You are interested in evaluating two competing theories of academic success. On the one hand, “Agility Theory” states that higher education is easiest absorbed by young people, and that older students are simply incapable of performing well in comparison. In contrast, “Wisdom Theory” states that the life experiences of older students give them a broader context within which to grasp the complexities of university study. In prior studies little evidence has been found for the linear effects of age on academic performance. Undaunted, you decide to test the theories based on data from graduate students at Iowa State University (ISU).

Your position is that the two theories are too simplistic in that each blindly accepts age as the sole cause of academic performance among university students. You believe that if a student is older due to having worked in an industry related to the subject matter s/he is studying at ISU, any “agility disadvantage” would likely be overcome by the student’s work experience in that industry. You set up a study to test the theories as well as this belief. From among all ISU undergraduate Electrical and Computer Engineering majors, you randomly sample 35 seniors who had a job in electrical or computer engineering prior to becoming an ISU undergraduate. You then obtain data about them on the following variables:

\begin{itemize}
  \item **G = GPA** (the student’s Grade Point Average on a scale from 0 to 4 points, where a 4.0 GPA means that the student earned a grade of A in every course taken)
  \item **A = AGE** (the student’s age in years)
  \item **T = TIME** (the number of years that the student worked in electrical or computer engineering prior to becoming an ISU undergraduate)
  \item **C = CONTEXT** (the student’s response to the question, “My pre-ISU work experience provided me with a useful context for my studies at ISU,” with values ranging from 1=strongly agree to 5=strongly disagree)
\end{itemize}

Your data on these variables are as follows:

\[
\begin{array}{cccccc}
\text{Correlation Coefficients} & \text{GPA (G)} & \text{AGE (A)} & \text{TIME (T)} & \text{CONTEXT (C)} & \text{Mean} & \text{Standard Deviation} \\
\hline
\text{GPA (G)} & 1.0 & 0.0 & 0.6 & 0.0 & 3.2 & 0.31 \\
\text{AGE (A)} & 0.0 & 1.0 & 0.7 & -0.7 & 28.0 & 4.4 \\
\text{TIME (T)} & 0.6 & 0.7 & 1.0 & 0.02 & 5.7 & 2.5 \\
\text{CONTEXT (C)} & 0.0 & -0.7 & 0.02 & 1.0 & 3.3 & 1.6 \\
\end{array}
\]

Note that consistent with prior studies, GPA and age are uncorrelated (i.e., \(r_{\text{GA}}=0\)).
a. Find the unstandardized regression equation (i.e., constant and 2 slopes) for the regression of GPA (G) on age (A) and prior time spent working in electrical or computer engineering (T). (Show your work.) [weight 5]

\[ \hat{b}_A = \hat{\beta}_A \frac{\hat{\sigma}_G}{\hat{\sigma}_A} = \frac{r_{GA} - r_{GT}r_{AT}}{1 - r_{AT}^2} \hat{\sigma}_G = \frac{0 - (.6 \times .7)}{1 - .7^2} \times \frac{.31}{4.4} = \frac{- .1302}{2.244} = -.058 \]

\[ \hat{b}_T = \hat{\beta}_T \frac{\hat{\sigma}_G}{\hat{\sigma}_T} = \frac{r_{GT} - r_{GA}r_{AT}}{1 - r_{AT}^2} \hat{\sigma}_G = \frac{.6 - (.0 \times .7)}{1 - .7^2} \times \frac{.31}{2.5} = \frac{.186}{1.275} = .146 \]

\[ \hat{a} = \bar{G} - \hat{b}_A \bar{A} - \hat{b}_T \bar{T} = 3.2 + (.058 \times 28) - (.146 \times 5.7) = 3.2 + 1.624 - .832 = 3.992 \]

The unstandardized regression equation is thus as follows:

\[ \hat{G} = 3.992 - .058A + .146T \]
b. Referring to the regression model obtained in part a, what proportion of the variance in G is explained by A and T? (Hints: Find only a single proportion, and show how it was calculated.)

\[
R^2_{G,AT} = \frac{r^2_{GT} + r^2_{AT} - 2r_{GT}r_{AT}}{1 - r^2_{AT}} = \frac{(0)^2 + (.6)^2 - 2(0 \times .6 \times .7)}{1 - .7^2} = \frac{.36}{.51} = .706
\]

c. Is the proportion found in part b significantly large at the .05 significance level? (Show your work.)

\[H_o: \beta_{G,AT} = 0 \quad H_A: \beta_{G,AT} > 0\]

\[
F_{32,0.05} < 3.32 = F^2_{32,0.05} = 38.422 = F^2_{32}
\]

Rejection rule: Reject \( H_o \) if \( F^2_{32} > 3.32 \).

Since 38.422 > 3.32, there is statistically significant evidence at the .05 significance level that a significantly large proportion of the variance in G is explained by A and T.
d. Based on the results found in part a, express in words the meaning of the partial slope associated with the effects of age on GPA. 

The unit of analysis in this study is the ISU senior with an Electrical and Computer Engineering major, who had a job in electrical or computer engineering prior to becoming an ISU undergraduate (hereafter referred to as a “senior”). Age here measures seniors’ lack of “mental agility.” Accordingly, my results indicate that one would estimate a drop of .058 in a senior’s grade point average for every additional one year of her or his age, after adjusting this age for the number of years that the student worked in electrical or computer engineering prior to becoming an ISU undergraduate (i.e., after removing whatever lack of agility may have resulted from the number of years she or he had aged while working in engineering).
e. Based on the slope discussed in part d, do you have statistically significant evidence that academic success declines with age? (Use the .05 significance level, show your work, and clearly state your conclusion.) [weight 6]

\[ H_o : b_A = 0 \]
\[ H_A : b_A < 0 \]

\[ SS_{TOTAL} = SS_G = \hat{\sigma}_G^2 (n - 1) = (31)^2 (35 - 1) = 3.267 \]

\[ SS_A = \sigma_A^2 (n - 1) = (4.4)^2 (35 - 1) = 658.24 \]

\[ SS_{ERROR} = (1 - R_{G,AT}^2) SS_{TOTAL} = (1 - .706) \times 3.267 = .96 \]

\[ MSE = \frac{SS_{ERROR}}{n - k - 1} = \frac{.96}{35 - 2 - 1} = .03 \]

\[ \hat{\sigma}^2_{b_A} = \frac{MSE}{SS_A (1 - r_{AT}^2)} = \frac{.03}{658.24 (1 - [.7]^2)} = .00009 \]

\[ z \approx t_{32} = \frac{\hat{b}_A}{\hat{\sigma}_{\hat{b}_A}} = \frac{-.058}{\sqrt{.00009}} = -6.114 \]

\[ z_{.05} = 1.645 \]

Rejection rule: Reject \( H_o \) if \( z < -1.645 \).

Since \(-6.114 < -1.645\), there is statistically significant evidence at the .05 significance level that academic success declines with age.
f. Turning to the variable, CONTEXT (C), you first examine its association with the other two independent variables. Of the variance in context (C) that is not associated with the linear effects of age (A), what proportion is associated with prior time spent working in electrical or computer engineering (T)? (Show your work.)

\[
r_{CT,A}^2 = \frac{(r_{CT} - r_{CA}r_{TA})^2}{(1 - r_{CA}^2)(1 - r_{TA}^2)} = \frac{[.02 - (-.7 \times .7)]^2}{(1 - .7^2)(1 - (-.7)^2)} = \frac{(.51)^2}{(.51)^2} = 1
\]

Note that all of the variance in C is explained by A and T.

g. Reflecting on your research project as a whole, you decide that one of the assumptions underlying multiple regression would be violated if you were to regress GPA (G) on age (A), context (C), and prior time spent working in electrical or computer engineering (T). What assumption is this? Why do you believe the assumption would be violated? [weight 1]

The design matrix, X, would be singular (i.e., \( \det[X] = 0 \)) if three columns of the matrix included my data on A, C, and T. Thus the assumption that X is full column rank would be violated, with consequences that the \( X^T X \) matrix could not be inverted and that neither slopes nor their standard errors could be calculated.