Instructor: Brian Kim
E-mail: kimbrian@umd.edu
Office: 1218S LeFrak Hall
Office Hours: Mon/Wed 2:00 PM – 3:00 PM, Tues 1:00 PM – 2:00 PM, and by appointment

Advances in hardware and software technologies have led to a rapid increase in the amount of data collected, with no end in sight. Decision making in the coming decades will depend, to an even greater extent, on extracting meaning and knowledge from all that data. In this class we focus on one branch of statistics, inferential statistics, to help us reason about data. By gathering datasets, formulating proper statistical analyses and executing these analyses, information professionals play a significant role in bridging the gap between raw data and decision making.

This course will introduce basic concepts in data analytics including study design, measure construction, data exploration, hypothesis testing, and statistical analysis. The course also provides an overview of commonly used data manipulation and analytic tools. Through homework assignments, projects, and in-class activities, you will practice working with these techniques and develop statistical reasoning skills.

LEARNING OBJECTIVES

After completing this course, you will be able to:

- Select and evaluate various types of data to use in decision making;
- Use prescriptive and descriptive analyses to reach defensible, data-driven conclusions;
- Select and apply appropriate statistical methods;
- Use R for basic data manipulation and analysis;
- Critically evaluate data analyses and develop strategies for making better decisions.

COURSE MATERIALS

Technology

The following is necessary for you to successfully complete the homework, labs, exams, and projects for this course.

Required:

- We will have in-class lab sections during which you will need a laptop computer. If you do not have a laptop, please contact me after the first class.
- You must install and use R. R programming language and software is free and available online (https://www.r-project.org/). R Studio is also highly recommended; R studio is an integrated development environment for R (http://web.cs.ucla.edu/~gulzar/rstudio/). I have posted links to R tutorials on Elms/Canvas.

Readings:

Completing the required reading for the class is essential to understanding the core statistical concepts. In order to learn, you must review the material multiple times. The required reading listed below in the course schedule is given for the main textbook used in the class. There are a few advanced topics that are beyond the course textbook, readings for these sections will be posted on Elms/Canvas.

Required:

- Diez, D., Barr, C., and Çetinkaya-Rundel, M. (2015), *OpenIntro Statistics*, 3rd ed., OpenIntro. https://www.openintro.org/stat/textbook.php
  - Available as a free pdf download, color hardcover of black & white paperback for purchase, and free interactive website.
- Online Statistics Education: A Multimedia Course of Study (http://onlinestatbook.com/). Project Leader: David M. Lane, Rice University.
  - This is a free, online stats book available via HTML, pdf, and e-pub.

Optional:

For many different reasons, not everyone likes the required textbook in a quantitative methods course. Alternatively, some of you may wish to further explore some of the course topics and how to use statistical software. The following books have been recommended or are used by my colleagues, and they address many of the same course topics from varied writing approaches. I will not provide copies of these books, but all of them should be available through a wide range of bookstores.

- Imai, K. (2017), *Quantitative Social Science: An Introduction*, Princeton Press. ISBN-13: 9780691175461
- Salkind, N.J. (2013), *Statistics for People Who (Think They) Hate Statistics*, *5th ed.,* Sage Publications, Inc. ISBN-13: 978-1452277714
- Hinton, P. (2004) *Statistics Explained: A Guide for Social Science Students*, *2nd ed,*. Routledge. ISBN-13: 978-0415332859
- Kremelberg, D. (2011) *Practical Statistics: A Quick and Easy Guide to IBM SPSS, Stata, and Other Statistical Software*, Sage Publications, Inc. ISBN-13: 978-1412974943

Additionally, if you want to learn more about using R for data science, I highly recommend this book:

- Wickham, H. and Grolemund, G. (2017) *R for Data Science*, O'Reilly Media. ISBN-13: 978-1491910399. http://r4ds.had.co.nz/index.html
  - This is a free book available in HTML and a physical copy is available for purchase

COURSE ACTIVITIES

Homework

There will be a total of 8 homework assignments. These assignments are meant to assess your mastery of the topics and techniques covered in class. You will be given approximately 5 textbook problems that cover material from the last few lectures. Completed assignments will be submitted via Elms/Canvas due by 10:00am on the scheduled due date.

Labs

Most Fridays, we will have a lab section for which you will have to bring a laptop to class. These lab assignments are designed to teach you more about how R works within the context of the topics being taught in class. You will work on the labs in class and have until Friday the next week to turn them in, completing any unfinished work at home. Completed labs will be submitted via Elms/Canvas due by 10:00am on the scheduled due date.

Projects

There will be a total of 3 projects. These are your opportunity to apply concepts learned in class to real problems and data sets from ideas and hypotheses you develop and test. These assignments will be narrative reports with statistical analysis. You may work with your colleagues to figure out the underlying concepts and problem-solving processes but are expected to work *individually* to write your own project paper. Completed assignments will be submitted via Elms/Canvas. The first two projects will be due one week from the assigned date at 10:00 am on the scheduled due date. Project 3 will be a "Final Project" and graded with greater scrutiny than previous projects.

Exams

There will be a midterm and a cumulative final exam. These exams provide an opportunity for you to test your understanding of the concepts, techniques, and problems associated with statistical reasoning. In order to learn and understand the material fully it is important to review and revisit it multiple times. The first midterm is taken in-class, the second midterm is an online, take-home exam. The online midterm exam will be timed, but you will be able to choose when you start the exam from anytime within the 2-day scheduled timeframe.

Grading

16% Homework (8)	each 2.0%
24% Labs (12)	each 2.0%
30% Projects (3)	
Short Projects (2)	each 5.0%
Final Project	20.0%
30% Exams	
Midterm	10.0%
Cumulative Final	20.0%

Grades will be assigned based on the total percent earned, using the following rubric. Grades will be rounded to the nearest 10th of a percent. Please come and talk to me early if you think that there might be a problem.

A   90.0-100% (A- 90.0-92.9%)
B   80.0-89.9% (B+ 87.0-89.9%); B- 80.0-82.9%)
C   70.0-79.9% (C+ 77.0-79.9%; C- 70.0-72.9%)
D   60.0-69.9% (D+ 67.0-69.9%; D- 60.0-62.9%)
F   0-59.9%

COURSE POLICIES

Excused Absences

If an assignment due date or scheduled exam is a religious holiday for you, please let me know at least one week in advance, so an alternate due date can be set. Missed exams with an excused absence must be made up within 2 weeks of the original deadline. Missed exams without a documented, excused absence cannot be made up and will receive a score of 0.

Late Work

Timely submission of the completed assignments is essential. The due date of each assignment will be stated clearly in the assignment description. Late homework and labs will have a maximum score of 1 point, while late projects will be subject to the following penalty:

Late Penalty Schedule

Up to 1 day	$1 \leq 2$ days	$2 \leq 3$ days	$3 \leq 4$ days	$4 \leq 5$ days	$5 \leq 6$ days	> 6 days
10%	15%	20%	25%	30%	40%	50%

Re-grading

Fairness in giving grades is very important to me, at the same time both our time is best spent on helping you learn the material. Re-grading of assignments and exams must be turned in within one week of receiving the graded work. They must be submitted as a written document in which you include the graded work, an explanation of what you believe was miss-graded, and an explanation for why you think it should be given a different score. For any re-grade requests, the entire assignment will be regarded and your score may go up or down.

Other Policies

Other policies relevant to undergraduate courses are found here:
http://ugst.umd.edu/courserelatedpolicies.html. Topics that are addressed in these various policies include academic integrity, student and instructor conduct, accessibility and accommodations, attendance and excused absences, grades and appeals, copyright and intellectual property.

OFFICE HOURS

Please visit me during office hours. This is an opportunity to ask questions about the material covered in the reading materials or in lecture. If you are having trouble in the course please talk

to me as soon as possible. If you do poorly or lower than you expected on the first exam, it is imperative that you come to office hours so that we can figure out the problem early.

ACADEMIC DISHONESTY

Cheating in any form (copying, falsifying signatures, plagiarism, etc.) will not be tolerated. It will result in a referral to the Office of Student Conduct irrespective of scope and circumstances, as required by university rules and regulations. There are severe consequences of academic misconduct, some of which are permanent and reflected on the student's transcript. If you have any questions regarding the University's policies on scholastic dishonesty, please see http://osc.umd.edu/OSC/Default.aspx.

It is very important that you complete your own assignments, and do not share files (excluding raw data), partial work or final work.

University of Maryland Code of Academic Integrity

The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This Code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course. It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity or the Student Honor Council, please visit http://shc.umd.edu/SHC/Default.aspx.

ACCOMMODATIONS

Please come and see me as soon as possible if you think you might need any special accommodations for disabilities. In addition, please contact the Disability Support Services (301-314-7682 or http://www.counseling.umd.edu/DSS/). Disability Support Services will work with us to help create appropriate academic accommodations for any qualified students with disabilities. If you experience psychological distress during the course of the semester you can get professional help at the Counseling Center (301-314-7651 or http://www.counseling.umd.edu/.

TENTATIVE COURSE SCHEDULE

Week	Day	Topics	Readings	Assignments (Homework, Project, Labs)
1	Mon 8/27	Syllabus, Data Analytics Workflow	Diez et al. 1.1-1.3 Open Statistics Education (OSE) pp. 11-14	Who You Are (Due by midnight)
	Wed 8/29	Variables, Study Design, and Sampling	Diez et al. 1.4-1.5 OSE pp. 20-28	Homework 1 assigned.
	Fri 8/31	(Lab 1) Introduction to R – Basic Data Management		
2	Mon 9/3	Labor Day		
	Wed 9/5	Descriptive Statistics and graphs	Diez et al. 1.6-1.7 OSE pp. 15-19	Homework 1 due. Homework 2 assigned.
	Fri 9/7	(Lab 2) Using R for descriptives and graphs		Lab 1 due.
3	Mon 9/10	Descriptive Statistics and graphs	OSE pp. 86-114	
	Wed 9/12	Normal distribution and z-scores	Diez et al. 3.1	Homework 2 due. Project 1 assigned.
	Fri 9/14	(Lab 3) Using the Normal distribution		Lab 2 due.
4	Mon 9/17	Other Distributions	Diez et al. 3.3-3.5	
	Wed 9/19	Sampling Distribution and the Central Limit Theorem	Diez et al. 4.1	Project 1 due. Homework 3 assigned.
	Fri 9/21	(Lab 4) CLT and Sampling Distribution part 1		Lab 3 due.

5	Mon 9/24	Inferential Stats	Diez et al. 4.2-4.3	
	Wed 9/26	Inferential Stats	Diez et al. 4.4-4.5 Aschwanden, C. *Science Isn't Broken*.	Homework 3 due. Homework 4 assigned.
	Fri 9/28	(Lab 5) CLT and Sampling Distribution part 2		Lab 4 due.
6	Mon 10/1	One sample t-tests	Diez et al. 5.1	
	Wed 10/3	Confidence Intervals		Homework 4 due.
	Fri 10/5	(Lab 6) Hypothesis Testing in R		Lab 5 due.
7	Mon 10/8	Two Sample t-tests	Diez et al. 5.2-5.3	Project 2 assigned.
	Wed 10/10	Power	Diez et al. 5.4	Homework 5 assigned.
	Fri 10/12	(Lab 7) Hypothesis Testing and Power in R		Lab 6 due.
8	Mon 10/15	One-sample z-test	Diez et al. 6.1	Project 2 due.
	Wed 10/17	Two-sample z-test	Diez et al. 6.2	Homework 5 due.
	Fri 10/19	(Lab 8) Hypothesis Testing with Proportions		Lab 7 due.
9	Mon 10/22	Midterm Review	See Elms/Canvas	
	Wed 10/24	Midterm		
	Fri 10/26	Midterm answers	See Elms/Canvas	Lab 8 due.

10	Mon 10/29	Understanding Chi-Square	Diez et al. 6.3	
	Wed 10/31	Chi Square Test of Independence	Diez et al. 6.4	Homework 6 assigned.
	Fri 11/2	(Lab 9) Chi-Square		
11	Mon 11/5	One-way ANOVA	Diez et al. 5.5 OSE pp. 515-531	
	Wed 11/7	One-way ANOVA		Homework 6 due.
	Fri 11/9	(Lab 10) One-way ANOVA		
12	Mon 11/12	Factorial ANOVA	OSE pp. 532-552	Final Project assigned.
	Wed 11/14	Factorial ANOVA		Homework 7 assigned.
	Fri 11/16	(Lab 11) Factorial ANOVA		
13	Mon 11/19	Final Project Discussion		Final Project proposals due.
	Wed 11/21	THANKSGIVING		Homework 7 due.
	Fri 11/23	THANKSGIVING		
14	Mon 11/26	Scatterplots and Correlations	Diez et al. 7.1	
	Wed 11/28	Simple Linear Regression	Diez et al. 7.2-7.4	Homework 8 assigned.
	Fri 11/30	Multiple Regression	Diez et al. 8.1	
15	Mon 12/3	(Lab 12) Simple and Multiple Linear Regression		

	Wed 12/5	Multiple Regression	Diez et al. 8.2-8.3	Homework 8 due.
	Fri 12/7	Final Project Peer Critiques		Final Project Initial Draft due.
16	Mon 12/10	Review	See Elms/Canvas	Lab 12 due.
Final Project due December 14				
Final Exam due December 18				

This schedule is for planning purposes and may change. See Elms/Canvas for current information and deadlines.

UMD Calendar Fall Semester 2018 Dates

First Day of Classes	August 27 (Monday)
Labor Day	September 3 (Monday)
Thanksgiving Recess	November 21-25 (Wednesday-Sunday)
Last Day of Classes	December 10 (Monday)
Reading Day	December 11 (Tuesday)
Final Exams	December 12-18 (Wednesday-Tuesday)
Commencement - Main Ceremony	December 18 (Tuesday)
Commencement - College/Department Ceremonies	December 19 (Wednesday)