

## 12.3: ANOVA Table

All of our sources of variability fit together in meaningful, interpretable ways as we saw above, and the easiest way to do this is to organize them into a table. The ANOVA table, shown in Table 12.3.1, is how we calculate our test statistic.

Table 12.3.1: ANOVA Table

Source	$SS$	$df$	$MS$	$F$
Between	$SS_B$	$k - 1$	$\frac{SS_B}{df_B}$	$\frac{MS_B}{MS_W}$
Within	$SS_W$	$N - k$	$\frac{SS_W}{df_W}$	
Total	$SS_T$	$N - 1$		

The first column of the ANOVA table, labeled “Source”, indicates which of our sources of variability we are using: between groups, within groups, or total. The second column, labeled “SS”, contains our values for the sums of squares that we learned to calculate above. As noted previously, calculating these by hand takes too long, and so the formulas are not presented in Table 12.3.1. However, remember that the Total is the sum of the other two, in case you are only given two  $SS$  values and need to calculate the third.

The next column, labeled “ $df$ ”, is our degrees of freedom. As with the sums of squares, there is a different  $df$  for each group, and the formulas are presented in the table. Notice that the total degrees of freedom,  $N - 1$ , is the same as it was for our regular variance. This matches the  $SS_T$  formulation to again indicate that we are simply taking our familiar variance term and breaking it up into difference sources. Also remember that the capital  $N$  in the  $df$  calculations refers to the overall sample size, not a specific group sample size. Notice that the total row for degrees of freedom, just like for sums of squares, is just the Between and Within rows added together. If you take  $N - k + k - 1$ , then the “ $-k$ ” and “ $+k$ ” portions will cancel out, and you are left with  $N - 1$ . This is a convenient way to quickly check your calculations.

The third column, labeled “ $MS$ ”, is our Mean Squares for each source of variance. A “mean square” is just another way to say variability. Each mean square is calculated by dividing the sum of squares by its corresponding degrees of freedom. Notice that we do this for the Between row and the Within row, but not for the Total row. There are two reasons for this. First, our Total Mean Square would just be the variance in the full dataset (put together the formulas to see this for yourself), so it would not be new information. Second, the Mean Square values for Between and Within would not add up to equal the Mean Square Total because they are divided by different denominators. This is in contrast to the first two columns, where the Total row was both the conceptual total (i.e. the overall variance and degrees of freedom) and the literal total of the other two rows.

The final column in the ANOVA table, labeled “ $F$ ”, is our test statistic for ANOVA. The  $F$  statistic, just like a  $t$ - or  $z$ -statistic, is compared to a critical value to see whether we can reject for fail to reject a null hypothesis. Thus, although the calculations look different for ANOVA, we are still doing the same thing that we did in all of Unit 2. We are simply using a new type of data to test our hypotheses. We will see what these hypotheses look like shortly, but first, we must take a moment to address why we are doing our calculations this way.

This page titled 12.3: ANOVA Table is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Foster et al. (University of Missouri’s Affordable and Open Access Educational Resources Initiative) via source content that was edited to the style and standards of the LibreTexts platform.

- 11.3: ANOVA Table by Foster et al. is licensed CC BY-NC-SA 4.0. Original source: <https://irl.umsl.edu/oer/4>.