RANK CORRELATION

PURPOSE
Compute the Spearman rank correlation between two variables.

DESCRIPTION
The rank correlation is recommended in the following cases:

   1. When the underlying data does not have a meaningful numerical measure, but it can be ranked;

   2. When the relationship between the two variables is not linear (however, it should still be either monotonically increasing or monotonically decreasing);

   3. When the normality assumption for the two variables is not valid.

The rank correlation statistic is the distribution-free analog of the correlation statistic. It is the correlation statistic applied to a pair of ranked variables. That is, given a pair (X_i, Y_i) of variables, find the pair of ranked variables (X_{ranki}, Y_{ranki}). These rankings are performed separately for X and Y. Ties are replaced with an average rank. For example, if the seventh and eighth ranks are the same value, both are replaced with 7.5. The formula for the rank comovement coefficient is then:

$$r = \frac{\sum_{i=1}^{N} (X_{ranki} - \overline{X_{ranki}})(Y_{ranki} - \overline{Y_{ranki}})}{\sqrt{\sum_{i=1}^{N} (X_{ranki} - \overline{X_{ranki}})^2 \sum_{i=1}^{N} (Y_{ranki} - \overline{Y_{ranki}})^2}} \qquad \textbf{(EQ 2-9)}$$

A perfect relationship yields a rank correlation of +1 (or -1 for a negative relationship) and no relationship yields a rank correlation of 0. The two variables must have the same number of elements.

SYNTAX
LET <par> = RANK CORRELATION <y1> <y2>   <SUBSET/EXCEPT/FOR qualification>
where <y1> is the first response variable;
      <y2> is the second response variable;
      <par> is a parameter where the computed rank correlation is stored;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

EXAMPLES
LET A = RANK CORRELATION Y1 Y2
LET A = RANK CORRELATION Y1 Y2 SUBSET TAG > 2

DEFAULT
None

SYNONYMS
None

RELATED COMMANDS
CORRELATION	=	Compute the Pearson correlation coefficient between two variables.
AUTOCORRELATION	=	Compute the lag 1 autocorrelation of a variable.
RANK COVARIANCE	=	Compute the rank covariance between two variables.
CORRELATION PLOT	=	Generate an autocorrelation or cross-correlation plot.
COVARIANCE	=	Compute the covariance between two variables.
LINEAR SLOPE PLOT	=	Generate a linear slope versus subset plot.

APPLICATIONS
Exploratory Data Analysis

IMPLEMENTATION DATE
Pre-1987

PROGRAM
        SKIP 25
        READ SNAIL.DAT Y X
        LET A = CORRELATION Y X

        The computed correlation is 1.0.