Simple Correlation

Task 1: Correlation between Two Variables

Use the (General Social Survey.dat) data set to find the strength of the relationship between fathers' education level (highest year of school completed, father: paeduc) and mother's education level (highest year of school completed, mother: maeduc).
In the main menu bar go to:

- Analyze
  - Correlate
    - Bivariate (meaning 2 variables)…
      - Transfer the appropriate set of variables to the Variable box
        - The default options selected are Pearson Correlation Coefficient, 2 tailed significance test, flag significant correlations
          - OK

	paeduc	maeduc
paeduc	r=1.00	
maeduc		r=1.00

Is the correlation significant? Yes / No If yes, at what significance level? _____
How many people are in the data set? _____
What proportion of variance in maeduc is explained by paeduc? _____

Note about interpreting significant correlations: With larger samples, small correlations may be deemed significant because of the power. A better way of interpreting correlations is to consider the proportion of variance (r^2). For example, a correlation of 0.2 may be significant, but accounts for only 4 percent of the variance.

Scattterplot: The scatterplot enables you to see whether a correlation will accurately summarize the relationship between 2 variables. Correlations are appropriate only for linear relationships. The r will be an underestimation if the relationship is curvilinear. It is important to examine scatterplots when studying relationships between variables.

To produce a scatterplot for the pair of variables, in the main menu bar go to:

- Graphs
  - Chart Builder - OK
  - Select "Scatter" from the gallery
    - Select "Simple" or the first graph presented – running your mouse over each example graph will tell you what they are.
      - Select the variables from the list on the upper left and drag and drop the variable on the selected axis
        - Transfer maeduc to the Y-axis and paeduc to the X-axis
          - OK (The graph will then be entered into your viewer folder)

SPSS produces simple scatterplots this way. To obtain a line of best fit (more on this next lab)

- Double click on your graph
  - Chart Editor window will open
    - From the menu bar in the Chart Editor window select ELEMENTS - Fit Line at Total
      - OK
        - Close the Chart Editor window

Describe the relationship between the maeduc and paeduc. _____

Task 2: Correlations for a Subset of the Sample

Determine the relationship between education (educ) and mothers education (maeduc) for male students. Reduce your output. To select a subsample of students you need to select cases. In the main menu bar:

- Data
  - Select Cases
    - If condition is satisfied
      - If
        - Move Sex into empty box on the right and create statement specifying the gender of interest (i.e., sex = 1 will specify males)
          - Continue
            - OK

Now run the correlation (analyze, correlate, biverate) and produce the scatterplot.

Male respondents	Education	Mother's Education
Education		
Mother's Education		

What proportion of variance in Education is explained by Mother's Education for male students? _____

What do you conclude? _____

Before running further analyses, you need to unselect the cases (Data, Select Cases, All Cases, OK).