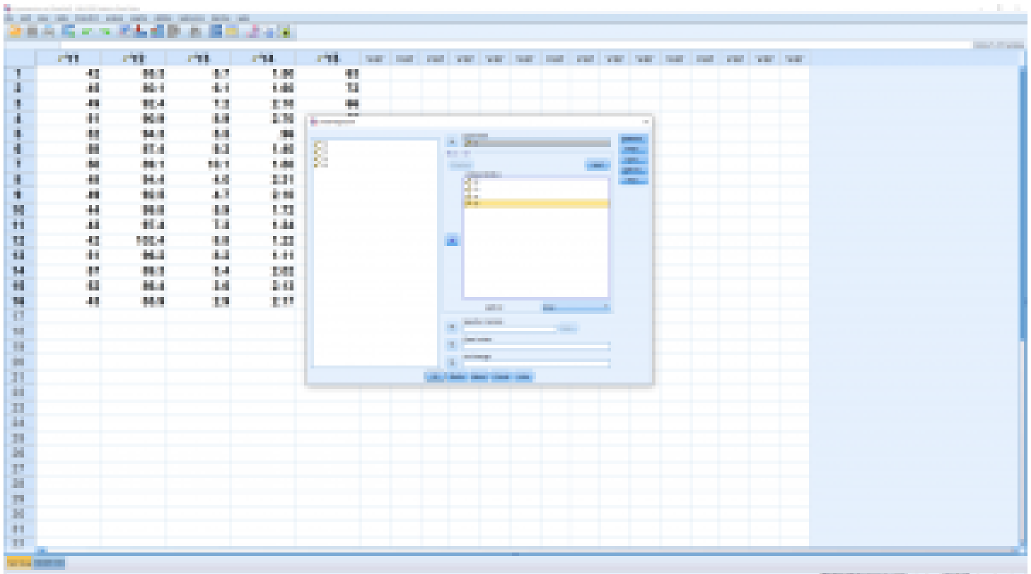
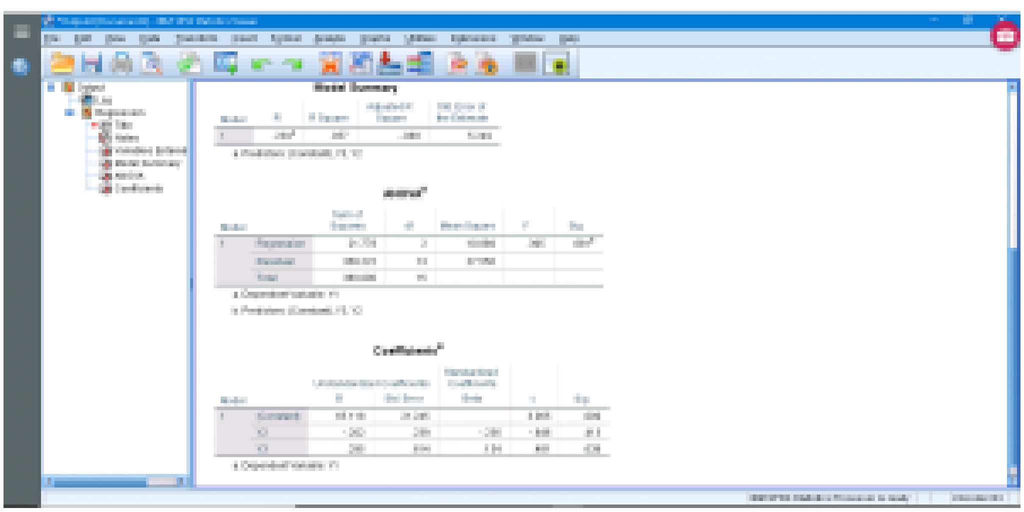
14.9: SPSS Lesson 12- Multiple Regression

Open "Hypertension.sav" from the Data Sets: It is very similar to the data file we used for demonstrating simple linear regression in SPSS but now we have more variables to choose from for independent variables. As before, we really should combine the strength variables but we'll pick y_2 and y_3. Let's pick age as a second independent variable, y_1. Pick Analyze → Regression → Linear and enter the independent and dependent variables :



SPSS screenshot © International Business Machines Corporation.

We will again ignore the submenus but note this time that they are to set up what is known as step-up and step-down analysis where independent variables are added or removed in an attempt to get a better fitting model by removing independent variables that are correlated with each other. The relevant output is (ignoring the table meant for step-up and step-down analysis) :



SPSS screenshot © International Business Machines Corporation.

The "Model Summary" table gives r, r^2 (here the model explains 5.7% of the variance of y), r^2_{adj} and s_{est} for multiple regression which we did not look at explicitly for multiple regression. The "ANOVA" table gives the test statistic F for the significance of r along with its p value, which is not significant here. Again, note that this is not the F we looked at in Section 14.10.2, notice the drastic difference in the degrees of freedom between for the two F values. But both do test the significance of the overall r. The models given by the "Coefficients" table are :

$$\begin{aligned} y &= b_0 + b_1 y_1 + b_2 y_2 \\ y &= 65.118 - 0.202 y_1 + 0.295 y_2 \end{aligned}$$

Note that the intercept is significant but the two slopes are not. If the variables were z-transformed first then we'd have:

$$\begin{aligned} z_y &= \beta_1 z_{y1} + \beta_2 z_{y2} \\ z_y &= -0.236 z_{y1} + 0.134 z_{y2} \end{aligned}$$

There is no way to get SPSS to plot the best fit plane through 3D scatterplot data.