Answers to Exam No. 2 on Topics from Chapters 2 & 3

1.
R^2 measures the fraction of the variance in the dependent variable explained by the model. If the constant term is not zero, R^2 is also equal to the square of the correlation between the observed value of the dependent variable (Y_t) and the predicted value $\hat{Y}_t$.

2.
The value of R^2 is given by 1 – (ESS/TSS). For the three models, the values are respectively, 0.165, 0.070, and 0.124. Because Model A has the highest R^2, using this criterion, it is the "best" model. The low values for R^2 indicate that the independent variables HSGPA, VSAT, and MSAT do not explain much of the variance in COLGPA.

In Section 3.5 we described how to test the model as a whole with an F-test.

3.
The null hypothesis is that X and Y are uncorrelated (that is, $r_{XY} = 0$) and the alternative is that they are correlated.

4.
The test statistic is $F_c = R^2 (n\text{-}2)/(1\text{-} R^2)$. In our example, $n = 427$ and the F-statistic for Model A is 84. Under the null hypothesis, this has an F-distribution with 1 d.f. for the numerator and 425 d.f. for the denominator.

5.
For a 1 percent-level of significance, the critical $F^*_{1,425}(0.01)$ is approximately 6.7. Because the calculated F_c is well above this, we reject the null hypothesis of lack of correlation between X and Y and conclude that they are correlated. This means that the model is significant overall.

6.
The null hypothesis is that a particular regression coefficient is zero. The alternative for a two-tailed test is that it is nonzero. The test statistic is the coefficient divided by the corresponding standard error. Under the null it has a t-distribution with 425 d.f. The critical $t^*_{425}(0.0025)$ is slightly above 2.807. If an observed t-value exceeds this (in absolute terms) we reject the null hypothesis and conclude that the coefficient is statistically significant. The calculated t-values are:
.

0.92058/0.20463	=	4.50
0.52417/0.05712	=	9.18
1.99740/0.14128	=	14.14
0.00157/0.00028	=	5.61
1.62845/0.15135	=	10.76
0.00204/0.00026	=	7.85

Because all the *t*-statistics exceed the critical value, every regression coefficient in every model is statistically significantly different from zero.

7.
A multiple regression model would combine the three models into one general model as

$$\text{COLGPA} = b_1 + b_2\text{HSGPA} + b_3\text{VSAT} + b_4\text{MSAT} + u$$