

Correlation

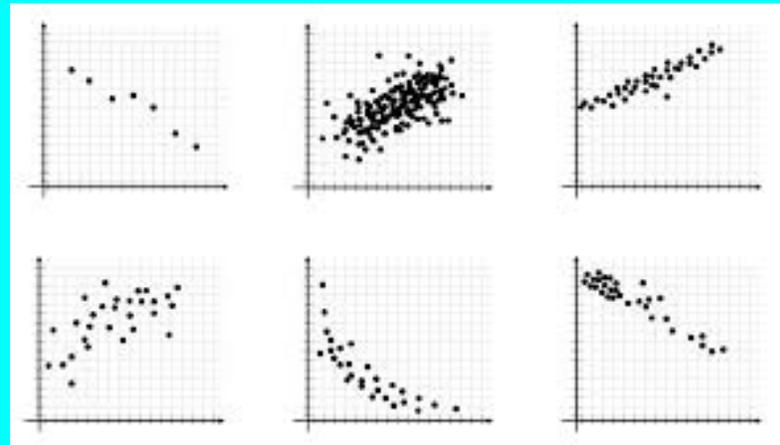
A correlation is a measure of the linear relationship between two variables. It is used when a researcher wishes to describe the strength and direction of the relationship between two normally continuous variables. The statistic obtained is Pearson's product-moment correlation (r), and SPSS also provides the statistical significance of r . In addition, if the researcher needs to explore the relationship between two variables while statistically controlling for a third variable, partial correlation can be used. This is useful when it is suspected that the relationship between two variables may be influenced, or confounded, by the impact of a third variable.

Correlations are a very useful research tool but they do not address the predictive power of variables. This task is left to regression. Regression is based on the idea that the researcher must first have some valid reasons for believing that there is a causal relationship between two or more variables. A well known example is the consumer demand for products and the level of income of consumers. If income increases then demand for normal goods such as cars, foreign travel will increase. In regression analysis, a predictive model needs to fit to both the data and the model. And then we can use the result to predict values of the dependent variable (DV) from one or more independent variables (IVs). In straight forward terms, simple regression seeks to predict an outcome from a single predictor; whereas multiple regression seeks to predict an outcome from several predictors.

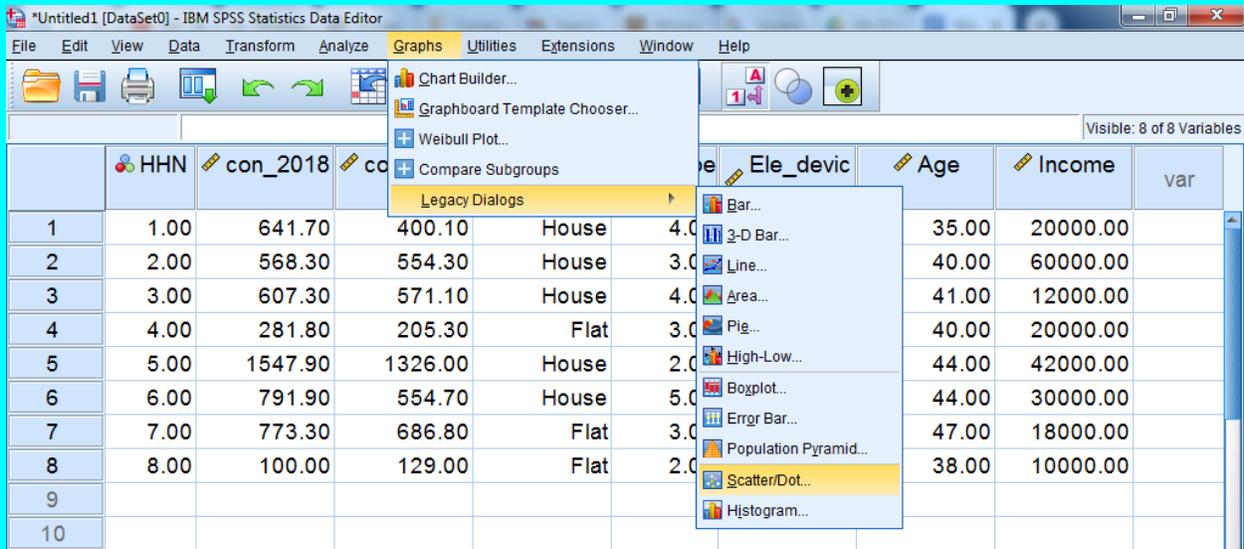
ScatterPlot

The relationship between two variables can be visualized by plotting the data in scatter diagram or scatter plot. A scatter-plot plots the data points for the two variables along the x and y axis taking the value of one variable along x axis and the value of another available along the y- axis. A scatter of points so formed is called the scatter plot. Such a plot exhibits patterns that indicate how variables are related. Using a scatter plot, we may detect whether the relationship between the variables are

- i) direct or inverse
- ii) Liner or curvilinear

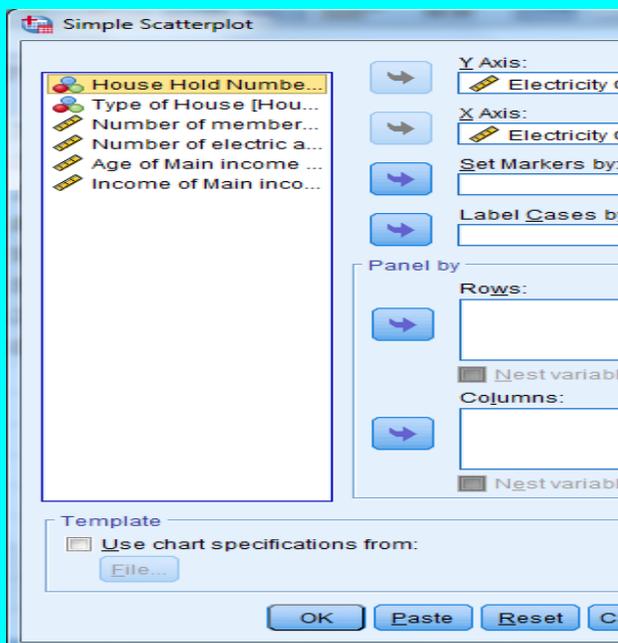
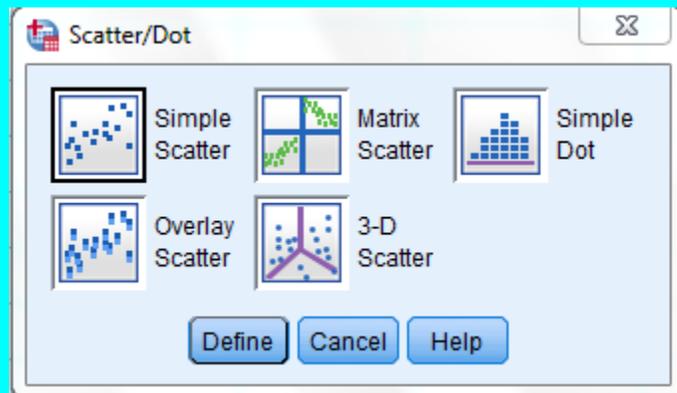


To draw the scatter plot, click on Graphs> Legacy Dialogs and then Scatter/Dot Then you will get the following window.

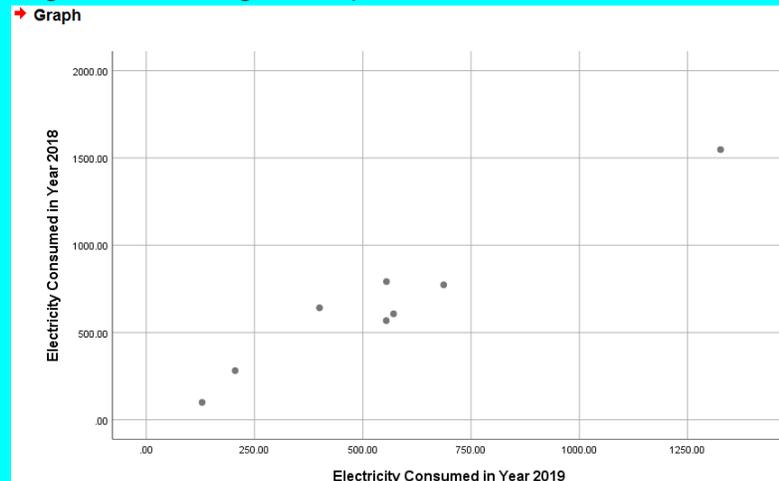


While clicking in Scatter/Dot you will get small window

Note that there are five scatterplot options. The Simple option graphs the relationship between two variables. The Matrix option is for two or more variables that you want graphed in every combination: variable is plotted with every other variable. Every combination is plotted twice so that each variable appears on both the X and Y axis. Make sure that the selected option is simple scatter.



Select the two variables which you want to see the relationship in Y axis and X axis. Click ok tab to get the following scatterplot

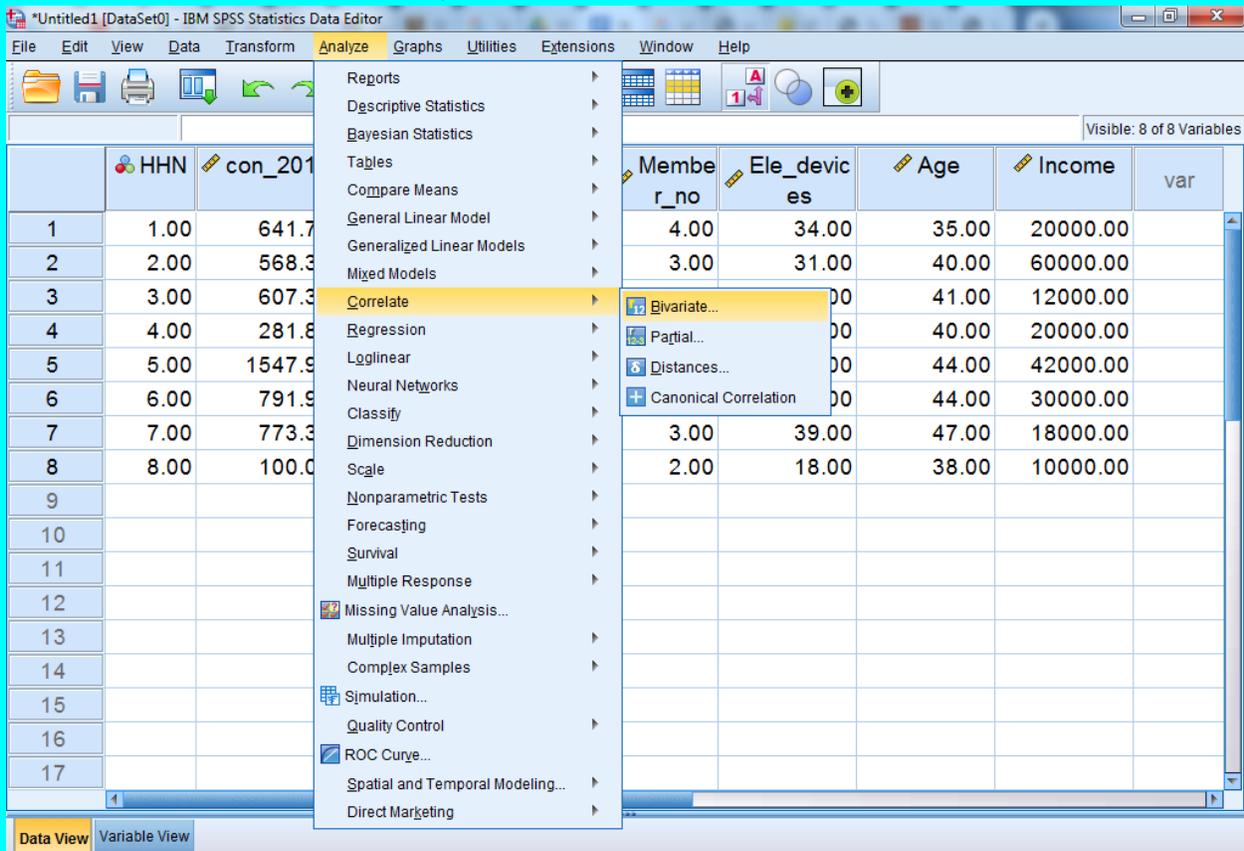


CORRELATION

Correlation analysis is useful when researchers are attempting to establish if a relationship exists between two variables. For example, a researcher might be interested in whether there is a relationship between the IQ level of students and their academic performance. It would be expected that high levels of IQ will be linked to high levels of academic performance. Such a relationship indicates a high degree of positive correlation between the IQ level of students and their academic performance. It is also possible to find correlation where the direction is inverse or negative. For example, a company making computer chips records a high level of defective products. They decide to invest considerable capital in improving their machines and processes. Over time it is seen that as the investment was increased the number of defective products declined. Therefore there is a negative correlation between investment and defective products. It is important to note that correlation provides evidence that there is a relationship between two variables. It does not, however, indicate that one variable causes the other. In other words, the correlation between variable A and B could be a result of A causing B, or B causing A, or there could be a third variable that causes both A and B. It is for this reason that researchers must always take into account the possibility of a third variable impacting on the observed variables. By using partial correlation it is possible to statistically control for these additional variables, which allows the possibility of a clearer, less contaminated, indication of the relationship between the two variables of interest. It is also important to understand the difference between a statistically significant correlation coefficient between the variables and what is of practical significance for the sample. When using large samples, even quite small correlation coefficients can reach statistical significance. For example, although it is statistically significant, the practical significance of a correlation of 0.09 is quite limited. In a case like this, the researcher should focus on the actual size of Pearson's r and the amount of shared variance between the two variables. To interpret the strength of the correlation coefficient, it is advisable to take into account other research that has been conducted in that particular area. SPSS can calculate two types of correlation. First it will give a simple bivariate correlation between two variables, also known as zero order correlation. Secondly, SPSS can explore the relationship between two variables, while controlling for another variable. This is called partial correlation. In this text, only

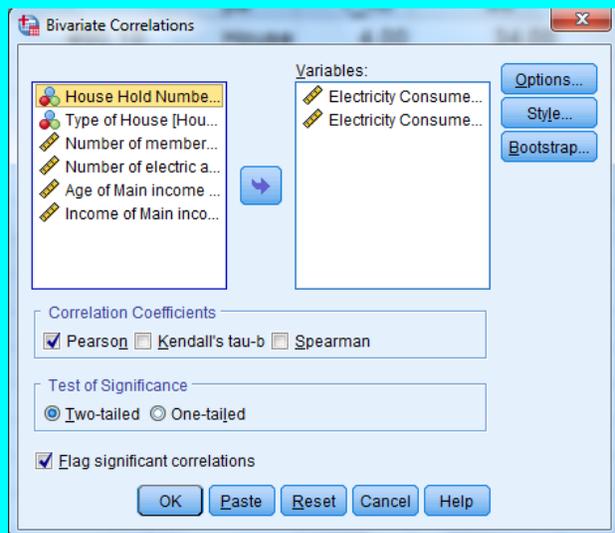
zero order correlation will be discussed. The two most popular correlations are: Spearman's and Pearson's product-moment correlation coefficients. The difference between them is that Pearson's product-moment correlation deals with interval or ratio data while Spearman rank-order can deal with ordinal data apart from interval and ratio data.

In data editor window, chose analyze, correlate and then bivariate



You will get the the following windows

Now you have to select the variables in the variables box which you want to see the relationship. Check to see person box is marked in the correlation coefficients box. Also mark the radio button on two tailed in the test of significance sub dialog box. If you are looking specially for the significance in one direction, you should mark one tailed instead. By default, the dialog box on flag significant correlations is ticked. Click the tab options... to open the Bivariate correlations: options dialog box. Select options on means and standard deviations and cross-product deviations and covariances if you wish to find



the means, standard deviation and covariance for the data. Click on continue and then ok to get the output table

➔ **Correlations**

Descriptive Statistics

	Mean	Std. Deviation	N
Electricity Consumed in Year 2018	664.0250	428.72362	8
Electricity Consumed in Year 2019	553.4125	366.41350	8

Correlations

		Electricity Consumed in Year 2018	Electricity Consumed in Year 2019
Electricity Consumed in Year 2018	Pearson Correlation	1	.975**
	Sig. (2-tailed)		.000
	Sum of Squares and Cross-products	1286627.615	1072382.318
	Covariance	183803.945	153197.474
	N	8	8
Electricity Consumed in Year 2019	Pearson Correlation	.975**	1
	Sig. (2-tailed)	.000	
	Sum of Squares and Cross-products	1072382.318	939811.969
	Covariance	153197.474	134258.853
	N	8	8

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