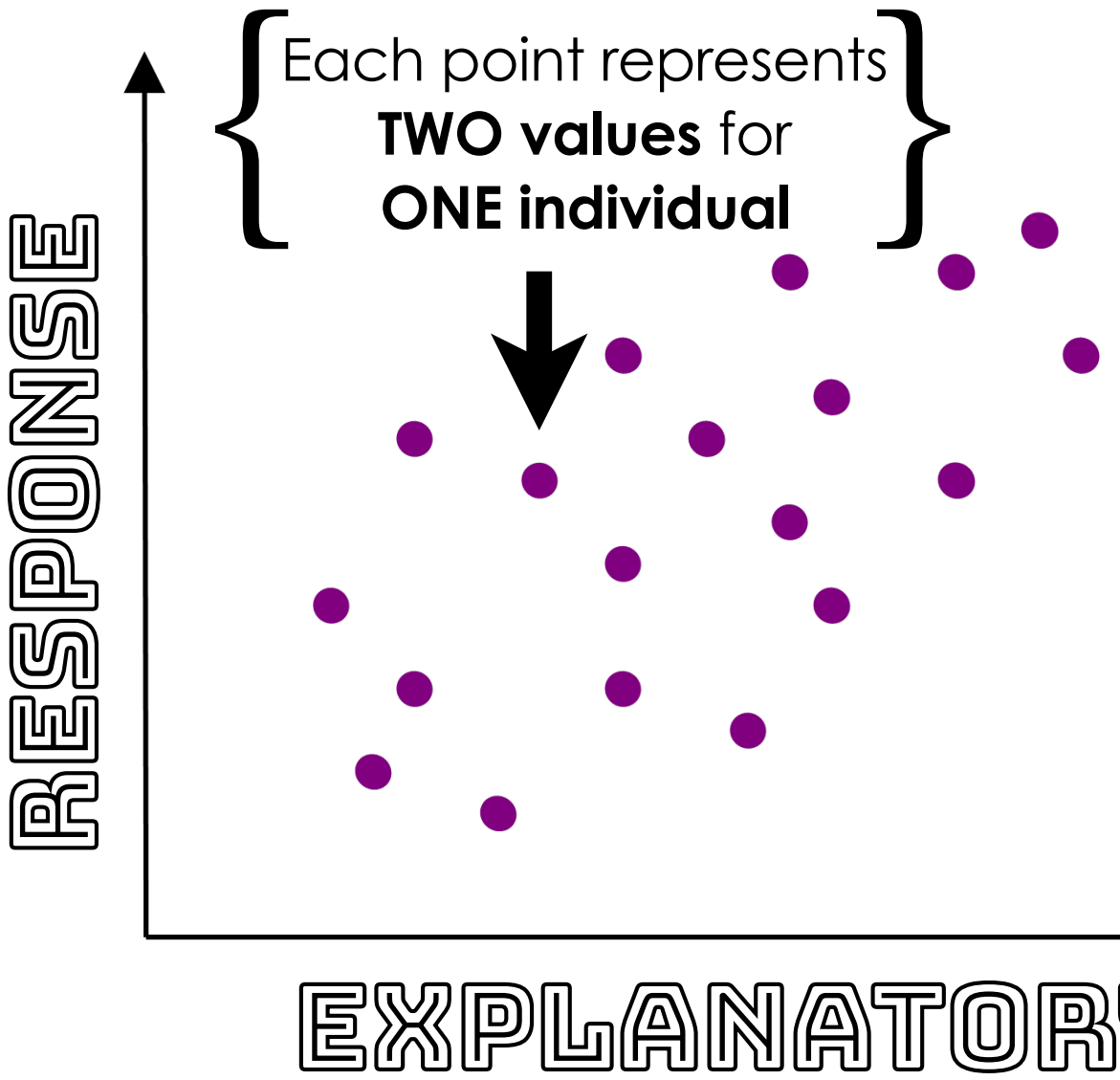
SCATTERPLOTS



Displays relationship between **TWO quantitative variables**

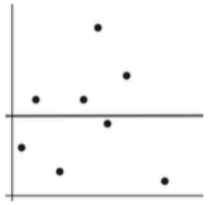
Describing Scatterplots

Direction

positive

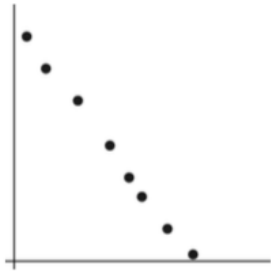
negative

none

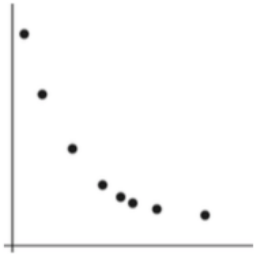


Form

linear

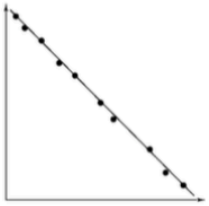


non linear

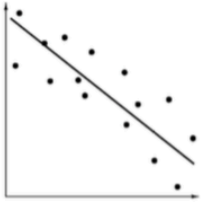


Strength

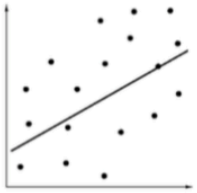
strong



moderate

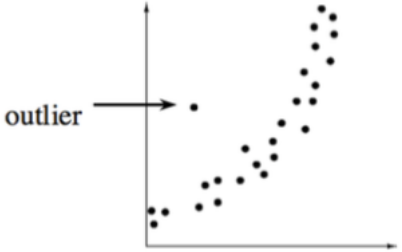


weak

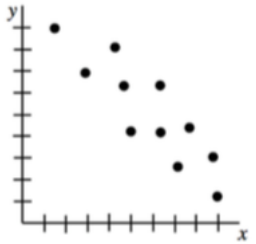


Outliers

outlier

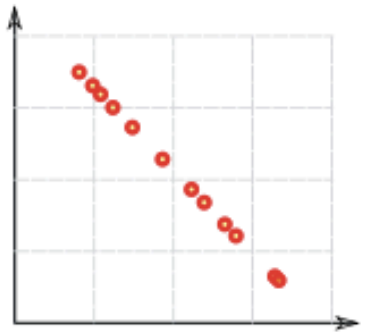


no outlier



Correlation

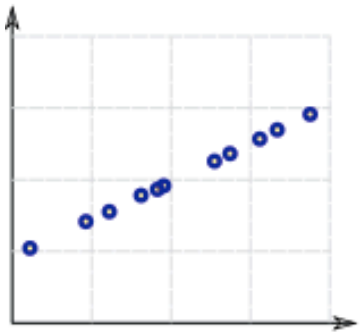
Perfect (–)



-1

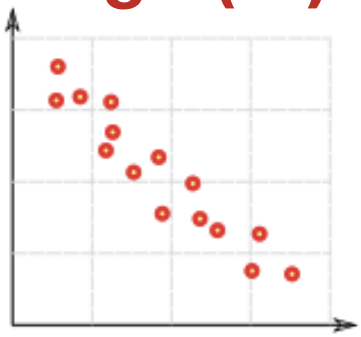
Describes **STRENGTH** & **DIRECTION**

Perfect (+)



1

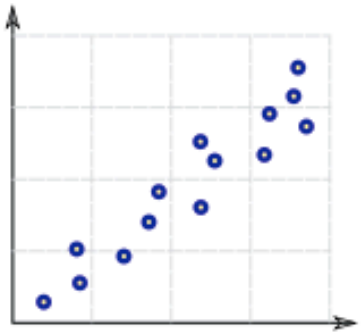
High (–)



-0.9

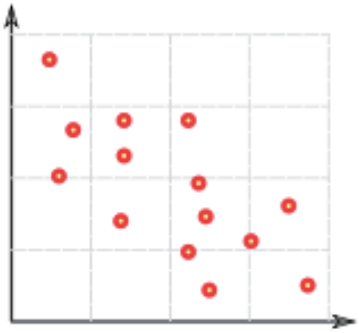
NOT resistant = greatly affected by outliers

High (+)



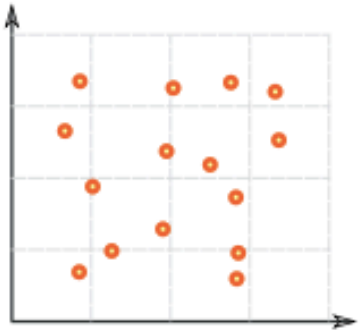
0.9

Low (–)



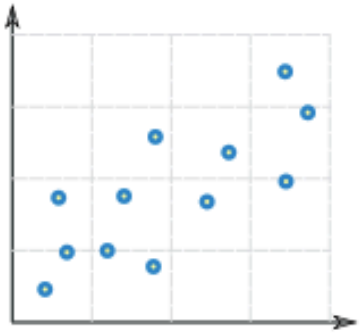
-0.5

None



0

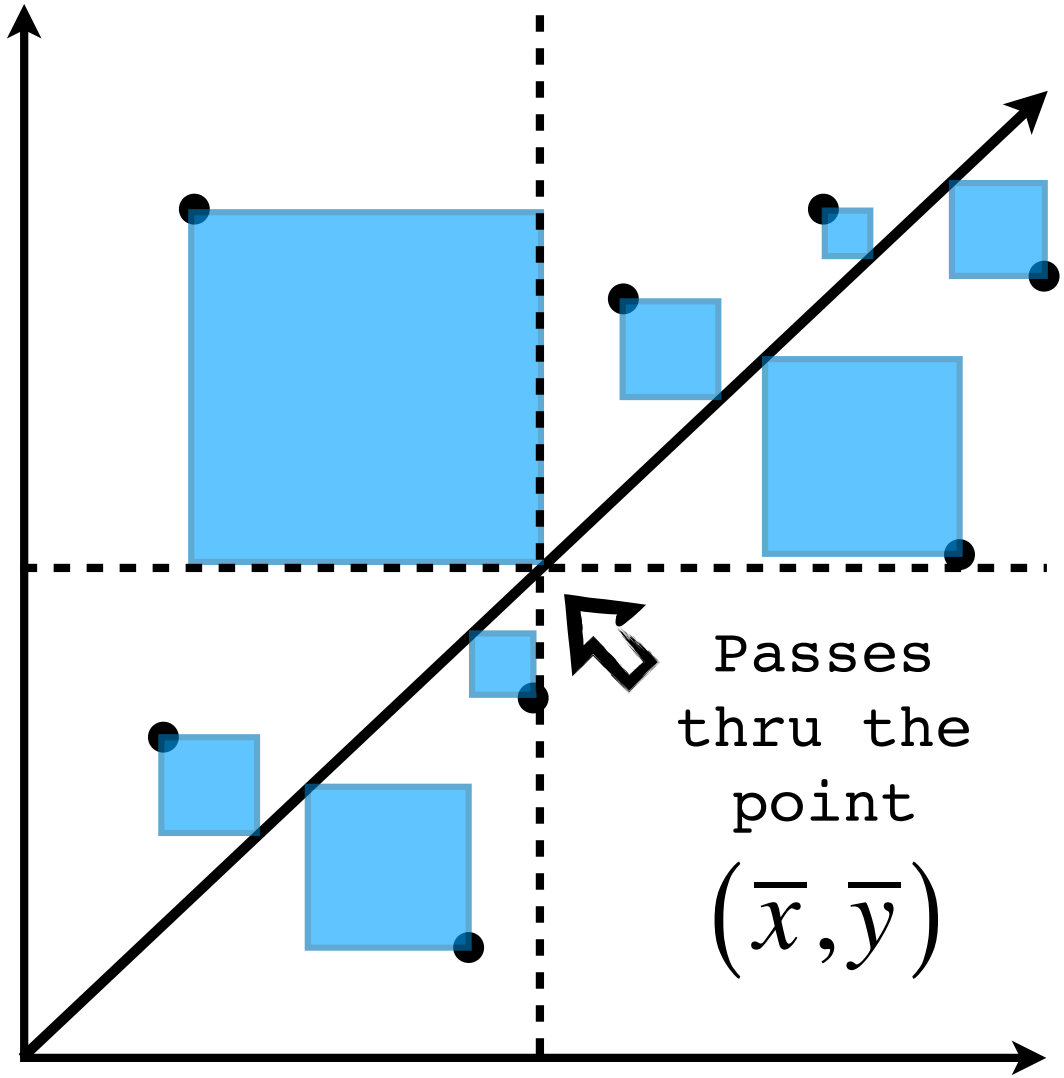
Low (+)



0.5

Least Squares Regression Line

(LSRL)



minimizes
sum of
squared
residuals

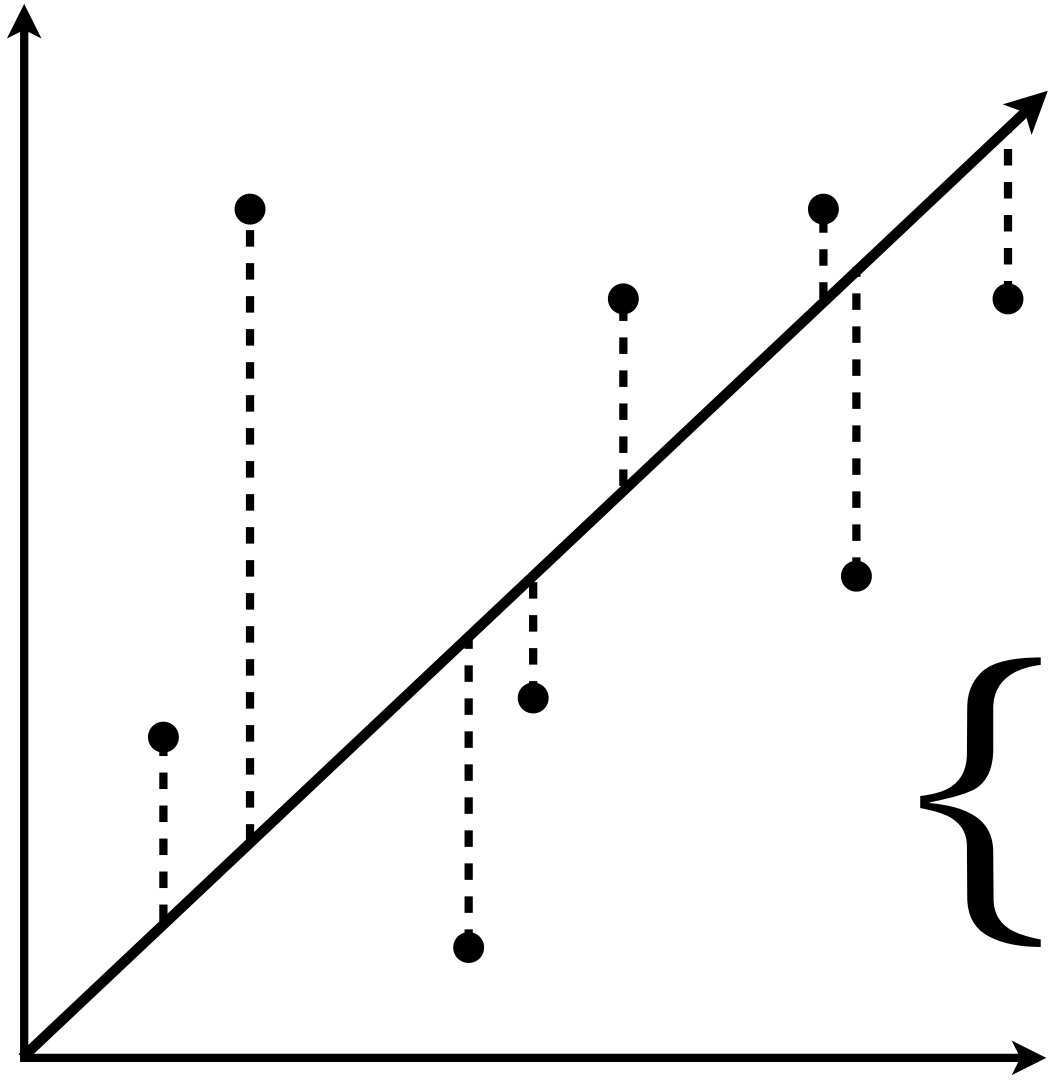
$$\hat{y} = a + bx$$

$a = y-$ intercept

$b =$ slope

Residuals

Prediction error



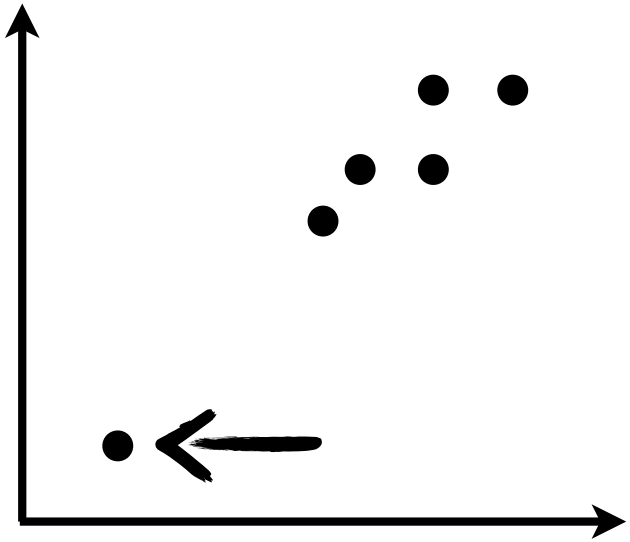
Observed – Predicted

$$y - \hat{y}$$

s = st. dev. of residuals

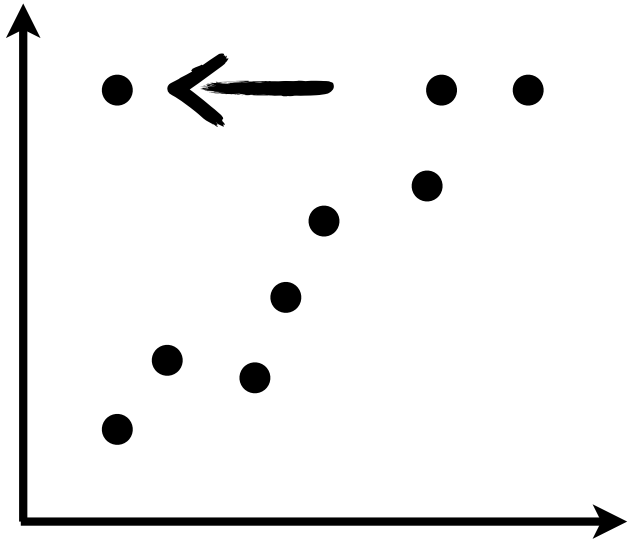
INTERPRETATION:
Actual value is typically [s] units from the predicted value

Unusual Points



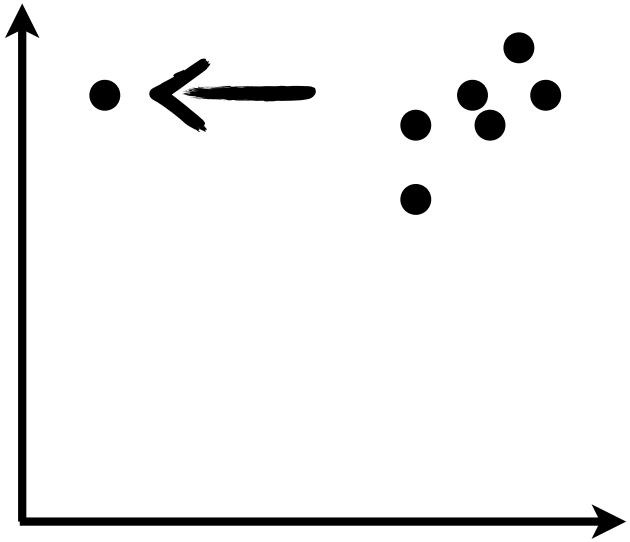
YES — High Leverage
NOT — Outlier
NOT — Influential

High Leverage
Very small/
large x-value



NOT — High Leverage
YES — Outlier
YES — Influential

Outlier
doesn't fit
overall pattern



YES — High Leverage
YES — Outlier
YES — Influential

Influential
Impacts the
LSRL