determined. The first is dfN[53] and is equal to $k - 1$ where k is the number of groups; dfD[54] equals ($N - k$), where N is the sum of the sample sizes ($N = n_1 + n_2 + n_3 \ldots + n_k$; these sample sizes do not need to be equal for each group). The F-test is always right-tailed.[55] For a summary of key ANOVA concepts see **Table 10**.

Table 10 Analysis of Variance Summary				
Source	**Sum of Squares**	**df**	**Mean Square**	**F**
Between	SS_B	dfN = k-1	MS_B	$\frac{MS_B}{MS_W}$
Within (error)	SS_W	dfD = N-k	MS_W	

SS_B means the Sum of Squares Between groups and is the numerator of the *mean squares between* (MS_B) formula. The denominator of MS_B is dfN. To obtain MS_B and allow for different sample sizes within groups, use the formula $s_B^2 = \frac{\sum n_i(\bar{X}_i - \bar{X}_{GM})^2}{k-1}$. The GM means the Grand Mean and can be identified by $\bar{\bar{X}}$ or $\bar{X}_{GM}$[56]. $\bar{X}_{GM}$ is computed by adding all the x values and dividing by the total number of values. (If the groups are the same size, you can take the means of the groups and divide by the number of groups.)

SS_W is the Sum of Squares Within groups and is the numerator of the *mean squares within* (MS_W) formula. The denominator of MS_W is dfD. To obtain MS_W using the variance, again allowing for different sample sizes within groups, use the formula[57]: $s_W^2 = \frac{\sum(n_i-1)s_i^2}{\sum(n_i-1)}$.

To obtain the F statistic, divide MS_B by MS_W. **Example 10.1** shows the computations relating to these key ANOVA concepts. After identifying the data, state the hypotheses. Determine the critical value using the F Distribution Table (see **Appendix C, Table C-5**); then determine the test value using the ANOVA table and formulas above. State your decision and summarize the findings.

Example 10.1 ANOVA
In a college program at a small private school, a census survey was taken of student preferences for the three teaching methods being used. It was theorized that all three methods were equally preferred by the students. **At α = .05, can it be concluded that the methods were equally preferred? At .01?**

[53] Variously referred to as v_1 or numerator.
[54] Variously referred to as v_2 or denominator.
[55] http://www.youtube.com/watch?v=CxeeqqyUAgE Dr A. G. Picciano
[56] The double bar may be confusing so using $\bar{X}_{GM}$ is often clearer.
[57] $\Sigma(n_i-1)$ is an equivalent form of (N - k), allowing for one df per sample.

Example 10.1 ANOVA

	Face to Face (f2f)	Online	Hybrid	
Freshmen	250	200	164	
Sophomore	220	280	200	
Junior	200	300	180	
Senior	182	240	200	
$\sum x =$	852	1020	744	$\sum\sum x = 2614$

Step 1. **State the hypotheses**. The three teaching methods are equally preferred.
H_0: $\mu_1 = \mu_2 = \mu_3$ H_A: At least one of the means is not equal.

Step 2. **Determine the critical values.** Use **Appendix C, Table C-5**, where v_1 and v_2 are used.
$k = 3$; N = 12. Use dfN = k -1 or 3 -1 = 2; $v_1 = 2$. dfD = N – k = 12 – 3 = 9; $v_2 = 9$.
At α = .05 & .01, $\boldsymbol{F_{CV.05} = 4.26}$ $\boldsymbol{F_{CV.01} = 8.02}$

Step 3. **Compute the test value.**

a. **Determine the means and variances of each group**.

Means: f2f: 852/4 = **213**; **Online:** 1020/4 = **255**; **Hybrid:** 744/4 = **186.**

Variances: $\sigma^2 = \frac{(\sum X_i - \mu)^2}{N}$

	f2f	sq.	Online	sq.	Hybrid	sq.
Freshmen	250-213 = 37	1369	200-255 = -55	3025	164-186 = -22	484
Sophomore	220-213 = 7	49	280-255 = 25	625	200-186 = 14	196
Junior	200-213 = -13	169	300-255 = 45	2025	180-186 = -6	36
Senior	182-213 = -31	961	240-255 = -15	225	200-186 = 14	196
$\sum$ sq. =		2548		5900		912
σ^2 =	2548/4 = 637		5900/4 = 1475		912/4 = 228	

b. **Identify the Grand Mean.** $\bar{X}_{GM} = (852 + 1020 + 744)/12 = \mathbf{218}$

c. **Determine the between-group variance**. $s_B^2 = \frac{\sum n_i(\bar{X}_i - \bar{X}_{GM})^2}{k-1}$

$$s_B^2 = \frac{4(213-218)^2 + 4(255-218)^2 + 4(186-218)^2}{3-1}$$

$$s_B^2 = \frac{4(25)+4(1369)+4(1024)}{2} = \frac{100+5476+4096}{2} = \frac{9672}{2} = \mathbf{4836}$$

d. **Determine the within-group variance.** $s_W^2 = \frac{\sum(n_i-1)s_i^2}{\sum(n_i-1)}$.

$$s_W^2 = \frac{(4-1)(637)+(4-1)(1475)+(4-1)(228)}{(4-1)+(4-1)+(4-1)} = \frac{1911+4425+684}{9} = \frac{7020}{9} = \mathbf{780}$$

e. **Find the test value**. $F = \frac{s_B^2}{s_W^2} = 4836/780 = \mathbf{6.2}$

Step 4. **Make the decision**.
6.2 > 4.26 Fail to accept the null hypothesis at alpha .05
6.2 < 8.02 Accept the null hypothesis at alpha .01

Example 10.1 ANOVA

Step 5. **Summarize the results**. At least one of the means is different at alpha .05. At alpha .01, the means are significantly equal. The numerator in Step 3c between groups is the Sum of Squares for between groups. The numerator in Step 3c for the within groups is the sum of squares for within groups. The Summary table for this is:

Source	Sum of Squares	df	Mean Square	F
Between	9672	2	4836	6.2
Within (error)	7020	9	780	

To review, dividing the Between Sum of Squares by df gives the Mean Square for Between. Doing the same for Within Group gives the Mean Square for Within. Then dividing the Mean Squares gives the F-test value. When finding the means are not equal, a *post hoc*, or after the fact, test should be conducted to determine which means are different. When the null is accepted, no post hoc test is needed.

10.2 Scheffé Test.

To find which means are different after an ANOVA test, a Scheffé test can be conducted. This takes two means at a time to compare so with three means, testing means $\bar{X}_1$ & $\bar{X}_2$, $\bar{X}_1$ & $\bar{X}_3$, and $\bar{X}_2$ & $\bar{X}_3$ are required. When there are more means, many more tests are required. The formula for the test is $F_s = \frac{(\bar{X}_i - \bar{X}_j)^2}{s_W^2\left[{}^1\!/_{n_i} + {}^1\!/_{n_j}\right]}$. The critical value is found by multiplying the critical F-test by k-1. This gives $F' = (k-1)(\mathrm{CV})$. When F_S is greater than F', reject the null hypothesis. **Example 10.2** demonstrates the use of the Scheffé test for the problem in Example 10.1 above.

Example 10.2 Scheffé Test

For $\bar{X}_1$ and $\bar{X}_2$, $F_S = \frac{(\bar{X}_1 - \bar{X}_2)^2}{s_W^2[{}^1\!/_{n_1} + {}^1\!/_{n_2}]} = \frac{(213-255)^2}{780[({}^1\!/_4)+({}^1\!/_4)]} = \frac{1764}{390} =$ **4.52**

For $\bar{X}_1$ and $\bar{X}_3$, $F_S = \frac{(\bar{X}_1 - \bar{X}_3)^2}{s_W^2[{}^1\!/_{n_1} + {}^1\!/_{n_3}]} = \frac{(213-186)^2}{780[({}^1\!/_4)+({}^1\!/_4)]} = \frac{729}{390} =$ **1.87**

For $\bar{X}_2$ and $\bar{X}_3$, $F_S = \frac{(\bar{X}_2 - \bar{X}_3)^2}{s_W^2[{}^1\!/_{n_2} + {}^1\!/_{n_3}]} = \frac{(255-186)^2}{780[({}^1\!/_4)+({}^1\!/_4)]} = \frac{4761}{390} =$ **12.2**

Critical values (found in **Example 10.1**) are $\boldsymbol{F_{CV.05}}$ **= 4.26** and $\boldsymbol{F_{CV.01}}$ **= 8.02**; apply the formula $F' = (k-1)(\mathrm{CV})$. F' = (3-1) 4.26 and F' = (3-1) 8.02. This gives the results as **8.52** and **16.04**, respectively.

For $\bar{X}_2$ and $\bar{X}_3$, F_S equals 12.2, which is the only F_s that exceeds F' at the .05 level of confidence. Thus, there is a difference between the Online and