

Multiple Regression in SPSS

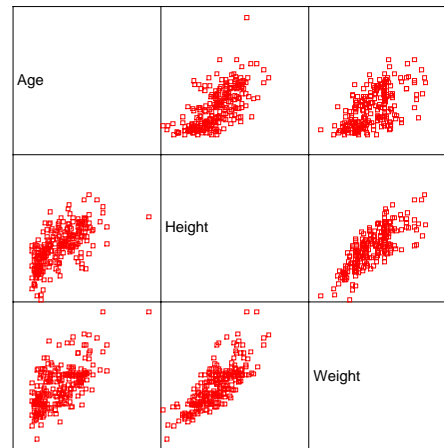
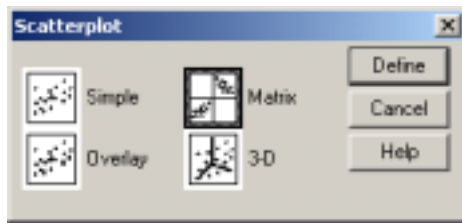
Data: WeightbyAgeHeight.sav

Goals:

- Examine relation between weight (response) and age and height (explanatory)
- Model checking
- Predict weight

I. View the Data with a Scatter Plot

To create a scatter plot, click through **Graphs\Scatter\Simple\Define**. A Scatterplot dialog box will appear. Click **Matrix** option and then click **Define** button. The Simple Scatterplot dialog box will appear. Click **age**, **weight** and **height** variables into matrix variables box.



II. Regression Analysis

To perform the regression, click on **Analyze\Regression\Linear**. Place **weight** in the **Dependent** box and place **age** and **height** in the **Independent** box. Click the **Statistics** button to select the **collinearity diagnostics** and click **Continue**. To obtain the 95% confidence interval for the slope, click on the **Statistics** button at the bottom and then put a check in the box for **Confidence Intervals**. Hit **Continue** and then hit **OK**.

The independent variables (**age**, **height**) are useful in predicting the dependent variable (**weight**) when the level of significance (P-value labeled with **Sig.** on the Output) is below 0.05.

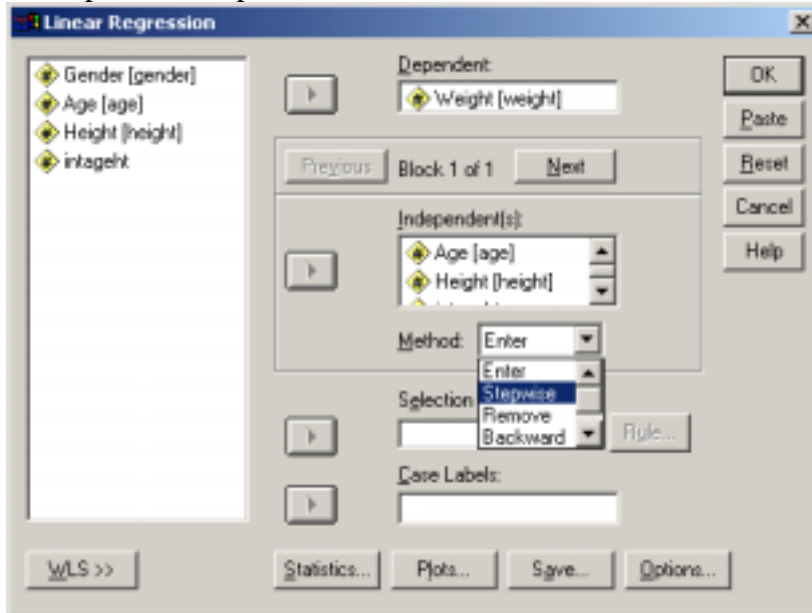
Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
|-------|------------|-----------------------------|------------|---------------------------|---------|------|-------------------------|-------|
| | | B | Std. Error | Beta | | | Tolerance | VIF |
| 1 | (Constant) | -127.820 | 12.099 | | -10.565 | .000 | | |
| | Age | .240 | .055 | .228 | 4.360 | .000 | .579 | 1.727 |
| | Height | 3.090 | .257 | .627 | 12.008 | .000 | .579 | 1.727 |

^a. Dependent Variable: Weight

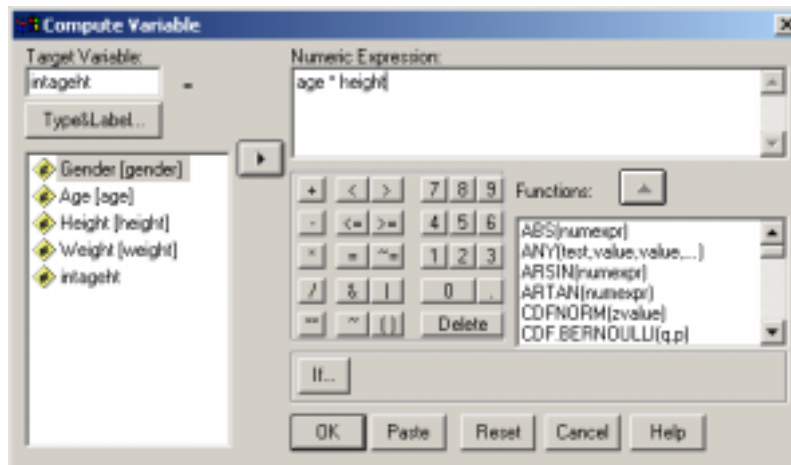
Stepwise Regression

To perform stepwise regression for automatically selecting significant variables, check the **Method** drop down list and choose the desired one and click OK. SPSS will produce an output table to present the final model with a coefficients table.



Interaction Term

To examine the interaction between age and height variables, first create the interaction variable (**intageht**). Click **Transform\Compute**, and in the compute variable dialog box, enter a name for the interaction term, **intageht**. Click and enter **age** variable and click "*" for multiplying and then click and enter **height** variable in the Numeric Expression box and then click **OK**.



To perform the regression, click on **Analyze\Regression\Linear**. Place **weight** in the **Dependent** box and place **age**, **height** and **intageht** in the **Independent** box. Click the **Statistics** button to select the **collinearity diagnostics** and click **Continue**, and then hit **OK**. In the coefficients table, VIF are all greater than 10 which implies collinearity.

Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
|-------|------------|-----------------------------|------------|---------------------------|--------|------|-------------------------|---------|
| | | B | Std. Error | Beta | | | Tolerance | VIF |
| 1 | (Constant) | 66.996 | 106.189 | | .631 | .529 | | |
| | Age | -.973 | .660 | -.923 | -1.476 | .141 | .004 | 250.009 |
| | Height | -3.13E-02 | 1.710 | -.006 | -.018 | .985 | .013 | 77.016 |
| | INTAGEHT | 1.936E-02 | .010 | 1.636 | 1.847 | .066 | .002 | 501.996 |

a. Dependent Variable: Weight

If categorical variables are to be included in the model, the indicator variables will need to be created. **Recode** under transform option is one of the methods to generate the indicator variables.

III. Residual Analysis

To assess the fit of the model, when performing the regression, also click on the **Save** button at the bottom of the dialogue box. In the boxes labeled **Predicted Values** and **Residuals**, click **Unstandardized** in both boxes. In the Hit **Continue** and then hit **OK**. There will now be a column labeled **res_1** and a column labeled **pre_1** in the data window. To make a residual plot, click **Graphs\Scatter\Simple\Define** and put **res_1** in the **Y-axis** box and again put **pre_1** in the **X-axis** box. Hit **OK**. The goal of a residual plot is to see a random scatter of residuals. The normality test in the Explore... option can be used to check for normality.

IV. Prediction Intervals

To calculate the mean prediction intervals and the individual prediction intervals, use the **Save** button that appears after clicking **Analyze\Regression\Linear**. Now in the box labeled **Prediction Values**, click on **Unstandardized**. This will give the predicted Y-values from the model. The data window will have a column labeled **pre_1**. For the prediction intervals, in the boxes near the bottom labeled **Prediction Intervals**, put check marks in front of **Mean** and **Individual**. In the data window, will now be columns, labeled **lmci_1**, **umci_1**, **lici_1**, and **uici_1**. The first two columns are for the lower and upper bounds for the 95% mean prediction interval. (We are 95% sure that the average y for the given x values is within that interval.) The second two columns are the 95% prediction intervals. To do a prediction, simply enter the value of the predictor variable at the last row of the data sheet under the predictor variable and go through the model building. The procedure is similar to that of simple linear regression.