Multiple regression practice problems

1. Data taken from Howell (2002). "A number of years ago, the student association of a large university published an evaluation of several hundred courses taught during the preceding semester. Students in each course had completed a questionnaire in which they rated a number of different aspects of the course on a 5-point scale (1= very bad to 5=excellent)" (p. 535). Five variables obtained are:

    a. overall – overall rating of the course.
    b. teach – rating of teaching skills of the instructor.
    c. exams – quality of tests and exams
    d. knowledge – rating of the instructor's knowledge of the material
    e. grade - student's anticipated grade for the course (1=F to 5=A)
    f. enroll – enrollment for the course

overall	teach	exams	knowledg	grade	enroll
3.4	3.8	3.8	4.5	3.5	21
2.9	2.8	3.2	3.8	3.2	50
2.6	2.2	1.9	3.9	2.8	800
3.8	3.5	3.5	4.1	3.3	221
3.0	3.2	2.8	3.5	3.2	7
2.5	2.7	3.8	4.2	3.2	108
3.9	4.1	3.8	4.5	3.6	54
4.3	4.2	4.1	4.7	4.0	99
3.8	3.7	3.6	4.1	3.0	51
3.4	3.7	3.6	4.1	3.1	47
2.8	3.3	3.5	3.9	3.0	73
2.9	2.2	3.3	3.9	3.3	25
4.1	4.1	3.6	4.0	3.2	37
2.7	3.1	3.8	4.1	3.4	83
3.9	2.9	3.8	4.5	3.7	70

1. Enter the variables **teach**, **exams**, **knowledg**, **grade**, and **enroll** into a multiple regression model predicting scores for **overall**. What proportion of variability is accounted for? What is the regression equation using unstandardized coefficients? Does the model account for a significant amount of variability? Why do you think so?
2. Use the Stepwise method to determine the regression equation when starting with the same predictor variables listed in Part 1. Please describe the steps SPSS went through in generating its regression equation. At each stage of the process list (a) the variable that was entered or removed from the equation (b) that variable's unique contribution, and (c) the R Square for the regression equation up to that point. Report the final version of the regression equation. What proportion of variability is accounted for by the final version of the regression equation.
3. Repeat Part 2, except use the Backward method (i.e., describe each step SPSS went through). Is the solution different from the one you got using the Stepwise method?
4. Does **teach** account for a significant amount of variability above and beyond that of **enroll**? What is the unique contribution of **teach**? Is this unique contribution significant? Why do you think so?
5. Generate the Venn diagram with **overall** as the criterion and **teach** and **enroll** as predictors.

2. Using the data set named **Lesson 33 Data File 1** from Howell (2002).

1. Enter the variables **arms**, **quads**, **injury**, and **age** into a multiple regression model predicting scores for **medindex**. What proportion of variability is accounted for? What is the regression equation using unstandardized coefficients? Does the model account for a significant amount of variability? Why do you think so?
2. Use the Stepwise method to determine the regression equation when starting with the same predictor variables listed in Part 1. Please describe the steps SPSS went through in generating its regression equation. At each stage of the process list (a) the variable that was entered or removed from the equation (b) that variable's unique contribution, and (c) the R Square for the regression equation up to that point. Report the final version of the regression equation. What proportion of variability is accounted for by the final version of the regression equation.
3. Repeat Part 2, except use the Backward method (i.e., describe each step SPSS went through). Is the solution different from the one you got using the Stepwise method?
4. Does **injury** account for a significant amount of variability above and beyond that of **age**? What is the unique contribution of **injury**? Is this unique contribution significant? Why do you think so?
5. Generate the Venn diagram with **medindex** as the criterion and **injury** and **age** as predictors.