**SAS Homework 2**

**Note: This homework is to be done in groups**.

Income and Expenditure Data. 100 Cross Section Observations

Source: Greene (1992)

MDR = Number of Derogatory Reports

Acc = Credit card application accepted (1=yes),

Age = Age in years+ 12ths of a year,

Income = Income, divided by 10,000 ,

Avgexp = Avg. monthly credit card expenditure,

Ownrent = OwnRent, individual owns (1) or rents (0) home.

Selfempl = Self employed (1=yes, 0=no)

We are interested in modeling “Avgexp” as a function of the other variables.

1. Find the basic statistics (means and frequencies) for each of the variables. Comment on what you learned.
2. Find the correlations between the independent variables. What did you learn? Which variables had the highest correlations?
3. Find the best regression model (use PROC REG) to explain average monthly credit card expenditure. Comment on all model results including (F-test, R-sq, Adjusted R-sq, coefficients, t-values and explain in plain English the meaning of the coefficients).
4. Which variable is the most important in explaining “AvgExp”? How did you find this? Rank all explanatory variables in terms of importance.
5. Test for the presence of nonlinearity effect for one X variable. Comment on what you find.
6. Test for one interaction effect between any two explanatory variables. Comment on what you find and the meaning of the interaction coefficient.
7. Perform diagnostics for multicollinearity (VIF, COLLIN) as discussed in class. Comment on your findings whether there is multicollinearity in the model.
8. Test for heteroscedasticity using White test (use PROC MODEL code). Estimate a weighted least squares model on this data if you find evidence of heteroscedasticity. Comment on how the WLS results are different compared to the OLS regression results.
9. Send your SAS program and output files by email. Remember to put your group number on each document.