



## Agenda

### Capsular course on Nutritional Epidemiology: Module 2 – Data Management Tools and Advanced Statistics

Joint Collaboration between National Centre for Excellence and Advanced Research on Diets (NCEARD), Lady Irwin College, International Institute for Population Sciences (IIPS) and UNICEF India

Date(s): May 1 – May 4, 2019

Venue: MPR-2, Department of Education, Lady Irwin College, 1 Sikandra Road, New Delhi - 110001

Resource person(s): Dr. Abhishek Singh, Dr. William Joe and Dr. Kaushalendra Kumar

Time	Session	Speaker/Moderator	Techniques and Learning Outcomes
<b>Day 1, Wednesday [May 1, 2019]</b>			
9: 00 AM – 9:30 AM	<b>Registration and Seating</b>		
9:30 AM-9:40 AM	Welcome note and brief about Lady Irwin College	Dr. Anupa Siddhu	
9:40 AM-9: 50 AM	A brief introduction about NCEARD and overview of the capsular course	Dr. Manisha Sabharwal	
9:50 AM-10.00 AM	Introduction of participants		
10:00 AM-11:00AM	Introduction to STATA software	Dr. Abhishek Singh	<b>What is STATA used for and how is this software is useful in research-</b> <ul style="list-style-type: none"><li>• Application and use of STATA in Social Sciences</li><li>• Comparison of STATA with other software</li><li>• STATA versions: IC, SE, MP</li><li>• Opening STATA software</li></ul>
11:00 AM – 11:30 AM	<b>Tea Break</b>		
11:30 AM-1:00 PM	Operating STATA and its Functionalities	Dr. Abhishek Singh & Dr. Kaushalendra Kumar	<b>Operating STATA and its functionalities-</b> <ul style="list-style-type: none"><li>• STATA Environment: command, results, review, variables, properties</li><li>• Menus: file, edit, data, graphics, statistics, user, window, help</li></ul>



			<ul style="list-style-type: none"> <li>• Button bar: open, save, print, log, new viewer, bring graph window to front, new-do file editor, data editor (edit/browse), variables manager, show more results, break</li> <li>• Do file: stata syntax</li> <li>• Log file: output and syntax saving</li> </ul>
1:00 PM-2:00 PM	<b>Lunch Break</b>		
2:00 PM-3:00 PM	Entering data in STATA	Dr. Kaushalendra Kumar	<b>Importing and exporting data in STATA from/to difference sources-</b> <ul style="list-style-type: none"> <li>• Working directory</li> <li>• Importing data: excel, CSV, TXT, SAS XPORT, SPSS</li> <li>• Exporting data: excel, CSV, TXT, SAS XPORT</li> </ul>
3:00 PM-3:30 PM	<b>Tea Break</b>		
3:30 PM-5:30 PM	Data management in STATA	Dr. Kaushalendra Kumar	<b>Preparing data for analysis, identifying and addressing outliers-</b> <ul style="list-style-type: none"> <li>• Cleaning the dataset: tabulate, missing values, renaming, replace and labeling</li> <li>• Describing the data: describe, summarize</li> <li>• Data structure: wide vs long</li> <li>• Looping in STATA: macros, while loops, foreach, forvalues</li> </ul>
<b>Feedback (Day 1)</b>			
<b>Time</b>	<b>Session</b>	<b>Speaker/Moderator</b>	<b>Techniques and Learning Outcomes</b>
<b>Day 2, Thursday [May 2, 2019]</b>			
9:00 AM-10:45 AM	Data manipulation - I	Dr. Kaushalendra Kumar	<b>Computation of new variables, recoding variables and applying logics-</b> <ul style="list-style-type: none"> <li>• Generating new variables: gen, egen, xtiles</li> <li>• Recoding variables</li> <li>• Conditional and Logical expressions: and, or, equal, greater/lower than</li> </ul>
10:45 AM- 11:00 AM	<b>Tea Break</b>		
11:00 AM-1:30 PM	Data manipulation - II	Dr. Kaushalendra Kumar	<b>Management of data file, deleting variables, sorting data-</b> <ul style="list-style-type: none"> <li>• Using preserve and restore</li> </ul>



			<ul style="list-style-type: none"> <li>• Keeping or dropping variables</li> <li>• Creating dummy variables</li> <li>• bysortcommand: gen new variable</li> </ul>
1:30 PM-2:30 PM	<b>Lunch Break</b>		
2:30 PM-3:45 PM	Data manipulation - III	Dr. William Joe	<b>Restructuring of data set</b> <ul style="list-style-type: none"> <li>• Data identifier and its uniqueness</li> <li>• Collapse command: generate new variable or data set</li> </ul>
3:45 PM-4:00 PM	<b>Tea Break</b>		
4:00 PM-5:30 PM	Data manipulation – IV	Dr. William Joe	<b>Merging files, aggregating files and restructuring of data</b> <ul style="list-style-type: none"> <li>• Combining datasets: Merge and Append two or more data set</li> <li>• Reshaping data: from long to wide and from wide to long</li> </ul>
<b>Feedback (Day 2)</b>			
<b>Time</b>	<b>Session</b>	<b>Speaker/Moderator</b>	<b>Techniques and Learning Outcomes</b>
<b>Day 3, Friday [May 3, 2019]</b>			
9:00 AM-11:00 AM	Basic statistical analysis - I	Dr. Kaushalendra Kumar	<b>Understanding nature and distribution of the variables, applying correlations &amp; identifying associations-</b> <ul style="list-style-type: none"> <li>• Summary statistics: mean, median, mode</li> <li>• Bivariate analysis: Cross tabulation and Chi-squared test</li> <li>• Trivariate table &amp; significance test</li> </ul>
10:45 AM-11:00 AM	<b>Tea Break</b>		
11:00 AM-1:00 PM	Basic statistical analysis - II	Dr. Abhishek Singh & Dr. Kaushalendra Kumar	<b>Applying statistical hypothesis tests in STATA-</b> <ul style="list-style-type: none"> <li>• Student t-test: Mean difference test</li> <li>• Correlation: correlation matrix or covariance matrix</li> </ul>
1:00 PM-2:00 PM	<b>Lunch Break</b>		
2:00 PM-3:30 PM	Graph - I	Dr. William Joe	<b>Plotting and interpreting graphs in STATA-</b> <ul style="list-style-type: none"> <li>• Histogram</li> <li>• Bar graph</li> <li>• Pie chart</li> </ul>



			<ul style="list-style-type: none"> <li>• Overlaying two graphs: twoway</li> <li>• Combining two graphs: graph combine</li> </ul>
3:30 PM-3:45 PM	<b>Tea Break</b>		
3:45 PM-5:30 PM	Graph - II	Dr. William Joe	<b>Plotting and interpreting scatter plots and regression graphs in STATA-</b> <ul style="list-style-type: none"> <li>• Scatter plot</li> <li>• Line plot</li> <li>• Fitted regression line</li> </ul>
<b>Feedback (Day 3)</b>			
<b>Time</b>	<b>Session</b>	<b>Speaker/Moderator</b>	<b>Techniques and Learning Outcomes</b>
<b>Day 4 [May 4, 2019]</b>			
9:00 AM-10:45 AM	Survey weight and survey set	Dr. Abhishek Singh & Dr. Kaushalendra Kumar	<b>Data adjustment by applying weights in STATA-</b> <ul style="list-style-type: none"> <li>• Analytical weight (aweight), frequency weight (fweight), sampling weight (pweight), importance weight (iweight)</li> <li>• Setting data into survey design and survey weight</li> </ul>
10:45 AM-11:00 AM	<b>Tea Break</b>		
11:00 AM-1:30 PM	Regression Analysis	Dr. Abhishek Singh & Dr. William Joe	<b>Identifying the relationship between outcome variable and one or more confounding variables through regression analysis in STATA-</b> <ul style="list-style-type: none"> <li>• Linear regression</li> <li>• Logistic regression</li> <li>• Multinomial logistic regression</li> </ul>
1:30 PM-2:30 PM	<b>Lunch Break</b>		
2:30 PM-3:00 PM	<b>Feedback (Day 4)</b>		
3:00 PM- 5:00 PM	<b>Valedictory function and award of certificates</b>		