

Web resources for Regression and Linear Modeling Text

All content copyright © Jason W. Osborne, 2016.

Thank you for your interest in my book! It was tremendous fun to write, and I hope you find it useful! Please do not hesitate to contact me at jason@jwosborne.com or jasonwosborne@gmail.com if you have comments on the book or any of my other work. Many resources for this book, including data sets, answer keys, syntax examples, etc. can also be found on my web site <http://jwosborne.com>. *If you are an instructor and would like access to some PowerPoint collections and/or mid-term assignment please do not hesitate to contact me.*¹

Data sets I used for this book:

1. **NHANES2001 data:** The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations.
 - a. Full documentation and description of survey etc. from the US Centers for Disease Control and Prevention (CDC): http://www.cdc.gov/nchs/nhanes/search/nhanes01_02.aspx
2. **NHIS2010 data:** The National Health Interview Survey (NHIS) is the principal source of information on the health of the civilian noninstitutionalized population of the United States and is one of the major data collection programs of the National Center for Health Statistics (NCHS) which is part of the Centers for Disease Control and Prevention (CDC).
 - a. Full documentation and description of survey, etc from CDC: http://www.cdc.gov/nchs/nhis/nhis_2010_data_release.htm
3. **AAUP data:** The AAUP dataset for the American Statistical Association's Statistical Graphics Section's 1995 Data Analysis Exposition. This dataset is taken from the March-April 1994 issue of *Academe*. Thanks to Maryse Eymonerie, Consultant to AAUP, for assistance in supplying the data. Faculty salary data are for the 1993-94 school year. You may wish to consult a copy of the special issue of *Academe* for more detailed descriptions of the variables.
 - a. Original source and documentation: <http://lib.stat.cmu.edu/datasets/>
 - b. Description of data source and set: <https://dl.dropboxusercontent.com/u/18489687/logistic/aaup%20document.txt>
4. **NELS88 data:**
 - a. Source and documentation: <http://nces.ed.gov/surveys/nels88/>
5. **Nativity 2013 data** from US Centers for Disease Control (CDC), contains public data on births in US during 2013.
 - a. Full documentation and information about the vital statistics data available online from CDC: http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm
6. **Youth Risk Behavior Survey (YRBS)** from US Centers for Disease Control (CDC), contains public data on self-reported health and risk behaviors from high school aged students.
 - a. Full documentation and information about the vital statistics data available online from CDC: <http://www.cdc.gov/healthyyouth/data/yrbs/index.htm>
7. **National College Health Risk Behavior Survey of 1995 (NCHRBS95):** Also from the US Centers for Disease Control and Prevention.
 - a. Full documentation and information about the study at <http://www.cdc.gov/healthyyouth/data/yrbs/data.htm>
8. **Education Longitudinal Study of 2002 (NCES/IES):** A nationally representative longitudinal study of 10th graders in 2002, followed students for many years.
 - a. Full documentation and information at: <http://nces.ed.gov/surveys/els2002/>
9. **Cars.sav:** a data set distributed publicly with SPSS that has data on over 300 cars statistics such as engine displacement, weight, fuel efficiency, horsepower, etc. <https://dl.dropboxusercontent.com/u/18489687/GLMbook/Cars.sav>

¹ No guarantee they are any good, mind you...

Chapter 1 resources:

- NHANES 2001 data: http://wwwn.cdc.gov/nchs/nhanes/search/nhanes01_02.aspx
- Chapter 1 Answer Key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch01_answer%20key.docx

Chapter 2 resources:

- AAUP data: <https://dl.dropboxusercontent.com/u/18489687/GLMbook/AAUP.sav>
- Box-Cox syntax: <https://dl.dropboxusercontent.com/u/18489687/GLMbook/NEW-box-cox.sps>
- Grades data: <https://dl.dropboxusercontent.com/u/18489687/GLMbook/grades.sav>

Chapter 2 Answer Key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch02_answer%20key.docx

Chapter 2 SPSS guide: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch02_SPSS.pdf

Chapter 3 resources:

- NELS88 data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH03_NELS.SAV
- AAUP data: <https://dl.dropboxusercontent.com/u/18489687/GLMbook/AAUP.sav>
- NHANES data for exercise #3: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch03_NHANES.sav

Chapter 3 answer key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch03_answer_key.docx

SPSS guide: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH03_SPSS_guide.pdf

Chapter 4 resources:

1. NHIS2010 smoking and diabetes data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH04_SMOK_DIAB.sav
2. Exercise 2 NHIS data: <https://dl.dropboxusercontent.com/u/18489687/GLMbook/NHIS2010-sm-Ch03.sav>
3. Exercise 3 AAUP data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/AAUP_original.sav
4. Other resources:
 - a. Chapter 4 answer key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch04_answer_key.docx
 - b. Box-Cox syntax for transformation if necessary: Box-Cox syntax: <https://dl.dropboxusercontent.com/u/18489687/GLMbook/NEW-box-cox.sps>
 - c. SPSS guide: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH04_SPSS_guide.pdf

Chapter 5 resources

- NELS88 data for Chapter 5:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH05_NELS88.SAV
- NHIS data for diabetes/BMI analyses:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch05_NHIS2010.sav
- NELS88 Marijuana use data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch05_mj.sav
- Other resources:
 - Tutorial on how to convert from probabilities to odds to log odds and back:
<http://pareonline.net/getvn.asp?v=17&n=11>
 - Chapter 5 answer key:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch05_answer_key.docx
 - SPSS guide: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH05_SPSS_guide.pdf

Chapter 6 resources

- NELS88 Marijuana use data for this chapter:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch06_MJdata.sav
- NHIS_SM: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH06_NHIS_sm.sav
- ERRATA:
 - On page 141, the slope for eq. 6.2c should be -2.949 to match output
- Other resources:
 - Chapter 6 answer key:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch06_answer_key.docx
 - SPSS syntax example of how to clean data using binary logistic from Ex 6.2:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch06_ex2_Syntax1.sps

Chapter 7 resources

- AAUP data Chapter 7: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch07_AAUP.sav
- NHIS2010 diabetes, AGE, BMI:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch05_NHIS2010.sav
- NELS88 Marijuana use data for this chapter:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch07_MJdata.sav
- Natality 2013 data (3% random sample):
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Natality_2013_sm.sav
- Other resources:
 - Chapter 7 answer key:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch07_answer_key.docx

Chapter 8 resources

- NELS 88 data set from chapter examples involving OLS regression:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch08_NELS88.sav
- data for logistic regression example:
https://dl.dropboxusercontent.com/u/18489687/logistic/NELS_data_pull1_2.SAV
- MJ data set: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch07_MJdata.sav
- NHIS data for exercise #5:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH08_NHIS2010.sav

- Other resources:
 - Chapter 8 partial answer key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch08_answer_key.docx

Chapter 9 resources

- Chapter NELS88 OLS regression examples: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch09_NELS88_OLS.sav
- Logistic and multinomial logistic regression examples (NELS 88): https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH09_multinomial_logistic.sav
- Data for exercises:
 - #3, 4, 5: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH09_multinomial_logistic.sav
 - NHIS data for exercises #6, 7 : <https://dl.dropboxusercontent.com/u/18489687/logistic/NHIS2010-master.sav>
 - Natality data for exercise #8: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH09_Natality.sav
- link to excel spreadsheet with interaction figuring: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH09_interactions.xlsx
- Other links of use:
 - [Kristopher Preacher's web site](#) contains templates for testing simple slopes, and findings regions of significance, for both 2-way and 3-way interactions. It also includes options for hierarchical linear modelling (HLM) and latent curve analysis.
 - [Yung-ju Yang's web site](#) contains SAS macros to plot interaction effects and run the slope difference tests for three-way interactions
 - [Cameron Brick's web site](#) contains instructions on how to plot a three-way interaction and test for differences between slopes in Stata
- link to SPSS syntax from this chapter: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch09_syntax.sps
- Chapter 9 partial answer key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch09_answer_key.docx

Chapter 10 resources

- NELS data for logistic regression example: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH010_logistic.sav
- Exercise 3 AAUP data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/AAUP_original.sav
- Exercise 4: smoking and BMI predicting diabetes: https://dl.dropboxusercontent.com/u/18489687/GLMbook/Ch10_NHIS_ex4.sav
- Chapter 10 answer key: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH10_answer_key.docx

Chapter 11 resources

- National College Health Risk Behavior Survey of 1995 data used in chapter: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH11_nchrbs95.sav
- Gabel lower back pain data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH11_Gabel.sav
- ELS 2002 data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH11_ELS2002_sm.sav

Chapter 12 resources

- Data from the chapter loglinear example:
https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH12_yrbs2013.sav
- ELS 2002 data: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH11_ELS2002_sm.sav
-

Chapter 13 resources

- Free student version of HLM: <http://www.ssicentral.com/hlm/downloads.html>
- Examples and help using HLM software: <http://www.ssicentral.com/hlm/examples.html>
- Data from the chapter HLM logistic regression example in .mdm format:
https://dl.dropboxusercontent.com/u/18489687/logistic/HLM_logistic_example_nol12resid
- Exercise #3 (another HLM logistic example): Download the following data files:
 - Level 1: https://dl.dropboxusercontent.com/u/18489687/logistic/Ch13_11_N200.sav
 - Level 2: https://dl.dropboxusercontent.com/u/18489687/logistic/CH13_12_N200.sav

Chapter 14 resources

- Data for MNAR examples: https://dl.dropboxusercontent.com/u/18489687/GLMbook/CH14_MNAR.sav

Chapter 15 resources

- SPSS bootstrap macro for OLS regression examples from the chapter
- SPSS bootstrap macro for Poisson regression examples from the chapter