1. (13 pts) (a) Run the following two regressions and give the ANOVA table for each
(no other results are needed):

   (i) predict satisfaction using age and anxiety.

   (ii) predict satisfaction using age, severity, and anxiety.

Take the difference between the SSEs for the two analyses and construct the F statistic
(general linear test statistic – see Lecture 13) for testing the null hypothesis that the
coefficient of severity is zero, in the model with all three predictors. What are the
degrees of freedom for this test statistic? (You do not need to conduct the test, just give
the test-statistic and df).

(b) Use the TEST statement in PROC REG to obtain the same test statistic. Give the
statistic, degrees of freedom, P-value, and conclusion. Name one other way you could
have conducted this test using SAS output.

(c) If you were to test the null hypothesis that the coefficients for both severity and
anxiety are zero, in the model with all three predictors, what would the degrees of
freedom be? (You do not need to conduct the test, just give the df).

2. (10 pts) In the regression which predicts satisfaction using age, severity, and anxiety
use the SS1, SS2, PCORR1, PCORR2, and VIF options after the MODEL statement
in PROC REG (with variables in this order).

(a) Using the results from the variance inflation factors, does multicollinearity seem to be
a problem for this model?

(b) Add up the Type I sums of squares for the three predictors (do NOT include the
intercept). Do the same for the Type II sum of squares (same as Type III here). Do
either of these sum to the model (regression) sum of squares? Are there any
predictors for which Type I and Type II sums of squares is the same? Explain why.

(c) Use the results from the partial coefficient of determination to describe which
variable contributes the most to the reduction in variation in satisfaction remaining
after the other two variables are included in the model. Which of the types (I or II)
are you using to determine this? Do either of the two types add up to the R^2 for the
full model?
3. (12 pts) (a) Use PROC GLM (test for Type I SS) to test whether the regression coefficient for anxiety is zero, in a model with age.

(b) Run regression models to predict satisfaction using the following sets of predictors. Give the Adjusted $R^2$. (Note: You do not need to show any other results of these models – only Adjusted $R^2$).

   i. Age
   ii. Age, Anxiety
   iii. Age, Severity, Anxiety

Which of these models yields the highest Adjusted $R^2$? Would selecting this model make sense in light of your results in 1(b), 2(c) & 3(a)?

(c) Obtain the residual plots for the model you selected in part (b). (Here, you need plots of residuals vs. fitted values, residuals vs. each predictor, and a normal probability plot). State whether there are any apparent model assumption violations or not, based on these plots.