

Workshop in Methods

Reproducible Results and the Workflow of Data Analysis

Dr. J. Scott Long

*Distinguished Professor and Chancellor's Professor of Sociology and Statistics
Indiana University*

Many disciplines are paying increasing attention to “reproducible results”. This is the idea other scientists should have access to your data so that they can reproduce the results from your published work. Producing reproducible results is critically important and highly dependent on your workflow of data analysis. This workflow encompasses the entire process of scientific research: Planning, documenting, and organizing your work; creating, labeling, naming, and verifying variables; performing and presenting statistical



analyses; preserving your work; and (perhaps, most important) producing replicable results. Most of our work in statistics classes focuses on estimating and interpreting models. In most “real world” research projects, these activities involve less than 10% of the total work. Professor Long’s talk is about the other 90% of the work. An efficient workflow saves time, introduces greater reliability into the steps of the analysis, and generates reproducible results.

Friday, January 22, 2016, 2:30-4pm

Social Science Research Commons Grand Hall

Woodburn Hall 200