Statistics 201 Homework Assignment for Mon, November 24th, due Wed, Dec 4th

There is a data set linked to the class webpage, called Student0405.txt (tab separated) or student0405.txt (comma separated), which contains data gathered in statistics classes in 2004 and 2005. The data set includes labels in the first row.

The two variables you will use are:
PartyDays = students' response when asked how many days a month they go to parties
ReligImp = response to the question "How important is religion in your life: Not, Fairly, Very?"

1. Create side by side boxplots comparing the PartyDays for the three Religious Importance groups.

2. Find and report the sample means, standard deviations and sample sizes for the three Religious Importance groups.

3. Using the boxplots in question 1 and the standard deviations in question 2, comment on whether or not it is appropriate to conduct an analysis of variance.

4. Write the population model and the null and alternative hypotheses for analysis of variance using both versions, as follows.
    a. Write the version of the model using group means. Define the parameters. Specify the conditions that accompany the model.
    b. Write the "factor effects" version of the model.
    c. Give the null and alternative hypotheses using the group means version, and then using the "factor effects" version.

5. Conduct a one-way analysis of variance to compare mean party days for the three religious importance groups, as follows:
    a. State null and alternative hypotheses in symbols and words *in the context of this situation*. Define the parameters used in your hypotheses.
    b. Show the ANOVA table produced by R.
    c. Give the test statistic and p-value for the test.
    d. State a conclusion in statistical terms *and* in the context of this situation.

6. Use the Tukey multiple comparisons procedure to determine which population means differ (if any). Show both numerical and graphical results. Write a summary stating your conclusions.