Name:_____
myID: _____

**Assignment 3**
**Part I**

1. Assume that you are comparing the GPAs of students from 3 departments, Engineering, Mathematics, and Physics. If the Mean GPA in Engineering = 3.37, and the F-ratio of the appropriate ANOVA statistic = 1.00, then what is the Mean GPA of:
    a. Students in Mathematics?
    b. Students in Physics?

    Since the F-ratio is 1, the means of all three groups are the same, which is 3.37.

§ Read the following and answer questions 2-3:

The Dean of the College of Education has asked you to assist with her research. She is comparing traditional teaching techniques for reading comprehension to two other methods. To this end, she exposed 3 groups of children (10 per group) to each of 3 teaching techniques, and measured their reading skills after a set period of time. You analyzed the data with the appropriate ANOVA test, and were on your way over to her office with the SPSS printout, when you spilled coffee all over your printout. The accident made your ANOVA summary table unreadable, except for only a few key items. Below is the result of this unfortunate waste of Starbucks. You could always delay your meeting with the Dean and re-run the analysis, but knowing how the Dean HATES missed meetings, reconstruct the ANOVA table with your knowledge of the components of the table and state the null and alternative hypothesis.

2. In symbols and words, what are the null and alternative hypotheses?
    The null hypothesis is $\mu_1 = \mu_2 = \mu_3$
    The alternative hypothesis is not all μ's are equal.
3. Complete the following ANOVA table:

| | SS | df | MS | F | P (α=.05) |
|---|---|---|---|---|---|
| Between Groups | 24 | 2 | 12 | 6 | 0.005 |
| Within Groups | 54 | 27 | 2 | | |
| Total | 78 | 29 | | | |

$df_{between\text{-}group}=3-1=2$
$MS_{between}=SS_{between}/df_{between}=24/2=12$
$df_{total}=3*10-1=29$, so $df_{within}= df_{total}- df_{between}=29-2=27$
$SS_{within}=27*2=54$
$SS_{total}=SS_{within}+SS_{between}=54+24=78$
$F(2,27)_{\alpha=0.05}=3.35$
$F= MS_{between}/MS_{within}=12/2=6$

4. What is your decision regarding the hypotheses (please interpret your conclusion)?
    The obtained F value (6) is bigger than the critical F value at α=.05. So we reject the null hypothesis and conclude that there are significant differences in the group means of children's reading skills.

5. (Exercise 3.1 in Keppel & Wickens). Find the critical values of F for the following situations:
   a. F(4, 30) at α=.05;
   b. F(1, 120) at α=.001;
   c. A=7, n=5, α=.10;
   d. A=3, n=9, α=.01.

   The critical values of F can be found in the F table in Appendix A.1. A particular value of F in this table is specified by: (1) the numerator degrees of freedom, that is, $df_{num}$=a-1; (2) the denominator degrees of freedom, that is, $df_{denom}$=a(n-1); (3) the proportion of area to the right of an ordinate, that is, α value.
   a. From the F table, we found F(4,30)=2.69 at α=.05;
   b. F(1,120)=11.38 at α=.001;
   c. $df_{num}$=7-1=6, $df_{denom}$=7*(5-1)=28, F(6,28)=2.00 at α=.10;
   d. $df_{num}$=3-1=2, $df_{denom}$=3*(9-1)=24, F(2,24)=5.61 at α=.01.

6. Using Excel (or some other statistical package) compute the following exact p-values:
   a. F = 10.45, $df_A$ = 4, $df_{S/A}$ = 30
   b. F = 3.00, $df_A$ = 1, $df_{S/A}$ = 120
   c. F = 5.26, $df_A$ = 6, $df_{S/A}$ = 29
   d. F = 3.00 $df_A$ = 2, $df_{S/A}$ = 25

   a.F = 10.45, $df_A$ = 4, $df_{S/A}$ = 30 ,(=fdist(10.45,4,30)); p=0.00002
   b. F = 3.00, $df_A$ = 1, $df_{S/A}$ = 120, p=0.0516; (=fdist(3.00,1,120)); p=0.086
   c. F = 5.26, $df_A$ = 6, $df_{S/A}$ = 29, p=0.00015; (=fdist(5.26,6,29)); p=0.0009
   d.F = 3.00 $df_A$ = 2, $df_{S/A}$ = 25, p=0.054803; (=fdist(3.00,2,25)); p=0.068

7. (Exercise 3.4 in Keppel & Wickens). Occasionally you will want to reconstruct parts of someone else's analysis, either to verify some of the calculations or to conduct additional tests. This is relatively simple if you have available the treatment means and some measures of variability such as the group standard deviations. Suppose you wanted to compare the results of the experiment reported in Table 3.5 (p. 51) with another study that included only the first three conditions (4 hr., 12 hr., and 20 hr.). Perform an analysis of variance on these three groups, using only the means and standard deviations. Complete the entries in the following ANOVA table:

| | SS | df | MS | F | P (α=.05) |
|---|---|---|---|---|---|
| Between Groups | 1970.168 | 2 | 985.184 | 7.193 | 0.005237 |
| Within Groups | 1232.748 | 9 | 136.972 | | |
| Total | 3202.916 | 11 | | | |

Hours without sleep

| 4 hr. | 12 hr. | 20 hr. |
|---|---|---|
| $a_1$ | $a_2$ | $a_3$ |
| 37 | 36 | 43 |
| 22 | 45 | 75 |
| 22 | 47 | 66 |
| 25 | 23 | 46 |
| Mean ($\overline{Y_j}$) 26.50 | 37.75 | 57.50 |
| Variance ($S_j^2$) 51.000 | 119.583 | 240.333 |
| Standard Deviation ($S_j$) 7.141 | 10.935 | 15.503 |

To perform an analysis of variance on these three groups, we need to know $MS_A$, $MS_{S/A}$, then compare the ratio between $MS_A$ and $MS_{S/A}$ to get the F value. Finally we made a decision on the hypothesis testing by comparing the obtained F value with the critical F value in the F table.

$$F = \frac{MS_A}{MS_{S/A}} = \frac{SS_A / df_A}{MS_{S/A}} \text{ where } SS_A = \sum n(\overline{Y_j} - \overline{Y_T})^2 \text{ and } MS_{S/A} = \frac{\sum S_j^2}{a}$$

For equal sample size, $SS_A = \sum n(\overline{Y_j} - \overline{Y_T})^2 = n\sum(\overline{Y_j} - \overline{Y_T})^2$

Put the values from Table 3.5 into the above equations:

$\overline{Y_T}$ =(26.50+37.75+57.50)/3=40.583

$$SS_A = \sum n(\overline{Y_j} - \overline{Y_T})^2 = n\sum (\overline{Y_j} - \overline{Y_T})^2$$

$$= 4*[(26.50-40.583)^2 + (37.75-40.583)^2 + (57.50-40.583)^2]$$

$$== 4*(198.331 + 8.026 + 286.185)$$

$$= 1970.168$$

$$MS_A = \frac{SS_A}{df_A} = \frac{1970.168}{2} = 985.184$$

$$MS_{S/A} = \frac{\sum S_j^2}{a} = \frac{51.000 + 119.583 + 240.333}{3} = 136.972$$

$$F = \frac{MS_A}{MS_{S/A}} = \frac{985.184}{136.972} = 7.193$$

The critical F(2,9) at α=.05 is 4.26, and 8.02 at α=.01.

The obtained F value is 4.26<7.193. So we rejected the null hypothesis at α=.05 and conclude that there are differences in the means of the three groups. The obtained F value is 7.193<8.02, so we failed to reject the null hypothesis at α=.01 and conclude that there are not significant differences in the group means.

## Part II (SPSS)

A researcher believes that the type of assignments given to students makes a difference in how well the student knows the material. She decides to randomly select 10 students to only have assignments that are project-based, 10 students to only have assignments that require pure memorization, and 10 students to only have assignments in the form of essay writing (all for a psychology class). After a semester of each group being taught with assignments only allowed for their group, a final exam is administered to score how well the students know the material.

The data is below.

| Project-Based | Pure Memorization | Essay Writing |
|---|---|---|
| 96 | 82 | 92 |
| 90 | 89 | 93 |
| 92 | 91 | 94 |
| 88 | 88 | 96 |
| 98 | 90 | 86 |
| 92 | 93 | 91 |
| 93 | 82 | 96 |
| 97 | 85 | 87 |
| 90 | 81 | 91 |
| 85 | 83 | 85 |

1. Write the null hypothesis in symbols and words.

   The null hypothesis is that the group means for each of the three groups are equivalent.
   $H_0: \mu_1 = \mu_2 = \mu_3$

2. How many degrees of freedom are there for both the between group sums of squares and the within group sums of squares?

   There are (A-1)= (3-1) = 2 degrees of freedom for the between group sums of squares. There are (N-1) = (30-1) = 29 degrees of freedom for the total sums of squares. The within group sums of squares equals the difference of these two values: (29-2) = 27.

3. Run in SPSS a one-way analysis of variance. What is your decision regarding your hypothesis test? Provide and explain the evidence you have to support your decision.

   We would reject the null hypothesis that the group means are equal at the $\alpha$=0.05 significance level ($F(2,27)=5.44$, $p=0.01$).
   The probability of observing an F statistic this high when the null is true is 0.01. This p-value is less than the alpha value so we reject the null hypothesis.

4. Looking at the ANOVA table, what numbers would you use to calculate the F value?

F = (Between group mean sums of squares)/(Within group mean sums of squares) = 92.63/17.04 = 5.44.

**ANOVA**

score

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 185.267 | 2 | 92.633 | 5.435 | .010 |
| Within Groups | 460.200 | 27 | 17.044 | | |
| Total | 645.467 | 29 | | | |