

Measuring bivariate relationships – measurement data

Measurement data is the most mathematically flexible information we can use in quantitative analysis. Ideally, we like to use mean-based measures. But if skew is a problem, there are a whole slew of rank-order measures that can be used instead. Measurement data is most often examined using correlation and regression. In fact, these are the most commonly analytic tools used in the social sciences. There are many different “flavors” of correlation and regression and we will only cover the main ones here. Correlation and regression are fairly sophisticated but, as we will go over many times, the fact that they are sophisticated means that they are more susceptible to problems and errors. I'd like you to learn to be a good critical reader of correlation and regression as well as responsible users of these techniques.

- Visualizing a linear relationship
 - Scatterplots and the method of least squares
 - Non-linear trends
- Measuring a linear relationship
 - Roughly normally distributed variables
 - Amount of variation explain/predicted
 - Amount of variation left unexplained/unpredicted
 - Heavily skewed variables
 - Variables with bimodal or otherwise problematic distributions
- Making further use of the least-squares line
 - Equation of the line
 - Slope
 - Intercept
 - Error
 - Coefficient of determination (R squared)
- Adding in more independent variables
- Correlation does not imply causation, regression does