Missing data is ubiquitous in applied social science research. Survey questions often go unanswered, businesses may be exempted from reporting requirements, and countries do not share their national statistics. Missing data is a serious problem and traditional “solutions” such dropping all observations with incomplete data raise questions of bias, validity and sample selection.

Various ways to address these problems will be discussed, including multiple imputation, maximum likelihood, and Heckman’s solution for non-random sample selection for continuous and binary outcome variables. Software emphasized will be Stata and Amelia II (e.g., for time-series-cross-sectional data sets).

Participants should have a basic knowledge of data manipulation and analysis in Stata (or similar software).

Instructor: Fred Boehmke
Political Science, University of Iowa

Where: B&E, 148, 105
When: May 23 & 24, 9am to 4:30pm, with a break for lunch (12-12:45pm).
Who is eligible? Graduate students and faculty at the University of Kentucky can enroll for free. You must register (http://qipsr.as.uky.edu/) for both the morning and afternoon sessions. Seating for the computer lab is limited to 50 seats.

The two-day workshop will discuss forms of missing data, problems and software solutions. There will be lectures in the morning session (in B&E 148) and computer lab sessions in the afternoon (in B&E 105). Emphasis is on acquiring a practical understanding for applied researchers.

Supported by QIPSR, Statistics, Agricultural Economics, and the Center for Poverty Research, as well as Sociology, Political Science and the College of Arts & Sciences.

http://qipsr.as.uky.edu/