

PubH 6617 - 001
Practical Methods for Secondary Data Analysis
Fall - 2017

Credits: 3
Meeting Days: Tuesdays and Thursdays
Meeting Time: 8:15-9:30 am
Meeting Place: Mayo C381
Instructor: J. Michael Oakes, PhD
Office Address: 431 West Bank Office Building (WBOB)
Office Phone: 612-624-6855
E-mail: oakes007@umn.edu
Office Hours: By appointment only; Contact Libby @ fortn002@umn.edu

I. Course Description

To introduce masters and doctoral students to methods for finding, transferring and processing existing data sources. The course emphasizes practical approaches to *pre-statistical* data processing and analysis with Stata statistical software on a PC with a MS Windows operating system, though Macintosh devices work very similarly. Advantages and limitations of several national data resources are discussed.

II. Course Prerequisites

Prospective students must be matriculated graduate or professional students. It is best if students have passed a graduate level (bio)statistics course but is not required. Instructor reserves right to waive requirements.

III. Course Goals and Objectives

Upon completion of this course the student should be able to:

- Better appreciate the effort/steps required to take a raw data set and produce an analytic data set
- Master the idea of Workflow for data processing and analysis
- Understand the crucial role of scientific transparency and gain ability to write clear, transparent code
- Master procedures for analytic project backup and related security concerns

- Appreciate the vast number and variety of existing data resources potentially useful in public health investigations, as well as methods for finding and exploiting these data.
- Understand basic and moderately advanced data structures, including the binary and hexadecimal number systems, flat-files, relational and hierarchical data resources.
- Be able to read into Stata at least moderately complex ASCII data
- Be able to exploit existing data conversion software.
- Be aware of central issues in complex sampling designs
- Be familiar with U.S. Census data
- Be aware of the National Health Interview Survey

IV. Methods of Instruction and Work Expectations

Class sessions will include lectures, group discussions, and interactive computer exercises. Guest presenters might include professionals with particular expertise in a given subject area.

This is a 3-credit course. Since 1 credit is roughly equal to 3 hours of effort per week, it follows that about 6 hours of homework per week should be expected. This means about 90 hours of homework per semester.

Students must complete all of the following to earn a passing grade in the course:

- Satisfactory completion of all assigned readings and class participation:* Assigned readings/websites must be completed before the class in which they will be discussed. Readings are critical to useful class discussions.
- Satisfactory completion of computer homework assignments:* Homework assignments will include essential programming and data manipulation techniques. Assignments are due at the start of class on the day they are due. Late homework will not be accepted. Homework must be presented in a neat and careful manner, as one would present a professional report. Detailed directions will be offered.
- Satisfactory completion of two in-class examinations.* The exams will cover all material from the course up to that data, and will be cumulative.

V. Course Text and Readings

- (a) While NOT required, five texts are recommended for more advanced students.

Long, J. Scott. 2009. *The Workflow of Data Analysis using Stata*. Stata Press

Boslaugh, Sarah. 2007. *Secondary Data Sources for Public Health*. New York: Cambridge University Press.

Long, J. Scott and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata, Second Edition*. College Station, TX: Stata Press. (a 'must-have' for folks using logistic regression and Poisson models) [there is an updated edition now]

Berk, Richard. 2004. *Regression Analysis. A constructive critique*. Thousand Oaks, CA: Sage Publications (a superb text outlining the limits of regression analysis)

(b) **Students must have access to a computer with Stata 15** and have internet access.

| | | |
|--|----------|----------|
| 6 month lease of "Stata/IC" | \$45.00 | |
| 12 month lease of "Stata/IC" | \$89.00 | Optional |
| Perpetual license for "Stata/IC" | \$198.00 | Optional |

For directions on how to purchase Stata, <https://www.stata.com/order/new/edu/gradplans/student-pricing/>

VI. *Planned Readings and Tutorials*

WEEK 1: INTRODUCTION

Getting Stata Manual (GSM) Chapters 1-5

WEEK 2: WORKFLOW, DO-FILES AND LOGS

Review <http://www.ats.ucla.edu/stat/stata/modules/syntax.htm>

WEEK 3: LABELING STUFF; UTILITIES

Long – Chapter 2 Planning, organizing, and documenting

Surf and Get (very!) familiar with UCLA Stata website

<http://www.ats.ucla.edu/stat/stata/>

Review www.stattransfer.com

WEEK 4: CREATING NEW VARIABLES, FUNCTIONS, DATES

Cox, Nicholas J. 2002. "Speaking Stata: On getting functions to do the work." *Stata Journal* 2:411-427.

Cox, NJ. 2002. "Speaking Stata: On numbers and strings." *Stata Journal* 2:314-329.

UCLA dates: <http://www.ats.ucla.edu/stat/stata/modules/dates.htm>

Elapsed dates: <http://stata.com/support/faqs/data/dateseq.html>

WEEK 5: BASIC DATASET MANIPULATION

By group processing: Cox, Nicholas J. 2002. "Speaking Stata: How to Move Step by: Step." *Stata Journal* 2:86-102.

Review "Reading Hierarchical Data" <http://www.stata.com/support/faqs/data/hier.html>

Review <http://www.ats.ucla.edu/stat/stata/modules/collapse.htm>

Review <http://www.ats.ucla.edu/stat/stata/modules/combine.htm>

WEEK 6: LOOPS & GRAPHS

Review <http://www.ats.ucla.edu/stat/stata/modules/acrossvars.htm>

Review entire UCLA Graphics module <http://www.ats.ucla.edu/stat/stata/modules/>

WEEK 7: EXAM

WEEK 8: INFERENCE AND REGRESSION

No assigned readings; Lecture notes only

WEEK 9: REGRESSION IN STATA; INTRO TO NHIS

Review <http://www.cdc.gov/nchs/nhis.htm>

WEEK 10: NHIS

Review <http://www.ihis.us/ihis/>

WEEK 11: ENCRYPTION & DATA SECURITY

Oakes, J. Michael. 2002. "Risks and Wrongs in Social Science Research: An Evaluator's Guide to the IRB." *Evaluation Review* 26:443-478.

Surf <http://www.irb.umn.edu/>

Surf about file encryption at: <http://computer.howstuffworks.com/encryption1.htm>

WEEK 12: CENSUS (AGGREGATE)

Wallman, Katherine. 1998. "Data on Race and Ethnicity: Revising the Federal Standard." *The American Statistician* 52:31-35.

WEEK 13: CENSUS (AGGREGATE)

Review <http://www.ipums.umn.edu>

WEEK 14: IPUMS

No new assigned reading.

WEEK 15: EXAM

No new assigned reading.

VII. Course Outline/Weekly Schedule

Note that the following schedule is tentative and is subject to change, even substantial change, as student abilities and interests emerge. The nature of this course requires that the following schedule not be binding.

WEEK 1: INTRODUCTION

(9/5 – 9/7) Introduction to Stata – window basics, especially online help; Introduction to Stata and hard-drives, folders/directories

WEEK 2: WORKFLOW, DO-FILES AND LOGS, BASICS

(9/12) Introduction to do-files

(9/14) Basic programming commands

WEEK 3: LABELING STUFF; UTILITIES & TRANSFERRING DATA BETWEEN PROGRAM

(9/19) do-file comments, variables and data labels, and notes

(9/21) Utilities (Winzip, Stat-Transfer); MS Excel into Stata

WEEK 4: CREATING NEW VARIABLES; FUNCTIONS; DATES

(9/26) GUEST PROFESSOR - Missing values; generate; functions

(9/28) egen; dates

WEEK 5: BASIC DATASET MANIPULATION

(10/3) NO CLASS, Vlog: Sorting; Merging; Appending; Collapsing

(10/5) Selective processing (if, in); Hierarchical data (by-group processing);

WEEK 6: LOOPS & GRAPHS

(10/10) foreach loops

(10/12) NO CLASS, Vlog: Graphs and plots

WEEK 7: MID-TERM EXAM

(10/17) NO CLASS, Vlog: Catch up before m

(10/19) Exam review

WEEK 8: INFERENCE AND REGRESSION ANALYSIS

(10/24) Exam #1

(10/26) Statistical inference

WEEK 9: REGRESSION IN STATA; INTRODUCTION TO NHIS

(10/31) Regression in Stata and

(11/2) presentation of results

WEEK 10: NHIS

(11/7) Introduction to the National Health Interview Survey (NHIS)

(11/9) Analysis of the National Health Interview Survey (NHIS)

WEEK 11: ENCRYPTION; REVIEW STUFF

(11/14) Encryption and data security

(11/16) Review and catch up

WEEK 12: US CENSUS DATA (AGGREGATE)

(11/21) Intro to Census: Basic Data Structure, Files, Hierarchy, Imputation

(11/23) No Class (Gobble, Gobble...)

WEEK 13: US CENSUS DATA (AGGREGATE)

(11/28) Manipulation and Analysis with Aggregate Census 2000

(11/30) Manipulation and Analysis with Aggregate Census 2000

WEEK 14: IPUMS

(12/5) Introduction to IPUMS

(12/7) IPUMS analysis

WEEK 15: REVIEW AND EXAM

(12/12) Review of everything

(12/14) Final Exam

VII. Evaluation and Grading

Letter grades and associated points are awarded in this course as follows below. These will appear in the student's official transcript. See <http://www1.umn.edu/usenate/policies/gradingpolicy.html> for details.

Pass/Fail Grading: An alternative to traditional A-F scale grades is the S/N grading scale. The "S" grade does not carry points but credit will count toward completion of student's degree if permitted by college or program. An "N" is given for student's exercising the S/N grading option but who fail to meet minimum course requirements. Students may change grading options without written permission as specified by the University and without penalty during the initial registration period or during the first two weeks of the semester. The grading option may not be changed after the second week of the term.

Grades will not be normed within the class; it is possible for all students to receive an "A."

| Grade Mapping | | | |
|----------------------|---------------------|---------------------|--|
| Class Points | Letter Grade | Grade Points | Interpretation (relative to requirements) |
| >94 | A | 4.00 | Outstanding |
| 90-94 | A- | 3.67 | |
| 87-89 | B+ | 3.33 | |
| 83-86 | B | 3.00 | Significantly Above |
| 80-82 | B- | 2.67 | |
| 77-79 | C+ | 2.33 | |
| 73-76 | C | 2.00 | Satisfactory |

| | | | |
|-------|-----|------|-----------------------------|
| 70-72 | C- | 1.67 | Less than Satisfactory |
| 70+ | S | - | At Least Satisfactory |
| - | N/F | - | Not Satisfactory, No Credit |
| - | I | - | Incomplete |

The maximum number of points a student can earn is 100. **Class participation** demonstrating comprehension of assigned readings and notes will account for **20 points (20%)** of the final grade. One half a point per week will be given for active participation, such as asking probing questions. Students will receive a 0.0 or 0.5 per week only. This means that class attendance is necessary but not sufficient to earn the maximum allowable points. **Four homework assignments** demonstrating knowledge and ability to perform required tasks will account for **20 points (20%)** of the final grade (5% each). **Three in class exams** will be given and worth **60 points (60%)** of the final grade (20% each). Numeric grades will map to letter grades as per the table above.

For additional information, please refer to:

<http://policy.umn.edu/Policies/Education/Education/GRADINGTRANSCRIPTS.html>.

Course Evaluation

The SPH will collect student course evaluations electronically using a software system called CoursEval: www.sph.umn.edu/courseeval. The system will send email notifications to students when they can access and complete their course evaluations. Students who complete their course evaluations promptly will be able to access their final grades just as soon as the faculty member renders the grade in SPHGrades: www.sph.umn.edu/grades. All students will have access to their final grades through OneStop two weeks after the last day of the semester regardless of whether they completed their course evaluation or not. Student feedback on course content and faculty teaching skills are an important means for improving our work. Please take the time to complete a course evaluation for each of the courses for which you are registered.

Incomplete Contracts

A grade of incomplete "I" shall be assigned at the discretion of the instructor when, due to extraordinary circumstances (e.g., documented illness or hospitalization, death in family, etc.), the student was prevented from completing the work of the course on time. The assignment of an "I" requires that a contract be initiated and completed by the student before the last official day of class, and signed by both the student and instructor. If an incomplete is deemed appropriate by the instructor, the student in consultation with the instructor, will specify the time and manner in which the student will complete course requirements. Extension for completion of the work will not exceed one year (or earlier if designated by the student's college). For more information and to initiate an incomplete contract, students should go to SPHGrades at: www.sph.umn.edu/grades.

University of Minnesota Uniform Grading and Transcript Policy

A link to the policy can be found at onestop.umn.edu.

VIII. Other Course Information and Policies

Grade Option Change (if applicable):

For full-semester courses, students may change their grade option, if applicable, through the second week of the semester. Grade option change deadlines for other terms (i.e. summer and half-semester courses) can be found at onestop.umn.edu.

Course Withdrawal:

Students should refer to the Refund and Drop/Add Deadlines for the particular term at onestop.umn.edu for information and deadlines for withdrawing from a course. As a courtesy, students should notify their instructor and, if applicable, advisor of their intent to withdraw.

Students wishing to withdraw from a course after the noted final deadline for a particular term must contact the School of Public Health Office of Admissions and Student Resources at sph-ssc@umn.edu for further information.

Student Conduct Code:

The University seeks an environment that promotes academic achievement and integrity, that is protective of free inquiry, and that serves the educational mission of the University. Similarly, the University seeks a community that is free from violence, threats, and intimidation; that is respectful of the rights, opportunities,

and welfare of students, faculty, staff, and guests of the University; and that does not threaten the physical or mental health or safety of members of the University community.

As a student at the University you are expected adhere to Board of Regents Policy: *Student Conduct Code*. To review the Student Conduct Code, please see:

http://regents.umn.edu/sites/default/files/policies/Student_Conduct_Code.pdf.

Note that the conduct code specifically addresses disruptive classroom conduct, which means "engaging in behavior that substantially or repeatedly interrupts either the instructor's ability to teach or student learning. The classroom extends to any setting where a student is engaged in work toward academic credit or satisfaction of program-based requirements or related activities."

Use of Personal Electronic Devices in the Classroom:

Using personal electronic devices in the classroom setting can hinder instruction and learning, not only for the student using the device but also for other students in the class. To this end, the University establishes the right of each faculty member to determine if and how personal electronic devices are allowed to be used in the classroom. For complete information, please reference:

<http://policy.umn.edu/Policies/Education/Education/STUDENTRESP.html>.

Scholastic Dishonesty:

You are expected to do your own academic work and cite sources as necessary. Failing to do so is scholastic dishonesty. Scholastic dishonesty means plagiarizing; cheating on assignments or examinations; engaging in unauthorized collaboration on academic work; taking, acquiring, or using test materials without faculty permission; submitting false or incomplete records of academic achievement; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement; altering, forging, or misusing a University academic record; or fabricating or falsifying data, research procedures, or data analysis. (Student Conduct Code:

http://regents.umn.edu/sites/default/files/policies/Student_Conduct_Code.pdf) If it is determined that a student has cheated, he or she may be given an "F" or an "N" for the course, and may face additional sanctions from the University. For additional information, please see:

<http://policy.umn.edu/Policies/Education/Education/INSTRUCTORRESP.html>.

The Office for Community Standards has compiled a useful list of Frequently Asked Questions pertaining to scholastic dishonesty <https://communitystandards.umn.edu>. If you have additional questions, please clarify with your instructor for the course. Your instructor can respond to your specific questions regarding what would constitute scholastic dishonesty in the context of a particular class-e.g., whether collaboration on assignments is permitted, requirements and methods for citing sources, if electronic aids are permitted or prohibited during an exam.

Makeup Work for Legitimate Absences:

Students will not be penalized for absence during the semester due to unavoidable or legitimate circumstances. Such circumstances include verified illness, participation in intercollegiate athletic events, subpoenas, jury duty, military service, bereavement, and religious observances. Such circumstances do not include voting in local, state, or national elections. For complete information, please see:

<http://policy.umn.edu/Policies/Education/Education/MAKEUPWORK.html>.

Appropriate Student Use of Class Notes and Course Materials:

Taking notes is a means of recording information but more importantly of personally absorbing and integrating the educational experience. However, broadly disseminating class notes beyond the classroom community or accepting compensation for taking and distributing classroom notes undermines instructor interests in their intellectual work product while not substantially furthering instructor and student interests in effective learning. Such actions violate shared norms and standards of the academic community. For additional information, please see: <http://policy.umn.edu/Policies/Education/Education/STUDENTRESP.html>.

Sexual Harassment:

"Sexual harassment" means unwelcome sexual advances, requests for sexual favors, and/or other verbal or physical conduct of a sexual nature. Such conduct has the purpose or effect of unreasonably interfering with an individual's work or academic performance or creating an intimidating, hostile, or offensive working or academic environment in any University activity or program. Such behavior is not acceptable in the

University setting. For additional information, please consult Board of Regents Policy:
<http://regents.umn.edu/sites/default/files/policies/SexHarassment.pdf>

Equity, Diversity, Equal Opportunity, and Affirmative Action:

The University will provide equal access to and opportunity in its programs and facilities, without regard to race, color, creed, religion, national origin, gender, age, marital status, disability, public assistance status, veteran status, sexual orientation, gender identity, or gender expression. For more information, please consult Board of Regents Policy:
http://regents.umn.edu/sites/default/files/policies/Equity_Diversity_EO_AA.pdf.

Disability Accommodations:

The University of Minnesota is committed to providing equitable access to learning opportunities for all students. The Disability Resource Center Student Services is the campus office that collaborates with students who have disabilities to provide and/or arrange reasonable accommodations.

If you have, or think you may have, a disability (e.g., mental health, attentional, learning, chronic health, sensory, or physical), please contact DRC at 612-626-1333 or drc@umn.edu to arrange a confidential discussion regarding equitable access and reasonable accommodations.

If you are registered with DS and have a current letter requesting reasonable accommodations, please contact your instructor as early in the semester as possible to discuss how the accommodations will be applied in the course.

For more information, please see the DS website, <https://diversity.umn.edu/disability/>.

Mental Health and Stress Management:

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance and may reduce your ability to participate in daily activities. University of Minnesota services are available to assist you. You can learn more about the broad range of confidential mental health services available on campus via the Student Mental Health Website: <http://www.mentalhealth.umn.edu>.

The Office of Student Affairs at the University of Minnesota:

The Office for Student Affairs provides services, programs, and facilities that advance student success, inspire students to make life-long positive contributions to society, promote an inclusive environment, and enrich the University of Minnesota community.

Units within the Office for Student Affairs include, the Aurora Center for Advocacy & Education, Boynton Health Service, Central Career Initiatives (CCE, CDes, CFANS), Leadership Education and Development – Undergraduate Programs (LEAD-UP), the Office for Fraternity and Sorority Life, **the Office for Community Standards**, the Office for Student Engagement, the Parent Program, Recreational Sports, Student and Community Relations, the Student Conflict Resolution Center, the Student Parent HELP Center, Student Unions & Activities, University Counseling & Consulting Services, and University Student Legal Service.

For more information, please see the Office of Student Affairs at <https://osa.umn.edu/>.

Academic Freedom and Responsibility, for courses that involve students in research

Academic freedom is a cornerstone of the University. Within the scope and content of the course as defined by the instructor, it includes the freedom to discuss relevant matters in the classroom and conduct relevant research. Along with this freedom comes responsibility. Students are encouraged to develop the capacity for critical judgment and to engage in a sustained and independent search for truth. Students are free to take reasoned exception to the views offered in any course of study and to reserve judgment about matters of opinion, but they are responsible for learning the content of any course of study for which they are enrolled.* When conducting research, pertinent institutional approvals must be obtained and the research must be consistent with University policies.

Reports of concerns about academic freedom are taken seriously, and there are individuals and offices available for help. Contact the instructor, the Department Chair, your adviser, the associate dean of the

college, (Dr Kristin Anderson, SPH Dean of Student Affairs), or the Vice Provost for Faculty and Academic Affairs in the Office of the Provost.

** Language adapted from the American Association of University Professors "Joint Statement on Rights and Freedoms of Students".*

Student Academic Success Services (SASS): <http://www.sass.umn.edu>:

Students who wish to improve their academic performance may find assistance from Student Academic Support Services. While tutoring and advising are not offered, SASS provides resources such as individual consultations, workshops, and self-help materials.

Template update 8/2017