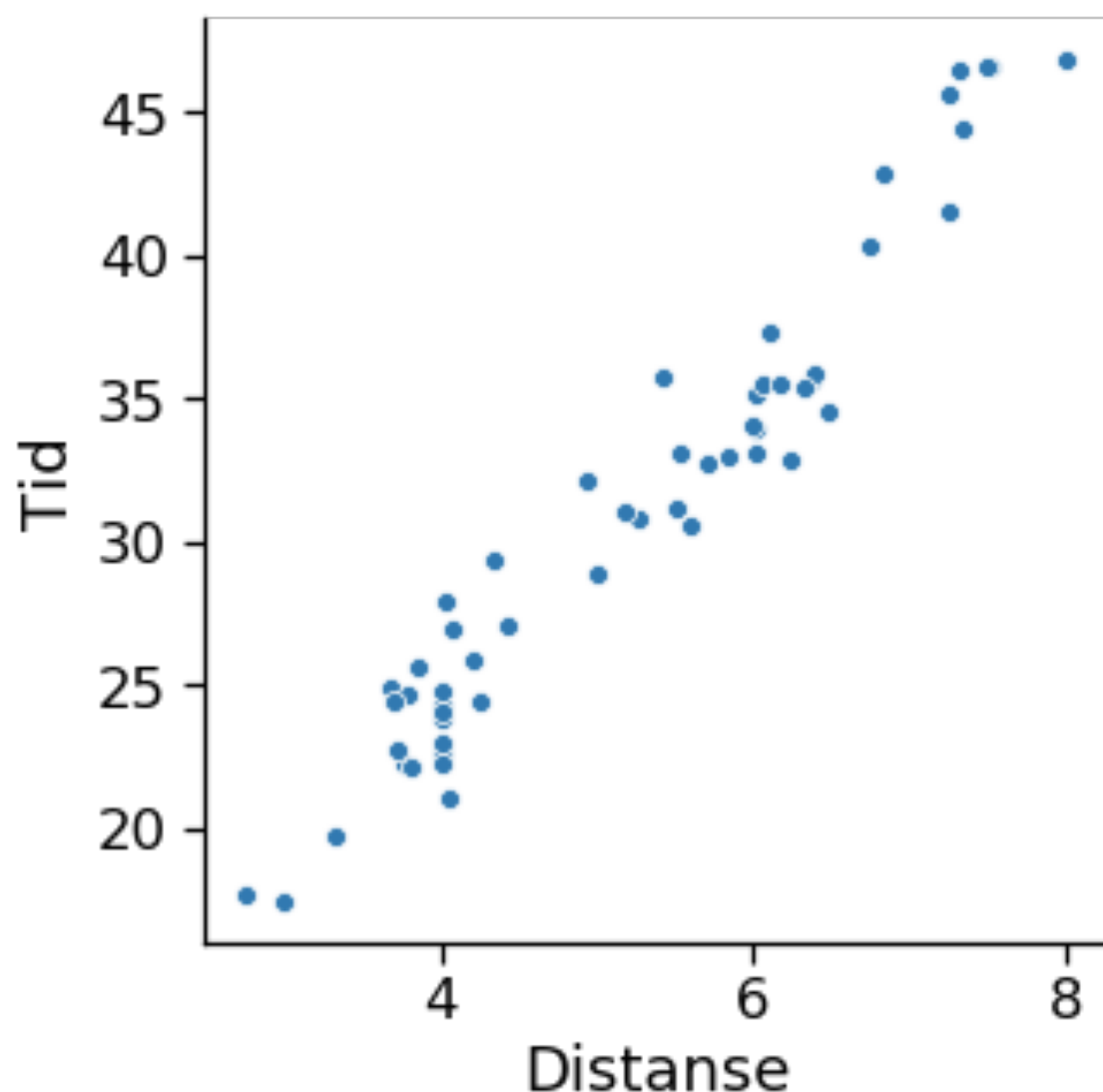


Data: 55 løpeturer lastet ned fra connect.garmin.com

Juni - Oktober 2020

Parvise observasjoner $(x_1, y_1), (x_2, y_2), \dots, (x_{55}, y_{55})$ **Anta: uavhengige!**

Distanse (km) $\curvearrowright$ $\curvearrowleft$ Tid (min)



$$\text{Cov}(X, Y) = E \left((X - \mu_x)(Y - \mu_y) \right)$$

$$\hat{\text{Cov}}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

$$\hat{\text{Cov}}(X, Y) \approx 10.5 \quad \text{Cov}(X, Y) > 0$$

Korrelasjon:

$$R = \frac{\hat{\text{Cov}}(X, Y)}{s_x \cdot s_y} \quad R \approx 0.97$$

Eksempel: Treningsdata



Data: 55 løpeturer lastet ned fra connect.garmin.com

Juni - Oktober 2020

Parvise observasjoner (x_1, y_1) , (x_2, y_2) , ..., (x_{55}, y_{55}) **Anta: uavhengige!**

Distanse (km) $\curvearrowright$ $\curvearrowleft$ Tid (min)

