# Statistics 2: More Advanced Linear Models

2021-07-15

Calculating, Interpreting, and Extending Linear Models

- Linear models minimize the squared distance between the points of two variables

- Interpretation
    - For every 1 additional meter in height, weight increases by 61.27 kg
    - The y-intercept is often meaningless, especially in biological data. Here's it's when you weigh 0 kg, your height is -39.06 meters

- The additional value here, whether than just asking if these variables are significantly related is that you can use new measurements of X to predict what Y will be
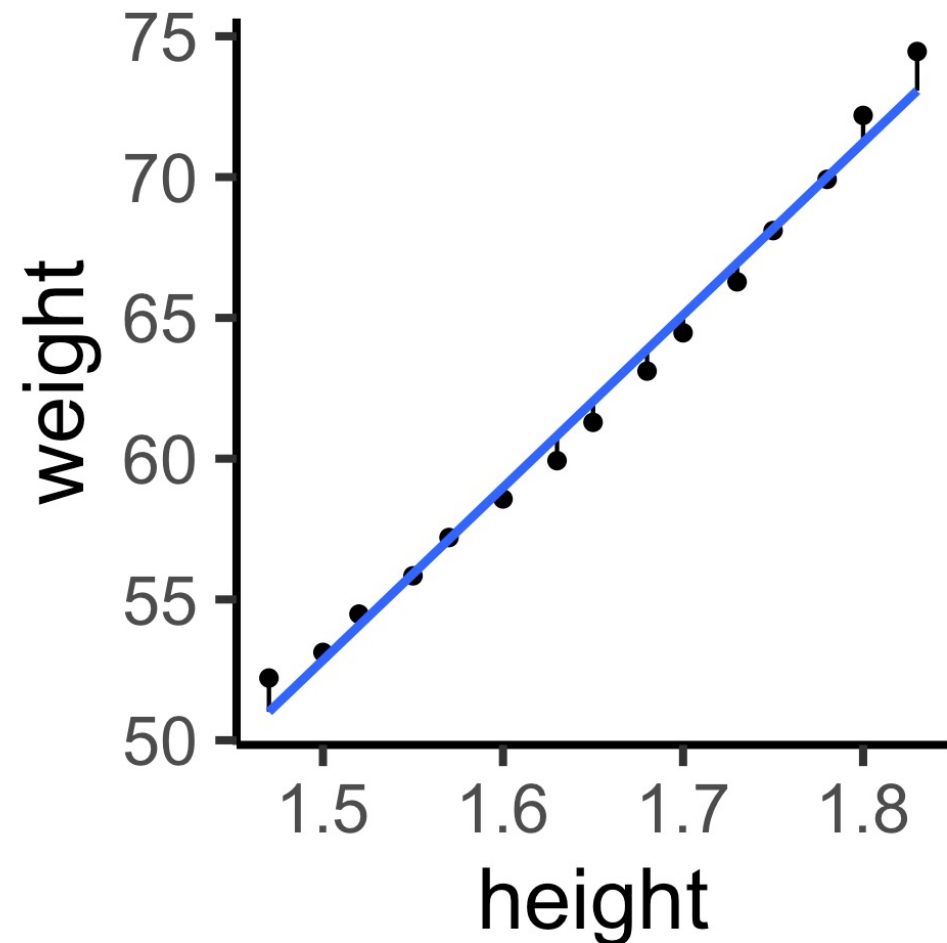
How is a linear model calculated for more than one variable?



3D multiple linear regression model
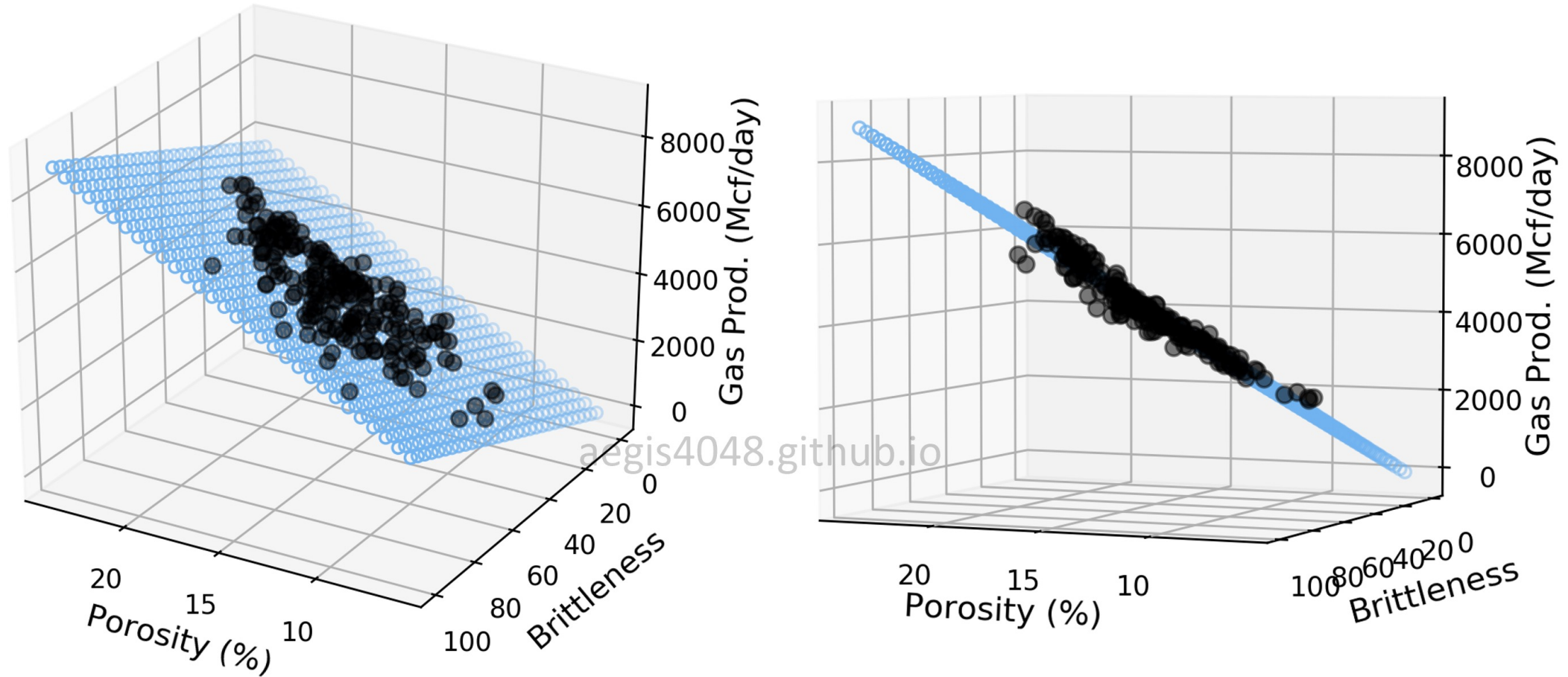
Additive Linear Models

- These are for when you have no relationship between your variables/when your variables are totally independent

- By using multiple independent variables that each explain some variation in Y, you can get a better prediction

- This is the most common type of multiple regression and what you should use by default
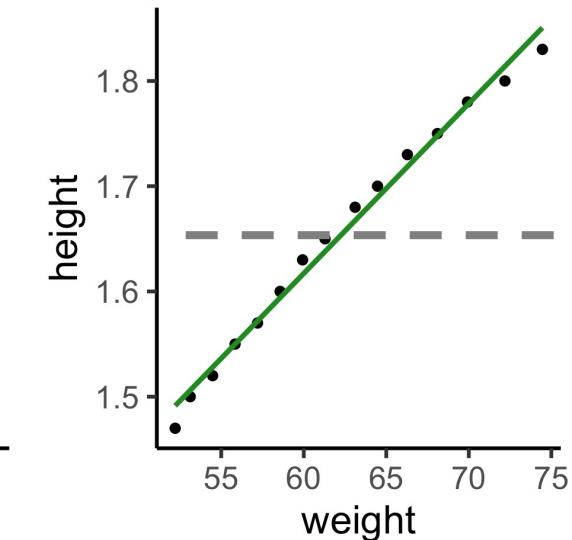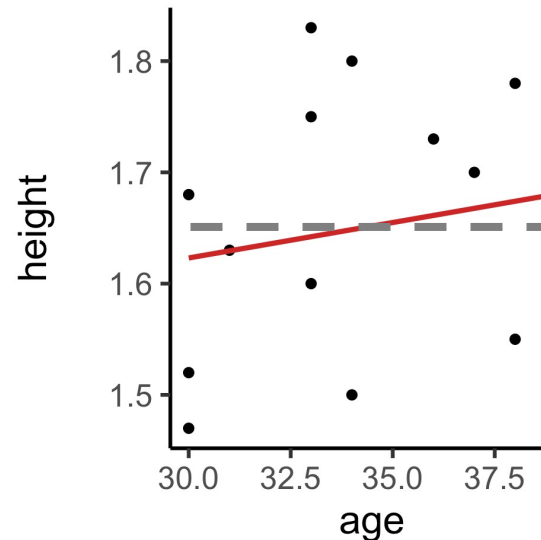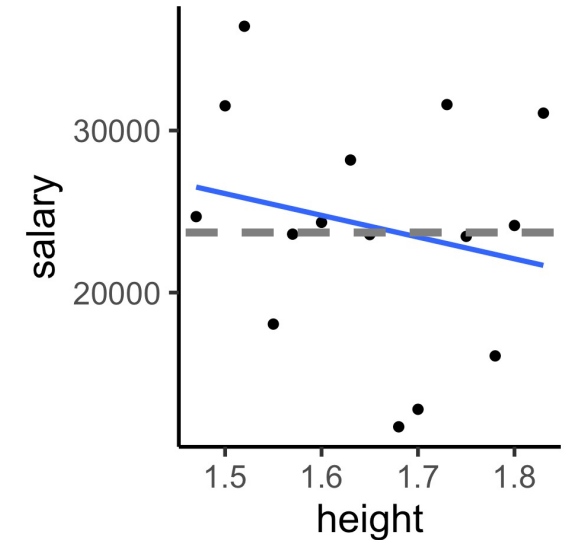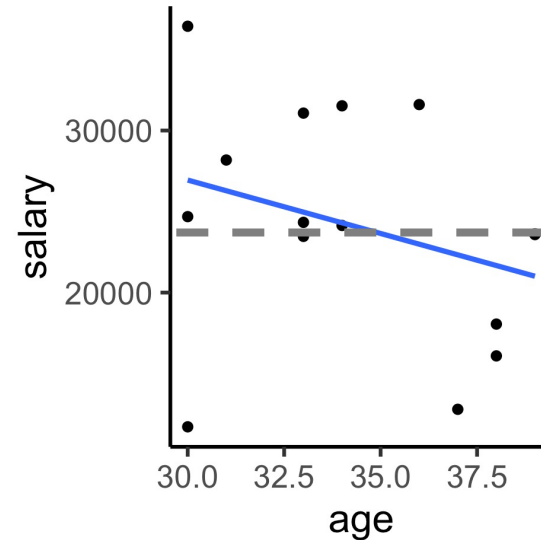
Interaction Linear Models

- You would use an interaction model when the values of your X variables depend on each other

- For example, height and weight. Tall people must weigh more, and short people must weigh less

- You may not know whether your variables are independent, so you will sometimes need to test both additive and interaction models
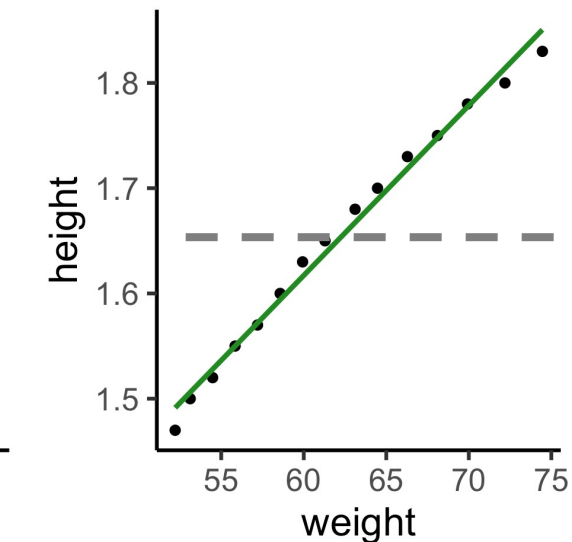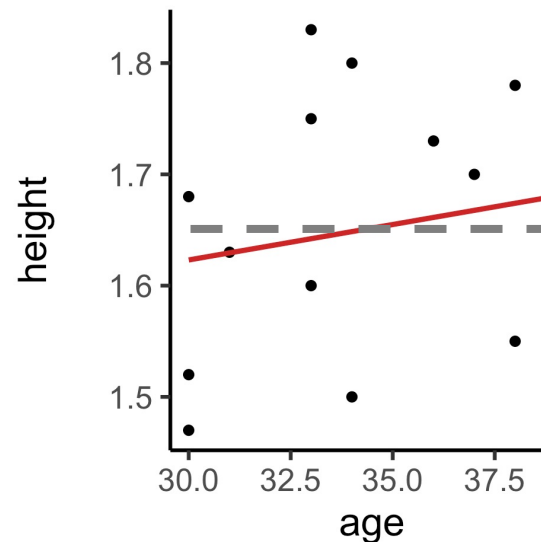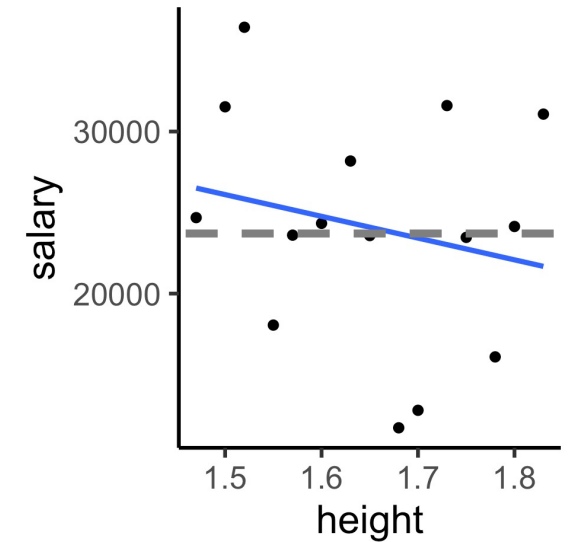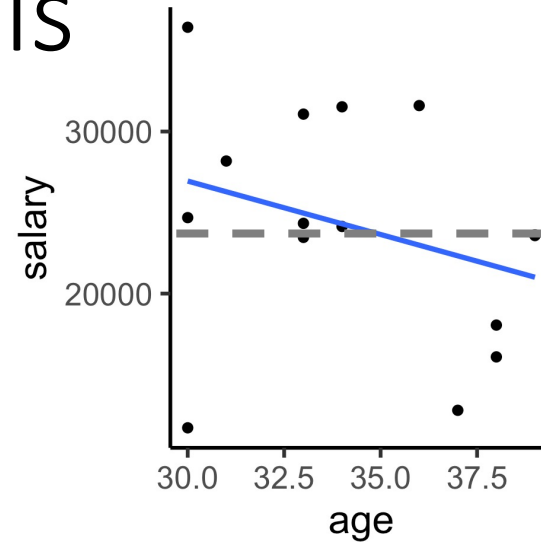
How do you pick the best model?

R^2

- R^2 explains the goodness-of-fit of a model
- It represents the percent variation the X variable explains in the Y variable
- Good models are close to 1 or -1, while bad models are close to 0
- Importantly it increases with each additional variable
- For example, for height vs weight, R^2 = 0.98, so weight explains 98% of the variation in height

Akaike Information Criterion (AIC)

- Also quantifies the goodness-of-fit of a model
- AIC selects for the model that explains the greatest amount of variation with the fewest variables (attempts to control for adding in additional variables)
- The lower the AIC the better
- For example, AIC for height vs weight is -84.47 but for salary vs age it's 312.22
- If AICs for models are similar, take the model with fewer variables