Lab Assignment No. 4: Multivariate Analysis of Variance
PSGA7816 Multivariate Statistical Analysis

In completing your assignment you are expected to be specific and analytical in your responses. Simply producing tables or output is **not sufficient** and if not accompanied by any explanation or discussion will not be given **any** credit. Please append the full syntax you used to compute your answers. Your typed assignment answers should include the products of your computations (tables or charts) pasted into the body of your paper with appropriate remarks and comments.

No e-mail submissions will be accepted. Late assignments will be penalized 2 points per day. The dataset for the assignment (crudeoil.xls, which reports chemical levels in crude oil that is siphoned off at certain zones) will be e-mailed to students. Homework 4 is due **March 12, 2009 at 1:30 PM**.

1) What are the bivariate correlations for your data? What interpretations in the multivariate environment can you draw?

2) Provide the appropriate (most informative for further multivariate analyses) univariate measures of central tendency.

3) What is the conceptual difference between a Hotelling's T^2 and a Multivariate Analysis of Variance (MANOVA)?

4) Assess whether your data is multivariate normally distributed and meets basic assumptions (Mardia's coefficients, Henze-Zirkler T, χ^2 Q-Q plot, and Bartlett's test).

5) What hypotheses will your one-way MANOVA be testing?

6) Run a one-way MANOVA. Report and discuss Wilk's Lambda and Roy's Maximum (or Largest) Root.

7) What specific conclusions can you draw from the MANOVA results and the univariate ANOVA follow-up tests?

8) Using the simultaneous confidence intervals for the least square means what further conclusions can you make about the specific mean differences of the individual chemicals at each zone?

9) Look at the distribution of residuals (by themselves and grouped by zone). Discuss their corresponding distributions.

10) What would be your overall final assessment of the data? What overarching conclusions would you make regarding the siphoning of crude oil and the zone's from which it is taken?

Extra Credit: *Verify Wilk's Lambda using either SPSS, SAS or Excel*