Pairwise Comparison Exercises (in lieu of those of Section 8.2)

Answer the questions below related to pairwise comparisons. When the exercise is along the same lines as one in the textbook, the exercise number will be given. Nevertheless, answer the questions as written on *this sheet* using the information provided *here*; do not do the exercise as it is written in the textbook.

1. (This one mirrors Exercise 8.46.) The **StudentSurvey** data set contains categorical variable *Award* and quantitative variable *Pulse*.

	Award	mean	sd	n
1	Academy	70.51613	12.35818	31
2	Nobel	72.21477	13.09093	149
3	Olympic	67.25275	10.97067	182

One can carry out 1-way ANOVA on these variables.

(a) What groups/populations are being compared? That is, what are the values of the explanatory variable?

(b) State null and alternative hypotheses for the 1-way ANOVA test.

(c) The ANOVA table appears below. Use an F-distribution to obtain a P-value corresponding to the test statistic. Why is it valid to obtain the P-value this way?

```
Analysis of Variance Table

Response: Pulse
           Df Sum Sq Mean Sq F value
Award       2   2047 1023.62  7.1039
Residuals 359  51729  144.09
```

(d) State your conclusion, in conte

(e) When the F-statistic of 1-way ANOVA is significant, we do pairwise comparisons. Below are 95% confidence intervals for the difference in (population) group means, constructed using the TukeyHSD approach. Use the computer output to determine which pairs of means have a difference that is significant at the 5% level.

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

$Award
                       diff        lwr        upr
Nobel-Academy      1.698636 [-3.878263   7.275535]
Olympic-Academy   -3.263382 [-8.752515   2.225752]
Olympic-Nobel     -4.962018  -8.083176  -1.840860
```

2. (This one mirrors Exercise 8.52, and extends the work done in Exercise 8.32.) The 1-way ANOVA test in Exercise 8.32 leads to a significant result, one that leads to rejection of the null hypothesis. So, which group means are different? Use the TukeyHSD pairwise comparison output below to determine which difference in means are significant at the (adjusted) 5% level.

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

$Calcium
                  diff         lwr        upr
Low-High     10.333333  [1.219540  19.4471264] ← 0 not inside
Medium-High   0.500000   -8.613793   9.6137931 ← 0 inside
Medium-Low   -9.833333 [-18.947126  -0.7195402] ← 0 not inside
```

Questions:

- How many values does our categorical var. take? Ans. 3, Low/Med./High
- Do you know any nos. in the ANOVA table?

	df	SS	MS	F
Groups	2			
Error				
Total				

- If we assume the F-stat. was significant (led to rejection of H_0), then doing the TukeyHSD is a sensible follow-up step. What is its aim? What conclusion does it suggest?

Conclude: $\mu_{Low} \neq \mu_{High}$

$\mu_{Medium} \neq \mu_{Low}$

t-score (standardized) for r

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{p. 529}$$

$$= \frac{0.944\sqrt{6-4}}{\sqrt{1-(0.944)^2}} = \frac{2(0.944)}{\sqrt{1-0.944^2}} = 5.722$$