
CLRES-2005: Computer Methods for Clinical Research

Date and Times: see below

Course Instructors:
Mark S. Roberts, MD, MPP
robertsm@upmc.edu 692-4826

Course Objectives:

The purpose of this course is to provide a concentrated introduction in the use of computers in the analysis of data from patient-oriented research. Specifically, participants will learn:

- Principles of database design
- Descriptive statistics and exploratory data analysis
- Graphical display of data
- Methods for data manipulation, transformation, merging and case selection
- Use of STATA™ software for conducting data analysis

Location:

Lecture: Parkvale 305 A and B

Lab: Parkvale 305A and B as well as Parkvale 222 (computer lab)

Course Requirements:

The primary requirement for the course will be a series of computer-based exercises. Direct, hands-on completion of the exercises is required.

Computer-based problem sets: 100%

Course Mechanics:

4 - 1.5 hour lectures; 4 - 2.5 hour computer lab sessions, over 3 consecutive days. The course will consist of a combination of didactic sessions followed by computer-lab based exercises designed to develop competence in the use of computers for data manipulation and analysis. Grading will be based on completion of the problem sets.

Teaching Fellows: (Get to know them!!!!!!)

Jesse Hsu yeh7@pitt.edu

Jonathan Yabes yabesjp@upmc.edu

Xinxin Dong xid11@pitt.edu

Emmanuel Sampene ems120@pitt.edu

Recommended textbook:

No text required, “getting Started with STATA” recommended

Recommended Software:

Stata™ 10

Session 1a	7/01/09 1-2:00	Lecture: Principles of databases, data acquisition and storage	Roberts
---------------	----------------	---	---------

Concepts and Topics:

The basic structural forms of database design will be described, including flat-file (row/column), hierarchical and relational. Advantages and disadvantages of the various database types will be presented. Data manipulation, delimiters, record length and other database concepts will be defined and discussed. Data are stored in variables, and the major types of variables (numeric, character, categorical) will be discussed. Basic variable definitions, categorical, ordinal, continuous measures

Required Reading: (prior to session) None

Session 1b	07/01/09 2:00-5:00 pm	Lab: Data entry and manipulation	Teaching Fellows
---------------	--------------------------	---	---------------------

Concepts and Topics:

During this laboratory session, the participants will complete the following exercises:

- Create a simple flat-file database in Stata with keyboard entry
- Inputting an existing dataset with delimiters to create a Stata dataset
- Create data labels for variables and codes
- Listing datasets and checking for errors
- File transfer programs and changing from one file format to another

Required Reading: (prior to session) None

Assignment: In-class exercise

Session 2a	07/02/09 8:30-9:30	Lecture: Descriptive statistics: summarizing variables	Roberts
---------------	-----------------------	---	---------

Concepts and Topics:

This lecture will define and describe measures that are used to characterize and summarize data. Measures of central tendency (mean, median, mode), spread (range, standard deviation, standard error) and shape (skewness and kurtosis) will be defined and presented.

Required Reading: Fundamentals of Biostatistics, Rosner, Chapter 2.

Session 2b	07/02/09 9:30-12:00	Lab: Descriptive Statistics & Basic Graphics in Stata	Teaching Fellows
---------------	------------------------	--	---------------------

Concepts and Topics:

Using Stata to produce measures of central tendency (mean, median, mode), spread (range, standard deviation, standard error) and shape (skewness and kurtosis). Multiple methods of producing tabular output in Stata. Using graphics available in Stata to display data and check for errors.

Assignment: In class assignment. Students will take a new dataset and use Stata to describe and summarize the data.

Session 3a	07/02/09 1:00-2:00	Lecture: Data transformation and manipulation	Roberts
---------------	-----------------------	--	---------

Concepts and Topics:

Data is often not in the form that is required for the analysis of a particular problem. Methods and rational for transforming data will be presented. Techniques of simple transformations will be discussed, as well as methods for dichotomizing or polychotomizing continuous data. Uses of scatter, dot, and box plots and probability density and cumulative density functions for data exploration and display will be reviewed.

Required Reading: (prior to session) none

Assignment

Session 3b	07/02/09 2:00-5:00	Lab: Data recoding and transformation & more graphics	Teaching Fellows
---------------	-----------------------	--	---------------------

Concepts and Topics: Techniques of simple transformations will be presented, as well as methods for dichotomizing or polychotomizing continuous data. Calculation of time between events. Changing string variables into numeric variables. More complex graphics.

Required Reading: Bring the Stata graphics manual with you to the computer lab.

Assignment: In class problem set

Session 4a	07/06/09 1:00-2:00	Lecture: Levels of analysis, merging files, alternative statistical packages and introduction to course datasets	Roberts
---------------	-----------------------	---	---------

Concepts and Topics:

Introduction to clustered data, multilevel data, keyed variables, and the impact of file design on type of analysis possible, efficient file structure for repeated (longitudinal) data. Alternative statistical packages available.

Required Reading: (prior to session) None

Session 4b	07/06/09 2:00-5:00	Lab: Merging files, changing file structure, analysis of real data	Teaching Fellows
---------------	-----------------------	---	---------------------

Concepts and Topics:

Merging data files to add variables, add cases, and use of keyed files (look-up tables) to add variables. Changing file structures to facilitate analysis. Looping procedures in Stata, writing and saving command files in Stata. Links to word processing of graphs and output files from Stata. Comparison of the advantages and disadvantages of SPSS and STATA

Required Reading: (prior to session) None

Assignment: In class problem-set