

## WORKSHOP IN METHODS

# Workflow of Data Preparation

## Dr. Bianca Manago

**Assistant Professor of Sociology  
Vanderbilt University**

Before conducting analyses, we all know that we need to “clean the data”, but what exactly does that mean? What steps are involved and in what order? How do we decide what needs to be done? Data preparation involves all the steps that occur between data collection and analysis (e.g., merging, appending, labeling, data analytics, cross-validation, constructing/re-constructing variables for analysis, identifying missing data). This seminar will provide a general framework for approaching these processes. The framework informs decisions about an ideal order in which data cleaning should be conducted to represent data both accurately and fully. This framework also delves into some of the trickier issues. For example, when you come across anomalous, vague, or missing data – what kinds of things should you consider? I will also provide guidance for ensuring that your findings are reproducible. Finally, I will discuss how to prepare data for analysis as efficiently as possible.

*Bianca Manago is an Assistant Professor in the Department of Sociology at Vanderbilt University. Her research addresses questions at the intersection of research methods, medical sociology, inequality, and social psychology. Her research has been funded by the National Science Foundation and the American Sociological Association.*



**Friday, February 3, 2023  
2-3:30PM**

**Social Science Research  
Commons Grand Hall  
(Woodburn Hall 200) and  
Zoom**

**Link to Register:  
[go.iu.edu/wim](https://go.iu.edu/wim)**

