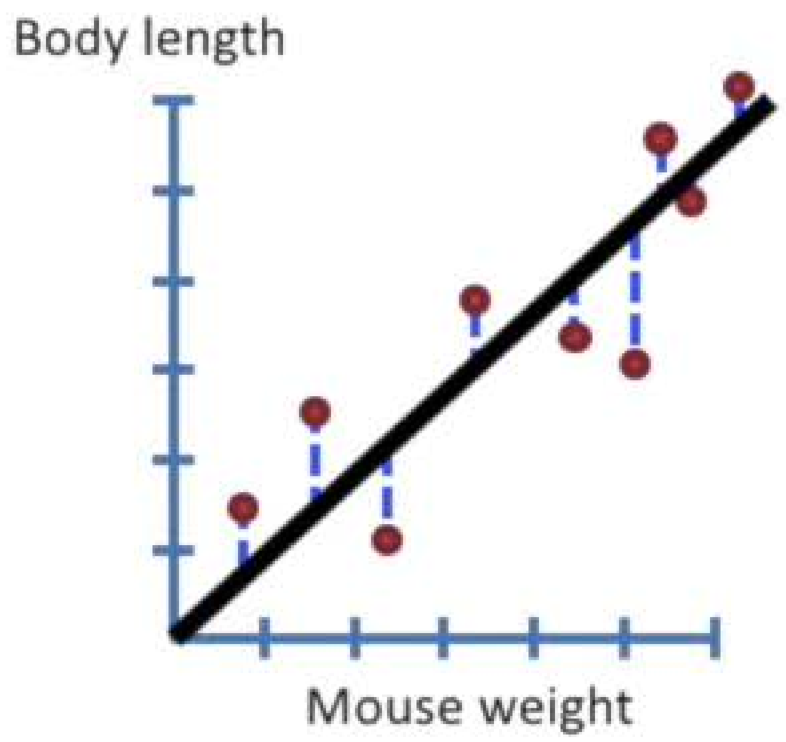
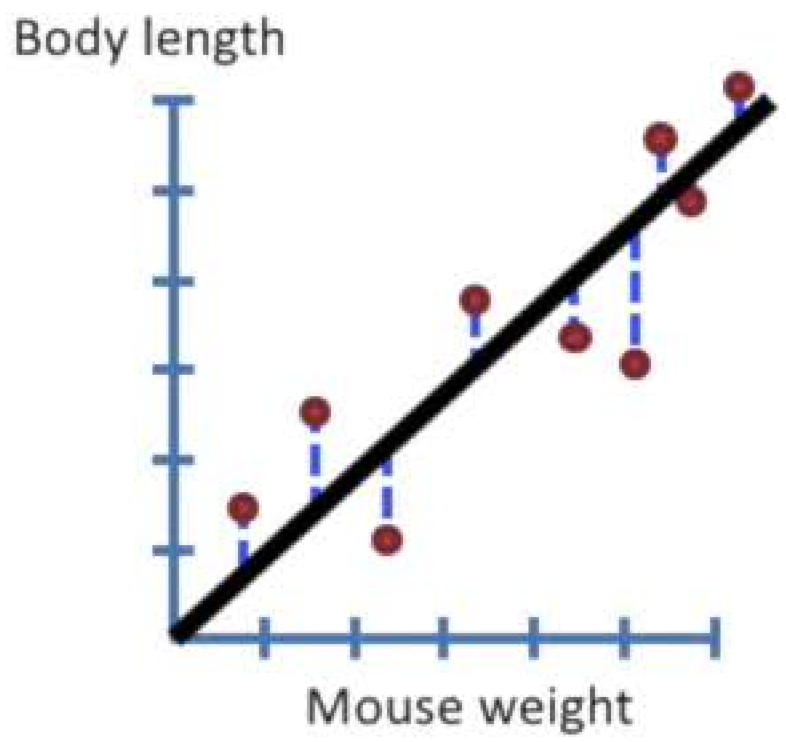
Simple regression



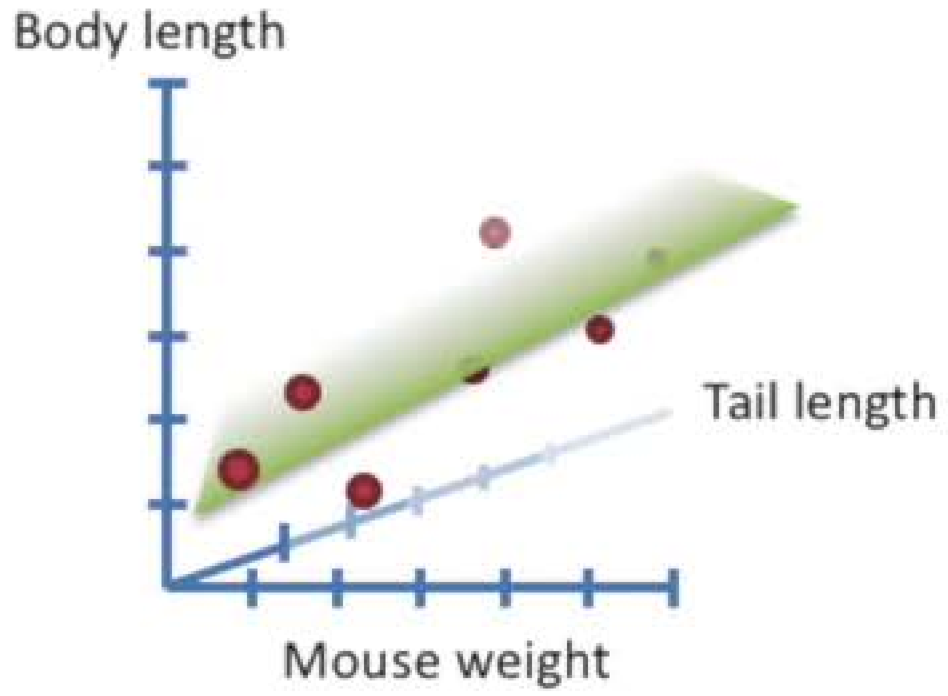
y = y-intercept + slope x

Simple regression



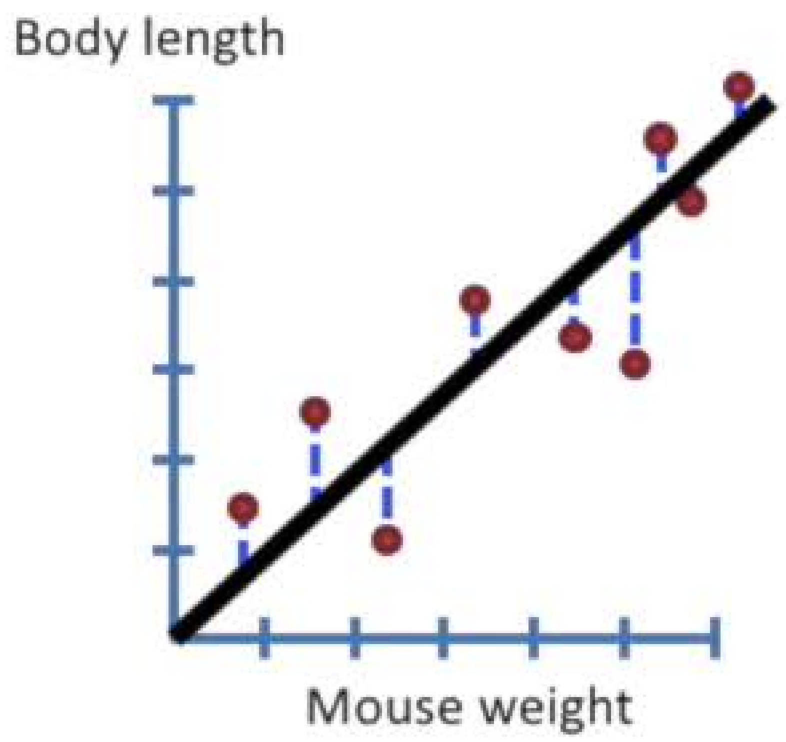
y = y-intercept + slope x

Multiple regression



y = y-intercept + slope x + slope z

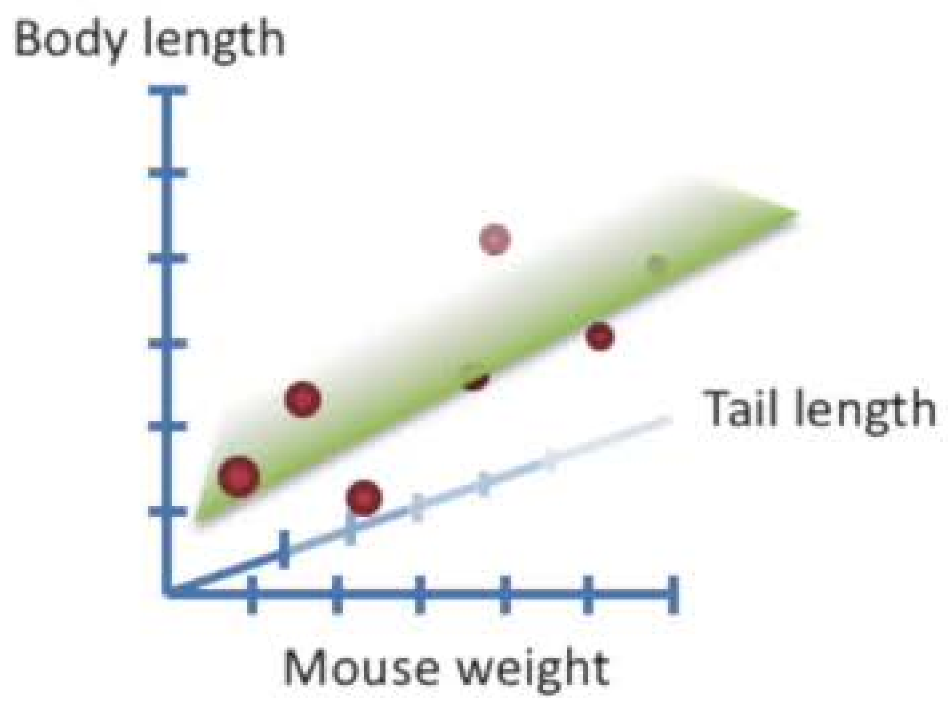
Simple regression



y = y-intercept + slope x

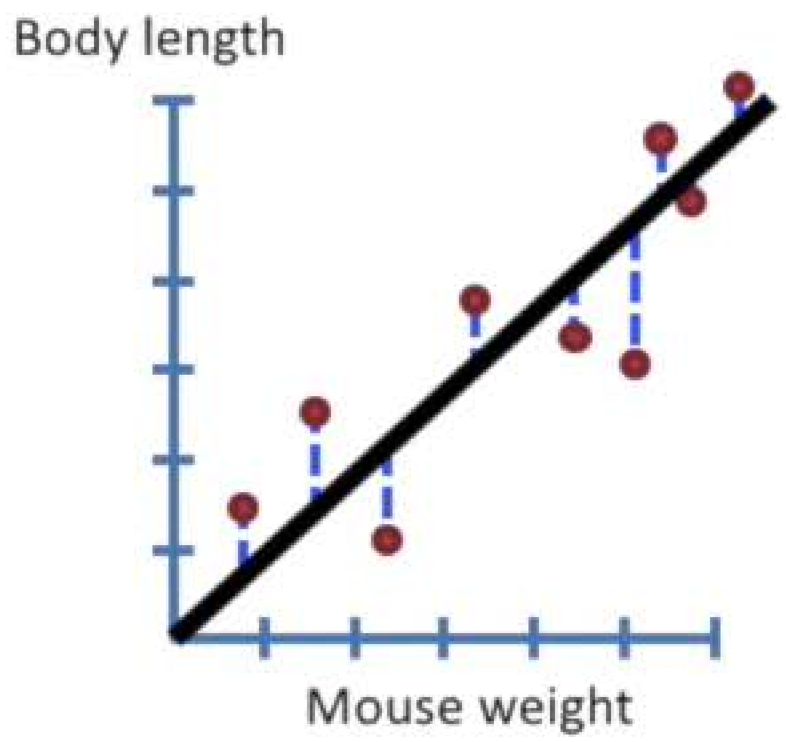
Calculating R^2 is the same for both simple and multiple regression

Multiple regression



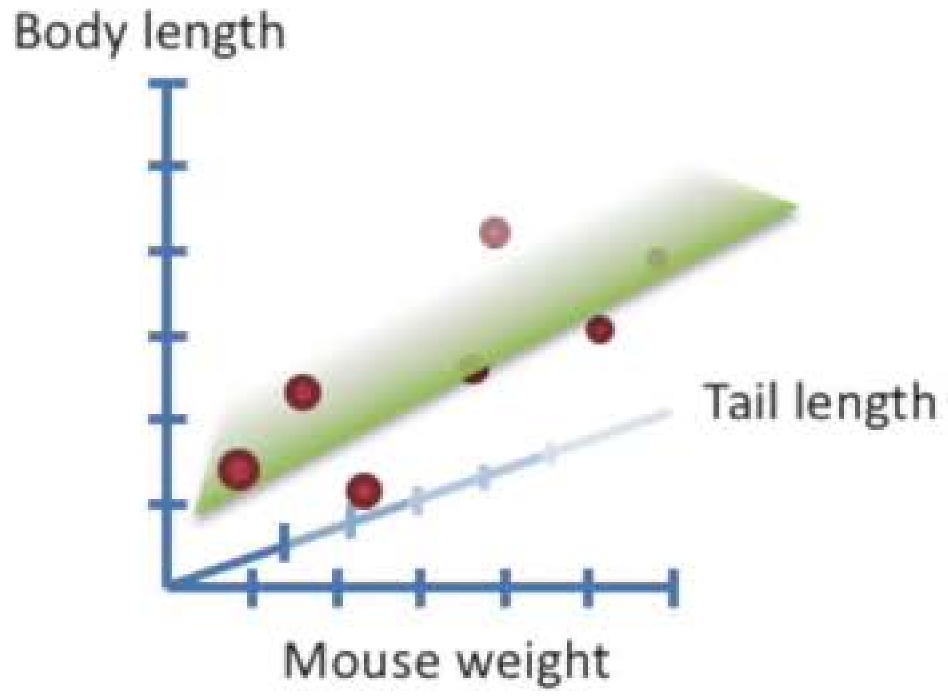
y = y-intercept + slope x + slope z

Simple regression



y = y-intercept + slope x

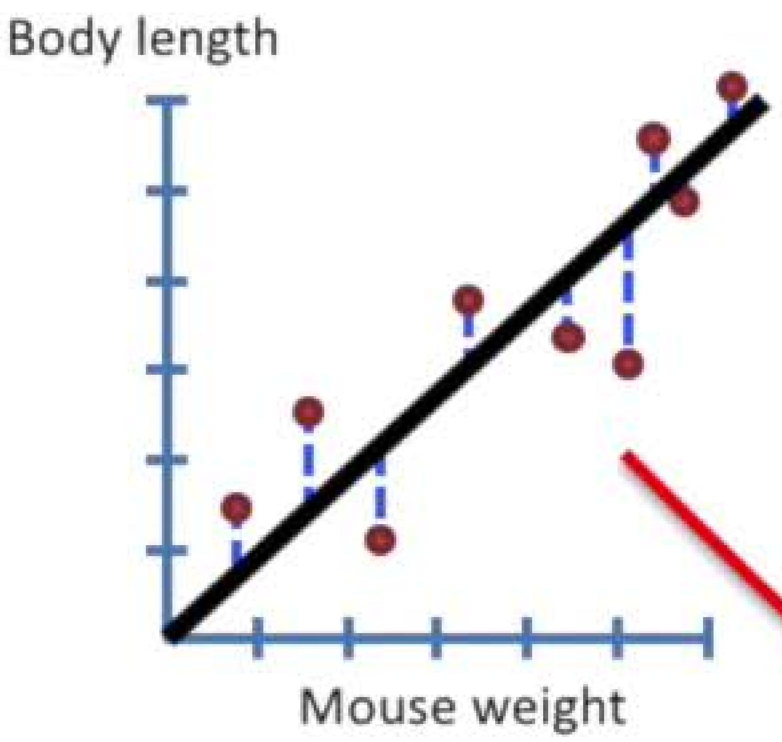
Multiple regression



y = y-intercept + slope x + slope z

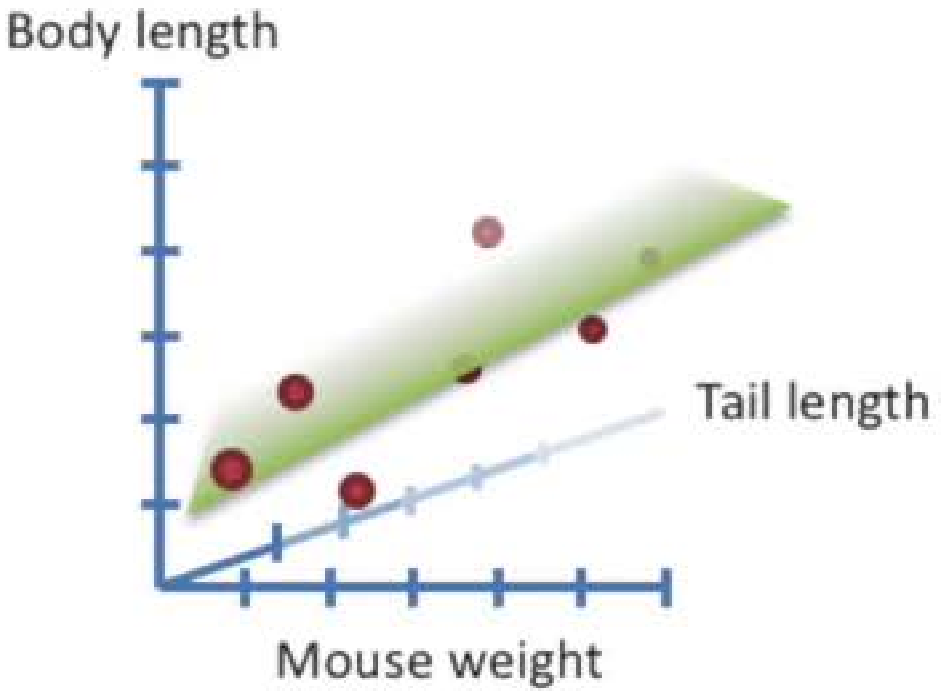
$$R^2 = \frac{\text{SS(mean)} - \text{SS(fit)}}{\text{SS(mean)}}$$

Simple regression



y = y-intercept + slope x

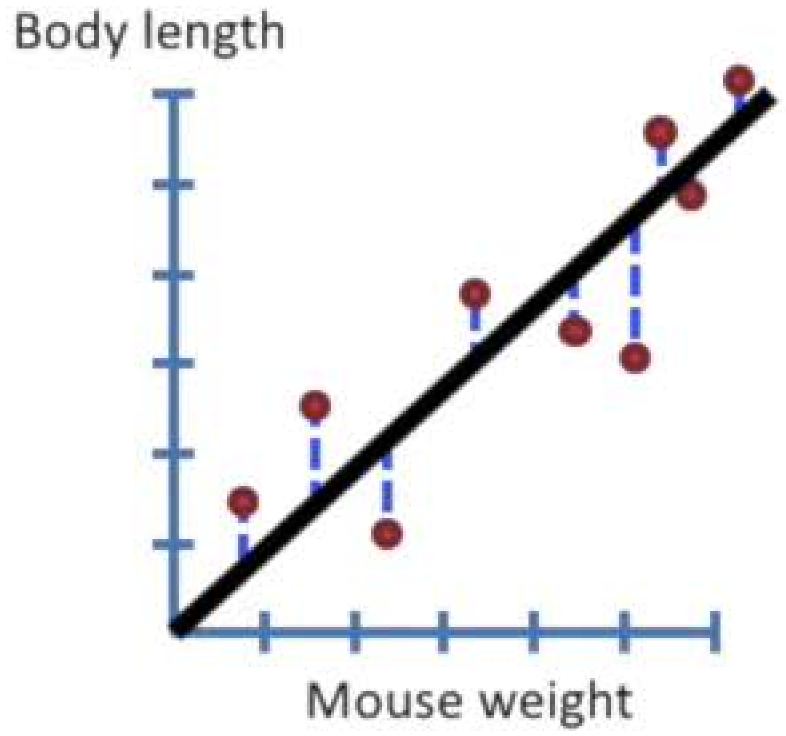
Multiple regression



y = y-intercept + slope x + slope z

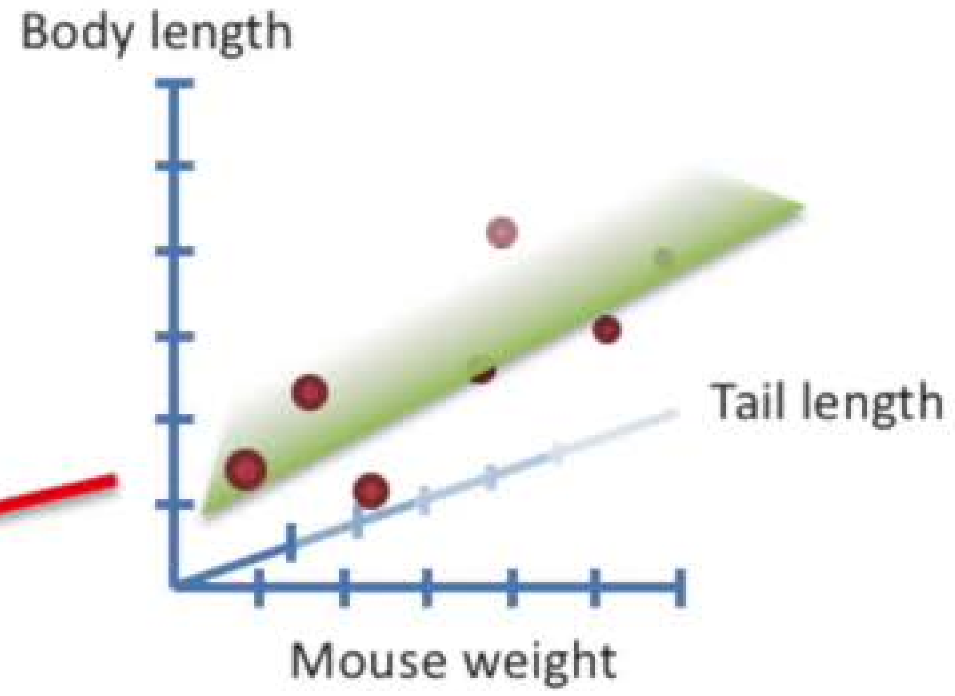
$$R^2 = \frac{\text{SS(mean)} - \text{SS(fit)}}{\text{SS(mean)}}$$

Simple regression



y = y-intercept + slope x

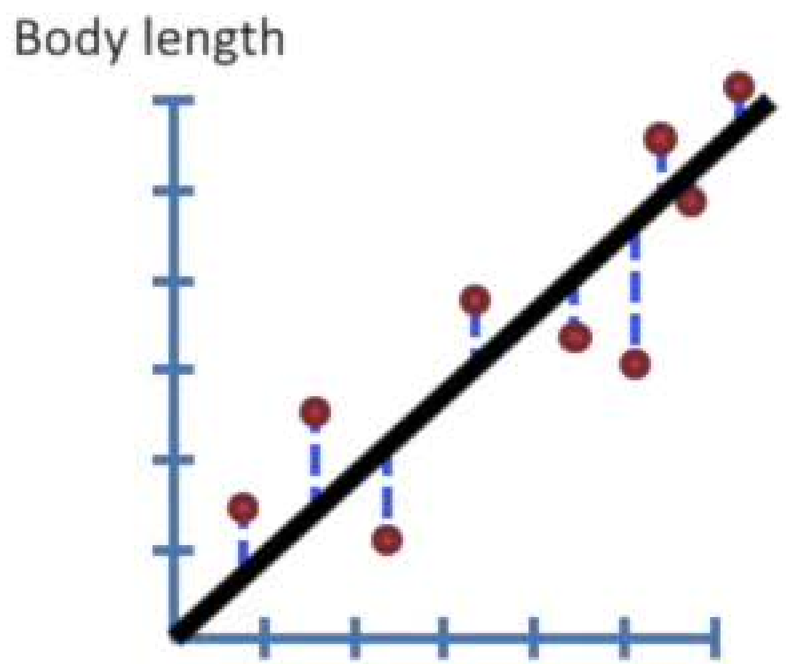
Multiple regression



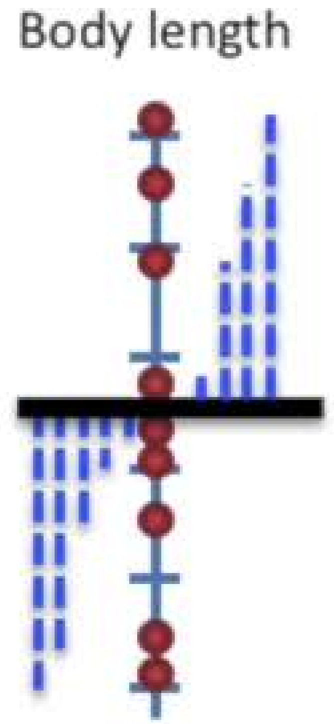
y = y-intercept + slope x + slope z

$$R^2 = \frac{\text{SS(mean)} - \text{SS(fit)}}{\text{SS(mean)}}$$

Simple regression



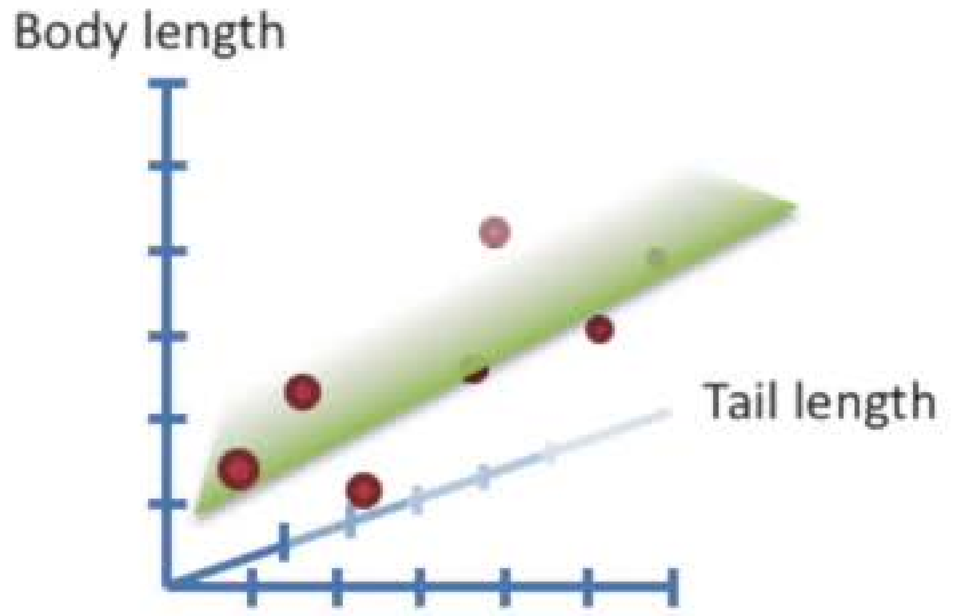
y = y-intercept + slope x



y = y-intercept

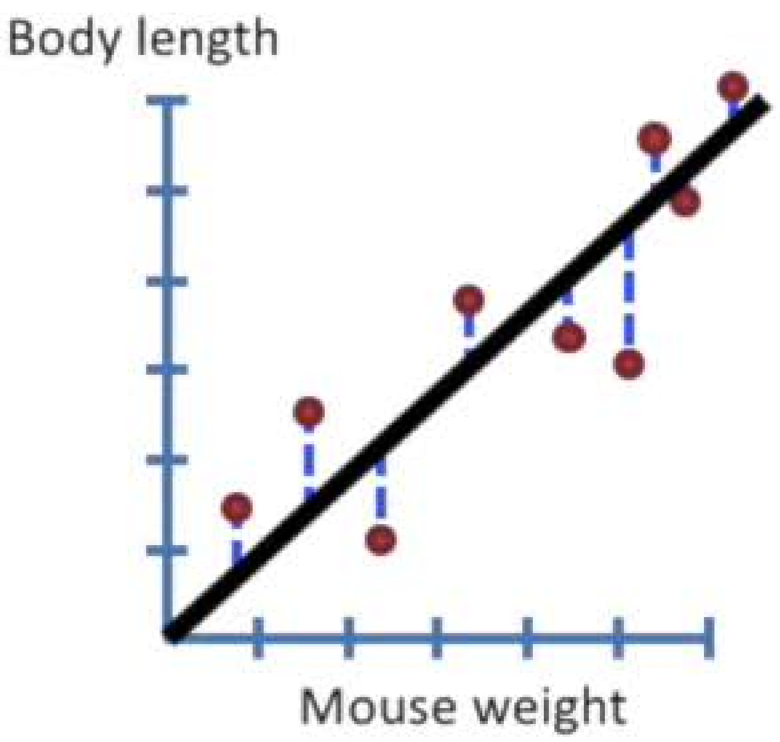
$$R^2 = \frac{SS(\text{mean}) - SS(\text{fit})}{SS(\text{mean})}$$

Multiple regression

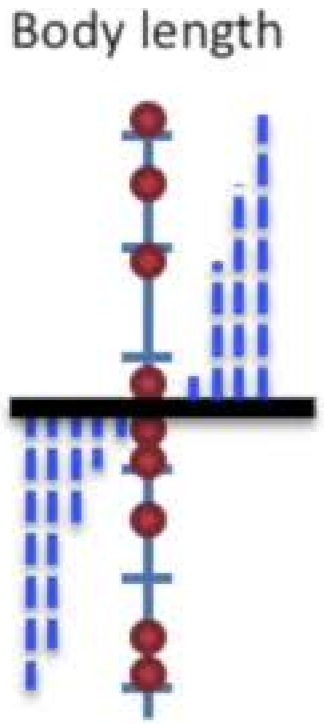


y = y-intercept + slope x + slope z

Simple regression

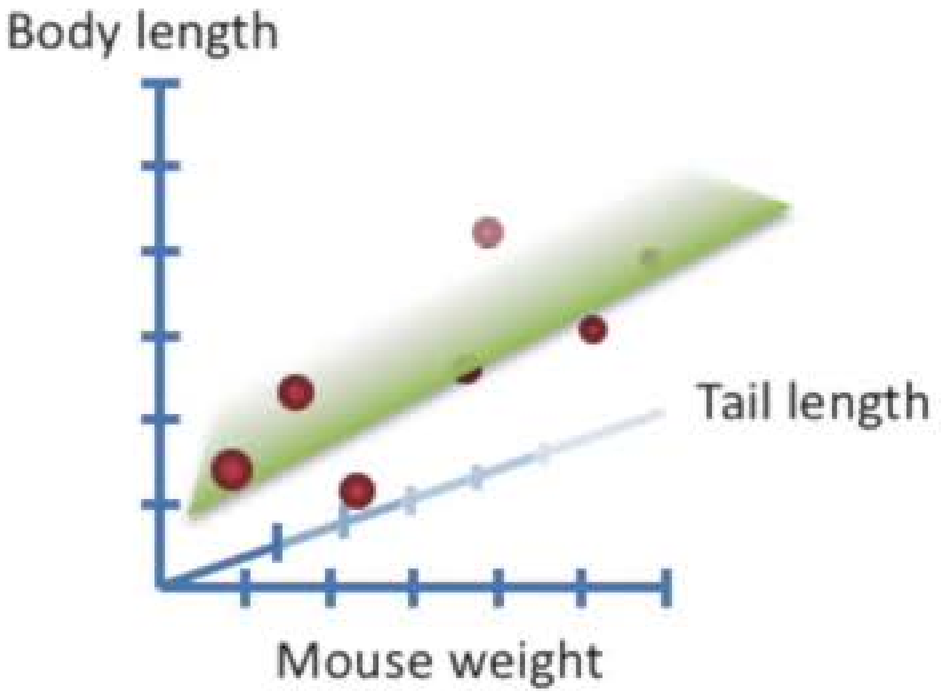


y = y-intercept + slope x



y = y-intercept

Multiple regression

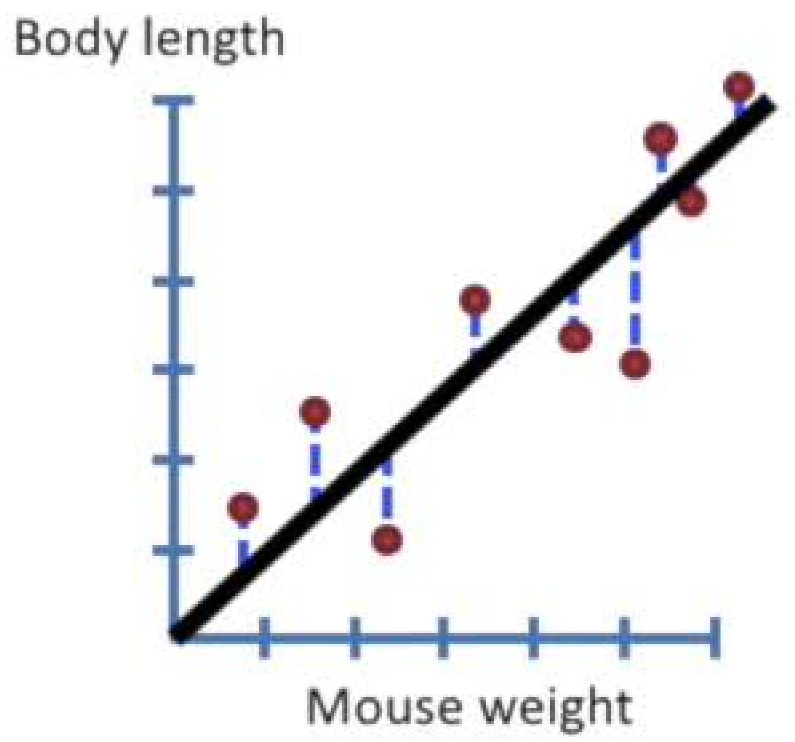


y = y-intercept + slope x + slope z

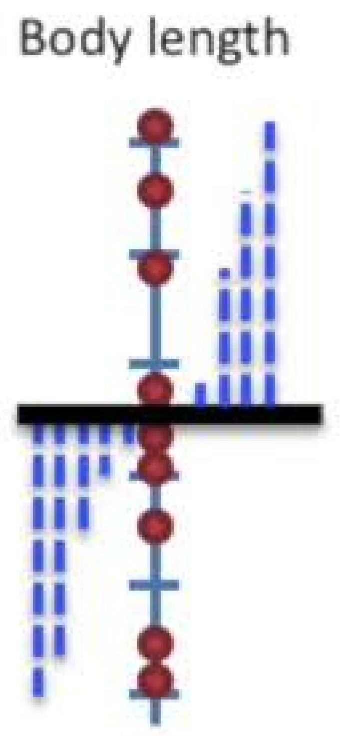
$$R^2 = \frac{\text{SS(mean)} - \text{SS(fit)}}{\text{SS(mean)}}$$

For multiple regression, you adjust R^2 to compensate for the additional parameters in the equation.

Simple regression

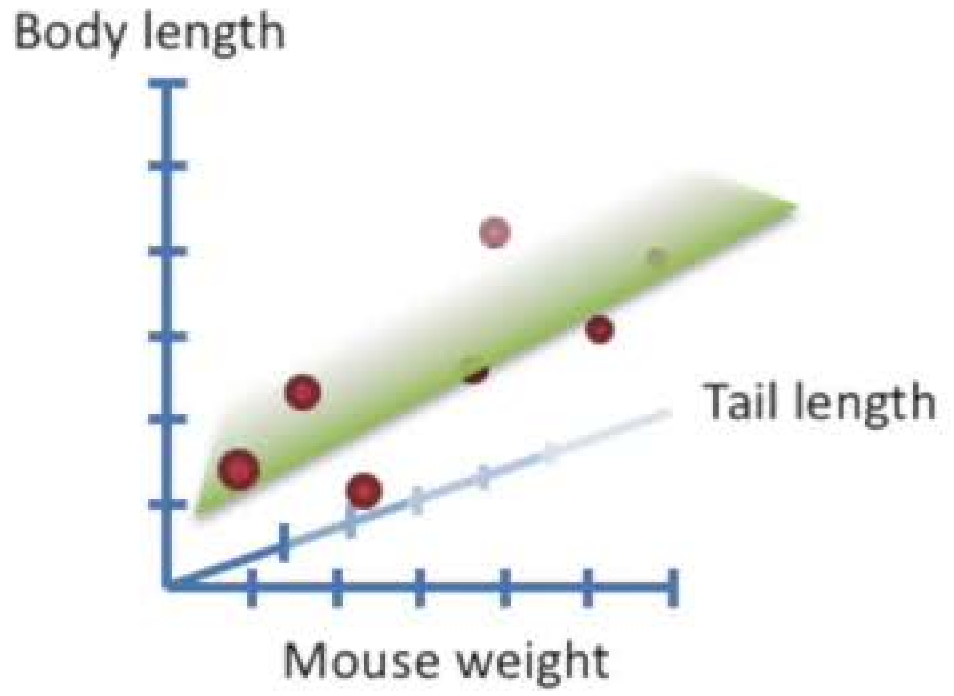


y = y-intercept + slope x



y = y-intercept

Multiple regression

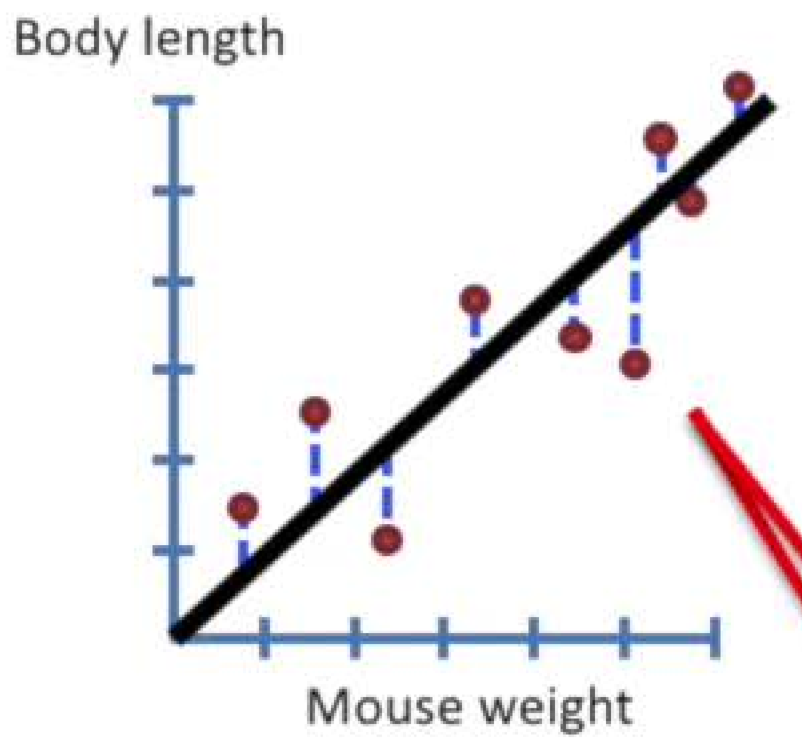


y = y-intercept + slope x + slope z

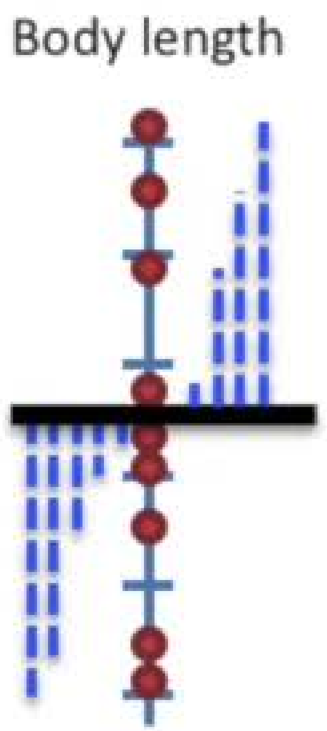
$$F = \frac{\text{SS(mean)} - \text{SS(fit)} / (p_{\text{fit}} - p_{\text{mean}})}{\text{SS(fit)} / (n - p_{\text{fit}})}$$

Calculating F and the p-value is pretty much the same…

Simple regression

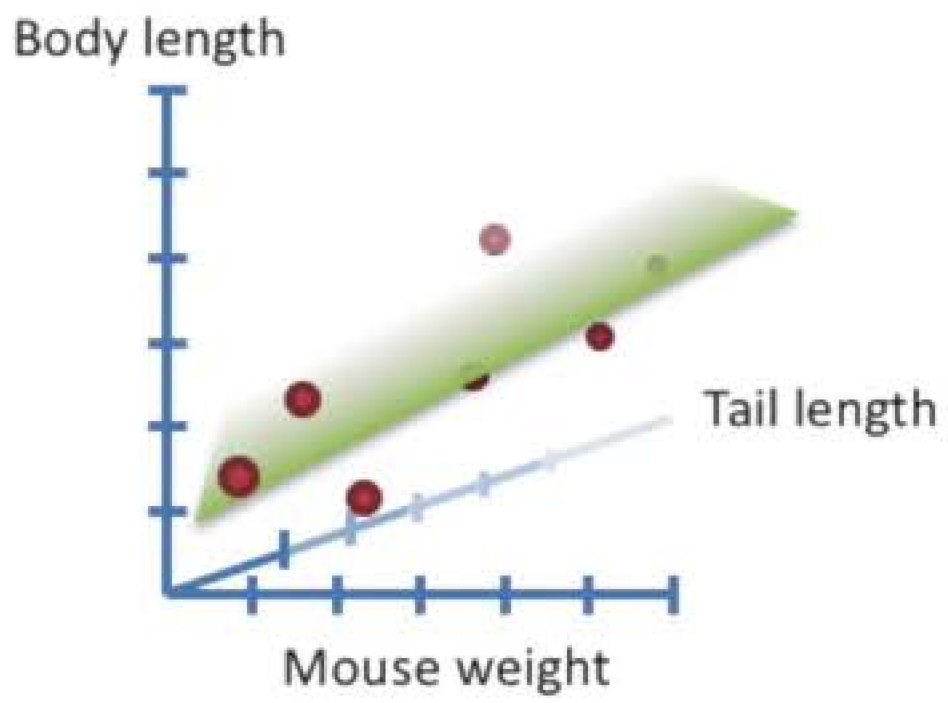


y = y-intercept + slope x



y = y-intercept

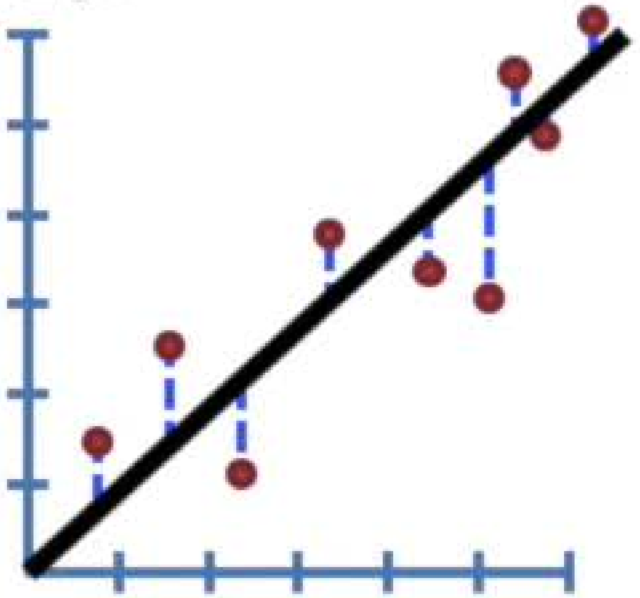
Multiple regression



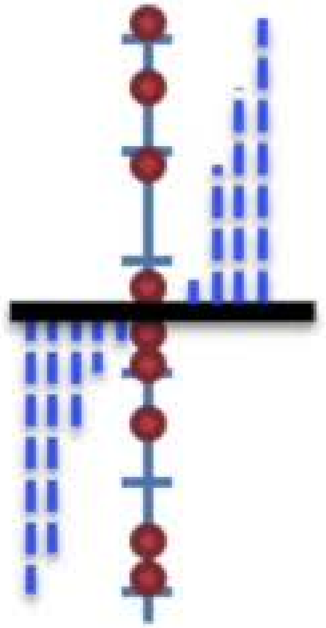
y = y-intercept + slope x + slope z

$$F = \frac{\text{SS(mean)} - \text{SS(fit)} / (p_{\text{fit}} - p_{\text{mean}})}{\text{SS(fit)} / (n - p_{\text{fit}})}$$

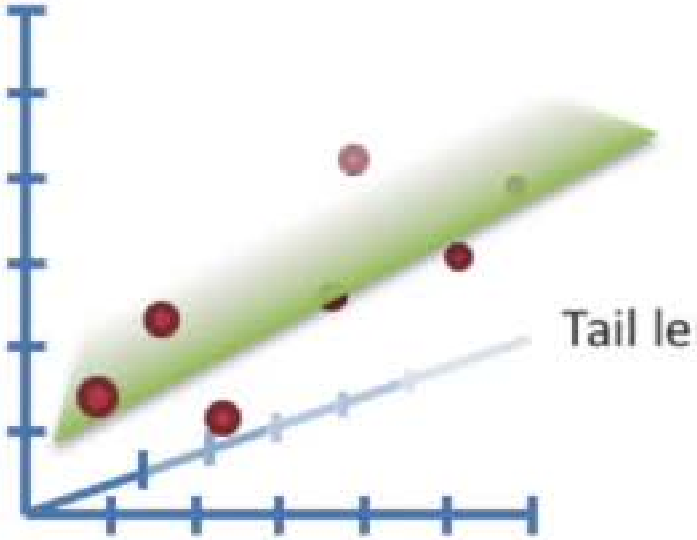
Simple regression



y = y-intercept + slope x



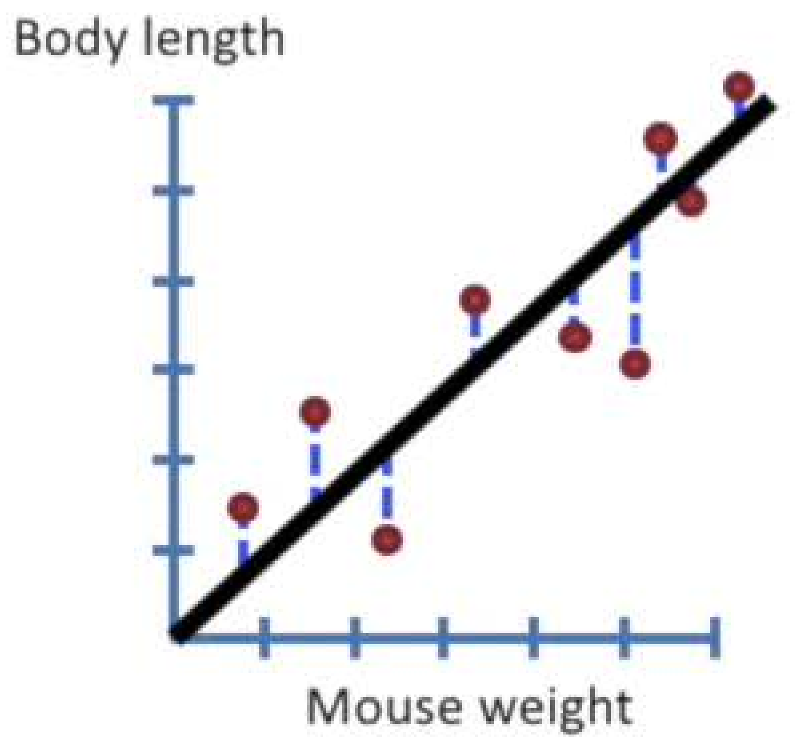
Multiple regression



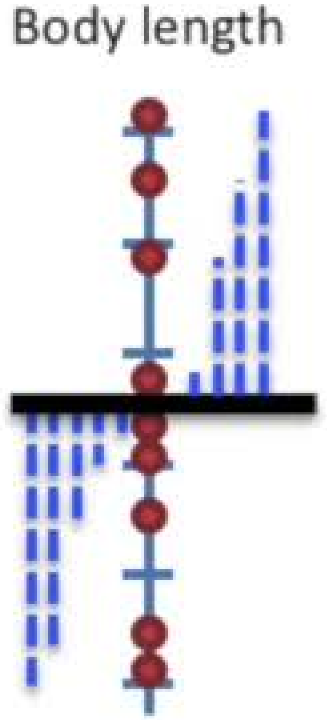
y = y-intercept + slope x + slope z

$$F = \frac{\text{SS(mean)} - \text{SS(fit)} / (p_{\text{fit}} - p_{\text{mean}})}{\text{SS(fit)} / (n - p_{\text{fit}})}$$

Simple regression

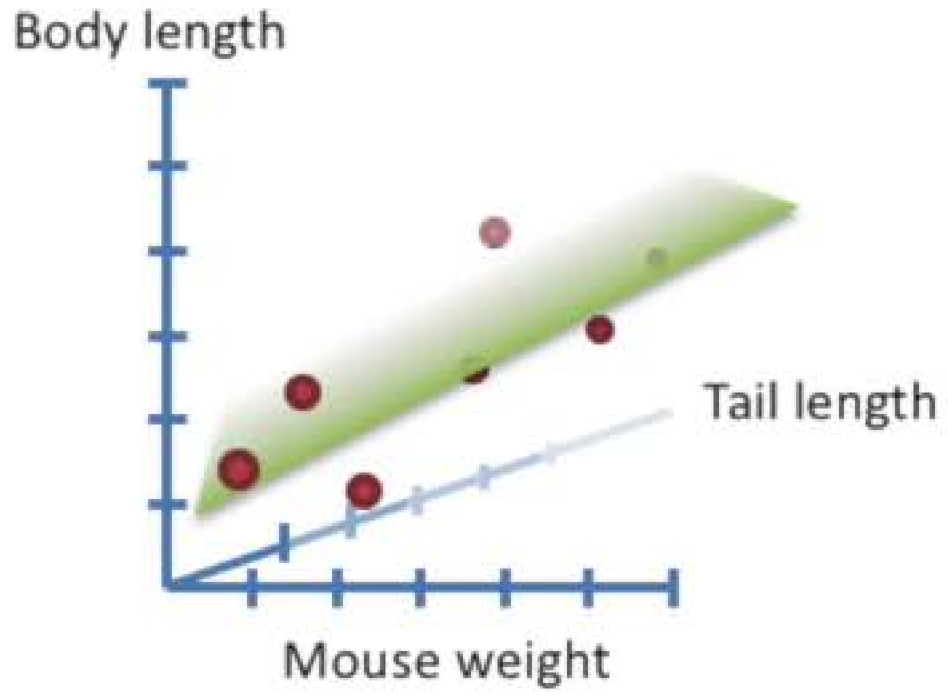


y = y-intercept + slope x



y = y-intercept

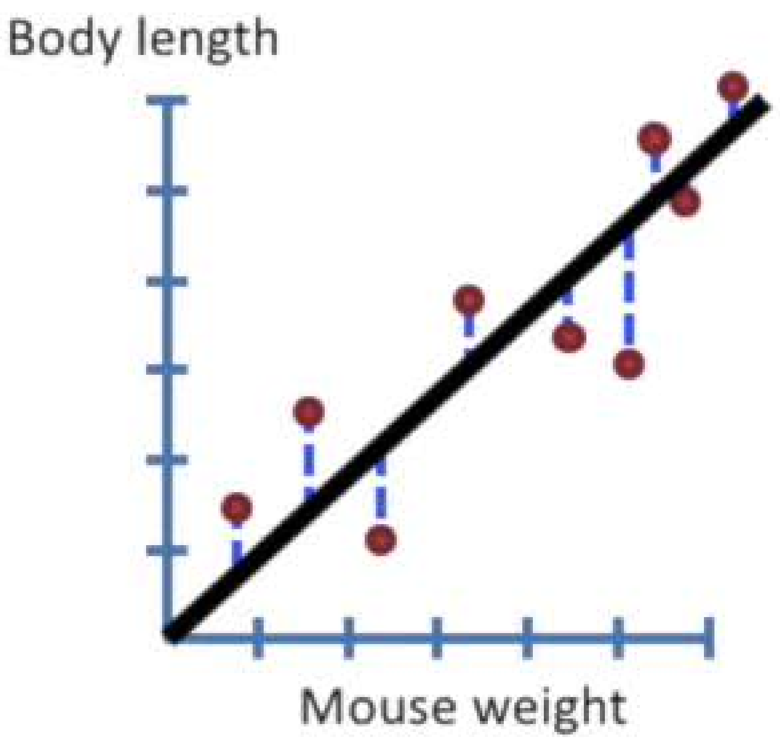
Multiple regression



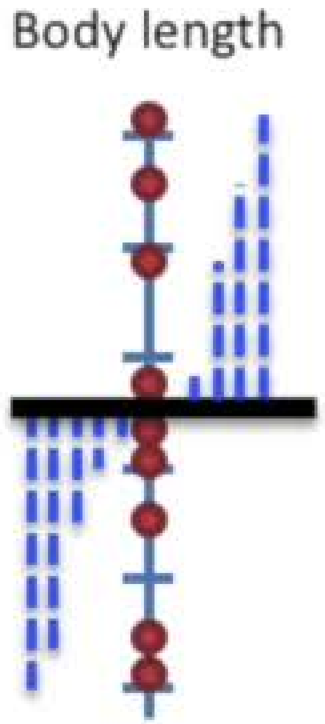
y = y-intercept + slope x + slope z

$$F = \frac{SS(\text{mean}) - SS(\text{fit}) \,/\, (p_{\text{fit}} - p_{\text{mean}})}{SS(\text{fit}) \,/\, (n - p_{\text{fit}})}$$

Simple regression



y = y-intercept + slope x

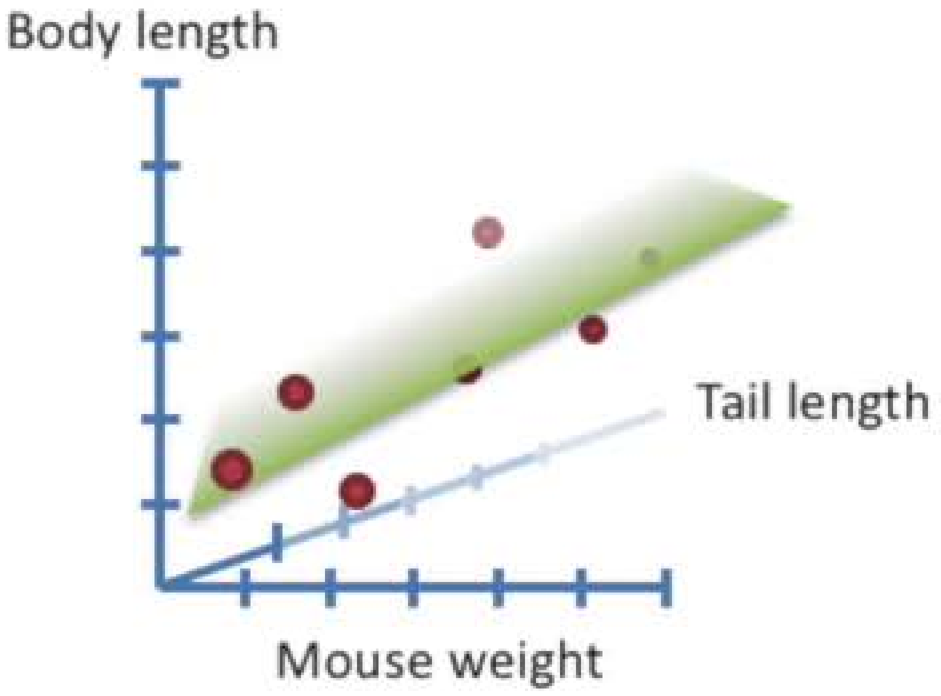


y = y-intercept

$p_{\text{fit}} = 2$

$$F = \frac{SS(\text{mean}) - SS(\text{fit}) / (p_{\text{fit}} - p_{\text{mean}})}{SS(\text{fit}) / (n - p_{\text{fit}})}$$

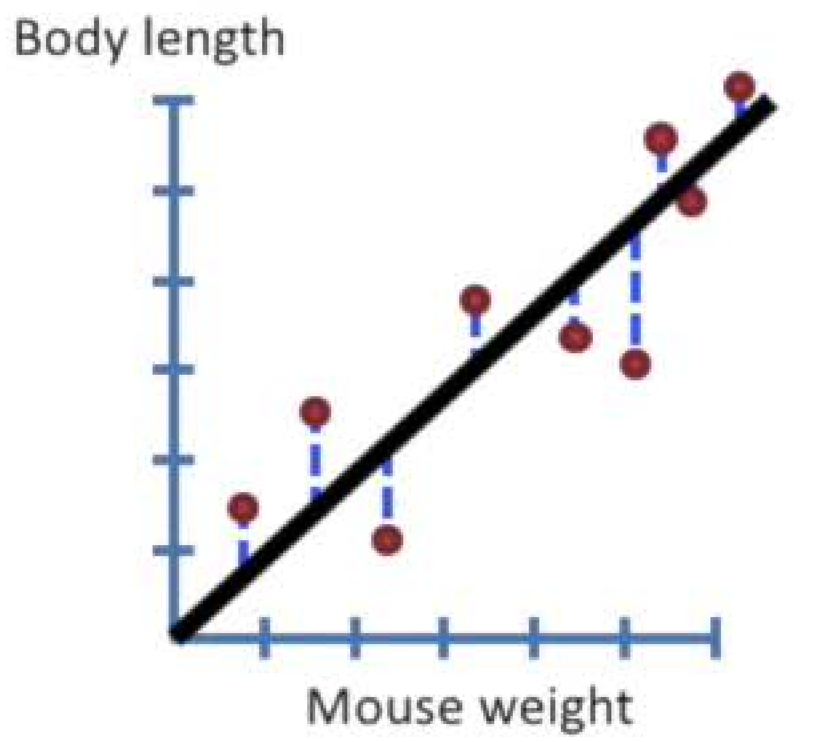
Multiple regression

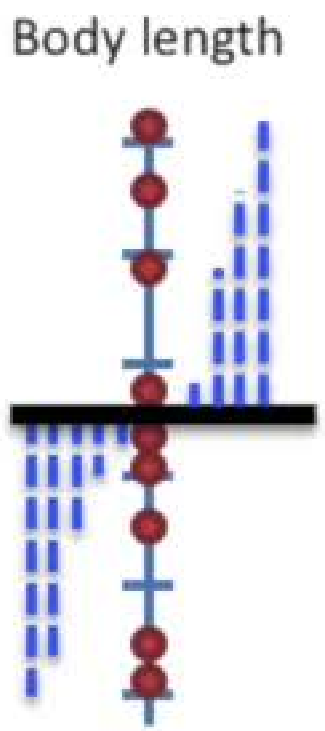


y = y-intercept + slope x + slope z

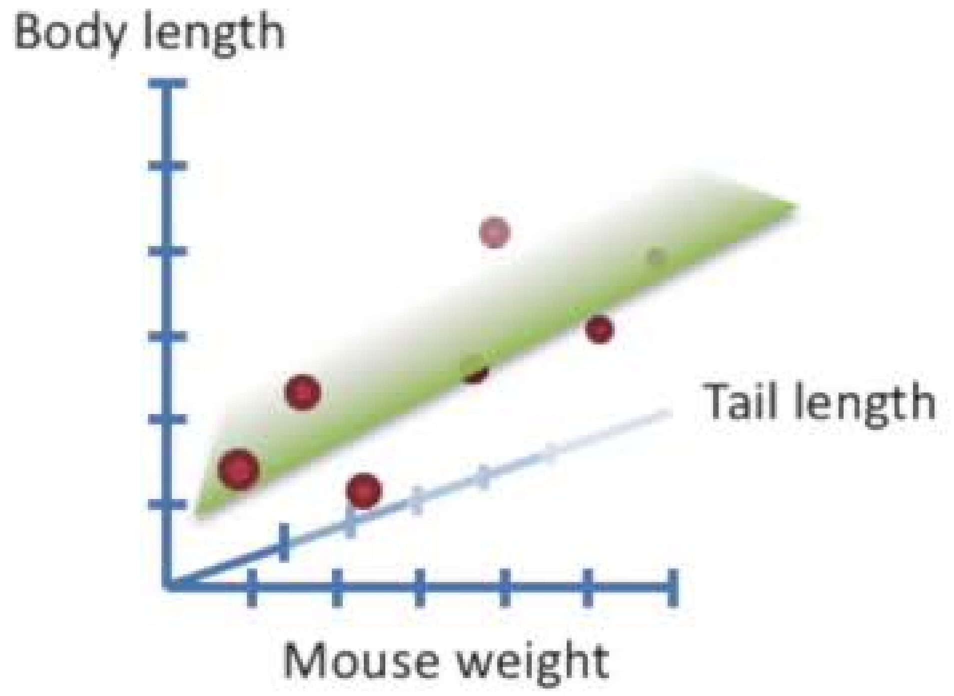
Simple regression

Multiple regression





y = y-intercept



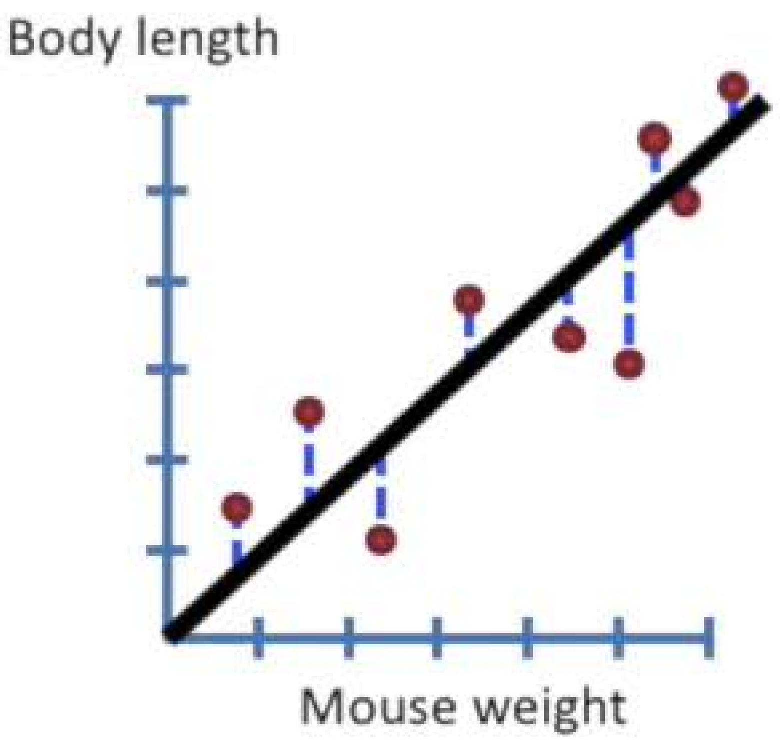
y = y-intercept + slope x

y = y-intercept + slope x + slope z

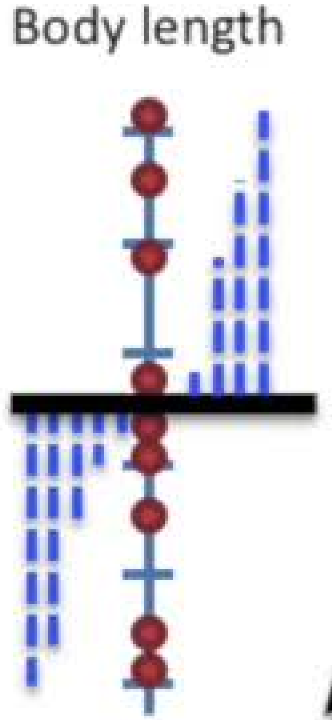
$$F = \frac{\text{SS(mean)} - \text{SS(fit)} / (p_{\text{fit}} - p_{\text{mean}})}{\text{SS(fit)} / (n - p_{\text{fit}})}$$

$p_{\text{fit}} = 3$

Simple regression



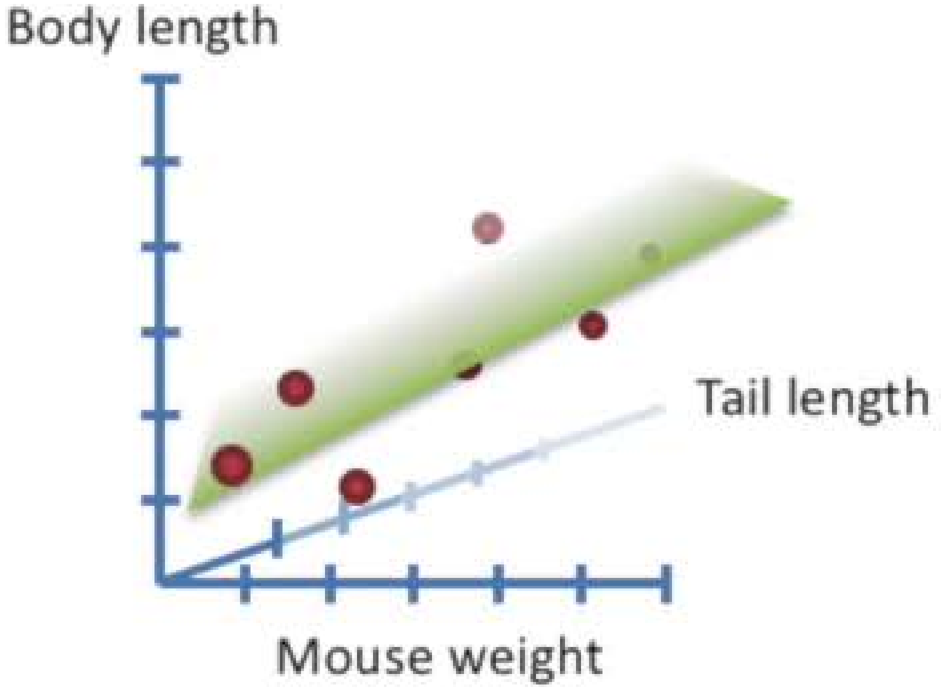
y = y-intercept + slope x



$p_{\text{mean}} = 1$

y = y-intercept

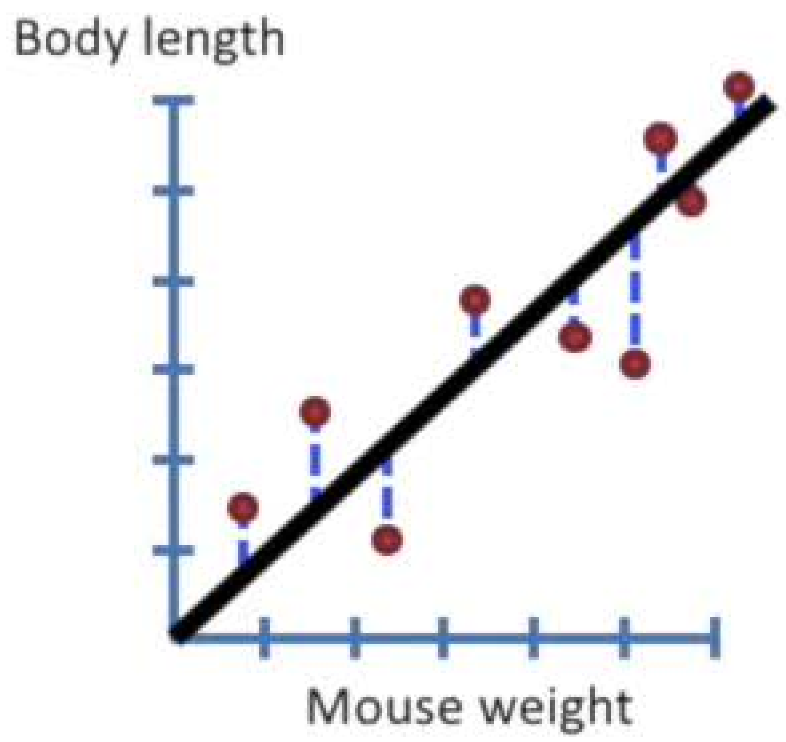
Multiple regression



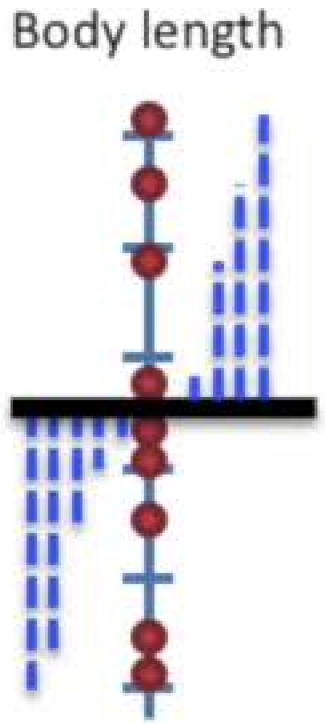
y = y-intercept + slope x + slope z

$$F = \frac{\text{SS(mean)} - \text{SS(fit)} / (p_{\text{fit}} - p_{\text{mean}})}{\text{SS(fit)} / (n - p_{\text{fit}})}$$

Simple regression

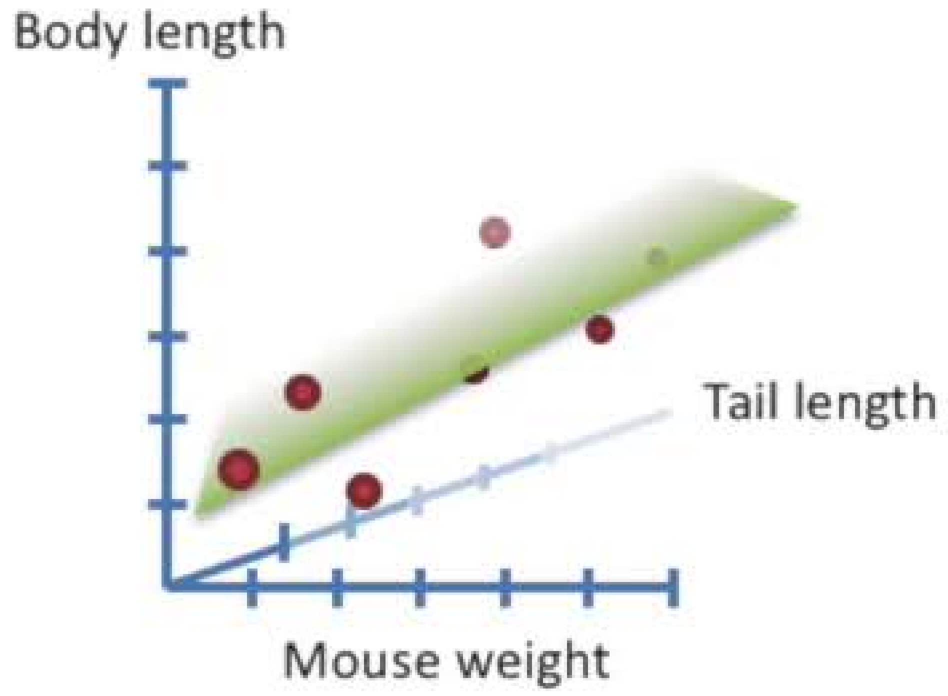


y = y-intercept + slope x



y = y-intercept

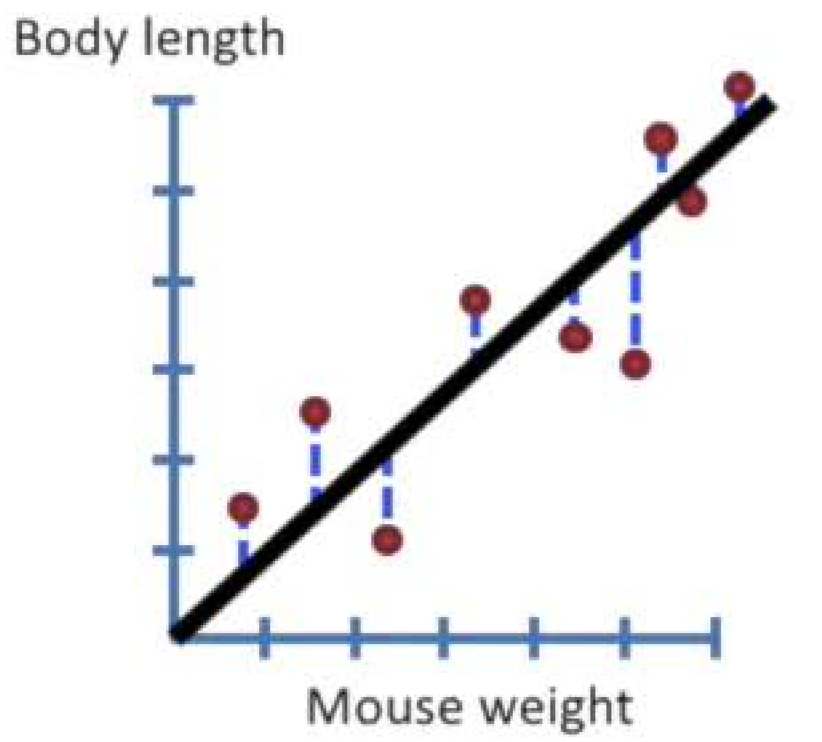
Multiple regression



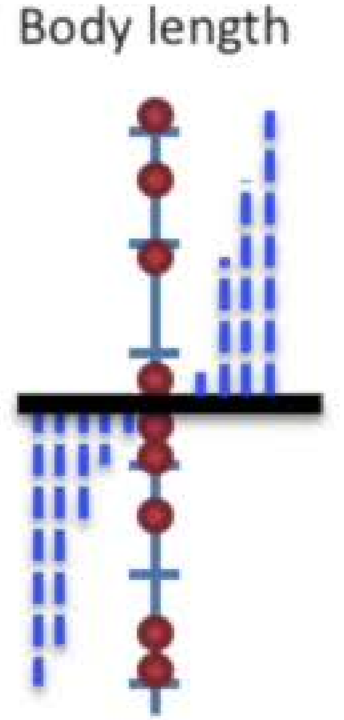
y = y-intercept + slope x + slope z

So far we have compared this simple regression to the mean...

Simple regression

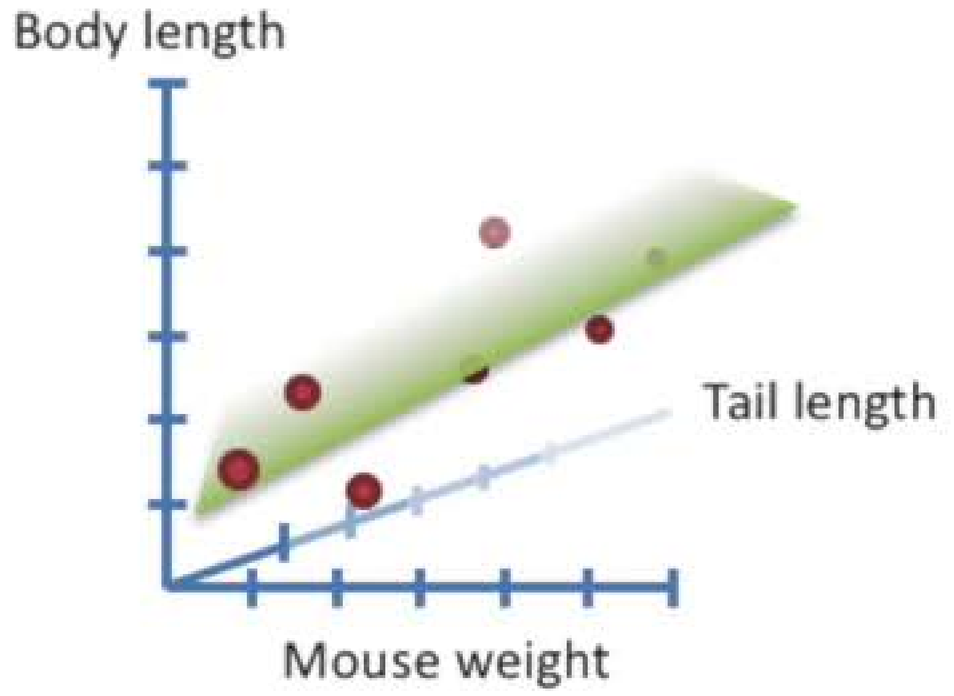


y = y-intercept + slope x



y = y-intercept

Multiple regression

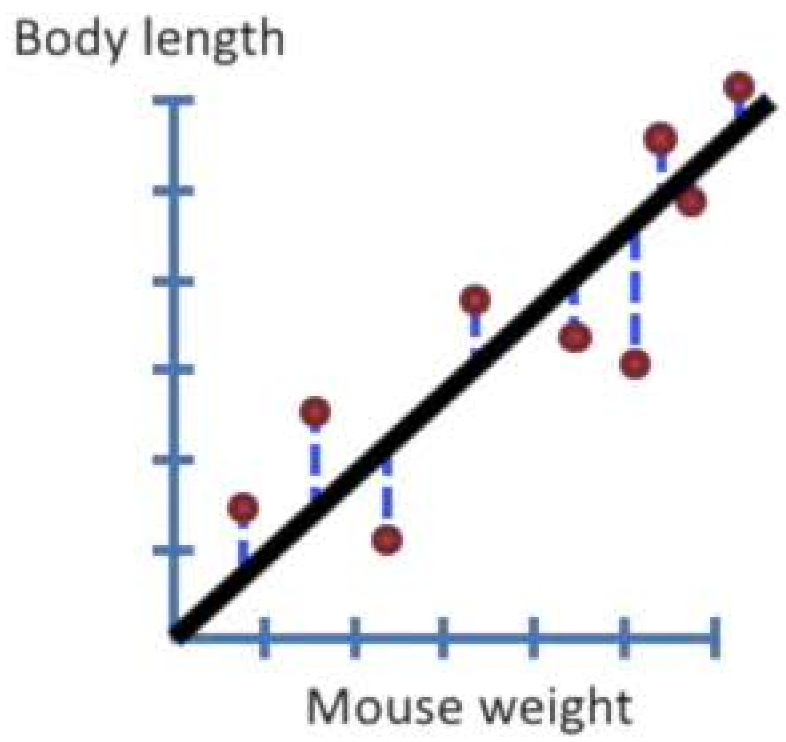


y = y-intercept + slope x + slope z

So far we have compared this simple regression to the mean…

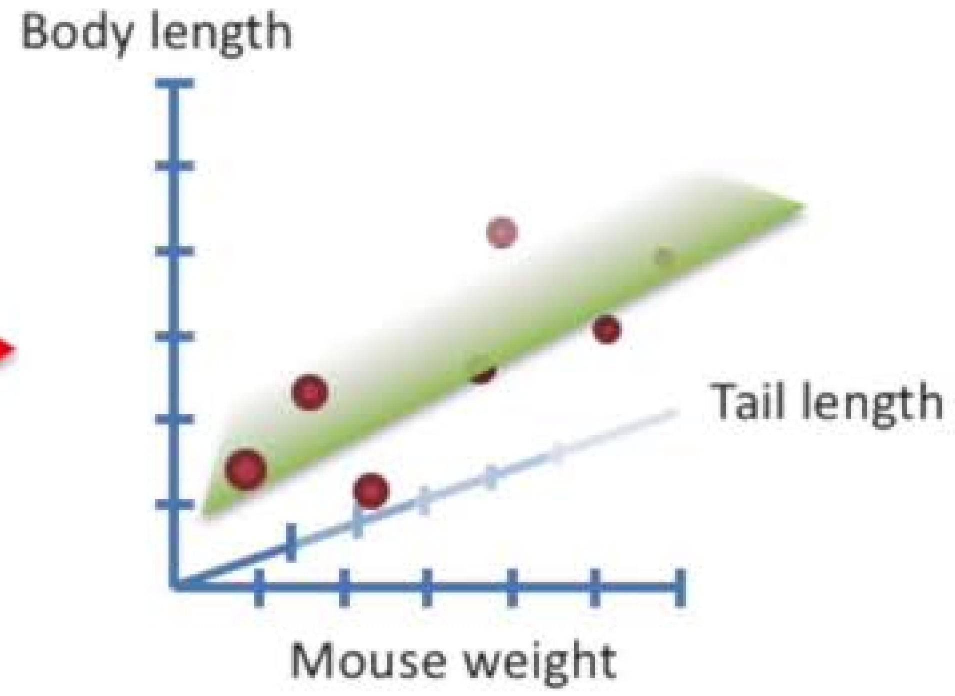
…and this multiple regression to the mean…

Simple regression



y = y-intercept + slope x

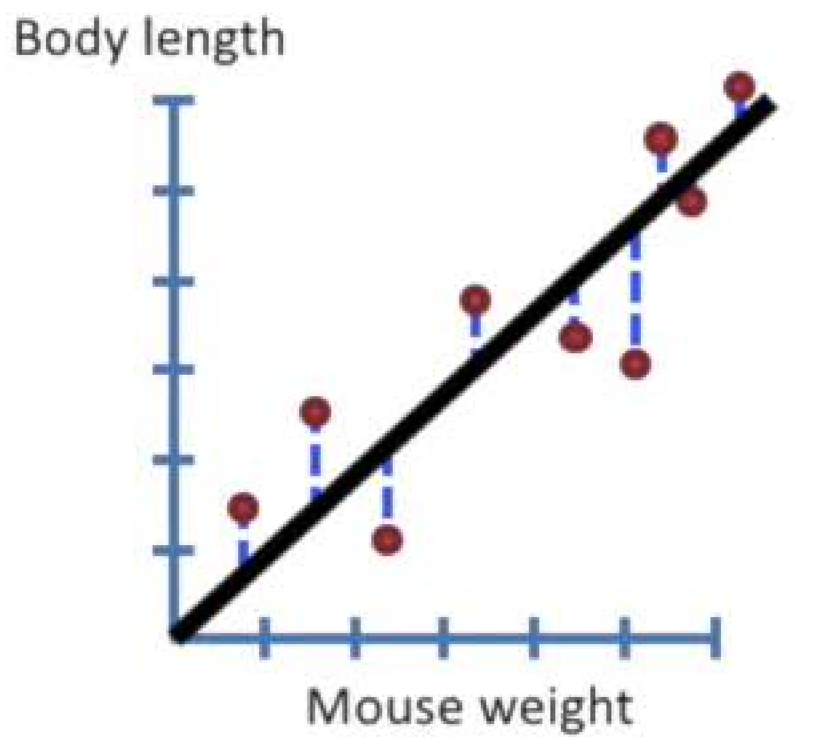
Multiple regression



y = y-intercept + slope x + slope z

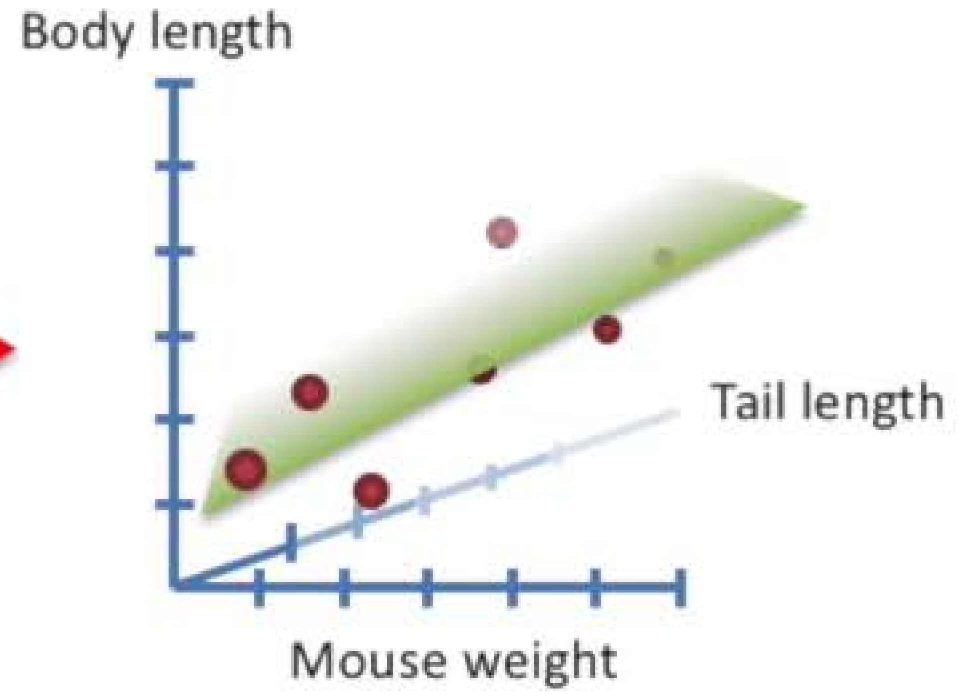
...but we can compare them to each other!

Simple regression



y = y-intercept + slope *x*

Multiple regression

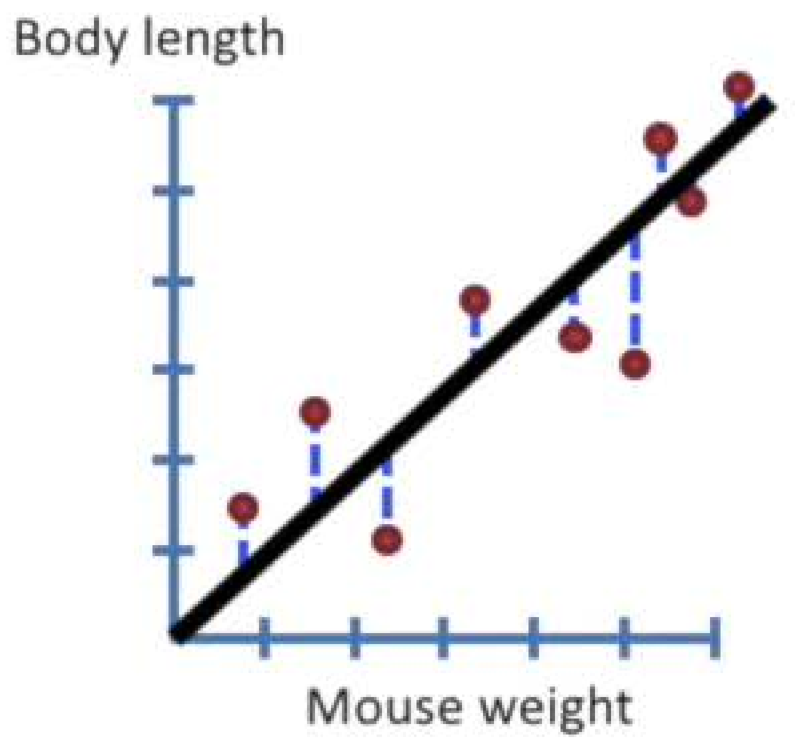


y = y-intercept + slope *x* + slope *z*

...but we can compare them to each other!

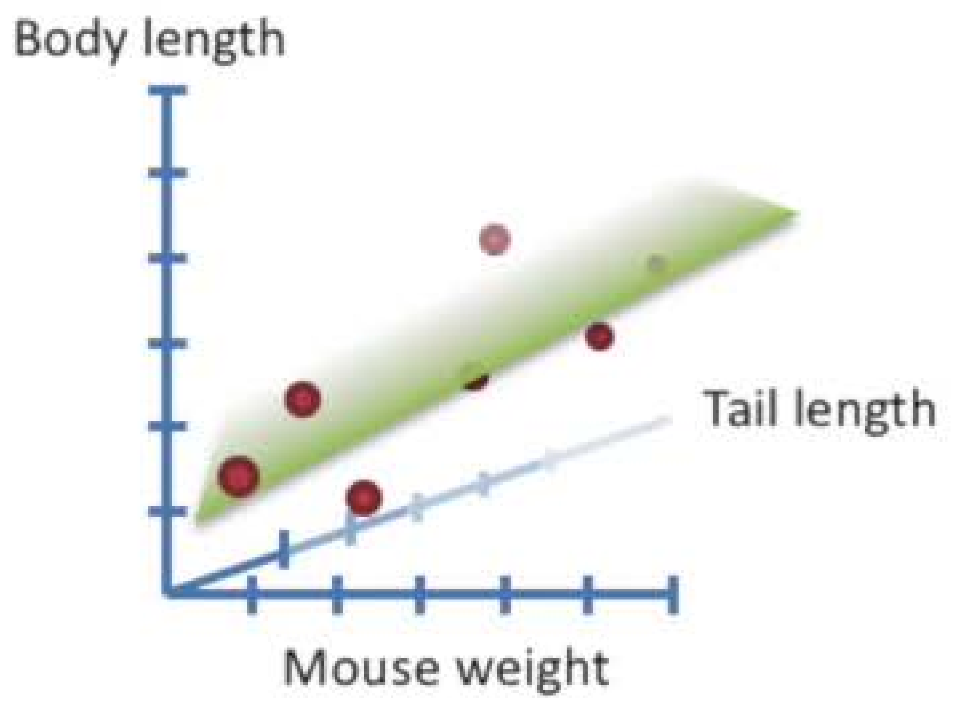
This will tell us if it is worth the time and trouble to collect the Tail Length data because we will compare a fit without it (the simple regression) to a fit with it (the multiple regression).

Simple regression



y = y-intercept + slope x

Multiple regression

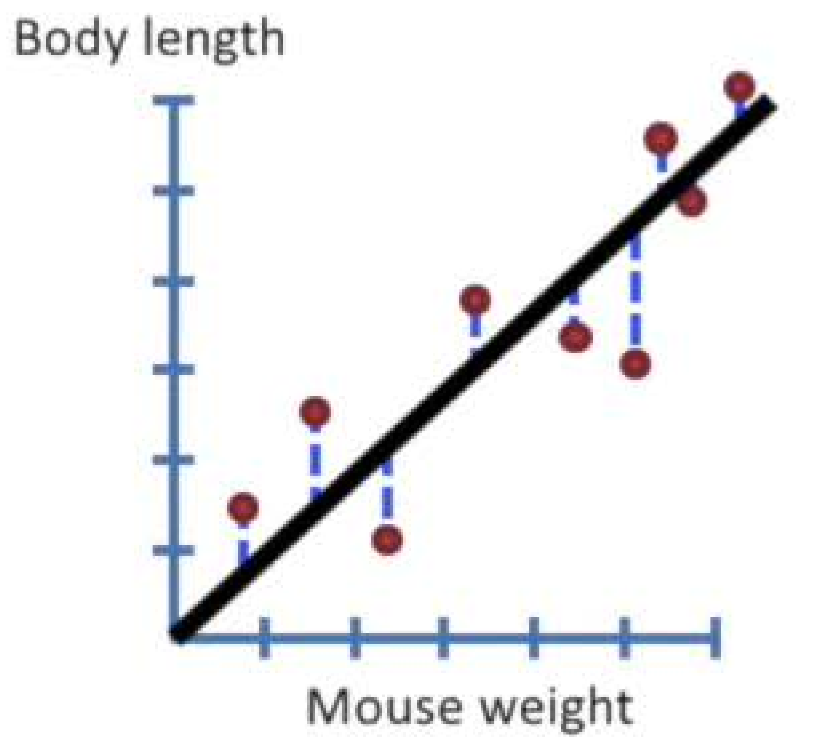


y = y-intercept + slope x + slope z

Calculating the F-value is the exact same as before, only this time we replace the "mean" stuff...

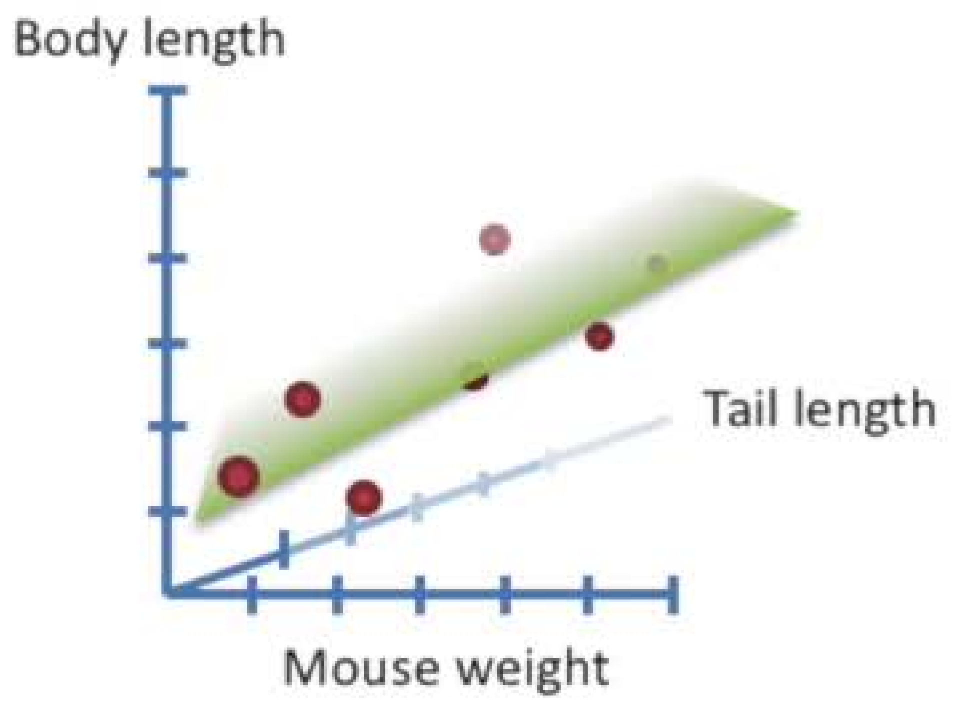
$$F = \frac{SS(\text{mean}) - SS(\text{fit}) / (p_{\text{fit}} - p_{\text{mean}})}{SS(\text{fit}) / (n - p_{\text{fit}})}$$

Simple regression



y = y-intercept + slope x

Multiple regression



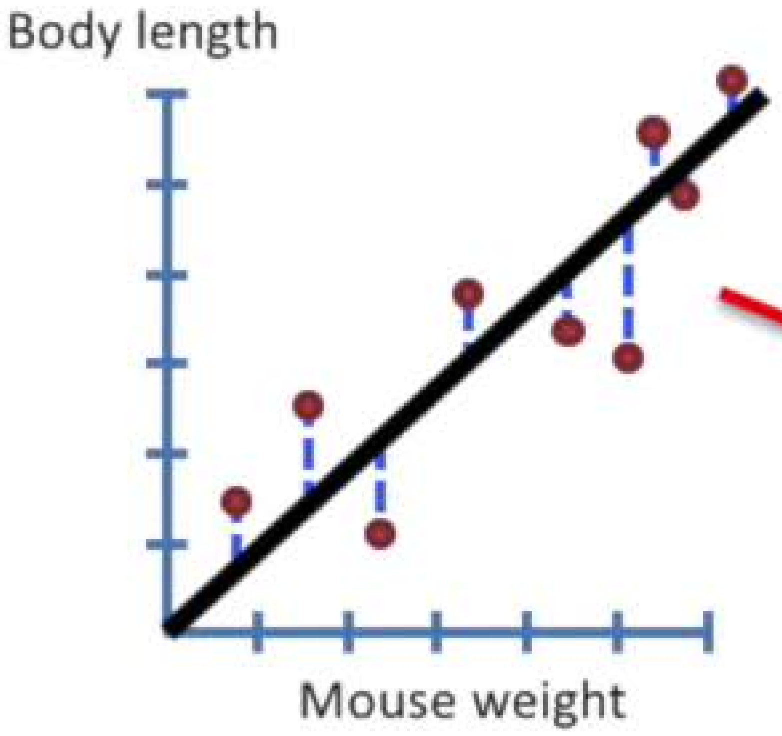
y = y-intercept + slope x + slope z

Calculating the F-value is the exact same as before, only this time we replace the "mean" stuff…

…with with simple regression stuff.

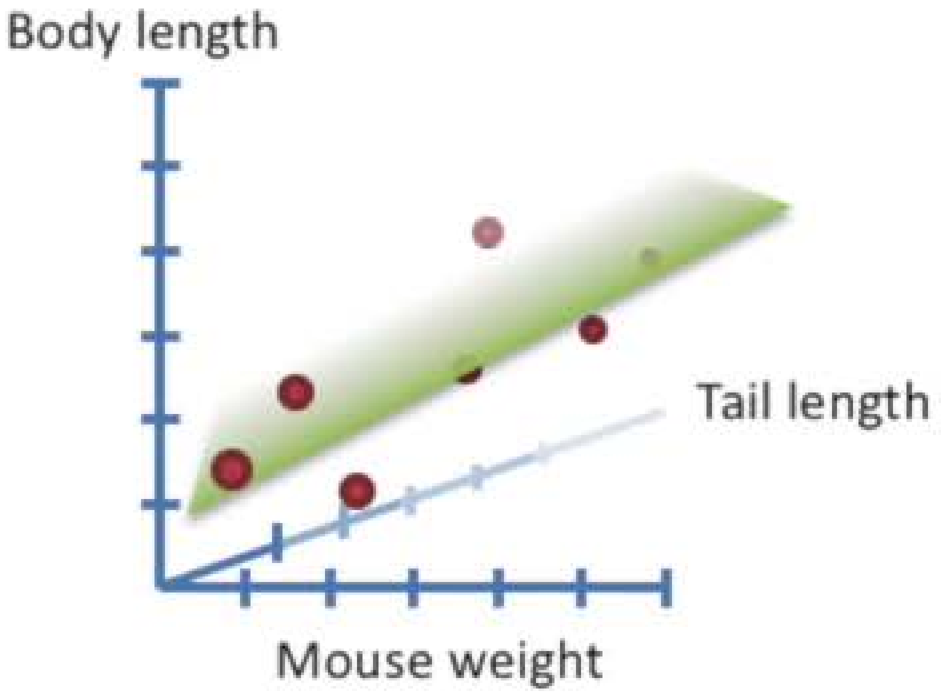
$$F = \frac{\text{SS(simple)} - \text{SS(multiple)} / (p_{\text{multiple}} - p_{\text{simple}})}{\text{SS(multiple)} / (n - p_{\text{multiple}})}$$

Simple regression



y = y-intercept + slope x

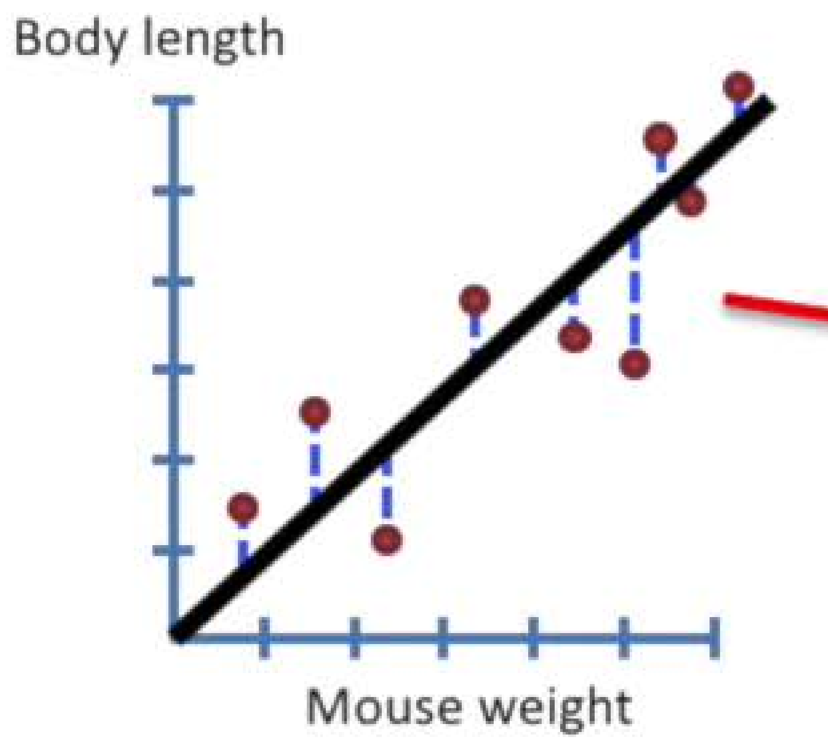
Multiple regression



y = y-intercept + slope x + slope z

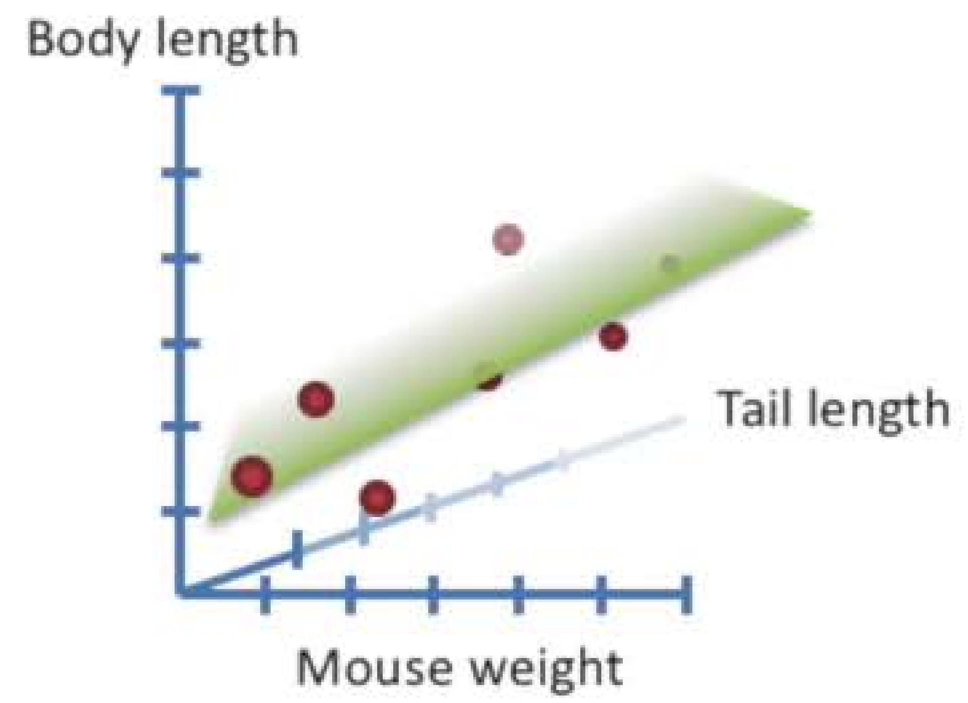
$$F = \frac{\text{SS(simple)} - \text{SS(multiple)} / (p_{\text{multiple}} - p_{\text{simple}})}{\text{SS(multiple)} / (n - p_{\text{multiple}})}$$

Simple regression



y = y-intercept + slope x

Multiple regression

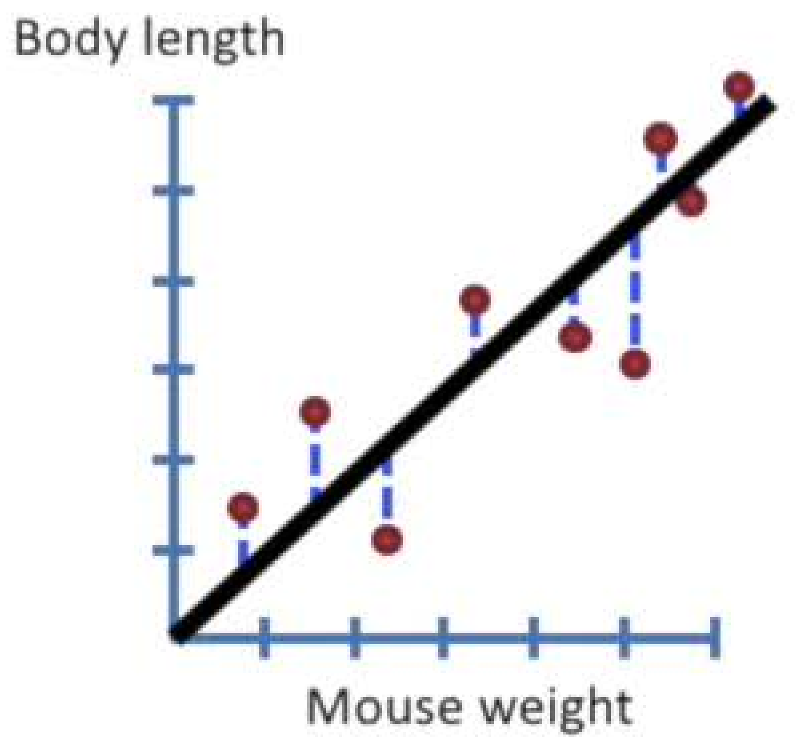


y = y-intercept + slope x + slope z

$p_{\text{simple}} = 2$

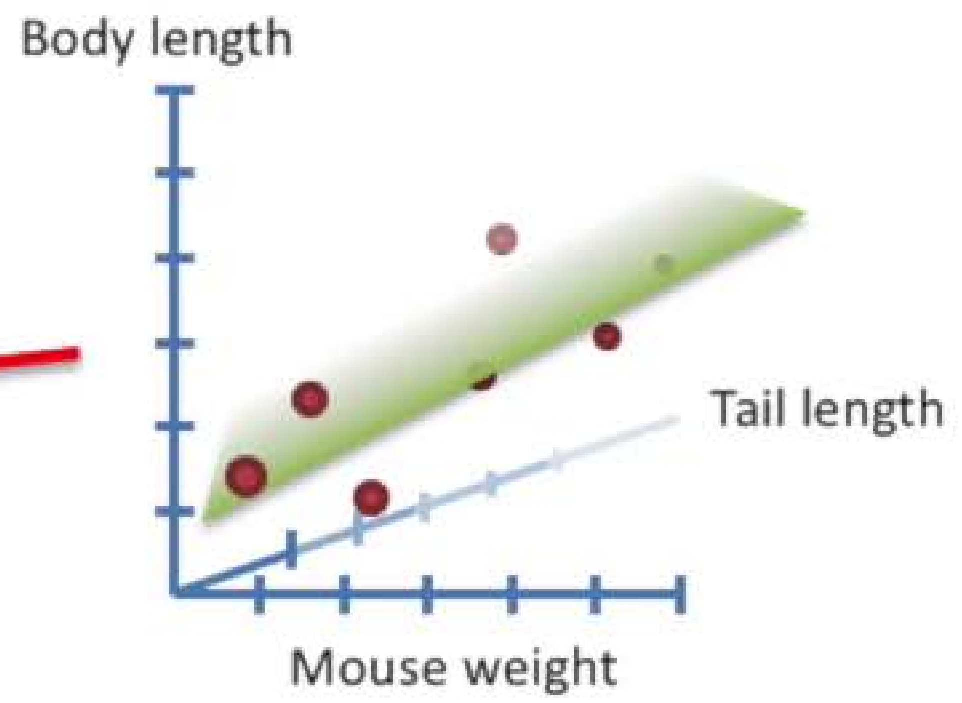
$$F = \frac{\text{SS(simple)} - \text{SS(multiple)} / (p_{\text{multiple}} - p_{\text{simple}})}{\text{SS(multiple)} / (n - p_{\text{multiple}})}$$

Simple regression



y = y-intercept + slope x

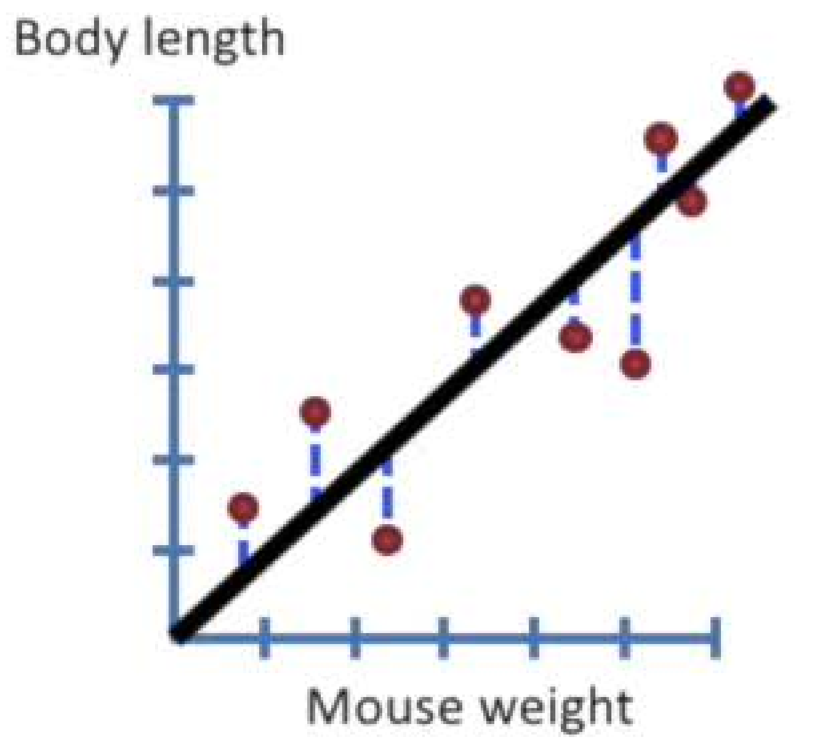
Multiple regression



y = y-intercept + slope x + slope z

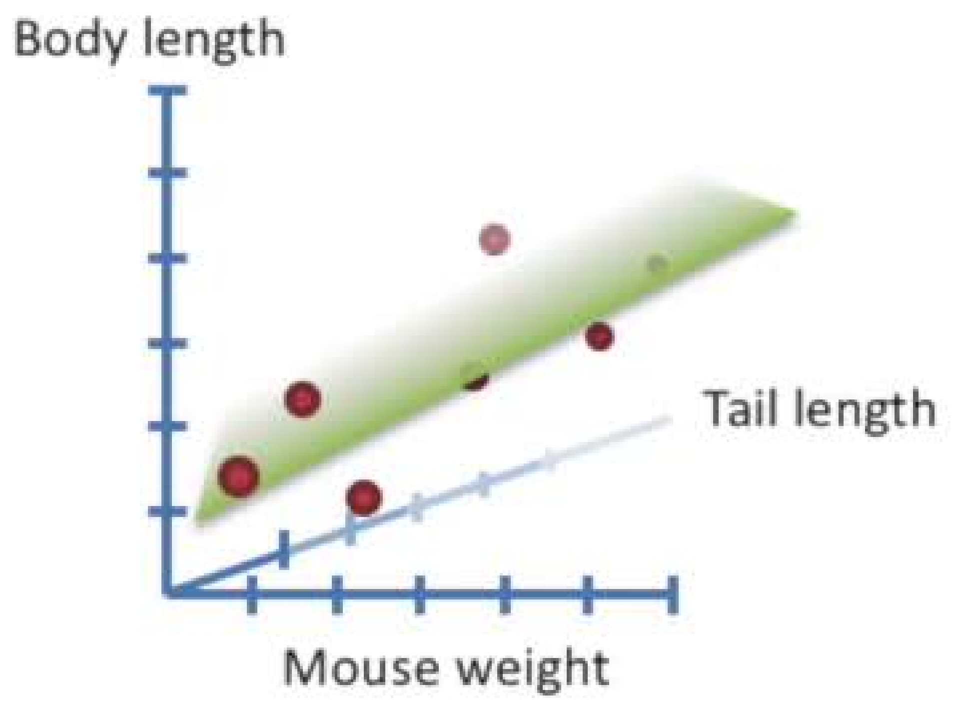
$$F = \frac{\text{SS(simple)} - \text{SS(multiple)} / (p_{\text{multiple}} - p_{\text{simple}})}{\text{SS(multiple)} / (n - p_{\text{multiple}})}$$

Simple regression



y = y-intercept + slope x

Multiple regression

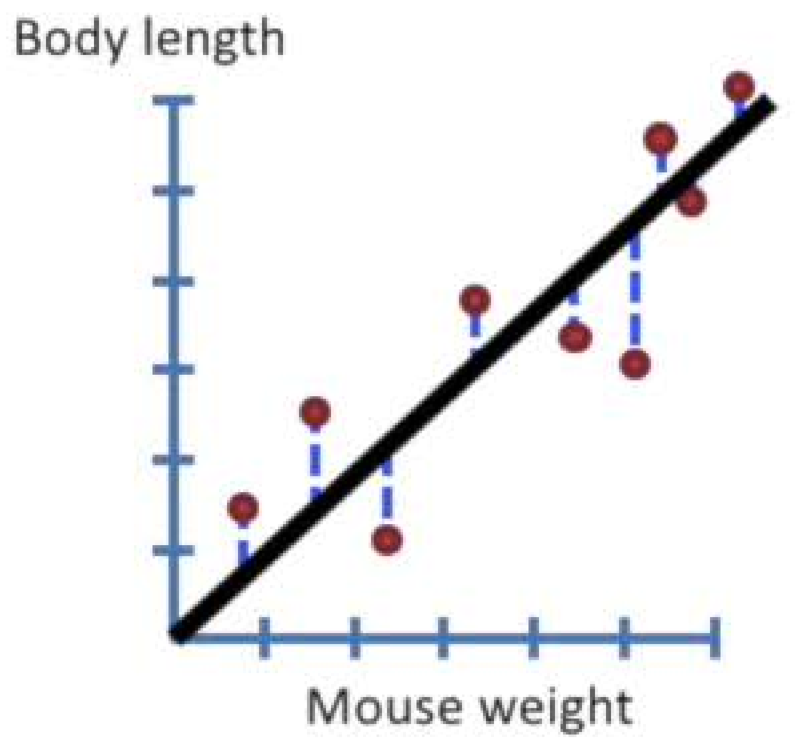


y = y-intercept + slope x + slope z

$$F = \frac{SS(\text{simple}) - SS(\text{multiple}) / (p_{\text{multiple}} - p_{\text{simple}})}{SS(\text{multiple}) / (n - p_{\text{multiple}})}$$

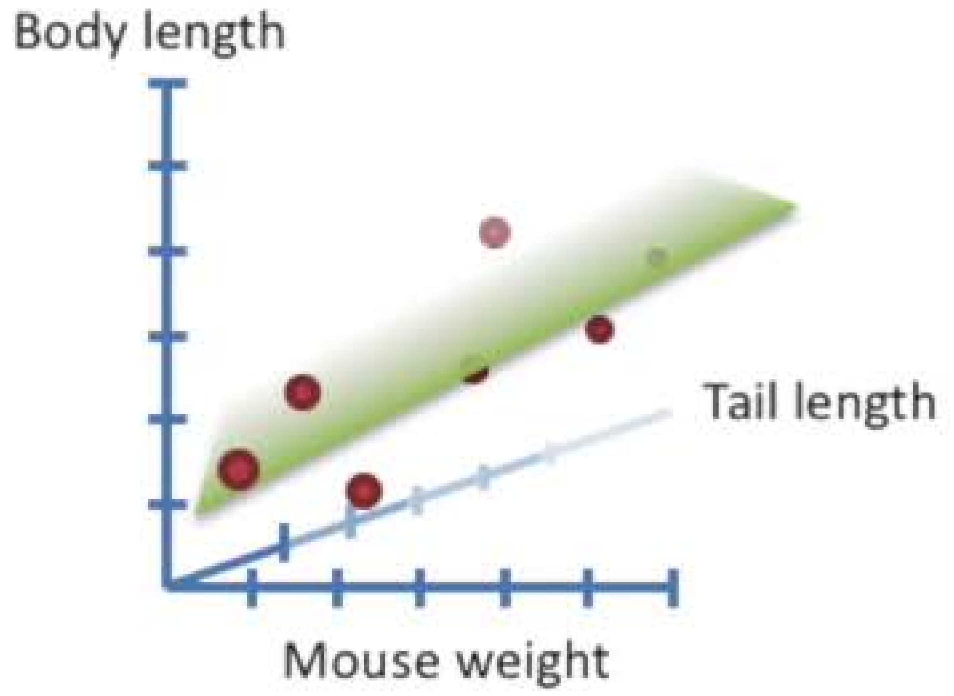
$p_{\text{multiple}} = 3$

Simple regression



y = y-intercept + slope x

Multiple regression



y = y-intercept + slope x + slope z

If the difference in R^2 values between the simple and multiple regressions is "big" and the p-value is "small", then adding Tail Length to the model is worth the trouble.