PIERIAN DATA

EXERCISE #5 - ANOVA – ANSWER KEY

1. What does ANOVA stand for? How is it used?

   ANOVA stands for Analysis of Variance. It is used to determine the probability that two samples come from populations that have the same variance.
   It can also help determine the probability that three or more samples have come from the same population.

2. What is the name of the probability distribution used with ANOVA?

   The F-Distribution

3. We already have a test to determine the probability that two samples come from the same population (the Students t-Test). Why would we use ANOVA?

   For each additional two-sample test, the level of confidence decreases.
   Testing three samples involves three unique comparison tests (AB, AC, BC).
   If each test has a confidence level of 0.95, then the resulting confidence level is (0.95)(0.95)(0.95) or 0.857.

ONE-WAY ANOVA

4. A real estate firm wants to know if there is a statistical difference in sales between the months of February, March and April. The following table shows sales figures from three divisions for those months. Using ANOVA, see if there is support for the null hypothesis that all three months were equal.

	Feb	Mar	Apr
	13	21	17
	16	19	19
	16	20	21
Sum:	45	60	57
Mean:	15	20	19
		Overall Mean:	18

Sum of Squares Groups (SSG)		
9	4	1
	Sum:	14
	Sum x3 rows:	42
	df Groups:	2

Sum of Squares Error (SSE)		
4	1	4
1	1	0
1	0	4
	Sum:	16
	df Error:	6

$$MSG = \frac{SSG}{df_{Groups}} = \frac{42}{2} = \mathbf{21}$$

$$F = \frac{MSG}{MSE} = \frac{21}{2.67} = \mathbf{7.875}$$

$$MSE = \frac{SSE}{df_{Error}} = \frac{16}{6} = \mathbf{2.67}$$

$$F_{critical}(0.05,2,6) = \mathbf{5.143}$$

Based on the results, what is our conclusion?

Since our calculated F value falls to the right of $F_{critical}$, we would reject the null hypothesis that sales were equal in February, March and April.