BIO
STATISTICS
COLLABORATION
OF AUSTRALIA

Study Guide

Data Management and Statistical Computing (DMC)

Semester 1, 2017

Prepared by:
Jennie Louise
School of Public Health
The University of Adelaide

THE UNIVERSITY
of ADELAIDE

CONTENTS

Academic Co-ordinator

Dr Jennie Louise

Data, Design and Statistics Service
Adelaide Health Technology Assessment
School of Public Health
The University of Adelaide
Adelaide SA 5005

Email: jennie.louise@adelaide.edu.au

Other Contacts

If you have trouble contacting the academic coordinator/academic staff, or have any other queries, please contact:
Erica Jobling

Executive Officer
Biostatistics Collaboration of Australia
BCA c/o NHMRC Clinical Trials Centre
Locked Bag 77
Camperdown NSW 1450

Phone: 02 9562 5076
Email: bca@ctc.usyd.edu.au

Welcome Letter

Welcome to Data Management and Statistical Computing (DMC). In this course we will develop statistical computing skills essential for managing and analysing data in health and medicine. This course provides an introduction to SAS and Stata, with the aim of giving you a foundation to build upon in your further studies and in your biostatistical career.

This course is delivered through the eLearning site at the University of Sydney. All course content other than readings will be uploaded to eLearning, including assignments and supplementary material, and discussions of material will take place on the Discussion Board. There is currently an Introductions thread on Discussion Board; please use this thread to introduce yourself to the rest of the class.

This course requires access to two statistical software packages: SAS and Stata. You should organise access to these as soon as possible. Instructions for purchasing a license, or accessing the web-based SAS on Demand, are given in the BCA Textbook and Software Guide.

If you have any questions or issues, please contact me by email at the address above.

I hope you enjoy the course!

Jennie Louise
January 2017

Course Objectives

This course aims to provide students with skills to undertake moderate to high level data management, manipulation, and analysis. On completion of this unit, students should:

1. Be able to undertake data manipulation and management using two major statistical software packages (Stata and SAS);

2. Be able to appropriately display and summarise data using statistical software;

3. Understand how to check and clean data;

4. Be able to link data files through unique and non-unique identifiers;

5. Have fundamental programming skills for efficient use of statistical software;

6. Understand key principles of confidentiality and privacy in data storage, management and analysis.

Course Content

This course consists of three modules:

*   Module 1: The basics. Importing and exporting data; recoding and formatting data; labelling variables and values; use of date data, displaying and summarising data.
*   Module 2: Graphs, Data management and Statistical Quality Assurance Methods. Includes advanced graphics for production of publication-quality graphs.
*   Module 3: Data Management. Using functions to generate new variables; appending, merging and transposing data; programming skills including loops, arguments and programs/macros.

Each module requires approximately 4 weeks of study; the final week of semester will be left for revision, or to cover other issues which arise during the course.

Course material consists of (a) the notes which are provided for each module, (b) the text books and other required reading, and (c) further notes, code and data files which will be provided on eLearning.

Text Books

You should have access to the following textbooks:

Jull S, Frydenberg M. An Introduction to Stata for Health Researchers, 4th ed. Stata Press, 2014.
Cody RP, Smith JK. Applied Statistics and the SAS Programming Language, 5th ed. New Jersey: Prentice Hall, 2005.

Readings

In addition to the text books, various articles are set as required readings in each module. These cannot be uploaded to eLearning. You can access the articles through your university's library; further assistance in accessing readings will be given during the course.

Software

You should have access to the following software packages:
- Stata version 12 or later (the latest version is v14)
- SAS version 9.2 or later (the latest version is 9.4)

If you have not yet organised access to these packages, you should do so as soon as possible. This is a practical course which requires regular use of the relevant software; delays in gaining access to these packages may impact your ability to complete the course.

Information on how to purchase software, or arrange access via a university site license or SAS on Demand, can be found in the BCA Textbook and Software Guide.

Method of Delivery

DMC is taught as an online course. For information on eLearning, see the BCA Introduction to eLearning. For further assistance with eLearning, you can contact the eLearning Helpdesk.

Module notes, data files and other documents will be made available on eLearning. Assignments and course announcements will likewise be uploaded to eLearning.

Communication should generally be via the Discussion Board on eLearning (unless of a personal/confidential nature). You are encouraged to post questions, ideas, suggestions and discussions on eLearning. The Course Coordinator will monitor and respond to communication; however, you are encouraged to answer other students' questions or assist in solving problems.

Assessment

The assessment for this unit consists of three assignments:
- Assignment 1 will cover Module 1, and is worth 30% of the overall course mark. It is due before midnight (EST) on Monday 3rd April 2017.

- Assignment 2 will cover Module 2, and is worth 35% of the overall course mark. It is due before midnight (EST) on Monday 8th May 2017.
- Assignment 3 will cover Module 3, as well as Modules 1 and 2, and is worth 35% of the overall course mark. It is due before midnight (EST) on Monday 12th June 2017.

All assignments will be posted on eLearning three weeks before the due date. Individual feedback will be provided to each student; model solutions will also be provided once marked assignments have been returned. Summary statistics on results for the entire class will also be provided.

Assignments should be submitted via the assignment submission tool on eLearning; if you experience difficulties with this submission method, assignments can be submitted via email.

Extensions

For various reasons, you may sometimes experience difficulties in getting your assignments submitted on the due date. Requests for an extension for an assignment must be made **in advance of the due date for that assignment**. The normal grounds for an extension being granted are bereavement, personal illness or illness in a family member requiring you to exercise a significant carer role.
These requests must be made directly to Jennie Louise by email, and should include appropriate documentation (e.g. medical certificate). The time and date of the request will be noted, and a reply sent by email with the decision as to whether an extension has been granted and, if so, stating the length of the extension.
Length of extension: Extensions granted by Unit Coordinators will normally be no longer than three days.

Penalties for Late Submission

Assignments should be submitted no later than midnight EST on the due date. Submissions after this time will be penalised at a rate of 5% of the earned mark per day, up to a maximum of 50%. Submissions after the solutions have been posted on eLearning will not be awarded any marks.
For example, if your mark for an assignment is 40/50 but you submit it two days late, 10% of your mark will be deducted so your final mark will be 36/50.

Assignment Cover Sheet

Where assignment work is submitted online using the Assignment tool in eLearning, you will be able to indicate your compliance with the plagiarism guidelines and policy by electronic means. In this case, you **do not** need to complete the DMC 2017 Assignment Cover Sheet.

If you submit work by another method, then you **do** need to complete the DMC 2017 Assignment Cover Sheet, in which you will be asked to certify that the submission is your own work and that you have read the policy of the university at which you are enrolled (see Appendix 2). The cover sheet can also be downloaded from eLearning.
If you are posting your submission, please include the signed cover sheet in the envelope.

If you are submitting via email, please scan the signed cover sheet and submit this with your assignment, or fax the signed cover sheet to the number specified on the sheet.

Please refer to the BCA Assessment Guide for further information.

Course Timetable

Semester 1, 2017 will commence on Monday 6th March.

Week	Week Commencing	Module	Assessment
1	Monday 6th March	1	
2	Monday 13th March	1	Assignment 1 Available
3	Monday 20th March	1	
4	Monday 27th March	1	
5	Monday 3rd April	2	**Assignment 1 Due**
6	Monday 10th April	2	Assignment 2 Available
Monday 17th – Friday 21st April Mid-Semester Break			
7	Monday 24th April	2	
8	Monday 1st May	2	
9	Monday 8th May	3	**Assignment 2 Due**
10	Monday 15th May	3	
11	Monday 22nd May	3	Assignment 3 Available
12	Monday 29th May	3	
13	Monday 5th June	Revision	
	Monday 12th June		**Assignment 3 Due**