STATISTICS AND DATA ANALYSIS I: INTRODUCTION

ICPSR SUMMER PROGRAM, 2017
June 27-July 21, 3:00-5:00pm

COURSE INSTRUCTOR:

Niccole Pamphilis
University of Glasgow
Niccole.Pamphilis@glasgow.ac.uk

TEACHING ASSISTANT:

Muna Adem
Indiana University
munaadem@umail.iu.edu

AIMS AND OBJECTIVES:

This workshop provides a basic introduction to statistics, probability, and data analysis. Topics include data acquisition/management, classification, and summarization; basic probability; exploration of common distribution used in statistics; along with confidence intervals and hypothesis testing.

The format for this workshop consists of daily lectures supplemented by additional sessions in the computing labs. There will be frequent assignments in order to give workshop participants hands-on experience with the methods and techniques covered in the class. Most assignments will involve the use of statistical packages, which will be available for use in computer labs. Workshop participants will be given the opportunity to do their work using the statistical software packages that are most common in their respective academic disciplines, including SPSS, R, and Stata.

The only prerequisite for this course is familiarity with basic algebra. Participants who need a review of basic algebraic concepts are advised to enroll in the ICPSR Summer Program lecture course Mathematics for Social Scientists I along with this workshop.

TEXTBOOKS:

Recommended:

COURSE STRUCTURE:

Readings:

Daily readings are outlined below. It is recommended that readings are completed prior to our meetings to allow for questions to be asked during class. Readings are based on the textbook listed above. Though additional readings from other texts may be useful for additional details or alternative presentations of the material.

Daily Meetings:

Lectures will take place daily and will focus on the material noted in the readings below.

Labs:

Lab sessions will be held to provide you with an opportunity to become familiar working with the statistical software used in your field. Lab sessions will also provide an opportunity to work through assignments and ask questions. Lab times and locations will be provided during the 1st week of classes.

Assessment:

Four pieces of assessment will be graded and used to determine final grades for the course. The assessments will increase in difficulty as the course progresses. All four pieces will require the use of statistical software, so be sure to leave time to access the computer labs (if needed).

The first three pieces of assessment, which will be distributed at intervals over the next four weeks. Each person seeking a grade must submit their own piece of work assessment.

Each piece must be:

- Typed, organized, and stapled
- Answers must be presented in a clearly labelled manner
- Include an appendix with the code you used to locate your final results
- Clearly label your final answers
- Any graphs and visuals that are used should be placed in texts as part of your answer and not as a separate attachment
- Make sure to answer all parts of a question (i.e., prompts to interpret/ explain your answer)

The fourth and final piece of assessment will be a group project and
presentation. Information on this piece of assessment will be distributed during week two and the project will be due on the final day of the course.

GRADING:

Your final grade will be based on four pieces of assessment. Assessments vary in difficulty and are thus weighted differently to reflect this:

- Assessment 1: 15%
- Assessment 2: 25%
- Assessment 3: 35%
- Project and Presentation: 25%
SYLLABUS AND READINGS:

**WEEK 1**

27 June 2017  Introduction; What are data and statistics  
Recommended Reading:  
Herzog Chapters 1 and 3  
Suggested Reading:  
Weiss 1.1 and 2.1-2.2

28 June 2017  How do we measure the world around us

29 June 2017  LAB DAY: Using Secondary Data Sources  
Recommended Reading:  
Herzog Chapter 5

30 June 2017  Concept to Operationalisation & Lab 1  
Assignment 1 *Given*

**WEEK 2**

03 July 2017  What are we seeing? What does it tell us?  
Recommended Reading:  
Weiss Chapter 3

04 July 2017  NO CLASS

05 July 2017  Visualising data  
Recommended Reading:  
Herzog Chapter 11 and 12  
Suggested Reading:  
Bartholomew Chapter 1  
Weiss Chapter 1.3-1.5; 2.3-2.6  
Assignment 1 *DUE*

06 July 2017  Lab 2  
Assignment 2 *Given*

07 July 2017  How do we know if differences matter (probability)  
Recommended Reading:  
Bartholomew Chapter 8 and 9  
Suggested Reading:  
Weiss Chapter 4
**WEEK 3**

10 July 2017  The Normal Distribution and Z-scores  
Recommended Reading:  
Bartholomew Chapters 3 and 4  
Suggested Reading:  
Weiss Chapter 6.1-6.3 and Chapter 7

T-Distribution, One-Sample T-tests, Critical Values

11 July 2017  Approach  
Recommended Reading:  
Weiss Chapter 8  
Suggested Reading:  
Bartholomew Chapter 10  
Determining Differences by groups, Probability

Assignment 2 DUE

12 July 2017  Approach  
Recommended Reading:  
Weiss Chapter 9.3-9.6 and  
Chapter 10.1-10.2

13 July 2017  Differences beyond the mean and confidence intervals  
Recommended Reading:  
Weiss Chapter 10.3

Assignment 3 Given

14 July 2017  Lab 3

**WEEK 4**

17 July 2017  Understanding Relationships (discrete variables)  
Recommended Reading:  
Weiss Chapter 13  
Suggested Reading:  
Bartholomew Chapter 10

Assignment 3 DUE

18 July 2017  Measuring relationships  
Recommended Reading:  
Weiss Chapter

19 July 2017  Understanding Relationships, part II (continuous variables)  
Recommended Reading:  
Weiss Chapter 14

20 July 2017  Putting it all together; Lab 4

21 July 2017  Group Presentations  
Presentations DUE