

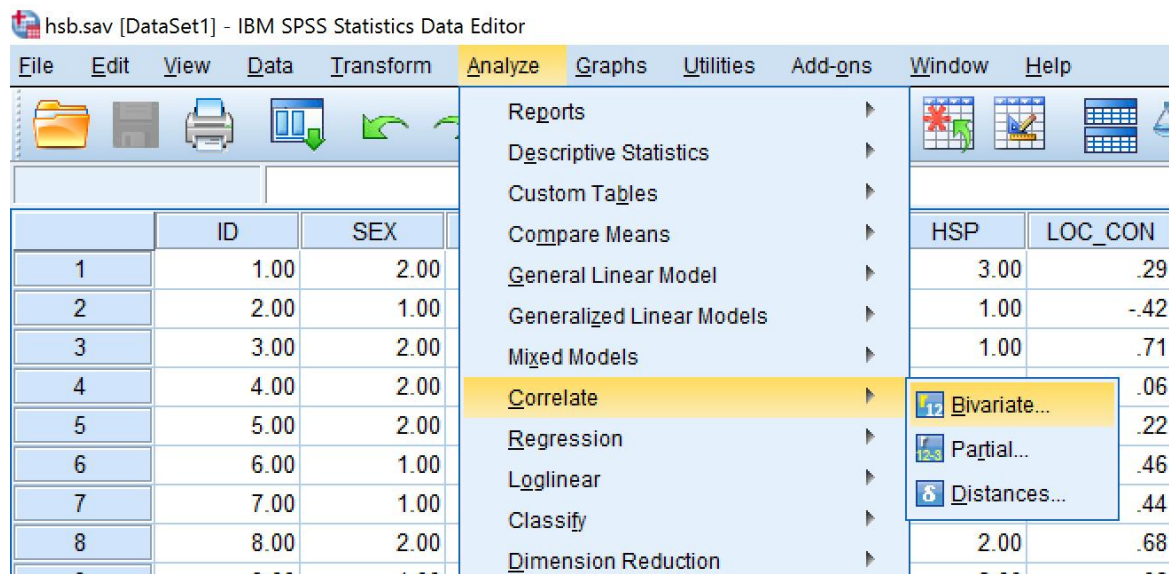
60080079 Introduction to Statistical Methods
Semester 2 2023-2024
Handout 2

A Brief Introduction to Simple Correlation, Regression, and Scatter Plot in SPSS

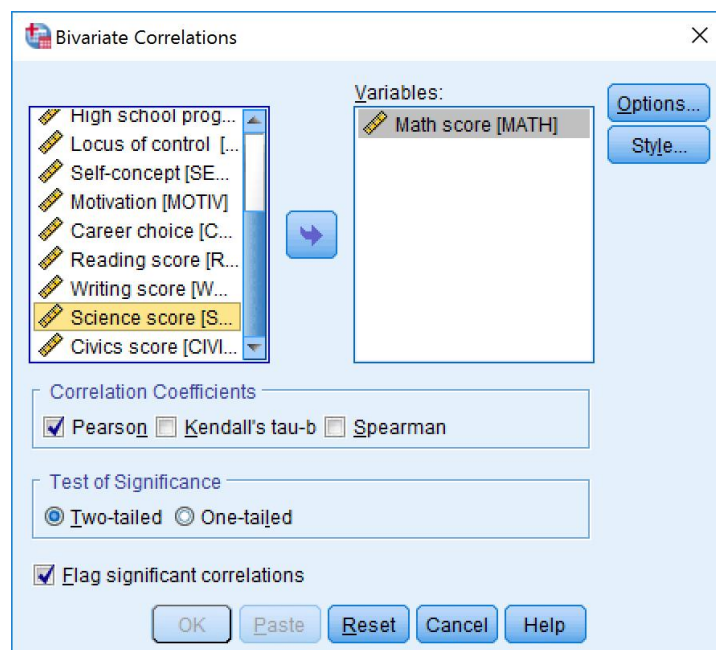
1. Correlation

Set up the data (i.e., open, cut-and-paste, type).

Analyze → Correlate → Bivariate.



Click the variables to be correlated in the Variables box. In our case, we are interested in correlated **Math** and **Science** scores.



With all the variables of interest in the box, hit **OK**.

We should get the following output:

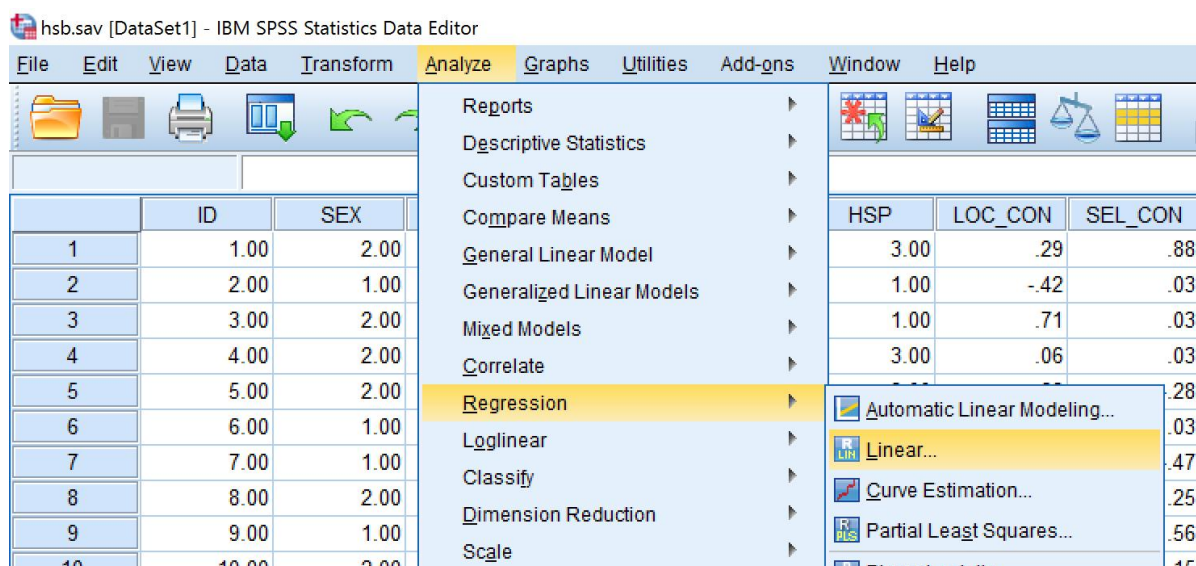
Correlations			
		Math score	Science score
Math score	Pearson Correlation	1	.650**
	Sig. (2-tailed)		.000
	N	600	600
Science score	Pearson Correlation	.650**	1
	Sig. (2-tailed)	.000	
	N	600	600

** . Correlation is significant at the 0.01 level (2-tailed).

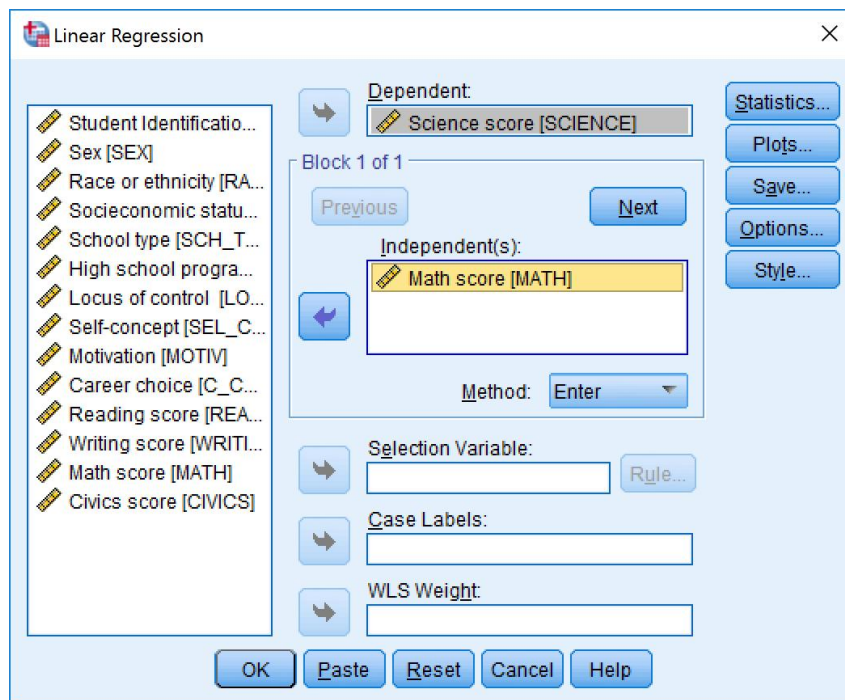
This output shows that the correlation between math and science scores is 0.650.

2. Simple Linear Regression

Analyze → Regression → Linear.



Click in the appropriate **Dependent** and **Independent** variables. In the example below, we want to investigate whether Science score can be predicted from Math score.



The image shows the 'Linear Regression' dialog box in SPSS. On the left is a list of variables: Student Identification..., Sex [SEX], Race or ethnicity [RA...], Socioeconomic statu..., School type [SCH_T...], High school progra..., Locus of control [LO...], Self-concept [SEL_C...], Motivation [MOTIV], Career choice [C_C...], Reading score [REA...], Writing score [WRITL...], Math score [MATH], and Civics score [CIVICS]. The 'Dependent' variable is 'Science score [SCIENCE]'. The 'Independent(s)' variable is 'Math score [MATH]'. The 'Method' is set to 'Enter'. On the right, there are buttons for 'Statistics...', 'Plots...', 'Save...', 'Options...', and 'Style...'. At the bottom are 'OK', 'Paste', 'Reset', 'Cancel', and 'Help' buttons.

Click **OK** to run the regression model.

Below are some output from this analysis.

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650 ^a	.422	.421	7.38616

a. Predictors: (Constant), Math score

Compare R from above with the correlation obtained earlier. Are they supposed to be the same or not?

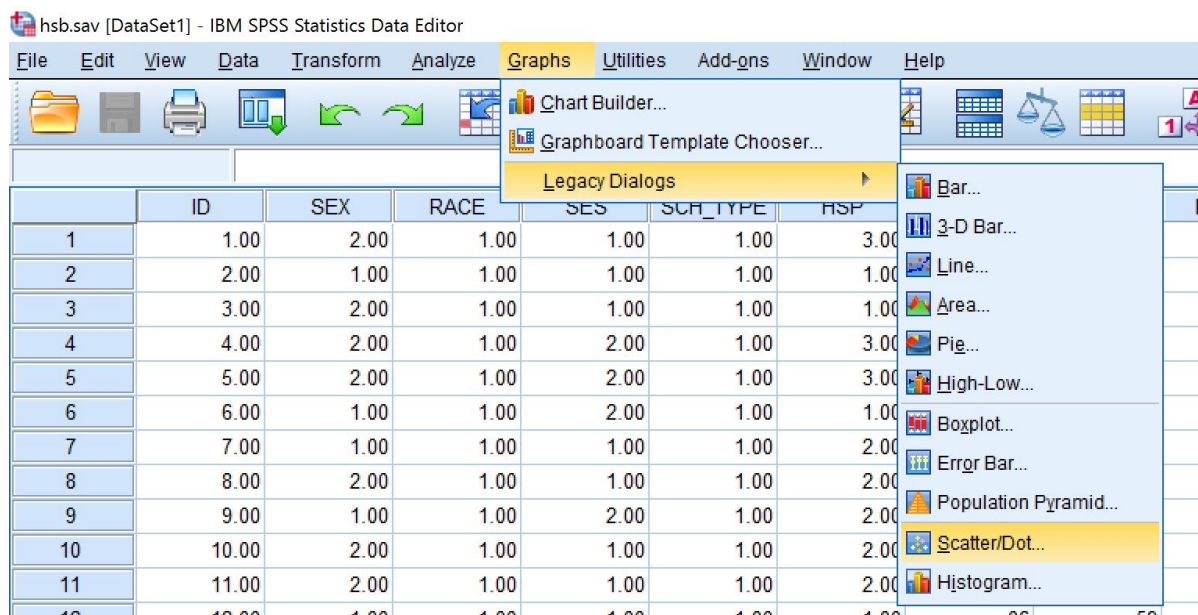
Coefficients ^a						
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17.044	1.689		10.090	.000
	Math score	.670	.032	.650	20.890	.000

a. Dependent Variable: Science score

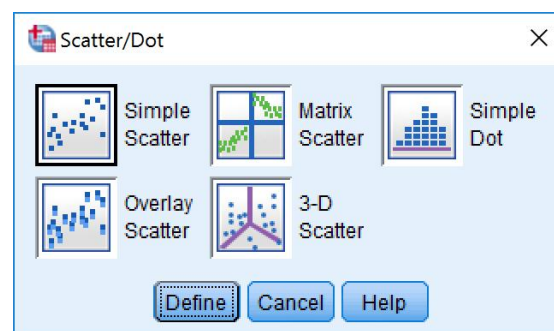
The B for constant (17.044) is the intercept and the B for Math score (0.670) is the slope.

3. Scatter Plot

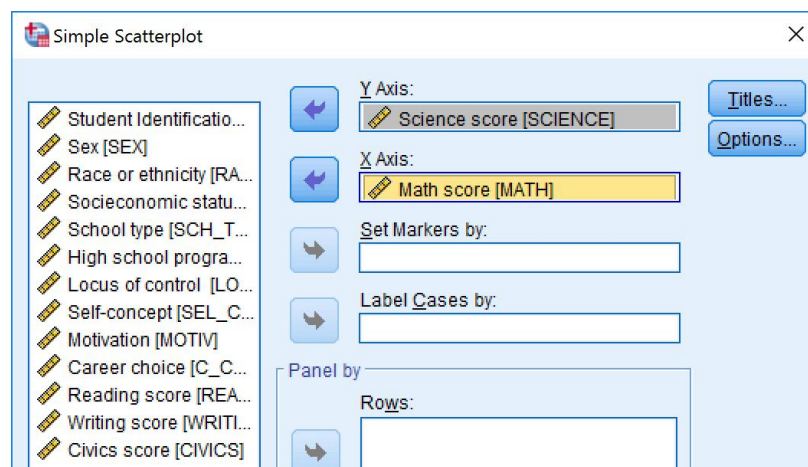
Graphs → Legacy Dialog → Scatter/Dot.



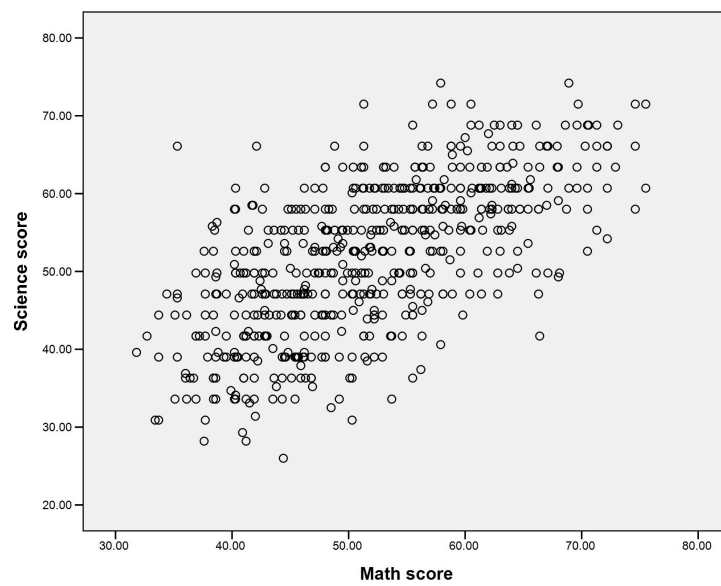
From the Scatter/Dot options, choose **Simple Scatter**, then hit **Define**.



Next, click in the appropriate variables for the X- and Y-axes. In our example, we want to predict Science from Math so **Science Score → Y Axis** and **Math Score → X Axis**.



You should get the following scatter plot as output:



The graph shows a clear positive relationship between math and science scores, and there are no obvious outliers.

What happens when you click in Sex in the **Set Markers by** box?