Synchronous
In person instruction:  
* Tuesdays and Thursdays 1:30-3:00 p.m.
* Starting Tuesday, September 6th 2022

Location:  
HHSL Computer Lab 510

Online help:  
Piazza account established and will be monitored by instructor

Instructor(s):  
Gail Rogers: Gail.Rogers@Tufts.edu (she/her/hers)

Office hours:  
by appointment (in office or online)
USDA-HNRCA, 711 Washington St, Boston MA
11th floor, Room # 1122
Phone: 617 556 3338

Teaching Asst.:  
Melanie Guirette Melanie.Guirette@tufts.edu

Graduate Credits:  
3.0
Prerequisites:  
None

Course Description:  
This semester long course will provide participants with sufficient knowledge to manage data from collection to analysis in SAS for Windows. Emphasis will be placed on the data life cycle, database structures, cleaning, manipulation, documentation, and dynamic reporting of data. Upon completion, students should be able to manage a small project from start to finish, providing well documented reproducible code to generate a report, codebook and data set ready for data analysis. Macro programming, structured query language (SQL), public data sets and reporting tools commonly used in nutrition analysis will be utilized during the semester. In-class exercises and homework assignments will allow students to acquire hands-on experience solving common data management tasks in SAS.

Course Objectives:  
Upon completion, students should be able to:

1) Create a folder, file and versioning system to organize both small and large projects
2) Build a simple REDCap database to collect data, and export such data into SAS
3) Use SAS data and procedure (proc) steps to import from and export to data of various formats and types
4) Merge, append and update SAS datasets
5) Use SAS functions and statements to summarize and define new variables within and across records
6) Use existing macros and create user defined macros to automate data processing and analysis
7) Use structured query language (SQL) to manipulate and aggregate data
8) Utilizing #3-7 above, perform basic cleaning tasks to prepare study data
9) Produce descriptive statistics and graphical displays in pdf, Word, and Excel for presentation
10) Transform data in preparation for statistical analyses
11) Create codebooks and data sets ready for analysis and public release
12) Be prepared to take the SAS certification test for basic programming

Texts or Materials:

https://tufts.primo.exlibrisgroup.com/permalink/01TUN_INST/1kc9gia/alma991018128213103851

This semester students will meet in the small computer lab 510 in the HHSL. We will use the computers there in conjunction with Tufts Box to work in SAS.

SAS is also available on the TTS Virtual Lab. Students will need to download and install the VMWare Horizon Client (Windows and Mac versions) to access the Virtual Lab. Please see https://access.tufts.edu/sas for more details.

For those on who want to additionally install SAS on their personal computer, SAS software version 9.3 or higher is required. Tufts has a site license with SAS and students may install a copy on their laptop. Please see https://access.tufts.edu/sas to get more information on downloading and installing the software. A typical installation can take up to an hour or more so please have your AC cord connected. Alternatively, students may also find SAS installed on the computers on the fifth floor of the HHSL library and PHPD student lounge. Please note that SAS may not work equally well on all operating systems, check for the compatibility at: http://support.sas.com/supportos/list.
Diversity statement
I believe that the diversity of student experience and perspective is essential to the deepening of knowledge in this course. I consider it part of my responsibility as an instructor to address the learning needs of all the students in this course. I will present material that are respectful of diversity: race, color, ethnicity, gender, age, disability, religious beliefs, political preference, sexual orientation, gender identity, socioeconomic status, citizenship, language, or national origin among other personal characteristics.

Accommodation of Disabilities:
While I intend to present material in a way that is inclusive and accessible, I realize this might not address all necessary accommodations needed by student with disabilities. Students with documented disabilities are entitled to academic accommodation appropriate to their needs. If you require accommodations for this course, please contact me confidentially prior to the end of the second week of classes.

Academic Conduct: Each student is responsible for upholding the highest standards of academic integrity, as specified in the Friedman School’s Policies and Procedures handbook (https://nutrition.tufts.edu/about/policies-and-procedures) and Tufts University policies (http://uss.tufts.edu/studentaffairs/judicialaffairs/Academic_Integrity.pdf). It is the responsibility of each student to understand and comply with these standards, as violations will be sanctioned by penalties ranging from failure on an assignment and the course to dismissal from the school.

Classroom Conduct: Students must notify the instructor in advance if you will be absent for the class. Also, inform the instructor in charge if you decide to leave the class early. Students who miss the class are responsible for downloading and reviewing the lecture materials and submitting exercises from the class within 2 weeks of class missed. More than 2 absences may result in a lower participation grade.
**Assessment and Grading:** 8 homework assignments (5%*8=40%)  
Participation in in-class programming exercises 5%  
Piazza blog post on one SAS white paper and use case 5%  
Mid-term SAS knowledge test (online, take home) 20%  
Final data management project 30%

<table>
<thead>
<tr>
<th>Grading Scale</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>94-100</td>
</tr>
<tr>
<td>A-</td>
<td>90-93</td>
</tr>
<tr>
<td>B+</td>
<td>87-89</td>
</tr>
<tr>
<td>B</td>
<td>84-86</td>
</tr>
<tr>
<td>B-</td>
<td>80-83</td>
</tr>
<tr>
<td>C+</td>
<td>77-79</td>
</tr>
<tr>
<td>C</td>
<td>73-76</td>
</tr>
<tr>
<td>C-</td>
<td>70-72</td>
</tr>
<tr>
<td>D+</td>
<td>67-69</td>
</tr>
<tr>
<td>D</td>
<td>60-66</td>
</tr>
<tr>
<td>F</td>
<td>0-50</td>
</tr>
</tbody>
</table>

**Assignments and Submission Instructions:**

Homework is due at the assigned deadline. Homework assignments shall be submitted to the instructor via Canvas along with corresponding SAS code, logs, and output prior to the deadline. Assignments received after their deadline will not be accepted or graded unless an extension is approved in advance. Students who are unable to complete an assignment or exam on time for any reason should notify the instructor by email prior to the deadline.

Please submit all files to CANVAS using the following file naming convention: HMK#_LastName_First initial_YYYYMMDD.sas (or *.log, *.mht, etc.). For example, the first homework I would submit would include:

HMK1_Rogers_G_20220915.sas  
HMK1_Rogers_G_20220915.log  
HMK1_Rogers_G_20220915.mht

Lab exercises should be submitted to the class BOX folder at the end of each class. This submission will count as your attendance for the class and will not be graded. Only the *.sas program needs to be submitted and should follow the naming convention of LAB#_LastName_FirstIniital_YYYMDD.sas
Course Schedule:
* This schedule is subject to modification at the instructor’s discretion.

<table>
<thead>
<tr>
<th>Class #</th>
<th>Topic</th>
<th>Assignment dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (Sep 06)</td>
<td>Data Life cycle</td>
<td></td>
</tr>
<tr>
<td>2. (Sep 08)</td>
<td>Database structures, Intro to REDCap</td>
<td></td>
</tr>
<tr>
<td>3. (Sep 13)</td>
<td>Building your own REDCap database</td>
<td>Finish building database</td>
</tr>
<tr>
<td>4. (Sep 15)</td>
<td>Getting started with SAS</td>
<td>Assign Homework #1</td>
</tr>
<tr>
<td>5. (Sep 20)</td>
<td>Working with your data 1</td>
<td></td>
</tr>
<tr>
<td>6. (Sep 22)</td>
<td>Working with your data 2</td>
<td></td>
</tr>
<tr>
<td>7. (Sep 27)</td>
<td>Working with your data 3</td>
<td>Homework #1 due, Assign #2</td>
</tr>
<tr>
<td>8. (Sep 29)</td>
<td>Programming styles and best practices</td>
<td></td>
</tr>
<tr>
<td>9. (Oct 04)</td>
<td>Modifying and combining SAS data sets 1</td>
<td>Homework #2 due, Assign #3</td>
</tr>
<tr>
<td>10. (Oct 06)</td>
<td>Modifying and combining SAS data sets 2</td>
<td></td>
</tr>
<tr>
<td>11. (Oct 11)</td>
<td>SQL 1</td>
<td></td>
</tr>
<tr>
<td>12. (Oct 13)</td>
<td>Getting external data into SAS 1</td>
<td>Homework #3 due, Assign #4</td>
</tr>
<tr>
<td>13. (Oct 18)</td>
<td>Getting external data into SAS 2</td>
<td></td>
</tr>
<tr>
<td>14. (Oct 20)</td>
<td>Transposing SAS data sets</td>
<td>Homework #4 Due</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Assign Take home SAS midterm</td>
</tr>
<tr>
<td>15. (Oct 25)</td>
<td>Working with Longitudinal data</td>
<td>Take home due</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homework #5 Assigned</td>
</tr>
<tr>
<td>16. (Oct 27)</td>
<td>SQL 2</td>
<td></td>
</tr>
<tr>
<td>17. (Nov 01)</td>
<td>Macros 1: Using SAS macro variables</td>
<td>Homework #5 due, Assign #6</td>
</tr>
<tr>
<td>18. (Nov 03)</td>
<td>Macros 2: User-defined macro programs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Nov 08) No class (Monday schedule)</td>
</tr>
<tr>
<td>19. (Nov 10)</td>
<td>Macros 3: Accessing and using</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>public macros</td>
</tr>
<tr>
<td>20. (Nov 15)</td>
<td>Visualizing data with ODS graphs</td>
<td>Homework #6 due, Assign #7</td>
</tr>
<tr>
<td>21. (Nov 17)</td>
<td>Data cleaning</td>
<td></td>
</tr>
<tr>
<td>22. (Nov 22)</td>
<td>Using basic statistical procedures</td>
<td>Homework #7 due, Assign #8</td>
</tr>
<tr>
<td>23. (Nov 24)</td>
<td>No class Thanksgiving</td>
<td></td>
</tr>
<tr>
<td>24. (Nov 29)</td>
<td>Using the Output Delivery System</td>
<td></td>
</tr>
<tr>
<td>25. (Dec 01)</td>
<td>Exporting to external sources</td>
<td></td>
</tr>
<tr>
<td>26. (Dec 06)</td>
<td>Creating codebooks</td>
<td>Homework #8 due</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Take home project assigned</td>
</tr>
<tr>
<td>27. (Dec 08)</td>
<td>Last Class: Debugging Programs/Review</td>
<td>Piazza blog entry, last day accepted</td>
</tr>
</tbody>
</table>

Dec 19th  Take home project due on Canvas 11:59 pm
Course Topics, Learning Objectives and Assignments

* This schedule is subject to modification at the instructor’s discretion.

Class # 1: (9/6) Data life cycle and file management overview, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Understand the phases of the data life cycle
2) Understand the management tasks associated with each phase
3) Prepare folder and file naming structure for projects

Preparation for class:
none

Assignments for this class:
Following the guidance from class, create the folder structure shell you will use for this class.

Class # 2: (9/8) Database structures, Intro to REDCap, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Identify the elements of a database and how to refer to them
2) Identify the different variable types and when to use them
3) Identify primary keys for individual tables and understand how they relate to database structure

Preparation for class:

https://redcap.vanderbilt.edu/consortium/videoplayer.php?video=redcap_overview03.mp4&title=Detailed+Overview+of+REDCap+%2814+min%29&text=This+14-minute+video+provides+a+thorough+overview+of+REDCap+and+much+of+its+functionality.+This+video+is+an+excellent+place+to+begin+learning+about+REDCap+and+what+it+is+capable+of.&referer=REDCAP_PUBLIC

Assignments for this class:
none
Class # 3: (9/13) *Building your own REDcap database, Gail Rogers*

**Learning objectives:**
Upon completion of this class, students will be able to:
1) Understand how to translate data collection tools into a database table and variables
2) Understand the difference between cross sectional and longitudinal data
3) Create a simple REDCap database (class project)
4) Enter data into REDCap
5) Export data from REDCap

**Preparation for class:**
None

**Assignments for this class:**
*Finish class database*

Class # 4: (9/15) *Getting started with SAS, Gail Rogers*

**Learning objectives:**
Upon completion of this class, students will be able to:
1) Open and run a basic SAS program
2) Properly document SAS programs
3) Understand the basic functions and options of the editor, log, output and results viewer windows
4) Use the contents procedure and basic print statements to explore data
5) Use SAS and other documented internet resources to find syntax help

**Preparation for class:**
*Chapter 1 of The Little SAS Book*

**Assignments for this class:**
*Homework # 1 assigned*

Class # 5: (9/20) *Working with your Data Part 1, Gail Rogers*

**Learning objectives:**
Upon completion of this class, students will be able to:
1) Create and redefine variables
2) Use selected SAS numeric and character functions
3) Use If /then statements
4) Use simple do loops for iterative processing
5) Subset data
Preparation for class:

Chapter 3 of The Little SAS Book

Assignments for this class:
Continue working on assignment #1

Class # 6: (9/22) Working with Your Data Part 2, Gail Rogers

Learning objectives:
Upon completion of this week, students will be able to:
1) Create labels and formats
2) Convert data from character to numeric and numeric to character
3) Use SAS date functions
4) Use simple arrays

Preparation for class:
Continue reading chapter 3 of The Little SAS Book

Assignments for this class:

Class # 7: (9/27) Working with Your Data Part 3, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Sort and print data
2) Use the WHERE statement to restrict output
3) Summarize data with proc TABULATE
4) Use proc REPORT to display data

Preparation for class:
Read chapter 4 of The Little SAS Book

Assignments for this class:
Assignment #1 due, Assignment #2 given

Class # 8: (9/29) Programming style and best practices, Gail Rogers

Learning objectives:
Upon completion of this week, students will be able to:
1) Understand the use of comments and white space in programming
2) Understand best practice for program organization
3) Use LIBNAME and the macro %let for relative paths
2) Use proper file naming conventions to attach output to code

Preparation for class:
Guidelines for coding SAS programs:
An Animated Guide: Coding Standards for SAS® Production Programs - PO02.pdf
https://www.lexjansen.com/nesug/nesug09/po/PO02.pdf

Code like it matters: Writing code that is readable and sharable

Assignments for this class:

Class # 9: (10/04) Modifying and Combining SAS data sets 1, Gail Rogers, HHSL 510

Learning objectives:
Upon completion of this class, students will be able to:
1) Create multiple data sets from one file
2) Append SAS data sets
3) Use data set options in concatenating files

Preparation for class:
Chapter 6 of The Little SAS Book

Assignments for this class:
Homework #2 due, Homework #3 assigned

Class # 10: (10/06) Modifying and Combining SAS data sets 2, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
2) Merge SAS data sets (1:1, inner and outer joins)
3) Use the IN= operator to track and control observations (left, right joins)
4) Merge files using multiple variables
5) Use data set options in merging files
Preparation for class:
Continue reading Chapter 6 of The Little SAS Book

Assignments for this class:
Continue assignment #3

Class # 11: (10/11) SQL 1, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Understand basic SQL syntax
2) Use SELECT and FROM clause to query a single table
3) Use SELECT and FROM clause to query multiple tables
4) Use SQL to combine and modify SAS data sets (inner, left, right and outer join, 1:1, 1:many) and CREATE tables
5) Understand the use of SQL to do table look-ups for large databases
6) Use SQL to insert data into existing files

Preparation for class:
The Little SAS book, Appendix: Coming to SAS from SQL
TBD

Assignments for this class:
Continue assignment #3

Class # 12: (10/13) Getting your data into SAS 1, Gail Rogers

Learning objectives:
Upon completion of this week, students will be able to:
1) Enter data into SAS using viewtable and datalines
2) Open other major statistical programming data files using LIBNAME, FILENAME, and engine statements
2) Use SAS data step and input statements to read in complex *.txt and *.CSV files using List input

Preparation for class:
Chapter 2 of The Little SAS Book

Assignments for this class:
Homework #3 due, homework #4 assigned
Class # 13: (10/18) *Getting your data into SAS II, Gail Rogers*

**Learning objectives:**
Upon completion of this week, students will be able to:
1) Open SAS, SPSS, and EXCEL files using libname, filename and engine statements
2) Use SAS procedures to import *.txt, *.csv, Excel and other external data
3) Use SAS data step and input statements to read in complex *.txt files

**Preparation for class:**
Continue reading *Chapter 2 of The Little SAS Book*

**Assignments for this class:**
*Continue assignment #4*

Class # 14: (10/20) *Transposing data, Gail Rogers*

**Learning objectives:**
Upon completion of this class, students will be able to:
1) Use the transpose procedure to convert data from long to wide and wide to long
2) Use data step procedure to aggregate, separate and reshape data in preparation for data analysis

**Preparation for class:**
*Continue reading Chapter 6 of The Little SAS Book.*

**Assignments for this class:**
*Homework #4 due, assign online/take home midterm*

Class # 15: (10/25) *Working with Longitudinal data, Gail Rogers*

**Learning objectives:**
Upon completion of this class, students will be able to:
1) Use shortcuts for variable naming
2) Identify first and last observations in a group
3) Use the Retain and Sum statements
4) Count the number of observations per by group
5) Calculate across observations

**Preparation for class:**

Intro to Longitudinal Data: A Grad Student “How-To” Paper (on canvas)

Assignments for this class:
Take home/online midterm due, Homework #5 assigned

Class # 16: (10/27) SQL 2, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Use SQL to aggregate wide and long data
2) Use SQL + functions to summarize wide and long data
3) Use SQL to count and group observations
4) Understand the overlap and differences in SQL vs. data step programming

Preparation for class:
TBD

Assignments for this class:
Continue assignment #5

Class # 17: (11/01) Macros 1, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Access and use built in macro variables supplied by SAS
2) Understand how to access guidance on using the macro variables
3) Understand the concept of a user defined macro

Preparation for class:
TBD

Assignments for this class:
Homework #5 due, homework #6 assigned

Class # 18: (11/03) Macros 2, Gail Rogers

Learning objectives:
Upon completion of this class, students will be able to:
1) Create simple user defined macros for repetitive tasks
2) Understand and trace logic of simple macros written by other programmers
3) Debug simple macros
Preparation for class:

*TBD*

Assignments for this class:

*Continue homework #6*

**No class on 11/08 – Monday’s schedule**

**Class # 19: (11/10) Macros, Gail Rogers**

**Learning objectives:**

Upon completion of this class, students will be able to:

1) Find, download public macros for nutrition
2) Incorporate the public macros into the SAS session
3) Use `%include`
4) Understand the difference between compiled and uncompiled macros

Preparation for class:

*TBD*

Assignments for this class:

*Continue homework #6*

**Class # 20: (11/15) Visualizing data with ODS graphs, Gail Rogers**

**Learning objectives:**

Upon completion of this week, students will be able to:

1) Use and understand the basics of Output Delivery System (ODS Graphics)
2) Create scatterplots, box plots, histograms, and series plots in SAS
3) Customize and save graphics output for use in other software packages

Preparation for class:

*The Little SAS book, Chapter 8*

Assignments for this class:

*Homework #6 due, homework #7 assigned*

**Class # 21: (11/17) Data cleaning, Gail Rogers**

**Learning objectives:**

Upon completion of this class, students will be able to:

1) Identify duplicate primary keys and records
2) Identify data errors with regards to plausibility of values and temporality of data
3) Identify outliers and missing data
4) Use previous exposure to SAS coding to automatize these tasks
5) Use graphing to help identify errors
6) Document data corrections
7) Understand what “cleaned data” means

Preparation for class:
   TBD

Assignments for this class:
   Continue homework #7

Class # 22: (11/22) Basic Statistical Procedures, Gail Rogers

Learning objectives:
   Upon completion of this class, students will be able to:
   1) Use the frequency, univariate, means and correlation procedures
   2) Do simple statistical tests using the T-test, ANOVA and Regression procedures
   3) Understand the output from the above procedures and how to customize output and save results in data steps as well as *.html, and *.rtf files

Preparation for class:
   Chapter 9 of The Little SAS Book
Assignments for this class:
   Homework #7 due, homework #8 assigned

No class Nov 22 – Thanksgiving break

Class # 23: (11/29) Using the Output Delivery System, Gail Rogers

Learning objectives:
   Upon completion of this class, students will be able to:
   1) Understand the basics of ODS
   2) Work with table and style templates
   3) Trace and select output tables
4) Create files for various destinations including SAS data set, *.PDF, *.RTF and *.HTML

Preparation for class:
Chapter 5 of The Little SAS Book

Assignments for this class:
Continue assignment #8

Class # 24: (12/01) Exporting data sets to external sources
Learning objectives:
Upon completion of this class, students will be able to:
1) Export data and formats to open source files with documentation (CSV file)
2) Export data to SPSS, Stata, R etc.
3) Understand the uses of proc EXPORT and ODS destinations

Preparation for class:
The Little SA book, Chapter 10

Assignments for this week:
Continue homework #8

Class # 25: (12/06) Creating codebooks
Learning objectives:
Upon completion of this class, students will be able to:
1) Understand what should be supplied in a codebook
2) Investigate and understand syntax for program coding that already exists
3) Use Proc codebook to generate a codebook in word and PDF

Preparation for class:
Reading of the documentation around proc codebook

Assignments for this week:
Homework # 8 is due, take-home project is assigned

Class # 24: (12/08) Debugging Programs, review
Learning objectives:
Upon completion of this class, students will be able to:
1) Identify common SAS errors and how to address them

Preparation for class:
Reading of the documentation around proc codebook

Assignments for this week: Piazza blog entry, last day accepted