

LAS792 Managing Research Data in the Social Sciences

2 or 3 credit hour course - Summer 2009

June 9 – June 24 (if taken for 2 Credits),

June 9 – July 3 (if taken for 3 Credits)

1:30-3:30pm Budig PC Lab (10 Budig Hall)

Line Number: 84135

Instructor: Larry Hoyle

Objectives:

This course is designed to teach the skills necessary to read, reorganize, transform, display, clean, archive, and export simple or complex data. The primary tool used in the class is the SAS system. Structured Query Language (SQL) in both SAS and Microsoft Access receives a strong emphasis.

Topics:

The SAS Environment – using SAS interactively

Reading Data

Reading More Complex Data - e.g. reading a Web page

Transforming Data

Reorganizing Data

Hierarchical Files –e.g. Census Public Use Microdata Samples (PUMS)

Joining and Concatenating Files SQL and Data steps

Designing your data structures

Custom output, e.g. Web pages, XML

Parameterizing Programs / Macros – Iterating procedures

More SAS Data Step Techniques, Perl regular expressions

Data Interchange – moving data among software packages, operating systems, XML

Data Cleaning

A brief tour of some other topics

- Archiving Data

- Survey design

- Multidimensional databases

- Online Analytical Processing (OLAP servers), Pivot Tables

- Visualizing data, Geographic Information Systems – GIS

- Data Warehousing

Course structure: approximately 50% lecture/demonstration and 50% hands-on lab.

Auditing Policy: Faculty and permanent staff are welcome to sit in as space permits.

Please contact Larry – LarryHoyle@ku.edu

	Tuesday, June 9,2009	Wednesday, June 10,2009	Thursday, June11,2009	Friday, June 12,2009
Read LSB ¹ Chapters 8 and 10 sometime during the first week	Overview & SAS Environment. LSB chap 1 pp. 2-27	Reading Data LSB chap 2 pp 30-49	Reading Data LSB chap 2 pp 50-73	Transforming Data LSB chap 3 pp 76-97
Monday, June15,2009	Tuesday, June16,2009	Wednesday, June17,2009	Thursday, June18,2009	Friday, June19,2009
Text - Fuzzy matches Perl Regular expressions text mining Cody SUGI paper 265-29.pdf Pless SUGI paper 043-29.pdf	Restructuring Data LSB chap 4 pp100-121	SQL SELECT ODBC SQL chaps 1,2 pp1-37. SQL chap 2 pp38-52 SQL chap 3 pp 54-68 Project – Reading Data Due	SQL Joining and Merging Data Indexing files Pums example SQL chap 3 pp 69-81 SQL chap 4 pp84-103	Joining and Merging Data With Data Steps LSB chap 6 pp 170-197
Monday, June 22, 2009	Tuesday, June 23,2009	Wednesday, June24, 2009	Thursday,June25, 2009	Friday, June26, 2009
SQL Set operations – Union, etc SQL chap 5 pp 106-125	SQL create, update, alter, integrity constraints	XML and Data Interchange, DDI (ICPSR) LSB chap 9 pp236-249 Palmer SUGI paper 036-29 SECOND ASSIGNMENT DUE	Custom Output Popups and drill down LSB chap 5 pp144-167 Zdeb SUGI papers 090-29.pdf, 251-29.pdf	updating data, audit trails, storage, archiving “TEST” (short in class data project) joining data
Monday, June 29, 2009	Tuesday, June 30, 2009	Wednesday, July 1, 2009	Thursday, July 2,2009	Friday, July 3, 2009
Parameterizing your programs LSB chap 7 pp200-213	Designing entering and storing Your Data Checking and Cleaning Your Data LSB chap 4 pp122-141 Cody	Data Warehousing OLAP and OLAP Servers, Pivot Tables	Visualizing data	Text Mining LAST ASSIGNMENT DUE
Monday July 6,2009 Optional Potpourri				

LSB – Delwiche, Lora D. and Slaughter, Susan J., *The Little SAS Book a primer, Third Edition* ISBN 1-59047-333-7. It is important that you use the third edition – it contains a lot of new material. (Recommended)

Slaughter, Susan J and.Delwiche, Lora D. *The Little SAS Book for Enterprise Guide® 4.1* ISBN 978-1-59994-089-2

(Recommended)

STEP – SAS Publishing *Step-by-Step Programming with Base SAS Software* ISBN 1-58025-791-7

(Recommended)

SQL – SAS Institute Inc., *SAS SQL Procedure User's Guide, Version 9* available free online at

http://support.sas.com/documentation/onlinedoc/91pdf/sasdoc_91/base_sqlproc_6992.pdf

(Available online)

Cody – Cody, Ron, *Cody's Data Cleaning Techniques Using SAS Software* ISBN 1-58025-600-7 (Recommended)

SAS User Group International (SUGI) & SAS Global Forum (2007 on) conference papers can be found online for free via <http://www.lexjansen.com/sugi/> Recommended links at the bottom of the page for special topics : ARRAYS, MACROS, ODS, PROC FORMAT, SQL, XML

Example papers:

Cody, Ron SUGI paper 265-29 An Introduction to Perl Regular Expressions in SAS 9

<http://www2.sas.com/proceedings/sugi29/265-29.pdf>

Palmer, Michael paper 036-29 XML in the DATA Step

<http://www2.sas.com/proceedings/sugi29/036-29.pdf>

Pless, Richard SUGI paper 043-29 An Introduction to Regular Expressions with Examples from Clinical Data

<http://www2.sas.com/proceedings/sugi29/043-29.pdf>

Zdeb, Mike SUGI paper 090-29 Pop-Ups, Drill-Downs, and Animation

<http://www2.sas.com/proceedings/sugi29/090-29.pdf>

Zdeb, Mike SUGI paper 251-29 The Basics of Map Creation with SAS/GRAPH®

<http://www2.sas.com/proceedings/sugi29/251-29.pdf>

Other Resources:

Delwiche, Lora D. and Slaughter, Susan J *Coming to SAS® from SPSS A Programming Approach*

http://support.sas.com/publishing/bbu/companion_site/62272.pdf

Other Course Notes

Course materials prepared by the instructor, together with the content of all lectures and review sessions presented by the instructor are the property of the instructor. Video and audio recording of lectures and review sessions without the consent of the instructor is prohibited. On request, the instructor may grant permission for students to audio tape lectures, on the condition that these audio tapes are only used as a study aid by the individual making the recording. Unless explicit permission is obtained from the instructor, recordings of lectures and review sessions may not be modified and must not be transferred or transmitted to any other person, whether or not that individual is enrolled in the course

Unless explicit permission is obtained from Larry Hoyle, materials on the course CD, server or web site may not be modified and must not be transferred or transmitted to any other person, whether or not that individual is enrolled in the course.

Students have found that reading the suggested material in advance is very helpful.