

# Social Statistics (Sociology 3811)

Zack W. Almquist  
Spring Semester, 2018

## Class Schedule

Lecture: TU 5:30-8:00 pm Blegen Hall 317  
Lab TH 5:00-5:40 & 6:00-7:40 pm Blegen Hall 440

**URL:** <https://canvas.umn.edu>

*Note:* Requires UMN login and registration in class to access.

**URL:** <http://www.github.com>

*Note:* Requires github login.

## Professor

**Name:** Zack W. Almquist  
**Office:** 1168 Social Science Building  
**Office Hours:** M 11:00AM - 1:00PM  
**Email:** almquist@umn.edu  
**Telephone:** 612-624-4300 (not recommended)

## Teaching Assistant

**Name:** Neeraj Rajasekar  
**Office:** Social Sciences Building 1146  
**Office Hours:** M 2:00-4:00pm  
**Email:** rajas011@umn.edu

## Course Objectives

SOC 3811 is an introduction to statistical thought with a focus on methods and techniques used in the social sciences. Note that many (if not all) of the methods and techniques discussed in this course can be equally applied to the physical or engineering sciences even

though the examples are drawn from the social sciences. Students are not expected to become expert statisticians, but they are expected to gain an understanding of how statistics can be used to address problems within the social sciences. Further, students will become knowledgeable and critical consumers of statistical information that appear in everyday life, such as in the media or work place. Last, the student will also be given the opportunity to acquire competency in basic data management and analysis tasks within the R statistical programming environment.

This course will cover basic methods for describing data, performing basic statistical tests and understanding conditional modeling approaches. Specifically we will discuss key metrics of data analysis such as the mean, median and mode, as well as expose the student to visualizations of data such as the histogram or bar-plot. Basics of statistical tests will be discussed, including basic exposure to hypothesis testing and statistical significance. In addition, basic exposure to statistical (or probability) modeling will be discussed with an emphasis on linear models.

The overall structure of this class is as follows: (1) introduction to and overview of data, (2) statistical inference, and (3) statistical models. Theme one will focus on introduction to the managing and manipulation of data, to the measuring and metrics used to understand data, and to the visualization of data. Theme two will focus on the introduction to key statistical concepts such as population and sample data and the basics of statistical inference. Finally, the third theme will focus on statistical (probability) models and basic methods for describing and modeling conditional distributions.

## Prerequisites

There are no formal prereqs for this course.

## Course Requirements

### Computers

It is not required that students bring their computers/laptops to lecture and lab (if one is owned), but it is *highly* recommended since lecture will make extensive use of the computer software R. Computer labs are available on campus, please consult with the office if you have trouble finding the various locations of campus computer labs.

### Readings

Weekly readings assignments can be found on the course syllabus. All readings are assumed to be completed before each lecture/seminar. You are expected to read over the class notes

each week and make sure you are familiar with the material as the course progresses. Questions are encouraged.

## Homework

Homework assignments will normally be administered on an approximately bi-weekly basis and will be due on every other Thursday at 10:00pm via canvas. Homework assignments are meant to achieve three results: (1) provide practice with the statistical concepts discussed in class and, and (2) provide practice with the computational and statistical programming language R and (3) provide a chance to demonstrate your mastery of material and highlight areas where more work is needed. You may work in a group, but all write-ups must be done independently. All collaborators should be appropriately cited in your write up and any detailed R code should also be provided.

Homework must be turned in via canvas and will only be accepted in pdf format. A zero grade will be given for any homework not submitted as a pdf. Your lowest homework grade will be dropped.

## Exams

To assess mastery of course material, three in class exams will be administered over the course of the semester. Each exam will be cumulative. The examinations will cover all material presented in lecture, lab, and assigned readings. In order to prepare for exams, students are advised to keep up-to-date on reading assignments and to attend lectures regularly. Your lowest exam grade will be dropped.

## Participation and labs

Individuals are expected to attend every course, to have completed every reading, and to participate with questions and discussion on each topic as presented. If you plan on missing any class period you are responsible for all material and for contacting the instructor in a timely manner. Labs will have an in-class practice activity that will be submitted via canvas at the end of every lab. Grading will be assigned based on participation and not mastery. Your lowest lab grade will be dropped.

## Grading

Homework:	30%
Participation:	10%
Labs:	10%
Exams:	50%

Lectures, readings, labs, and review sessions are provided for each student's benefit. It is the responsibility of the student to take advantage of these opportunities to acquire and demonstrate mastery of course material, so as to achieve his or her desired grade.

### Letter grade assignment

A	93%+
A-	90-92.99%
B+	87-89.99%
B	83-86.99%
B-	80-82.99%
C+	77-79.99%
C	73-76.99%
C-	70-72.99%
D	60-69.99%
F	<59.99%

### Required Texts

- David M Diez, Christopher D Barr and Mine Çetinkaya-Rundel. Introductory Statistics with Randomization and Simulation.  
[https://www.openintro.org/stat/textbook.php?stat\\_book=isrs](https://www.openintro.org/stat/textbook.php?stat_book=isrs).
- Golemund, Garrett and Hadley Wickham. (2017). R for Data Science. O'Reilly Press.  
<http://r4ds.had.co.nz/>.

### Readings

Be prepared to discuss all readings assigned at anytime in lecture/seminar.

### Required Software

#### Web browser

All software used in this course will be available in a web browser. You will access R and RStudio via <https://rstudio.cla.umn.edu> which will require your UMN login. All spreadsheet examples and requirements will be done in google sheets and are available via your UMN google account.

For those interested, details on how to install the statistical software on your own machine are made available below. **Not required.**

## Software (Not required)

We will be using the R statistical programming language. R can be downloaded at <http://www.r-project.org/>.

RStudio IDE (Integrated Development Environment) is a software application which facilitates interaction with the R statistical programming language. It is often preferred to the GUI (Graphic User Interface) made available through CRAN. You can download it at <http://www.rstudio.com/>.

A github account is not required, but can be useful. One can register for a github account at <https://github.com/>. You can find information about how github works with Rstudio at <http://z.umn.edu/rstudiogit>, and github maintains a quite good help-system at <https://help.github.com/>.

## Course Policies

### Missing Class, etc.

It is expected that each member of the class will attend every lecture/discussion. If there is an appropriate reason to miss class it is expected that the individual will email or discuss in person with the instructor at least one week in advance. For any medical issues please see the UMN website for university policies.

### Cheating, etc.

All work is assumed to be your own and all individuals are expected to follow the university policy on cheating and misconduct. If you have any questions please consult the UMN website for university policies.

## Class Structure

This class meets for two and half hours one day a week for approximately 15 weeks. Each week the course will be divided into two pieces: (1) lecture, and (2) lab. Lecture will focus on introducing the necessary statistical material and basic exposure to modern statistical software. The lab will focus on introducing students spreadsheets and R.

## Assignments and Readings

### Theme: Data

#### Week 1 (01/16/18): Introduction to Data

- *Readings:*
  - ISRS: Sections 1.1-1.2
  - Watch this video <https://www.youtube.com/watch?v=1NWWQ2oxNho>
  - RDS 1.4-1.6
- *Lab* ( 01/18/18 ):
  - Due in Lab
- *Homework:*
  - Assignment 1

#### Week 2 (01/23/18): Data Collection

- *Readings:*
  - ISRS: Sections 1.3-1.5
- *Lab* ( 01/25/18 ):
  - Due in Lab
- *Homework:*
  - Due on 01/25/18 at 10:00pm

#### Week 3 (01/30/18): Describing Data

- *Readings:*
  - RDS: Chapter 4-5 and 12
- *Lab* ( 02/01/18 ):
  - Due in Lab
- *Homework:*
  - Assignment 2

#### **Week 4 (02/06/18): Visualizing Data**

- *Readings:*
  - RDS: Chapter 3
- *Lab ( 02/08/18 ):*
  - Due in Lab
- *Homework:*
  - Due on 02/08/18 at 10:00pm

#### **Week 5 (02/13/18): EXAM I**

- *Readings:*
  - NO READING
- *Lab ( 02/15/18 ):*
  - Due in Lab
- *Homework:*
  - None

#### **Theme: Inference**

#### **Week 6 (02/20/18): Introduction to Statistical Tests**

- *Readings:*
  - ISRS: Section 2.1-2.3
- *Lab ( 02/22/18 ):*
  - Due in Lab
- *Homework:*
  - Assignment 3

### **Week 7 (02/27/18): Introduction to Statistical Tests**

- *Readings:*
  - ISRS: Section 2.4-2.8
- *Lab ( 03/01/18 ):*
  - Due in Lab
- *Homework:*
  - Due on 03/01/18 at 10:00pm

### **Week 8 (03/06/18): Statistical Significance: Categorical Data**

- *Readings:*
  - ISRS: Chapter 3
- *Lab ( 03/08/18 ):*
  - Due in Lab
- *Homework:*
  - Assignment 4

### **Week 9 (03/13/18):**

[*Spring Break*]

- NO CLASS

### **Week 10 (03/20/18): Statistical Significance: Numerical Data**

- *Readings:*
  - ISRS: Sections 4.4-4.5
- *Lab ( 03/22/18 ):*
  - Due in Lab
- *Homework:*
  - Due on 03/22/18 at 10:00pm



### **Week 11 (03/27/18): EXAM II**

- *Readings:*
  -
- *Lab* ( 03/29/18 ):
  - Due in Lab
- *Homework:*
  - None

### **Theme: Models**

### **Week 12 (04/03/18): Introduction to Models**

- *Readings:*
  - ISRS: Sections 5.1-5.3
- *Lab* ( 04/05/18 ):
  - Due in Lab
- *Homework:*
  - Assignment 5

### **Week 13 (04/10/18): Linear Models**

- *Readings:*
  - ISRS: Sections 5.4-5.5
- *Lab* ( 04/12/18 ):
  - Due in Lab
- *Homework:*
  - Due on 04/12/18 at 10:00pm

**Week 14 (04/17/18): Generalized: Linear Models: Logistic Regression**

- *Readings:*
  - Chapter 6
- *Lab ( 04/19/18 ):*
  - Due in Lab
- *Homework:*
  - Assignment 6

**Week 15 (04/24/18): Review**

- *Readings:*
  - No Reading
- *Lab ( 04/26/18 ):*
  - Due in Lab
- *Homework:*
  - Due on 04/26/18 at 10:00pm

**Week 16 (05/01/18): Exam III**

- *Readings:*
  - No Reading
- *Lab ( 05/03/18 ):*
  - Due in Lab
- *Homework:*
  - None

**Week 17 (05/08/18): Finals**

- NO FINAL