

\_\_\_(r<sup>2</sup> number as a %)\_\_\_ of the variation in \_\_\_(response variable)\_\_\_  
is explained by a linear regression on  
\_\_\_(explanatory variable).

- When the \_\_\_(explanatory variable)\_\_\_  
is 0 the \_\_\_(response variable)\_\_\_ IS PREDICTED to be \_\_\_(intercept #).

- For each increase of 1 \_\_\_(units)\_\_\_ in \_\_\_(explanatory variable)\_\_\_  
there is a \_\_\_(slope #)\_\_\_ increase in the \_\_\_(response variable)\_\_\_  
ON AVERAGE.

- Make a scatter plot
- Figure out explanatory vs response
- F,S,D
  - x and y outliers
- Run regression (write it properly)
- Know six facts about r
  - Positive is positive, negative is negative
  - -1 to 1, closer to 1s is stronger
  - Does not care about units
  - Correlation is the same if you switch x and y
  - Straight line only
  - Really affected by outliers
- Interpret r
  - o Direction
  - o Strength
- Interpret slope
- Interpret intercept
- Interpret  $r^2$
- Recognize to range of validity/extrapolation problem
- Use the line to predict, use the line "backwards"

Correlation describes

- the strength and
- direction of
- the **straight line** relation

- Positive is positive, negative is negative
- -1 to 1, closer to 1s is stronger
- Does not care about units
- Correlation is the same if you switch x and y
- Straight line only
- Really affected by outliers