There is a potential for confusion in examining the ANOVA table from a multiple regression fit. The ANOVA table (usually) *depends* on the order in which the variables are specified, whereas the regression estimates do not.

Consider (yet again) the second case study in Chapter 10, involving bats, birds, and energy expenditure. Compare the differences in the summaries of these two identical models.

```
> fit04 <- lm(log(energy) ~ log(mass) + type)
> fit05 <- lm(log(energy) ~ type + log(mass))
```

Here is a summary of the first fit.

```
> summary(fit04)

Call:
lm(formula = log(energy) ~ log(mass) + type)

Residuals:
     Min       1Q   Median       3Q      Max
-0.23224 -0.12199 -0.03637  0.12574  0.34457

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.47410    0.23902  -6.167 1.35e-05 ***
log(mass)    0.81496    0.04454  18.297 3.76e-12 ***
typeeBat    -0.02360    0.15760  -0.150    0.883
typenBat    -0.10226    0.11418  -0.896    0.384
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.186 on 16 degrees of freedom
Multiple R-Squared: 0.9815,        Adjusted R-squared: 0.9781
F-statistic: 283.6 on 3 and 16 DF,  p-value: 4.464e-14
```

Here is the summary for the second fit.

```
> summary(fit05)

Call:
lm(formula = log(energy) ~ type + log(mass))

Residuals:
     Min       1Q   Median       3Q      Max
-0.23224 -0.12199 -0.03637  0.12574  0.34457

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.47410    0.23902  -6.167 1.35e-05 ***
typeeBat    -0.02360    0.15760  -0.150    0.883
typenBat    -0.10226    0.11418  -0.896    0.384
log(mass)    0.81496    0.04454  18.297 3.76e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.186 on 16 degrees of freedom
Multiple R-Squared: 0.9815,        Adjusted R-squared: 0.9781
F-statistic: 283.6 on 3 and 16 DF,  p-value: 4.464e-14
```

Finally, here are both ANOVA tables.

```
> anova(fit04)

Analysis of Variance Table

Response: log(energy)
          Df  Sum Sq Mean Sq  F value    Pr(>F)
log(mass)  1 29.3919 29.3919 849.9108 2.691e-15 ***
type       2  0.0296  0.0148   0.4276    0.6593
Residuals 16  0.5533  0.0346
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> anova(fit05)

Analysis of Variance Table

Response: log(energy)
          Df  Sum Sq Mean Sq F value    Pr(>F)
type       2 17.8445  8.9222  258.00 6.694e-13 ***
log(mass)  1 11.5770 11.5770  334.77 3.758e-12 ***
Residuals 16  0.5533  0.0346
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Notice that the two summaries have exactly the same estimates, standard errors, t-statistics, and p-values, but differ only in the order in which they are presented. This is not the case for the ANOVA tables — here the test statistics and p-values are different. The ANOVA table presents the results from a series of tests that occur in a *sequential* order. In both ANOVA tables the sum of squares and degrees of freedom associated with the residuals are the same. The total sum of squares for the variables are the same as well but the ANOVA table partitions it differently. Notice that

$$29.3919 + 0.0296 = 17.8445 + 11.5770 = 29.4215$$

If you add the log($mass$) variable to the model with only the intercept, the residual sum of squares (RSS) reduces by 29.3919. After doing this if you add the variable *type*, the RSS decreases by only another 0.0296. The changes to the RSS are different if the variables are added in a different order, but the final RSS is the same for the final model.

The F-statistic is the ratio of two mean squares, or equivalently

$$F = \frac{(\text{Extra-sum-of-squares}) \Big/ (\text{Number of coefficients tested})}{\text{Estimate of } \sigma^2}$$

In the first ANOVA table, the F-statistic for the variable log($mass$) is found by computing the extra-sum-of-squares as the difference in the residual sum of squares between the reduced model with an intercept only and the intermediate model with an intercept and log($mass$), but the denominator uses the estimate of σ^2 from the full model with an intercept, log($mass$), and dummy variables for *type*.

```
> fitA1 <- lm(log(energy) ~ 1)
> fitA2 <- lm(log(energy) ~ log(mass))
> fitA3 <- lm(log(energy) ~ log(mass) + type)
> ssA <- sum(residuals(fitA2)^2) - sum(residuals(fitA1)^2)
> dfA <- df.residual(fitA1) - df.residual(fitA2)
> sigma2A <- ((summary(fitA3))$sigma)^2
> fstatA <- (ssA/dfA)/sigma2A
> ssA

[1] -29.39191

> dfA
```

```
[1] 1

> sigma2A

[1] 0.03458235

> fstatA

[1] -849.9108

> sqrt(fstatA)

[1] NaN
```

Notice that the square root of the F-statistic here is not the absolute value of the t-statistic from the summary of the regression and the p-value is different.

In contrast, the second ANOVA table computes the F statistic as described on page 281 of the textbook, in which the full model has all of the parameters and the reduced model has *type* and an intercept. In this case, the F-statistic is exactly what we want to test if log(*model*) should be included. Notice that the square root of the F-statistic is the absolute value of the t-statistic from the summary and that the p-values are identical.

```
> fitB1 <- lm(log(energy) ~ 1)
> fitB2 <- lm(log(energy) ~ type)
> fitB3 <- lm(log(energy) ~ log(mass) + type)
> ssB <- sum(residuals(fitB2)^2) - sum(residuals(fitB3)^2)
> dfB <- df.residual(fitB2) - df.residual(fitB3)
> sigma2B <- ((summary(fitB3))$sigma)^2
> fstatB <- (ssB/dfB)/sigma2B
> ssB

[1] 11.57700

> dfB

[1] 1

> sigma2B

[1] 0.03458235

> fstatB

[1] 334.7662

> sqrt(fstatB)

[1] 18.29662
```

The first ANOVA is appropriate to carry out an extra-sum-of-squares F-test for inclusion of *type* as a factor. The second ANOVA is appropriate to carry out the same test for log(*mass*).

In short, if you want to use the R function `anova` to carry out a partial F-test for a variable or factor in a multiple regression, it should be the *last* variable listed in the model. Or, you can carry out the models of interest and reconstruct the F statistic by hand from the output. If the variable is quantitative (or if there is only one degree of freedom), the p-value from the t-test in the summary will agree with the p-value from the correct ANOVA table. If the variable is categorical with more than two levels, there is no single p-value from a t-test that corresponds to the global test that all parameters associated with the level have value 0.