Missing Data and Multiple Imputation

**CLASS SESSIONS**
Friday, June 17, 2016 10:00 AM - 3:30 PM
The classroom location is in the Hammer Health Sciences Center- LL204
Hammer Health Sciences Library (HHS) is located at 701 W. 168th St. New York, NY 10032
Directions can be found here: https://www.cuepisummer.org/contactpage

**INSTRUCTOR**
Shakira Suglia
Associate Professor
Epidemiology
Mailman School of Public Health
e-mail: sfs2150@columbia.edu

**COURSE DESCRIPTION**
This course will highlight the advantages of multiple imputation as a strategy for addressing missing data, and provide guidance on best practices. Multiple imputation can help researchers take full advantage of their available data, preserve sample size in multivariable analyses with missing covariate data, and reduce bias. A hands-on exercise will allow participants to quantitatively describe and impute missing data in either SAS or Stata and interpret the results.
Tips for developing and documenting a multiple imputation strategy for publication will be discussed.

**PREREQUISITES**
Previous training in biostatistics or statistics including descriptive statistics and basic regression modeling required. Familiarity with SAS or Stata software will be strongly preferred, though if already proficient with SPSS or R participants may be able to use these for multiple imputation with limited instructor support.

**COURSE LEARNING OBJECTIVES**
By the end of the course, participants will be able to:

1. List and categorize causes of missing data commonly encountered in research
2. Describe the advantages and limitations of multiple imputation as compared with common alternatives including complete case analysis, a missing data indicator, and single imputation
3. Plan and document a strategy to use multiple imputation as a main approach or sensitivity analysis
4. Check and interpret multiple imputation results Access resources for further learning and practice
RECOMMENDED COURSE READING LIST


COURSE STRUCTURE
Class time is 5 hours total. The structure of the workshop will include a mixture of lecture and guided exercise with hands-on assistance from the instructor.

COURSE SCHEDULE

<table>
<thead>
<tr>
<th>Schedule</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00-10:45am</td>
<td>Defining the problem &amp; Missing data Mechanisms</td>
</tr>
<tr>
<td>10:50-12:00pm</td>
<td>Options for Analyses &amp; How to Implement Multiple Imputations</td>
</tr>
<tr>
<td>Lunch Break from 12:00-12:30pm</td>
<td></td>
</tr>
<tr>
<td>12:30-2:15pm</td>
<td>Missing Data and Multiple Imputation Exercise</td>
</tr>
<tr>
<td>2:15-3:00pm</td>
<td>Special Considerations for MI analyses</td>
</tr>
</tbody>
</table>