UNIVERSITY OF TENNESSEE
COLLEGE OF SOCIAL WORK

Analysis of Social Work Data II   Instructor: John G. Orme, Ph.D.
SW 606 (3 credit hours)   Office: 313 Henson Hall
Spring, 2018   Phone: 865-974-7503
Wednesday: 9:05 – 12:05 & 1:25 – 4:25   Office Hours: Thursday, 1:30-3:30
Room: Henson Hall, Room 418   e-mail: jorme@utk.edu
Instructor's Web Page:
http://www.csw.utk.edu/faculty/pages/orme/index.htm
Teaching Assistant: Aaron Brown, MSW, Ph.D. Student/Graduate Teaching Assistant
abrown89@tennessee.edu
Sarah Browning, Administrative Specialist: sbrowni2@utk.edu (865) 974-6481

Prerequisites

Successful completion of SW 605 or consent of the instructor.

Ph.D. Program Mission

The purpose of social work education at the doctoral level is to develop an understanding and appreciation of the scientific method so that graduates can improve and extend the knowledge base of social work practice, and inform social policy, through empirical research. Graduates of doctoral programs assume leadership roles in social work education, research, and practice.

The Ph.D. program in social work provides training in social and behavioral science research methods and the opportunity for students to apply these methods to social problems and social work practice areas. Doctoral research is directed toward developing knowledge that can be used by social work educators, practitioners, administrators, and policy makers. Graduates of the Ph.D. program are prepared to contribute to the improvement of the design and implementation of social services and to develop and disseminate social work knowledge through research and teaching.

Ph.D. Program Competencies

➢ Upon completion of the UT CSW Ph.D. program, students will be able to:
➢ Demonstrate expert knowledge in a focused substantive area relevant to social work.
➢ Conduct independent and original scientific research that advances knowledge in a substantive area.
➢ Communicate scientific findings in an effective way to a range of audiences (from lay persons to other scientists).
➢ Secure funding for a substantive research agenda.
➢ Teach students the knowledge, skills, and values they need to be proficient social workers in a substantive area.

Code of Conduct

It is the student's responsibility to have read the College of Social Work Ethical Academic and Professional Conduct Code that is in the College of Social Work Ph.D. Student Handbook (www.csw.utk.edu).

The Honor Statement

An essential feature of the University of Tennessee is a commitment to maintaining an atmosphere of intellectual integrity and academic honesty. As a student of the university, I pledge that I will neither knowingly give nor receive any inappropriate assistance in academic work, thus affirming my own personal commitment to honor and integrity.

University Civility Statement

Civility is genuine respect and regard for others: politeness, consideration, tact, good manners, graciousness, cordiality, affability, amiability and courteousness. Civility enhances academic freedom and integrity, and is a prerequisite to the free exchange of ideas and knowledge in the learning community. Our community consists of students, faculty, staff, alumni, and campus visitors. Community members affect each other's well-being and have a shared interest in creating and sustaining an environment where all community members and their points of view are valued and respected. Affirming the value of each member of the university community, the campus asks that all its members adhere to the principles of civility and community adopted by the campus (http://civility.utk.edu/)

Classroom Courtesy

You may bring your laptops to class for taking notes and doing appropriate classroom work. It is not appropriate to cruise the internet, check your email, etc., during class. The instructor reserves the right to disallow the use of laptops in class if students use them inappropriately (e.g. emails, or web surfing). All cell phones and smart phones should be silenced and put away during class and text messaging is strictly prohibited. If you have a special circumstance, talk to the instructor in advance. All mp3 players (iPods, etc.) are to be off and stored during class. This includes removal of all headphones, ear buds, and other listening devices from ears, head, and neck during class time.

Ethical Guidelines

The American Statistical Association's Ethical Guidelines for Statistical Practice are intended to help statistics practitioners make and communicate ethical decisions. Clients, employers, researchers, policymakers, journalists, and the public should be urged to expect statistical practice to be conducted in accordance with these guidelines and to object when it is not. While learning how to apply statistical theory to problems, students should be encouraged to use these guidelines, regardless of whether their target professional specialty will be "statistician." Employers, attorneys, and other clients of statistics practitioners have a responsibility to provide a moral environment that fosters the use of these ethical guidelines (American Statistical Association: http://www.amstat.org/about/ethicalguidelines.cfm).

Disability Services

If you need course adaptation or accommodations because of a documented disability or if you have emergency information to share, please contact the Office of Disability Services at 100 Dunford Hall at 974-6087. This will ensure that you are properly registered for services.

Dimensions of Diversity

The College of Social Work and the University of Tennessee welcome and honor all people. In accordance with the U.S. National Association of Social Workers (NASW) and the U.S. Council on Social Work Education (CSWE 2015 Educational Policy Statement), "the dimensions of diversity are understood as the intersectionality of multiple factors including" age, class, color, culture, mental or physical disability and ability, ethnicity, gender, gender expression, gender identity, immigration status, marital status, national origin, political ideology, race, regionality, religion and spirituality, sex, sexual orientation, and tribal sovereign status. The College values intellectual curiosity, pursuit of knowledge, and academic freedom and integrity. "A person's diverse life experiences may include oppression, poverty, marginalization, and alienation as well as privilege, power, and acclaim" (CSWE 2015 Educational Policy Statement). The College of Social Work promotes social justice and social change, and strives to end discrimination, oppression, poverty, and other forms of social injustice.

Office of Information Technology (OIT) (http://oit.utk.edu)

Phone: (865) 974-9900
Online: help.utk.edu
Walk-in: The Commons at Hodges Library
Contact OIT for help with:
➢ Your NetID/Password
➢ Your UT email account
➢ Connecting to the campus network
➢ Student computer support
➢ Technology consults
➢ Workshops
➢ Computer labs
➢ MyUTK
➢ Online@ut (Blackboard Learn)

Intercollegiate Graduate Statistics Program (IGSP)

The Intercollegiate Graduate Statistics Program (IGSP) (http://igsp.bus.utk.edu/about.asp) is a formal University of Tennessee academic program established to enable students to earn either a minor or an M.S. in Statistics simultaneously with a master's or doctoral degree in another department. Approved coursework taken to meet doctoral requirements in the student's home department may also be credited toward the M.S. in Statistics. Similarly, approved coursework in statistics taken to meet the requirements for a master's or doctoral degree in another department may also count toward the minor in Statistics. The program is open to graduate students in all departments which have an approved minor and/or M.S. joint major curriculum offered through the program. The program is administered by an Executive Committee, consisting of college representatives from all colleges with approved programs, with advisory input from the program faculty.

Course Description

Conceptual understanding of fixed-effects linear regression models and ability to specify, test, interpret, and present results of linear regression models. Extension of linear regression to discrete dependent variables in the framework of the generalized linear model.

Course Competencies

By the completion of this course, the students are expected to be able to demonstrate (through course activities, assignments, and/or exams):

Competencies:

➢ Use SPSS to input, manage, and analyze multiple regression (MR) models defined broadly to include linear regression and multiple regression with discrete dependent variables
➢ Understand and discuss the relationship among apparently diverse statistical methods in the context of the General and Generalized Linear Models
➢ Identify salient characteristics of variables (e.g., continuous or discrete, skewed, censored, reliability, validity), samples (e.g., unequal probability of selection requiring weighting, multilevel sampling), and designs (e.g., random assignment, cross-sectional or longitudinal) important in the selection and interpretation of multiple regression models
➢ Understand different types of effects that can be tested with multiple regression models (e.g., mediation, moderation, linear, curvilinear effects) and be able to test and interpret these different types of effects
➢ Formulate hypotheses suitable for null hypothesis significance testing with multiple regression models
➢ Select, compute, and interpret effect sizes and associated confidence intervals for multiple regression models
➢ Identify and examine the viability of assumptions underlying any and all statistical analyses used, potential consequences of violating these assumptions, and methods for handling important violations for multiple regression models
➢ Handle missing data appropriately and discuss strengths and limitations of different methods for handling missing data with multiple regression models
➢ Use best practices to communicate results of multiple regression models in an effective way to a range of audiences (from lay persons to other scientists) (i.e., know what to report and how to report it) (e.g., select, construct, and interpret tables and graphs to represent appropriate salient characteristics of variables and relationships among variables)
➢ Critically examine and interpret the use of multiple regression models reported in major social work and social science journals

Statistical Tutoring and Support

As graduate students, you are entitled to 10hrs of free statistical tutoring and support from Research Computing Support per semester. Keep in mind, however, that this service will not help you with homework assignments, quizzes, or tests. They will only answer general questions related to statistics and research support. For basic problems, you can send email to

stathelp@utk.edu or use the walk-in service at 200 Stokely Management Center, UT Knoxville. For complex problems, call them at 974-8975 or send email to stathelp@utk.edu.

Evaluation of the Course

At the end of the semester, a formal evaluative questionnaire regarding both the curriculum and my instruction will be administered. The information gathered from this process will be used to improve future courses and instruction. Additionally, I strongly encourage you to provide feedback to me during the semester either in person or anonymously—I am here to teach you and want to do it well!

Class Materials on dropbox.com

At the beginning of the semester you will receive an email from me notifying you that you can access all of the course materials in a folder on dropbox.com. Please download and save all of these materials within the first two weeks of class and use them from your computer, not the dropbox folder.

The dropbox.com folder contains each week's Power Point presentation, SPSS output, out-of-class exercise, and other materials. You are responsible for bringing this material to class each week. The folder also contains the course syllabi, the readings (not texts) listed below as required reading (in pdf format), and the data sets described below that you can use for the out-of-class exercises. Also, it contains the following, both of which provide a lot of information about using SPSS:

Field, A. (2013). *Discovering statistics using SPSS* (4rd ed.). Los Angeles: Sage. (Chapter 3)

Levesque, R. (2011). *SPSS programming and data management: A guide for SPSS and SAS users.* IBM.

http://www.spsstools.net/

Calculator

For the most part you will use SPSS to analyze data. However, you will need to work through some formulas with a calculator, and so you will need an inexpensive calculator. You cannot use a calculator on your phone or computer.

SPSS

SPSS is the primary statistical software that you will use in this course. You can purchase SPSS on CD from the computer bookstore in the University Center for a nominal cost. Or, you can download it (and other statistical software) free from the following website, although the download time can be pretty long sometimes:

https://webapps.utk.edu/oit/softwaredistribution/

Following is an especially good book detailing how to use SPSS for a wide range of statistical analyses:

Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). Los Angeles: Sage.

You also can use SPSS (and a lot of other statistical software) directly through the Citrix server. Please first read the instructions at:

https://oit.utk.edu/research/documentation/Pages/How-to-Use-Apps@UT.aspx

1. Go to apps.utk.edu (https://apps.utk.edu/vpn/index.html)
2. (https://apps.utk.edu/vpn/index.html)
3. Log-in with your user id and password
4. You may need to download the Citrix software
5. Once you've downloaded Citrix, you may need to change a browser setting
    a. Go to Tools/Add ons
    b. Locate the Citrix Server Plug-in and select "always allow"
6. From apps.utk.edu, click the "statistics and data mining" folder
7. Select the statistical program of your choice
8. It may take approximately 1 minute for the software to open

Save your data to the server and your analyses will run much faster than from your desktop or laptop! However, ALWAYS back up all of your files on your computer, not just on the server.

You are more than welcome to bring your computer to class and work through the SPSS exercises I will illustrate in class. Many students find this very useful.

Learning Style

You might find it useful to assess and think about your learning style:

http://www.educationplanner.org/students/self-assessments/learning-styles.shtml

SPSS Web Sites

The following web site provides a wealth of information about SPSS:

http://www.spsstools.net/

PROCESS for SPSS

PROCESS is a computational tool for examining moderation, mediation, and models that combine moderation and mediation (Hayes, 2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford. http://www.afhayes.com/ (Available online from Hodges library) (Chapter 4 & Appendix A in dropbox). We will use PROCESS in lectures 6 (moderation) and 7 (mediation).

Go to http://www.processmacro.org/download.html and download PROCESS and this will download a number of files. Read the following to learn how to install the SPSS PROCESS dialog for SPSS: dialoginstall.pdf. This document is in the folder with other materials downloaded when you download PROCESS. If you have installed PROCESS in SPSS you should be able to do this to select it:

> *Analyze*
>   *Regression*
>   *PROCESS, by Andrew F. Hayes* (http://www.afhayes.com/)

Sample Power 3.0

Sample Power 3.0 is a program for SPSS that allows you to do a statistical power analysis. (This isn't available for the MAC OS. See G*Power 3, below, for a free program that works on the MAC OS.) Please download this program:

1. Go to http://oit.utk.edu/software/
2. Click on the link under "Downloads"
3. Download "Instructions for Downloading SPSS [current version]" (Scroll to the bottom of this document to find authorization code for Sample Power 3.0).
4. Download Sample Power 3.0 (As you go through the Installation Wizard, it will ask you for the authorization code mentioned above).

G*Power 3

G*Power 3 covers statistical power analyses for many different statistical tests and it works with the MAC and Windows OS. G*Power 3 can be obtained free from:

http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3

Online Statistics Videos

You can go to the following web sites and watch videos on various topics in statistics. This is free, but you do have to sign up to view the videos.

Against All Odds: Inside Statistics
http://www.learner.org/resources/series65.html?pop=yes&vodid=252432&pid=139

Learning Math: Data Analysis, Statistics, and Probability
http://www.learner.org/resources/series158.html

Creative Heurisitcs
Types of Data: Nominal, Ordinal, Interval/ratio
http://youtu.be/hZxnzfnt5v8

Hypothesis tests, p-value
http://youtu.be/0zZYBALbZgg

Understanding the p-value
http://youtu.be/eyknGvncKLw

Important statistical concepts: significance, strength, association
http://www.youtube.com/watch?v=FG7xnWmZlPE&feature=share&list=PLD01983098F7E77B7

Sampling
http://youtu.be/be9e-Q-jC-0

Using SPSS
Paired Sample *t*-test
http://youtu.be/Qt-TlsY_b3k

Basics- Values and labels
http://www.youtube.com/watch?v=iZ8WAtEXZpw&feature=share&list=ULiZ8WAtEXZpw

Class Schedule

1/10   **NO CLASS (SSWR Conference)**

1/17   Lecture 1 in dropbox
Introduction to multiple regression
Keith (2015), Chapter 1, Appendix B
Orme & Combs-Orme (2009), Preface, & pp. 1-12

1/24   Lectures 2 & 3 in dropbox
Introduction to multiple regression
Keith (2015), Chapters 2 & 3
FIRST QUIZ

1/31   Lectures 2 & 3 in dropbox
Introduction to multiple regression—cont'd
Keith (2015), Chapters 4, 5

2/7   Lecture 4 in dropbox
Diagnosing and solving multiple regression problems
Modeling curvilinear relationships
Keith (2015), Chapter 8 (pp. 172-80) & Chapter 9 (pp. 182-203)
Orme & Combs-Orme (2009), pp. 21-29

2/14   Lecture 5 in dropbox
Categorical independent variables
Keith (2015), Chapter 6
http://www.ats.ucla.edu/stat/spss/webbooks/reg/chapter5/spssreg5.htm

2/21   Lecture 6 in dropbox
Moderation (Interactions)
Keith (2015), Chapters 7 & 8 (pp. 161-172)

2/28   Lecture 7 in dropbox
Mediation
Keith (2015), pp. 169-171, 273-274
Hayes (2013), Chapter 4

3/7   Lecture 8 in dropbox
Type I errors, statistical power, & sample size selection
Missing data
Keith (2015), Chapter 9 (203-212)
Cox et al. (2014)

3/14   **Spring Break**

3/21   Lecture 9 in dropbox
Introduction to the Generalized Linear Model
Orme & Combs-Orme (2009), Preface, Chapter 1 & Appendix B

3/28   Lectures 10 & 11 in dropbox
Poisson and negative binomial regression

Orme & Combs-Orme (2009), Chapter 5

4/4 Lectures 10 & 11 in dropbox
Poisson and negative binomial regression
Orme & Combs-Orme (2009), Chapter 5

4/11 Lectures 12 & 13 in dropbox
Binary logistic regression
Keith (2015), Chapter 10 (pp. 213-228)
Orme & Combs-Orme (2009), Chapter 2

4/18 Lectures 12 & 13 in dropbox
Binary logistic regression—cont'd
Orme & Combs-Orme (2009), Chapter 2

4/25 Final Exam Review
LAST QUIZ

5/2 *FINAL EXAM*

Course Requirements and Computation of Course Grade

In addition to the assignments listed below, each student is expected to read 100% of the assigned reading material and to come to each class prepared to discuss (i.e., ask and answer relevant questions) the reading material.

Evaluation

Weekly Quizzes	40%
Weekly Assignments	20%
Final Exam	40%

Here's an Excel worksheet that will calculate your final grade. Just double click on it and enter your grades.

SW 606 Grade	
Quiz	90
Out-of-Class	100
Exam	90
Final Grade	92

Quizzes

Each Wednesday at the beginning of class you will be given a 10-question quiz. Each quiz will cover both class and reading material from all preceding weeks, although the primary emphasis will be upon information from the previous week. These quizzes will provide both the instructor and students with ongoing feedback concerning knowledge acquisition. Such feedback will provide information necessary to take corrective measures by both the instructor and students to assure adequate learning. Quizzes will count for 40% of your grade.

Quizzes will be discussed after their completion, and graded quizzes will be returned on Wednesday after their completion. Your lowest quiz grade will be dropped, and the remaining

quiz grades will be averaged to get your final quiz grade.

Computer Assignments

Weekly computer assignments and brief summaries of the obtained results will be assigned each week and they will be due the following week. These assignments will correspond to the weekly topics. These assignments will count for 20% of your grade, and each assignment will be graded on a Pass/Fail (i.e., A/F) basis. ***COMPUTER ASSIGNMENTS THAT ARE TURNED IN LATE WILL BE ASSIGNED AN "F".***

I strongly advise you to do the computer assignments as soon after class as possible. Don't wait until the day before they are due.

Final Exam

A *comprehensive* final exam will be administered at the end of the semester. This exam will count for 40% of your grade.

Grading Scale

Final grades will be assigned as follows:

94-100   A
91-93   A-
88-90   B+
84-87   B
80-83   B-
76-79   C+
72-75   C
68-71   D
<68   F

Course Bibliography

Required Texts

Keith, T. Z. (2015). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling* (2nd ed.). New York: Routledge.

Orme, J. G., & Combs-Orme, T. (2009). *Multiple regression with discrete dependent variables*. New York, NY: Oxford University Press.

Required Readings

Cox, B. E., McIntosh, K., Reason, R. D., & Terenzini, P. T. (2014). Working with missing data in higher education research: A primer and real-world example. *The Review of Higher Education, 37*, 377-402. doi: 10.1353/rhe.2014.0026

Hayes, A. E. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford. (Chapter 4)

Selected Bibliography

Cohen, J. Cohen, P. West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Dattalo, P. (2008). *Determining sample size: Balancing power, precision, and practicality*. New York: Oxford University Press.

DeMaris, A. (2004). *Regression with social data: Modeling continuous and limited response variables.* Hoboken, NJ: John Wiley.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.

Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). Los Angeles: Sage.

Fox, J. (2008). *Applied regression analysis and generalized linear models (2nd ed.).* Thousand Oaks, CA: Sage.

Greene, W. H. (2011). *Econometric analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). New York: Routledge.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford.

Hilbe, J. M. (2007). *Negative binomial regression*. New York: Cambridge University Press.

Hosmer, D. W., & Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). New York: John Wiley & Sons.

Hosmer, D. W., & Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time to event data* (2nd ed.). New York: John Wiley & Sons.

Huck, S. W. (2009). *Statistical misconceptions*. New York: Psychology Press.

Kerlinger, F. N. (1999). *Foundations of behavioral research* (4th ed.). New York: Holt, Rinehart & Winston.

Little, T. D. (Ed.) (2013). *The Oxford handbook of quantitative methods*. New York: Oxford University Press.

Long, J. S., & Freese, J. (2015). *Regression models for categorical dependent variables using Stata* (3rd ed.). College Station, TX: Stata Press.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Lawrence Erlbaum.

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. Los Angeles, CA: Sage.

Osborne, J. (Ed.) (2008). *Best practices in quantitative methods*. Los Angeles, CA: Sage.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Fort Worth: Harcourt Brace College Publishers.

Raykov, T., & Marcoulides, G. A. (2008). *An introduction to multivariate analysis*. New York: Routledge Academic.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Stevens, J. P. (2009). Applied multivariate statistics for the social sciences (5th ed.). New York, NY: Psychology Press.

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Allyn & Bacon.

Tables[1]

---

[1] Here is a little hint about constructing tables. Centering columns on the decimal point is usually preferred. This can be done in Word using the decimal tab. Simply follow the following

The following web site and book provide valuable information about creating tables in APA style:

Nicol, A. A. M., & Pexman, P. M. (2010). *Presenting your findings: A practical guide for creating tables* (6th ed.). Washington, DC: American Psychological Association.

National Council on Family Relations: http://oregonstate.edu/~acock/tables/

Graphs and other Figures

The following book provides valuable information about creating graphs and other types of figures in APA style:

Nicol, A. A. M., & Pexman, P. M. (2010). *Displaying your findings: A practical guide for creating figures, posters, and presentations* (6th ed.). Washington, DC: American Psychological Association.

Teaching Statistics

Here are good web sites for those of you who end up teaching statistics (Don't laugh. I didn't set out to do this.):

Journal of Statistics Education
http://www.amstat.org/PUBLICATIONS/JSE/

Statistics Education Research Journal
http://www.stat.auckland.ac.nz/~iase/publications.php?show=serj

Hulsizer, M., & Woolf, L. M. (2009). *A guide to teaching statistics: Innovations and best practices*. Hoboken, NJ: John Wiley.
http://www.webster.edu/teachstats/

Consortium (C) for the Advancement (A) of Undergraduate (U) Statistics (S) Education (E)
http://www.causeweb.org/

Available Data Sets

The following provides a description of a few of the data sets contained in your dropbox folder.

directions for each column (http://oregonstate.edu/~acock/tables/).

1. Show the ruler below the toolbars on the top of your screen in your Word document.
2. Identify the left tab symbol on the extreme left of the ruler. It looks like an L.
3. Click on this symbol three times. You will be scrolling through various tabs. The 3rd time, the decimal tab will be visible. It looks like an upside-down T with a dot on the right hand side.
4. Highlight the entire column of your table.
5. Click the location in the ruler that reflects where you want the decimal for that column to be.
6. That's it. Unless you already have other tabs active in that column or have used the center text command, all numbers should be aligned at the decimal around that point.

General Social Survey

The following web page provides a detailed description of the General Social Survey.

http://www.norc.org/GSS+Website/

This data set contains data for the cross-sectional portion of the 2008 General Social Survey. The data set is named gss2008.sav. The codebook and questionnaires for this data set are also contained in the GSS2008 folder.

Grades

This data set contains quiz grades, final exam grades, and final course grades for students who completed either SW 605 or SW 606 (1993 – 2015).

The data set is named Stat Grades.sav.

Foster Parent Applicant Study

This is a study of 161 foster family applicants. The study design and variables are described in Orme et al. (2004). The data set is named Orme.sav and the article by Orme et al. (2004) is contained on in the Articles folder.

Bank Study

This data set and its description were obtained from SPSS (Norusis, 1988, pp. 126-127).

This is a study of salaries at a midwestern bank involved in EEO litigation. Data are available for 474 individuals hired between 1969 and 1971. This file is excerpted from H. V. Roberts' *Statistical Bases in the Measurement of Employment Discrimination* in *Comparable Worth: Issues and Alternatives* (ed. E. Robert Livernash). Washington, D.C.: Equal Employment Advisory Council, 1980).

The data set is named bank.sav.

Employment Study

This data set and its description were obtained from Enders (2010).

These are hypothetical data for 480 employees on eight work-related variables: age (years), job tenure (years), female (0 = male, 1 = female), psychological well-being, job satisfaction, job performance, turnover (intention to turnover—0 = no, 1 = yes), and IQ.

The data set is named employment.sav.

Depression Study

This data set and its description were obtained from Afifi and Clark (1984, p. 3, 30-39).

The data for the depression study have been obtained from a complex, random, multiethnic sample of 1000 adult residents of Los Angeles County. The study was a panel or longitudinal design where the same respondents were interviewed four times between May 1979 and July 1980. About three-fourths of the respondents were reinterviewed for all four interviews. The fieldwork for the survey was conducted by professional interviewers from the Institute for Social Science Research at UCLA.

This ongoing research is an epidemiological study of depression and help-seeking behavior among free-living (noninstitutionalized) adults. The major objectives are to provide estimates of

the prevalence and incidence of depression and to identify causal factors and outcomes associated with this condition. The factors examined include demographic variables, life events stressors, physical health status, health care utilization, medication use, life-style and social support networks. The major instrument used for classifying depression is the Depression Index (CESD) of the National Institute of Mental Health, Center for Epidemiologic Studies. A discussion of this index and the resulting prevalence of depression in this sample are given in Frerichs, Aneshensel, and Clark (1981).

The longitudinal design of the study offers advantages for assessing causal priorities since the time sequence allows us to rule out certain potential causal links. Nonexperimental data of this type cannot directly be used to establish causal relationships, but models based on an explicit theoretical framework can be tested to determine if they are consistent with the data. An example of such model testing is given in Aneshensel and Frerichs (1982).

Only a subset of the factors measured on a sample of the respondents is included in order to keep the data set easily comprehensible. The data are from a subset of 294 respondents randomly chosen from the original 1000. Only data from the first time period are included.

The data set is named depress.sav.

Magnets and Pain Relief

Magnetic fields have been shown to have an effect on living tissue as early as the 1930's. Plants have been shown to have an improved growth rate when raised in a magnetic field (Mericle et al., 1964). More recently, doctors and physical therapists have used either static or fluctuating magnetic fields to aid in pain management, most commonly for broken bones. In the case study presented here, Carlos Vallbona and his colleagues sought to answer the question "Can the chronic pain experienced by postpolio patients be relieved by magnetic fields applied directly over an identified pain trigger point?" This data set and its description were was obtained from http://onlinestatbook.com/case_studies_rvls/magnets/index.html. (Go to this web site for additional information about this study.)

Patients experiencing post-polio pain syndrome were recruited. Half of the patients were treated with an active magnetic device and half were treated with an inactive device. The patient was asked to subjectively grade pain at the trigger point under palpitation on a scale from 0 to 10 (0 is the least pain, increasing to 10). All patients rated their pain before ("PretestPain") and after ("PosttestPain") application of the device.

Following the initial pain assessment, an envelope containing a device was randomly selected from the box containing active and inactive devices. This device was applied to the pain area for 45 minutes and then removed. The patient then evaluated his or her pain again at the region or trigger point. The treatment condition is indicated by the variable "Group;" subjects receiving treatment with the active magnet have a "1" on this variable; subjects treated with the inactive placebo have a "0."

The data set is named Magnets&Pain.sav and the Vallabona et al. (1997) article is contained in the Articles folder.

Instructor Reputation

How powerful are rumors? Frequently, students ask friends and/or look at instructor

evaluations to decide if a class is worth taking. Kelley (1950) found that instructor reputation has a profound impact on actual teaching ratings. Towler and Dipboye (1998) replicated and extended this study by asking (a) Does an instructor's prior reputation affect student ratings? and (b) Does the size of this effect depend on student characteristics. This case study presents only data relevant to the former question. This data set and its description were obtained from http://onlinestatbook.com/case_studies_rvls/magnets/index.html. (Go to this web site for additional information about this study.)

Kelley, H. H.(1950). The warm-cold variable in first impression of persons. Journal of Personality, *18*, 431-439.

Towler, A., & Dipboye, R. L. (1998). The effect of instructor reputation and need for cognition on student behavior (poster presented at American Psychological Society conference, May 1998).

Subjects were randomly assigned to one of two conditions. Before viewing the lecture, students were give a summary of the instructors prior teaching evaluations. There were two conditions: Charismatic instructor ("1") and Punitive instructor ("2").

Summary given in the "Charismatic instructor" condition:
Frequently at or near the top of the academic department in all teaching categories. Professor S was always lively and stimulating in class, and commanded respect from everyone. In class, she always encouraged students to express their ideas and opinions, however foolish or half-baked. Professor S was always innovative. She used differing teaching methods and frequently allowed students to experiment and be creative. Outside the classroom, Professor S was always approachable and treated students as individuals.

Summary given in the "Punitive instructor" condition:
Frequently near the bottom of the academic department in all important teaching categories. Professor S did not show an interest in students' progress or make any attempt to sustain student interest in the subject. When students asked questions in class, they were frequently told to find the answers for themselves. When students felt they had produced a good piece of work, very rarely were they given positive feedback. In fact, Professor S consistently seemed to grade students harder than other lecturers in the department.

Then all subjects watched the same twenty-minute lecture given by the **exact same** lecturer. Following the lecture, subjects rated the lecturer. Subjects answered three questions about the leadership qualities of the lecturer. A summary rating score was computed and used as the variable "rating" here.

The data set is named Instructor Reputation.sav.

Sesame Street Study
This data set and its description were obtained from Stevens (1990, pp. 232-239) and a data disk supplied by Stevens.

This is part of a large data set that evaluated the impact of the first year of the Sesame Street television series. Children were sampled from five different sites in the U.S. Age was coded in

months. Viewing category was coded from 1 (if child rarely watched the show) to 4 (if the show was viewed more than five times per week). Setting in which Sesame Street was viewed was coded as 1 for home and 2 for school. The next variable on the file was a treatment condition in which some children were encouraged to view Sesame Street (coded as 1) and others were not (coded as 2). The next set of variables on the file (variables 8 through 13) involves pretest scores on a variety of cognitive variables. After viewing the series, the children were posttested on these same variables; these are variables 14 through 19 on the file. The final variable on the file is a pretest measure of vocabulary maturity, as measured by the Peabody Picture Vocabulary Test. The first variable on the file is subject ID, and then comes SITE, SEX, AGE, viewing category, SETTING, and treatment condition (variable 7). The data set is named sesame.sav.

Stress Reduction and Smoking Relapse
This data set and its description were obtained from Stevens (2002, pp. 661-652). This study by Christine Wynd (*Journal of Advanced Nursing*, 1992, 204-302), investigated the effect of stress reduction in preventing smoking relapse. Subjects who had completed a local smoking cessation program were randomly assigned to either an experimental group or an attention placebo control group. The subjects in the treatment group were exposed to a three session treatment, and were measured on 3 variables at each session: creative imagery, stress and smoking rate. In the data listing I have abbreviated these variables as CRE, STR and SMR. The subjects in the experimental group are coded as 1 and those in the control group as 0.

The data set is named Wynd.sav and the article by Wynd is contained in the Articles folder.

Changes in Attitudes of Grade School Children Toward School Subjects
This data set and its description were obtained from Stevens (1990, pp. 246-248) and a data disk supplied by Stevens.

These are data collected on 189 third through sixth graders from a suburban midwestern public school. Sex is the first variable on the file, grade is the second variable and teacher is the third variable (there were five teachers). The children were measured on preference toward the following subjects: mathematics, language arts, science, reading and social studies prior to an intervention. These are variables 4 through 8 on the file. An intervention was employed with the teachers in an attempt to change the children's preferences (attitudes). The children were posttested on these subject preferences again 4 months later. These are the remaining 5 variables on the file.

The data set is named attitude.sav.

Keith Data

These are the data sets described and illustrated in Keith (2006) (see Appendix A).

References

Afifi, A. A. & Clark, V. (1984). *Computer-aided multivariate analysis*. Belmont, California: Lifetime Learning Publications.

Aneshensel, C. S. & Frerichs, R. R. (1982). Stress, support, and depression: A longitudinal causal model. *Journal of Community Psychology, 10,* 363-376.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.

Frerichs, R. R., Aneshensel, C. S. & Clark, V. A. (1981). Prevalence of depression in Los Angeles County. *American Journal of Epidemiology, 113,* 691-699.

Keith, T. Z. (2006). *Multiple regression and beyond*. Boston, MA: Allyn & Bacon.

Stevens, J. (2002). Applied multivariate statistics for the social sciences (4th ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stevens, J. (1990). *Intermediate statistics: A modern approach.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Additional Data Sets

The Inter-University Consortium for Political and Social Research (ICPSR) has a very rich array of data sets that you can use.

Christopher Eaker, Assistant Professor & Data Curation Librarian, can help you access the ICPSR data sets and many more existing data sets that might be of interest. Professor Eaker can be contacted at:

ceaker@utk.edu
(865) 974-4404
https://www.lib.utk.edu/scholar/

The following websites also have numerous public use data sets that you can use:

World Health Organization
http://www.who.int/research/en/

CDC Wide-ranging OnLine Data for Epidemiological Research
http://wonder.cdc.gov/

FedStats
http://www.fedstats.gov/

Centers for Disease Control and Prevention Public-Use Data Files
http://www.cdc.gov/nchs/data_access/ftpdata.htm

United States Department of Health and Human Services
http://aspe.os.dhhs.gov/datacncl/datadir/

Substance Abuse & Mental Health Data Archive
http://www.icpsr.umich.edu/SAMHDA/

Data.gov
http://www.data.gov/

National Center for Education Statistics
http://nces.ed.gov/surveys/

Roper Center for Public Opinion Research
http://www.ropercenter.uconn.edu/data_access.html

The Data and Story Library
http://lib.stat.cmu.edu/DASL/DataArchive.html

Carnegie Mellon University StatLib
http://lib.stat.cmu.edu/datasets/

Henry A. Murray Research Archive
http://www.murray.harvard.edu/frontpage

Sociometrics Data Archives
http://www.socio.com/dataarchives.htm

National Data Archive on Child Abuse and Neglect
http://www.ndacan.cornell.edu/NDACAN/AboutNDACAN.html

Center for AIDS Prevention Studies
http://www.caps.ucsf.edu/tools/data/VCT/

TheDataWeb
http://www.thedataweb.org/datasets.html

Rice Virtual Lab in Statistics
http://onlinestatbook.com/case_studies_rvls/index.html

Dr. Karl L. Wuench's SPSS-Data Page
http://core.ecu.edu/psyc/wuenschk/SPSS/SPSS-Data.htm

Project on Human Development in Chicago Neighborhoods
http://dvn.iq.harvard.edu/dvn/dv/mra/faces/study/StudyPage.xhtml?studyId=307&studyListingIndex=1_7f5088797bc253a292ee18fd7b70&rvn=2

Health Information and Trends Survey (HINTS): How Americans Find and Use Cancer Information
http://hints.cancer.gov/dataset.jsp

National Survey of Children's Health
http://nschdata.org/Content/Default.aspx

National Survey of Children with Special Health Care Needs
http://cshcndata.org/Content/Default.aspx

The National Longitudinal Study of Adolescent Health (Add Health)
http://www.cpc.unc.edu/projects/addhealth

National Ambulatory Medical Care Survey
http://www.cdc.gov/nchs/ahcd.htm

National Hospital Discharge Survey
http://www.cdc.gov/nchs/ahcd.htm

National Nursing Home Survey
http://www.cdc.gov/nchs/nnhs.htm

National Immunization Survey
http://www.cdc.gov/nis/datafiles.htm

Surveillance Epidemiology and End Results Program
http://seer.cancer.gov/data/options.html

Behavioral Risk Factor Surveillance System
http://www.cdc.gov/brfss/technical_infodata/surveydata.htm

Youth Risk Behavior Surveillance System
http://www.cdc.gov/HealthyYouth/yrbs/data/index.htm

National Health and Nutrition Examination Survey
http://www.cdc.gov/nchs/nhanes.htm

National Health Interview Survey
http://www.cdc.gov/nchs/nhis.htm

Joint Canada/United States Survey of Health
http://www.cdc.gov/nchs/nhis/jcush.htm

Longitudinal Studies of Aging
http://www.cdc.gov/nchs/lsoa.htm

State and Local Area Integrated Telephone Survey
http://www.cdc.gov/nchs/slaits.htm

National Vital Statistics System
http://www.cdc.gov/nchs/nvss.htm

National Maternal and Infant Health Survey
http://www.cdc.gov/nchs/nvss/nmihs.htm

Panel Study of Family Dynamics
http://psidonline.isr.umich.edu/default.aspx
Multiple sub-studies/ data sets:
http://psidonline.isr.umich.edu/Studies.aspx

Secondary Data Analysis

The following books are good sources for the secondary data analysis:

Boslaugh, S. (2007). *Secondary data sources for public health*. New York: Cambridge University
    Press.

Bulmer, M., Sturgis, P. J., & Allum, N. (2009). *Secondary analysis of survey data* (Volumes 1 – 4).
    Los Angeles: Sage.

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied survey data analysis*. Boca Raton, FL: CRC Press.

Smith, D. (2008). *Using secondary data in educational and social research*. Berkshire, England: Open University Press.

Vartanian, T. P. (2011). *Secondary data analysis*. New York: Oxford University Press.

Web Sites and Related Resources

The following provide an online statistics book with links to other statistics resources on the web, and more:

http://www.ats.ucla.edu/stat/
http://onlinestatbook.com/
http://davidmlane.com/hyperstat/
http://stattrek.com/default.aspx
http://cast.massey.ac.nz/core/index.html?collection=Public
http://www.jerrydallal.com/LHSP/LHSP.HTM
http://www.sjsu.edu/faculty/gerstman/StatPrimer/ (also contains numerous data sets)
http://www.sportsci.org/resource/stats/
http://www.visualstatistics.net/
http://faculty.vassar.edu/lowry/webtext.html
http://www.stat.berkeley.edu/users/stark/SticiGui/
http://www.bmj.com/collections/statsbk/
http://www.usi.edu/libarts/socio/stats.htm
http://www.seeingstatistics.com/
http://www2.chass.ncsu.edu/garson/pa765/statnote.htm
http://www.psychstat.missouristate.edu/introbook/sbk00.htm
http://www.itl.nist.gov/div898/handbook/eda/eda.htm

Glossary of Statistical Terms

The following provide glossaries of statistical terms that you might find useful.

http://www.statsoft.com/textbook/glosfra.html
http://www.geodata.soton.ac.uk/biology/lexstats.html
http://stattrek.com/Help/Glossary.aspx
http://www.animatedsoftware.com/elearning/Statistics%20Explained/glossary/se_glossary.html
http://dorakmt.tripod.com/mtd/glosstat.html
http://davidmlane.com/hyperstat/glossary.html
http://www.stats.gla.ac.uk/steps/glossary/alphabet.html
http://www.ablongman.com/html/abrami/glossary/glossary.html