Statistics-I. Continuous Evaluation (December 2021). Group ARA

We have the following data file (<https://www.uv.es/mperea/DBZ.jasp>) containing various variables from a sample of college students.

1. We want to examine whether there is “high\_school\_grade” and “first\_year\_grade” in the sample. The questions are. (1) is the relationship approximately linear?; (b) What is the value of the appropriate correlation index and what does it mean? (Copy/Paste from JASP)

2. We have a regression equation in which the dependent variable is "high\_school\_grade" and the predictors are “IQ\_test”, “Age”, and "stress". Please answer the following questions: (a) What proportion of variance of "high school grade" can explain the regression equation?; (b) Perhaps some (or all) of the predictors might not be particularly informative, what would be the predictors if we use a *Stepwise* procedure (and what would be the interpretation of the output)? (Copy/Paste from JASP)

3. We have the following probability density function: f(x)=d from 0<X<0.5. The questions are the following: (a) What is the value of f(0.25) and what does it mean?; and (b) What is the value of F(0.25) and what does it mean?

4. We want to select the top 1% of students with the best marks. Assuming that the distribution (in past years) of the marks consistently follows a normal distribution with mean 40 and standard deviation of 8, what is the value of the cutoff point that will separate these very talented students? (Copy/Paste from Excel)