Scatter plots

- bi-variate (2 characteristics / pieces of data)
- looking a relationship. (correlation)
- positive/neg correlation (slope)
- Generate Least Squares Regression Line. (specific type of trend line, line of best fit)

manatees killed

residual for an obs.

residual = $y_{obs} - y_{pred}$

$= y_{obs} - \hat{y}$

power boat regs

$M = 0.125P - 41.4$

$r = 0.94$

correlation coefficient

Contrast r and r^2

- r - correlation coefficient (tells how strong/weak the data is correlated)
  - positive correlation → as $x \uparrow$ $y \uparrow$
  - r does not have unit
  - $-1 \leq r \leq 1$

- $\begin{cases} 0 < |r| < 0.4 & \text{weak} \\ 0.4 \leq |r| < 0.7 & \text{moderate} \\ 0.7 \leq |r| \leq 1 & \text{strong.} \end{cases}$

$r^2 \rightarrow$ coefficient of determination

- measures the linearity of your data.

- $0 \leq r^2 \leq 1$

r^2 gives the % of your data that can be explained by assuming a linear model.

---

r^2 is not enough, you need to check the residual plot for patterns.