

Real data



=

Model 1



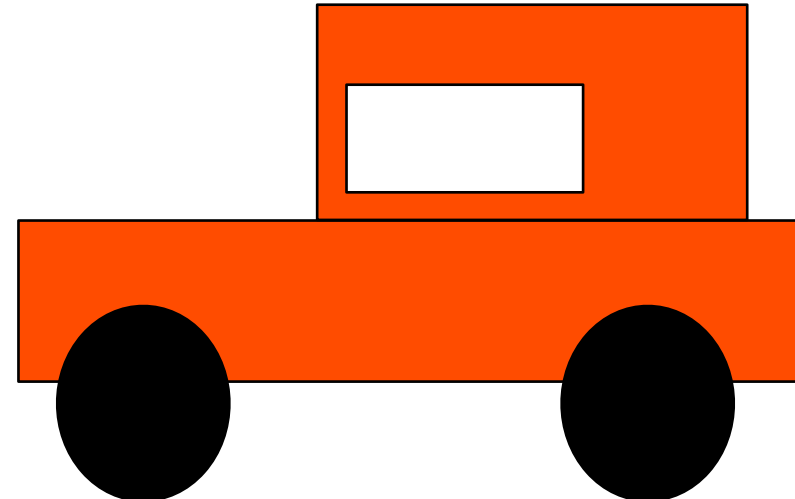
+ Error

Real data



=

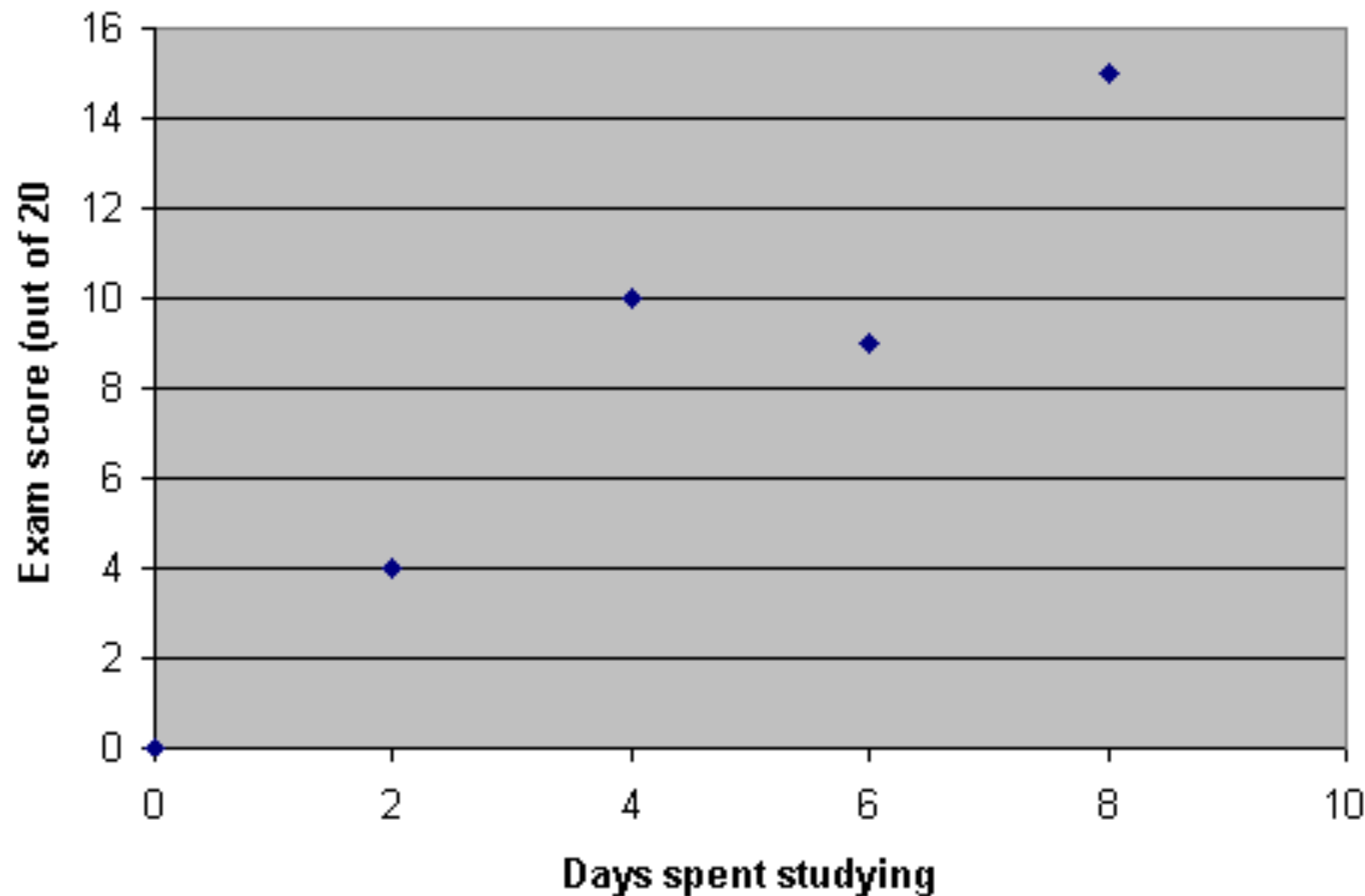
Model 2



+ Error

- We want to select the model which has the smallest error (aka model residuals)...

# Regression



We can plot data on exam performance and days spent studying.

Wouldn't it be helpful if we could draw a straight line such that if we know the value on one axis (x say), we could predict the value on the other (y say) ?