## Course Information

**Term:** Summer  
**Credits:** 1

### Class Meetings

There is 1 required in-class session.

**Check CourseWeb for your group assignment.**

**You do not need to attend any of the lab sessions; they are optional to help you if you need it.** You can do all or some of the lab assignments from your own computer if you wish.

**Location:** All groups meet in Parkvale 222

## Course Description

The purpose of this course is to provide a concentrated introduction to the use of computers in the analysis of data from patient-oriented research. Specifically, you 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 (version 15) software for conducting data analysis
Course Format

This course is offered in a hybrid (partially online/partially face-to-face) format.

There is one mandatory in-person session where the instructor will describe 1) conceptual issues in designing databases and data entry systems, and 2) the structure of how Stata works.

You will also 1) watch video lectures on your own time and 2) complete lab assignments (in Stata) which expand on the topics covered in the in-class lecture. The in-class lecture will also describe the connection between the material in the videos and labs.

Course Assignments:

- **Attend the 3-hour in-class session** which is at a time specific to your group. After, or while you are attending the class, you will need to fill out a short worksheet to reinforce the main points of the class.
- **Watch the video lessons** and complete the associated worksheets on each of the 4 modules:
  - Data Acquisition & Representation
  - Data Description
  - Data Manipulation
  - Merging Multiple Data Sets
- **Complete the 4 Stata lab assignments** on the following topics:
  - Data Entry and Manipulation
  - Descriptive Statistics & Basic Graphs in Stata
  - Data Recoding, Transformations & More Graphs
  - Levels of analysis, merging, appending, and reshaping data files

You can watch the videos either before or after the in-class session. Watching the videos before the in-class session may help you form questions. On the other hand, you may be able to go through the videos faster if you wait until after the in-class session, so it’s up to you. I would suggest watching them once before the in-class session to formulate any questions, and then once after, during which time you can complete the associated worksheet.

I would suggest completing the lab assignments last (i.e. after attending the in-class and completing the associated worksheet), and after the watching the videos and completing those worksheets. However, you can do the work in any order you would like.

All assignments are due 1 week after the course ends.

The hybrid format may be a different format than you are familiar with, but this format has some clear advantages. Specifically, it allows you to:

- Engage with the content at your own pace and on your own schedule
- Rewind, re-watch, or jump ahead as necessary
- Use class time for interaction and clarification
- Spend less total time in class
Required Materials

Readings: There is no required textbook, or readings, for this course.

Software: Stata 15 will be available on the lab computers. Directions on how to purchase Stata through the University of Pittsburgh can be found here: http://technology.pitt.edu/category/software-for-students

Course Requirements

Your final letter grade will be based on the following:

- In-class session: 5% for attendance, 5% for the corresponding worksheet
- Module worksheets/video lectures: 16% for 1st worksheet; 8% for each of the other 3 worksheets (40% total)
- 4 lab exercises: 12.5% each (50% total)

All assignments must be turned in 1 week after class ends. There will be a 10% penalty per day for assignments turned in late.

Grading

Final letter grades will be assigned according to the following grading scale:

A = 90-100  B = 80-89  C = 70-79  D = 60-69  F = 0-59

Help

Please don’t hesitate to contact your instructor via email if you have any questions about course content. You can expect an answer within 24-48 hours.

If you have any questions about scheduling or technology (CourseWeb, GoToMeeting, the sign-in system for attendance, etc.) please contact the course administrator, Juliana Tambellini at: tambellinijm2@upmc.edu.

Course Policies

You are responsible for knowing and following these course policies. Please read them carefully.

Assignment Policy: You are expected to do the work for this course independently and turn it in by the designated deadline.

Attendance Policy: Attendance is mandatory for the first in-class session. If you miss that session, you will need to see another student in the class to get the notes to then complete the worksheet. Please sign in on the computer provided in suite lobby; this records your
attendance. If you encounter a problem with the sign-in system, please contact the course instructor(s) immediately. **Attendance at the other the lab sessions is optional.**

**Recording Policy**

ICRE-Produced Recordings: ICRE faculty and/or staff may video and/or audio record this course (hereby referred to as “Recordings”). By enrolling in this course, you hereby give the University of Pittsburgh and the Institute for Clinical Research Education, through its faculty, employees, agents, licenses or assigns, the irrevocable and worldwide right to use your name, voice, likeness and/or image in all forms and media (to include internet websites and online course website). You waive your right to inspect or approve the finished version(s) of the Recordings, including any copy that may be created in connection therewith. You understand that you will not be paid for your participation in the Recording and that you are not entitled to your own copy of the Recording. You understand that the University of Pittsburgh is not responsible for any unauthorized use of the Recording. You have read this syllabus and have no questions about the contents and are an adult over the age of 18.

Student-Produced Recordings: To ensure the free and open discussion of ideas, **students may not record classroom lectures, discussion and/or activities without advanced written permission of the instructor.** Any such recording properly approved in advance can be used solely for the student’s own private use.

**Academic Integrity**

You are expected to comply with the University of Pittsburgh’s Policy on Academic Integrity detailed here: http://www.provost.pitt.edu/info/ai1.html. Any student suspected of violating this obligation for any reason during the semester will be required to participate in the procedural process, initiated at the instructor level, as outlined in the University Guidelines on Academic Integrity. This may include, but is not limited to, the confiscation of the examination of any individual suspected of violating University Policy. Furthermore, no student may bring any unauthorized materials to an exam, including dictionaries and programmable calculators.

**Disabilities**

If you have a disability for which you are requesting an accommodation, you are encouraged to contact both your instructor and the Office of Disability Resources and Services, 216 William Pitt Union, 412-648-7890 / 412-3837355 (TTY), as early as possible in the term. Disability Resources and Services will verify your disability and determine reasonable accommodations for this course.
### Competencies addressed in this course:

<table>
<thead>
<tr>
<th>Competency</th>
<th>Module 1</th>
<th>Module 2</th>
<th>Module 3</th>
<th>Module 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management and Biomedical Informatics: Identify pertinent issues in the construction of effective data security and management plans for various research designs.</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Data Management and Biomedical Informatics: Organize datasets (variable display and structure) appropriately for given statistical techniques.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Applied Analytic Techniques: Describe appropriate data analysis plans for addressing specific research questions.</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Applied Analytic Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

The specific goals of the course are broken out below by each of the Modules in the video lectures. The lab assignments and in-class session will re-inforce and expand the same concepts as listed for the video lectures.

### Module 1: Data Acquisition and Representation

**At the conclusion of this session, you should be able to:**

- Create a simple flat-file in Stata with keyboard entry
- Input an existing dataset with delimiters to create a Stata dataset
- Create data labels for variables and codes
- List datasets and check for errors
- Use file transfer programs and change from one file format to another

### Module 2: Data Description

**At the conclusion of this session, you should be able to:**

- Use Stata to produce measures of central tendency (mean, median, mode), spread (range, standard deviation, standard error), and shape (skewness, kurtosis).
- Produce tabular output in Stata using multiple methods.
- Use graphics available in Stata to display data and check for errors
Module 3: Data Manipulation

At the conclusion of this session, you should be able to:

- Perform simple transformations
- Dichotomize or polychotomize continuous data
- Calculate time between events
- Change string variables into numeric variables
- Create more complex graphics

Module 4: Merging Multiple Data Sources

At the conclusion of this session, you should be able to:

- Merge data files to add variables, add cases, and use keyed files (look-up tables) to add variables
- Change file structures to facilitate analysis
- Write and save command files in Stata
- Link to word processing of graphs and output files from Stata