**Stat E-100:
Introduction to Quantitative Methods
for the Social Sciences and Humanities**

Tentative Syllabus for Fall Semester 2015

(Last Updated September 8, 2015)

**Please Note:** The syllabus is required reading for the course. You will be expected to understand the policies and assignments discussed in the syllabus.

**Overview:** This course introduces the basic concepts of data analysis and statistical computing, both increasingly used in the social sciences and the humanities. The emphasis is on the practical application of quantitative reasoning, visualization, and data analysis. The goal is to provide students pragmatic tools for assessing statistical claims and conducting their own basic statistical analyses. Topics covered include basic descriptive measures, measures of association, sampling and sample size estimation, and simple linear regression. Assignments are based on real-world data and problems in a wide range of fields in the social sciences and humanities, including psychology, sociology, education, and public health.

# I. Course Mechanics

## Instructor: Ethan Fosse (Ph.D., Harvard)

Office Hours and Location: To be determined

Email: efosse@fas.harvard.du or fosse.ethan@gmail.com

**Head Teaching Assistant:** Mark Ouchida (M.Ed., A.L.M.)

Office Hours and Location: To be determined

Email: mark\_ouchida@harvard.edu

**Course Lectures:** Thursdays, 5:30-7:30pm, Maxwell-Dworkin G115

For information on the location of Maxwell-Dworkin G115, please see the following:

<http://imsroombook.fas.harvard.edu/room.php?rm=rm0092>

**Weekly Sections:** There will be optional weekly sections during which teaching assistants will review topics from the lectures, review additional example problems, and answer questions about the weekly homework assignments. More details will be available as the semester begins.

**Teaching Assistant Office Hours:** Each teaching assistant will also hold weekly office hours. More details will be available as the semester begins.

**Assigned Teaching Assistant:** You will be assigned a teaching assistant for the duration of the course after the semester starts. You may attend any section (or all sections) that fit your schedule, but your assigned teaching assistant is your first contact for grading issues, section issues, and questions about the homework or exams.

**II. Course Resources**

**Course Website:** The course website will have additional information on the course mechanics, homework assignments, lecture materials, and the teaching staff. Also note that all exams will be taken entirely online and that problem sets will be submitted entirely through course website. It will be updated as the semester starts. The course website may be accessed here (make sure you log in so you can view the website):<https://canvas.harvard.edu/courses/4131>

**Required Textbook:** The required textbook is *OpenIntro Statistics, 3rd Edition* by David M. Dietz, Christopher D. Barr, and Mine Cetinkaya-Rundel published in 2015. It is available online for **free** as a pdf or for less than $10 from Amazon: <https://www.openintro.org/stat/textbook.php?stat_book=os>

**Statistical Software:** The course will use R with RStudio. R is the underlying programming language, while RStudio is a graphical user interface that makes working with R much easier. Both are free, open-source, and used widely by biostatisticians. To install R with RStudio, go to the following link and click on the installer for your computer’s operational system: <https://www.rstudio.com/products/rstudio/download/>

# III. Grading and Course Requirements for Undergraduates

The course requirements differ slightly for undergraduate and graduate students. The information in this section applies **only** to students enrolled as **undergraduate students**. If you are enrolled as an undergraduate, then your overall grade is based on the following:

* Regularly-assigned **problem sets** that will count for **30%** of your grade
* A **midterm exam** that will count for **30%** of your grade
* A **final exam** that will count for **40%** of your grade

**Problem Sets (30%):** Problem sets are assigned on a semi-weekly basis and submitted online through the course website. Details on the problem sets will be published on the course website as the semester starts. Late problem sets are not accepted for any reason. Working with other students on the homework is allowed and encouraged but only as long as you hand in your **own work** and do not simply copy the work of someone else.

**Midterm Exam (30%)**: The midterm exam will last 2 hours and will be **offered online** on the course website during a specified 24-hour time period. The midterm exam will consist of three parts: first, a set of multiple choice questions; second, a set of questions in which you will be asked to enter a numerical answer; finally, two written questions in which you will be prompted to critically evaluate a hypothetical analysis or presentation of data. The midterm exam is “open book” in that you can use any textbook, notes, problem sets, or lecture slides to help you answer the questions.

**Final Exam (40%):** The final exam will last 2 hours and will be **offered online** on the course website during a specified 24-hour time period. Although the final exam is cumulative, it will focus more on concepts covered after the midterm. The final exam will consist of three parts: first, a set of multiple choice questions; second, a set of questions in which you will be asked to enter a numerical answer; finally, two written questions in which you will be prompted to critically evaluate a hypothetical analysis or presentation of data. The final exam is “open book” in that you can use any textbook, notes, problem sets, or lecture slides to help you answer the questions.

# IV. Grading and Course Requirements for Graduate Students

The course requirements differ slightly for undergraduate and graduate students. The information in this section applies **only** to students enrolled as **graduate students**. If you are enrolled as a graduate student, then your overall grade is based on the following:

* Regularly-assigned **problem sets** that will count for **20%** of your grade
* A **midterm exam** that will count for **20%** of your grade
* A **final exam** that will count for **30%** of your grade
* A **final project** that will count for **30%** of your grade

**Problem Sets (20%):** Problem sets are assigned on a semi-weekly basis and submitted online through the course website. Details on the problem sets will be published on the course website as the semester starts. Late problem sets are not accepted for any reason. Working with other students on the homework is allowed and encouraged but only as long as you hand in your **own work** and do not simply copy the work of someone else.

**Midterm Exam (20%)**: The midterm exam will last 2 hours and will be **offered online** on the course website during a specified 24-hour time period. The midterm exam will consist of three parts: first, a set of multiple choice questions; second, a set of questions in which you will be asked to enter a numerical answer; finally, two written questions in which you will be prompted to critically evaluate a hypothetical analysis or presentation of data. The midterm exam is “open book” in that you can use any textbook, notes, problem sets, or lecture slides to help you answer the questions.

**Final Exam (30%):** The final exam will last 2 hours and will be **offered online** on the course website during a specified 24-hour time period. Although the final exam is cumulative, it will focus more on concepts covered after the midterm. The final exam will consist of three parts: first, a set of multiple choice questions; second, a set of questions in which you will be asked to enter a numerical answer; finally, two written questions in which you will be prompted to critically evaluate a hypothetical analysis or presentation of data. The final exam is “open book” in that you can use any textbook, notes, problem sets, or lecture slides to help you answer the questions.

**Final Project (30%):** The final project will consist of a short paper (no more than 10 pages) describing the analysis of a dataset using one the methods discussed in the course. The final project will be submitted online through the course website. Details on the final project will be given as the semester starts.

# V. Tentative Course Schedule

The tentative course schedule is given below. It is subject to change as the semester begins. The due dates for problem sets, assigned textbook readings, and lecture topics are subject to change.

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Lecture Topic** | **Textbook Readings**  | **Problem Sets Due** |
| Sept 3 | Introduction to Data | Ch. 1: Pp. 7-13, 15-20  | None |
| Sept 10 | Categorical Data | Ch. 1: Pp. 43-48  | None |
| Sept 17 | Numerical Data | Ch. 1: Pp. 28-37  | Problem Set #1  |
| Sept 24 | Probability Tables and Relative Risk  | Ch. 2: Pp. 76-101 | Problem Set #2 |
| Oct 1 | Correlation Analysis | Ch. 1: Pp. 26-27; Ch. 7: Pp. 338-339 | Problem Set #3 |
| Oct 8 | Simple Linear Regression  | Ch. 7: Pp. 331-338, 340-348 | Problem Set #4 |
| Oct 15 | Basics of Sampling | Ch.1: Pp. 20-23; Ch. 3: Pp. 127-141; Ch. 4: Pp. 168-174 | Problem Set #5 |
| Oct 22 | **MIDTERM EXAM** |  | None |
| Oct 29 | Sampling Distribution | Ch. 4: Pp. 175-197  | None |
| Nov 5 | Tests for Means | Ch.5: Pp. 219-239  | Problem Set #6 |
| Nov 12 | Tests for Proportions | Ch. 6: Pp. 274-286 | Problem Set #7 |
| Nov 19 | Tests for Contingency Tables  | Ch. 6: Pp. 297-302 | Problem Set #8 |
| Nov 26 | **Thanksgiving Break** |  | None |
| Dec 3 | Inferences for Correlation and Simple Linear Regression | Ch. 7: Pp. 351-355 | Problem Set #9 |
| Dec 10 | Course Review | None | Problem Set #10 |
| **Dec 17** | **FINAL EXAM** | None | None |

# VI. Other Course Policies

# Collaboration: You may discuss the problem sets with other students, but you must write the final answer yourself. Solutions prepared by copying or paraphrasing someone else’s work are not acceptable. All computer output you submit must come from work that you have done yourself.

# Academic Integrity: Harvard University expects students to understand and maintain high standards of academic integrity. Breaches of academic integrity are subject to review and may be grounds for disciplinary action. Please review the examples of violations of academic integrity at the following link: <https://www.extension.harvard.edu/resources-policies/student-conduct/academic-integrity>

# Students with Disabilities: Harvard University has made a commitment to creating an accessible academic and campus community. If you have a disability, we ensure that you have equal opportunity to participate in, contribute to, and benefit from our academic and residential programs. Additional information may be found here: <https://www.extension.harvard.edu/resources-policies/resources/disability-services-accessibility>