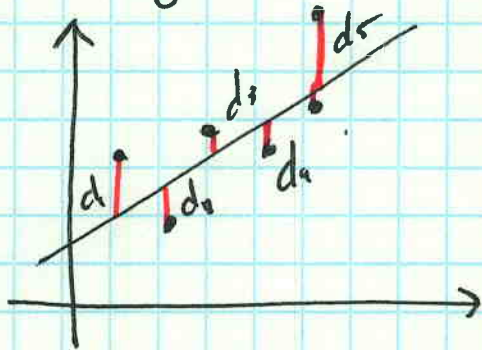


11.7 Day 2 Linear regression / Least Squares

Regression Line \rightarrow line of best fit.



Residual - distance between the data and the ~~the residual~~ Regression line.

Least Square Regression Line

Line has the smallest value when residuals are squared & added.

Least Squares Equation

$$\hat{y} = ax + b \quad a = r \frac{S_y}{S_x} \quad b = \bar{y} - a\bar{x}$$

r - correlation coefficient
 \bar{x}, \bar{y} , means S_x, S_y - Sample S.D.

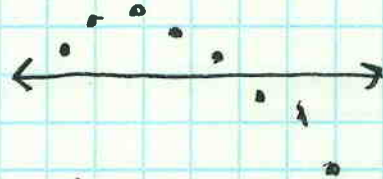
\hat{y} = predicted value

Residual Plot \rightarrow shows distance

of each data point above or below the regression line.



Random
good linear
model



\equiv magnifies
outliers

not a
good linear
model