

2. Select the data file named **Merged_PIRLS_Data.sav**.
3. As type of the analysis, select **PIRLS (Using Student Weights)**. The weight variable is automatically selected by the software. As this is an example for analysis on student level, the weight TOTWGT is selected by default. Additionally the variance estimation procedure is defined according to the technical specifications of the study.
4. From the **Statistic Type** dropdown menu, select **Linear Regression**.
5. Note that there are three options under the **Missing Data Option** dropdown menu – **Pairwise**, **Listwise** and **MeanSubstitution**¹⁷. Depending on how you want to treat the missing data, you might change it. For the time being, we will leave it as default (listwise). When choosing **Pairwise**, all available data are used in the analysis, when choosing **Listwise** only cases with complete data are used in the analysis, when choosing **MeanSubstitution** missing data will be replaced with the mean for the variable. We do not recommend the use of **MeanSubstitution** when entering categorical variables in your analysis. This option is only used to select cases based on the continuous variables. Cases with missing values in any of the categorical variables are deleted from the analysis file.
6. In the next steps, all variables for the analysis are selected:
 - As **Grouping Variables**, the software always selects variable IDCNTRY by default. No other variable needs to be added for this example.
 - Next the independent variables need to be identified. To activate this section, click into the area of the **Independent Variables** field. Now you will need to select variable ASBG04 as a categorical variable, select “Effect Coding”, 5 for the “Number of Categories”, and 3 as your reference category. As your continuous independent variable for the analysis choose ASBGSCR.
 - Click on the **Dependent Variable** field. Select variable ASBGSLR from the variable list and move it to the dependent variable field by pressing the right arrow button in this section or by double clicking on the variable name.
7. The weight variable is automatically defined by the software. As this is an example for analysis on student level, the weight TOTWGT is selected by default. For the correct weight and jackknifing variables, please refer to the technical documentation specific to the study.
8. Click on the **Define/Modify** button next to **Output Files** and specify the name of the output files. For our example we will use the name “LinearRegression?”. This filename will be used to create an SPSS file with the syntax to perform the analysis, a set of SPSS and Excel files with the statistics from the analysis (one with model statistics, one with descriptive statistics, and one with the linear regression coefficients), and the SPSS output file with summary statistics from the analysis. The suffixes **_Desc**, **_Model** or **_Coef** are added to the filename to identify the statistics contained in the corresponding file.
9. Click on the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window ready for execution. The syntax file should be executed by opening the **Run** menu of the syntax window and selecting the **All** menu option. Alternatively you can also submit the code for processing with the keystrokes **Ctrl+A** (to select all),

¹⁷ For information about how SPSS treats data under each of these options, please review the documentation for the **MISSING** subcommand within the **REGRESSION** command.