BROWN SCHOOL
WASHINGTON UNIVERSITY IN ST. LOUIS
MASTER OF PUBLIC HEALTH PROGRAM
SPRING 2019
Advanced data analysis
S55 MPH 5245

CREDIT HOURS: 3
GRADE: L/G

INSTRUCTORS: Kimberly J. Johnson, MPH, PhD and Kyle Pitzer
OFFICE: Goldfarb 237 (Instructor)

ROOM: Goldfarb 330

OFFICE HOURS: Kim Johnson: Tuesdays 12-1 pm G237 or by appointment

TIMES: Wed 9-12 and 2-5

PHONE (Instructor): 314-935-9154 (office), 314-809-9452 (cell)

E-MAILS: kijohnson@wustl.edu, kyleapitzer@wustl.edu

I. COURSE DOMAIN AND BOUNDARIES

Students will be provided with an introduction to theory and analytical techniques commonly employed in the practice of public health through mini-lectures and a hands-on experiential learning approach. The flow of the course will follow that of the format frequently used in the production of public health reports starting with preparing your data for analyses and ending with communicating your results. Students will work as individuals or teams to develop and conduct a project, choosing from different real-world public health datasets that are amenable to some of the most common types of analyses that students will encounter in their public health careers. Mini-lectures will address theory as well as practical tools and in class exercises will be focused on application of the theories and tools. Homework will be assigned that provides students with further opportunity to gain competency in strategies that can be employed to analyze each type of dataset. The course will also provide students with the opportunity to gain skills in data visualization strategies, conducting reproducible research, presenting projects, and effectively responding to constructive critiques.

Prerequisites:
- Foundations of Public Health: Epidemiology
- Foundations of Public Health: Biostatistics
- Applied Linear Modeling
- Basic competency in R or motivation to learn R in those with competency in SAS, SPSS, or STATA

II. MPH FOUNDATIONAL KNOWLEDGE AND COMPETENCIES ADDRESSED IN THIS COURSE:

A. Foundational Knowledge

1. Explain the role of quantitative and qualitative methods and sciences in describing and assessing a population's health.

B. Foundational Competencies (and or Specialization Competencies as applicable)

1. Describe preferred methodological alternatives to commonly used statistical methods when assumptions are not met.
2. Apply and interpret common statistical methods for inference (e.g., ANOVA, linear and logistic regression, survival analysis) found in public health studies.
3. Explain the principles of study design for identifying risk factors for outcomes (e.g, sampling strategies, data collection, study design trade-offs).

4. Describe principles and the application of key concepts from probability and inference (e.g., random variation, measurement error, confounding bias, effect modification) to colleagues without extensive statistical training.
5. Demonstrate an understanding of systematic biases (selection and information biases) that affect observational, quasi-experimental, and experimental studies.
6. Demonstrate an understanding of the components of reproducibile research.

C. Other Competencies

1. Communicate audience-appropriate public health content, both in writing and through oral presentation.

III. BROWN SCHOOL ACADEMIC POLICIES

Academic Integrity: If a faculty member or student suspects that academic or professional integrity has been violated, they are required to submit an Academic Integrity or Professional Integrity Violation form found on Inside Brown for review by the Assistant Dean of the program. The Assistant Dean or designated representative will aid in the investigation of the violation, which includes but is not limited to gathering relevant evidence; conversations with the instructor, student(s) involved, witnesses, and others as necessary. Depending on the seriousness of the case, the Assistant Dean may choose to refer the matter directly to the University Student Conduct Board. This referral procedure will generally be followed if it is believed that the penalty is likely to involve suspension or expulsion from the University. The Assistant Dean for the program or designated representative will offer to meet privately with the student(s) against whom the complaint has been made. It is the student's responsibility to familiarize themselves with the behaviors that constitute an academic integrity violation requiring referral.

Student Handbook 2018

Accommodations: If you have a learning, sensory, or physical disability or any other diagnosis that requires accommodations and/or assistance in lectures, reading, written assignments, and/or exam taking, please work with the Disability Resource Center, a University-wide resource that provides academic accommodations support and referrals. Students are required to provide a Verification of Individual Student Accommodation (VISA) letter to the instructor and are encouraged to work directly with the instructor to discuss specific course needs. The Director of Student Affairs can assist with coordination between the Disability Resource Center and the Brown School.

Pronouns: The Brown School embraces and promotes gender expansiveness as reflective of the lived experiences of many students, staff, faculty and members of our expanded community. The correct use of an individual's pronouns is a critical part of an individual's identity and of building an inclusive community. Students, faculty and staff are encouraged to use pronouns during introductions, are expected to use expressed pronouns of all Brown School community members, and are encouraged to apologize when mistakes are made. Educational resources are available at: https://campuslife.wustl.edu/lgbtqia/lgbt-resources/gender-pronouns/

English Language Proficiency: If your English language proficiency is such that you may need special assistance in lectures, reading, written assignments, and/or exam taking, please communicate these needs to your instructor who may refer you to the English Language Program (ELP). ELP is a University-wide resource that provides classes and academic English language support designed to increase non-native English speaking students' English language proficiency and to facilitate their academic success at Washington University. You may also find the Academic Assistance resources available through the Office for International Students and Scholars to be helpful.

Professional Use of Electronic Devices in the Classroom: Computers or other electronic devices, including "smart pens" (devices with an embedded computer and digital audio recorder that records the classroom

lecture/discussion and links that recording to the notes taken by the student), may be used by students at the discretion of the faculty member to support the learning activities in the classroom. These activities include taking notes and accessing course readings under discussion. If a student wishes to use a smart-pen or other electronic device to audio record lectures or class discussions, they must notify the instructor in advance of doing so. Permission to use recording devices is at the discretion of the instructor, unless this use is an accommodation approved by Disability Resources.

Nonacademic use of laptops and other devices and use of laptops or other devices for other coursework is distracting and seriously disrupts the learning process for other people in the classroom. Neither computers nor other electronic devices are to be used in the classroom during class for nonacademic reasons or for work on other coursework. Nonacademic use includes emailing, texting, social networking, playing games, instant messaging, and use of the Internet. Work on other coursework may include, but is not limited to, use of the Internet, writing papers, using statistical software, analyzing data, and working on quizzes or exams. The nonacademic use of cell phones during class time is prohibited, and they should be set on silent before class begins. In the case of an emergency, please step out of the room to take the call. The instructor has the right to hold students accountable for meeting these expectations, and failure to do so may result in a loss of participation or attendance points, a loss of the privilege of device use in the classroom, or being asked to leave the classroom.

Religious Holidays: The Brown School recognizes the individual student's choice in observing religious holidays that occur during periods when classes are scheduled. Students are encouraged to arrange with their instructors to make up work missed as a result of religious observance, and instructors are asked to make every reasonable effort to accommodate such requests.

IV. WASHINGTON UNIVERSITY ACADEMIC SUPPORT POLICIES

Accommodations based upon sexual assault: The University is committed to offering reasonable academic accommodations to students who are victims of sexual assault. Students are eligible for accommodation regardless of whether they seek criminal or disciplinary action. Depending on the specific nature of the allegation, such measures may include but are not limited to: implementation of a no-contact order, course/classroom assignment changes, and other academic support services and accommodations. If you need to request such accommodations, please direct your request to Kim Webb, Director of the Relationship and Sexual Violence Prevention Center, or Jen Durham Austin, Support Services Counselor. Both Kim Webb and Jen Durham Austin are confidential resources; however, requests for accommodations will be shared with the appropriate University administration and faculty. The University will maintain as confidential any accommodations or protective measures provided to an individual student so long as it does not impair the ability to provide such measures.

If a student comes to me to discuss or disclose an instance of sexual assault, sex discrimination, sexual harassment, dating violence, domestic violence or stalking, or if I otherwise observe or become aware of such an allegation, I will keep the information as private as I can, but as a faculty member of Washington University, I am required to immediately report it to my Department Chair or Dean or directly to Ms. Jessica Kennedy, the University's Title IX Director. If you would like to speak with Ms. Kennedy directly, she can be reached at (314) 935-3118, jwkennedy@wustl.edu, or by visiting the Title IX office in Umrath Hall. Additionally, you can report incidents or complaints to the Office of Student Conduct and Community Standards or by contacting WUPD at (314) 935-5555 or your local law enforcement agency. See: Title IX

You can also speak confidentially and learn more about available resources at the Relationship and Sexual Violence Prevention Center by calling (314) 935-3445 for an appointment or visiting the 4th floor of Seigle Hall. See: RSVP Center

Bias Reporting: The University has a process through which students, faculty, staff and community members who have experienced or witnessed incidents of bias, prejudice or discrimination against a student can report their experiences to the University's Bias Report and Support System (BRSS) team. See: brss.wustl.edu.

Mental Health: Mental Health Services' professional staff members work with students to resolve personal and interpersonal difficulties, many of which can affect the academic experience. These include conflicts with or worry about friends or family, concerns about eating or drinking patterns, and feelings of anxiety and depression.

Center for Diversity and Inclusion (CDI): The Center of Diversity and Inclusion (CDI) supports and advocates for undergraduate, graduate, and professional school students from underrepresented and/or marginalized populations, creates collaborative partnerships with campus and community partners, and promotes dialogue and social change. One of the CDI's strategic priorities is to cultivate and foster a supportive campus climate for students of all backgrounds, cultures, and identities. See: diversityinclusion.wustl.edu/

Additional Issues or Concerns: If you feel that you need additional supports in order to be successful in your time at Brown, beyond the mentioned accommodations, please contact Essie Rochman, Director of Student Affairs at erochman@wustl.edu. She can assist you in navigating a myriad of concerns. Her office is in Brown Hall, room 320.

VI. READINGS AND SOFTWARE REQUIREMENTS

Readings: The reading assignments for the course will include links to journal articles and other material provided on our class Github website: https://github.com/kijohnson/ADA_Spring_2019.

Software: Students are required to use R software for this course for analyses. Instructions for how to gain access is provided on the Github website:
https://github.com/kijohnson/ADA_Spring_2019/tree/master/Class%200

VII. ORGANIZATION OF THE COURSE

The format of the course will be lectures, examples, applied in class activities, and discussion. You will perform exercises that reinforce the epidemiologic and biostatistics principles and methods that are presented in class. Your mastery of the course objectives will be assessed through quizzes, homework assignments, and a final course project.

VIII. ROLE OF STUDENT AND INSTRUCTORS

Student Expectations and Requirements**:**

- You are expected to read any assigned materials prior to class.
- You should be prepared to discuss the assigned readings at the start of each class and to participate in class discussions.
- You must return your assignments on time.
- You are responsible for doing your own work.
- You are expected to check email regularly for any course announcements.
- If you have any issues or challenges with the class, please come to see me or the TA so that we can resolve them as quickly as possible.

Instructor Expectations and Requirements:

- We will come to class prepared, organized, and enthusiastic.
- We can be reached by phone, email, or during office hours/scheduled appointments.
- We retain the right to change the content and order of the lectures and exercises to meet the needs of students who are enrolled in the course.

IX. ASSIGNMENTS AND GRADING CRITERIA

Homework (40% of final grade). There will be 12 homework assignments (most will be relatively short). Students may drop 2 of the 12 assignments.

Final project (30% of final grade): Students will develop a data analysis project from conception to dissemination of results through a presentation and report using a real-world dataset. Students will find their own datasets. Examples of real-world datasets that could be used for this project include Surveillance, Epidemiology, and End Results (SEER: https://seer.cancer.gov/), Behavioral Risk Factor Surveillance System (BRFSS: https://www.cdc.gov/brfss/index.html), and National Health and Nutrition Examination Survey (NHANES: https://www.cdc.gov/nchs/nhanes/index.htm). Another source of datasets is the Inter-university Consortium for Political and Social Research (https://www.icpsr.umich.edu/icpsrweb/ICPSR/search/studies). Students are expected to apply the concepts used in class to their final project. The final project can be conducted by an individual or a team of no more than 3 members. For students choosing to work in teams, each team member will receive the same grade. It is therefore advised that you choose your team members with careful consideration as we will not take off points for individual team members who may cause other team members frustration! ☺ Components of the project include:

1. **Concept proposal (5 points)**
2. **Abstract (5 points)**
3. **Code for project on github site (5 points)**
4. **Final project presentation (10 points)**
5. **Responses to reviewers (5 points).** Reviewers will be composed of students in the class and/or 2-3 faculty/staff members who will provide written comments that the instructor will assemble and provide to the student/team. Responses are due within one week of the final presentation.

Midterm Exam (20% of final grade). There will be two parts to this exam: 1) take home analysis and 2) in class written exam that focuses on applying the lessons taught in the class through T/F, multiple choice, and/or short answer.

Participation and professionalism (10% of final grade). Participation will be assessed based on **attendance and professionalism.** For full credit, students are required to attend 14/15 classes to receive full attendance points. Attendance will be tracked. Permission to miss additional classes may be granted without the subsequent loss of points. Professionalism is considered in the final participation grade. Professionalism considerations include effective participation in team work and with the instructor to resolve learning issues, timely responses to email, participation in discussions, and attentiveness in class. We reserve the right to deduct up to 5 points from your total grade (out of 100 total points) for unprofessional classroom behavior. Some real-life examples of past student unprofessional classroom behavior:

- Online shopping, looking for housing, use of social media, general surfing, or texting while there is a lecture/discussion happening in the classroom. Dr. Johnson has actually observed students doing this while sitting behind them when there is a guest lecturer.
- Throwing a pencil when turning in an exam that a student felt was difficult.
- Loud sighs/expressions of frustration during class (Please ask questions to clarify confusion)
- Disruptive whispering, chatting, or laughing while the instructor is talking.
- Openly criticizing other instructors/classes/students while in the classroom.

Grading Scale:

95-100: A*	90-94: A-	88-89: B+	82-87: B
80-81: B-	77-79: C+	74-76, C	70-73: C-
69 & below, F			

***What constitutes A work:** As indicated in "Course Administration and Grading" (https://insidebrown.gwb.wustl.edu/students/MPH/Pages/AcademicResources.aspx), A's are given for outstanding work defined as "**exceeds expected level of performance for graduate study".** Not all students will receive A's in this course and extra credit assignments are not given to assist students in recovering lost points. Students who perform at the outstanding level in this class are typically those that

come to class prepared, start their homework well before the due date, and ask questions to clarify gaps in their understanding. If you are not performing at an A level of work, please know that we do not use grades to judge the overall quality of the person. There are many factors besides grades that contribute to success and we do realize that students have varying priorities and demands on their time. We view each student holistically and grades are just one part of the overall view.

X. Getting help with R and course content in general.

If you need help with understanding the course materials, there are several options including:

- Making an appointment with one or both of the instructors.
- Making an appointment with the StatLab that provides excellent tutoring with software and programming. Please visit the StatLab website for more information at https://insidebrown.gwb.wustl.edu/People/students/StudentSupportResources/Pages/StatLab.aspx
- Good websites for R issues:
    - Quick-R: https://www.statmethods.net/ .
    - UCLA Institute for Digital Research and Education (idre): https://stats.idre.ucla.edu/r/
    - Stackoverflow: https://stackoverflow.com/ (please see: https://stackoverflow.com/tour before posting a question on this site. Users of this site get very crabby when questions are not well-written (clearly telling what problem you are trying to solve and showing what you already tried), so make sure you read about what not to ask about at the tour link.
    - Google your error message or your "how do you…" question and often you will find a solution on the web. This is a completely normal part of coding, so embrace it!

XI. COURSE OUTLINE

Class Date	Readings	Topic (s)	Lecture	In class exercises	What is due?
1: 1/16	• None	Course overview (e.g., syllabus, Canvas), introduction to fellow students, Getting started	The project life cycle/R review (Kim Johnson)	Analysis of survey data	**HW1:** Survey responses and installation of R and R studio
2: 1/23	• Peng RD. Reproducible Research in Computational Science. *Science* Vol 334, 2011. Pg. 1226-1227 • Knuth DE. Literate Programming. http://www.literateprogramming.com/knuthweb.pdf (read pg 1 part A) • Analysis and Reporting Modules 3c and 3d Peng 1: https://www.youtube.com/watch?v=4rBX6r5emgQ Peng 2: https://www.youtube.com/watch?v=VOaN33aAcEw • Peng 3: https://www.youtube.com/watch?v=SppP5TXnmwc&t=7s	Reproducible research I	Reproducible research (Kim Johnson)	Paper review for reproducible methods Module 3: https://coding2share.github.io/ReproducibilityToolkit/	**HW2:** Survey analysis questions
3: 1/30	• Ram K. Git can facilitate greater reproducibility and increased transparency in science. Ram Source Code for Biology and Medicine 2013, 8:7 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3639880/ • Blischak JD, Davenport ER, Wilson G. A Quick Introduction to Version Control with Git and GitHub. PLoS Computational Biol. 2016 Jan 19;12(1):e1004668. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004668	Reproducible research II (brief introduction followed by Git Lab)	Git for version control (Kyle Pitzer)	In class peer review of HW2 Installing and using Git Bash and GitHub Des Module 5: https://coding2share.github.io/ReproducibilityToolkit/	**HW3:** Reproducible research assignment

4: 2/6	• The visual display of quantitative information / Edward R. Tufte. Edition. 2nd ed. Published Cheshire, Conn. : Graphics Press, c2001. *Read Chapter 1*	Data Visualization	Data visualization in R for scientific presentations (Kyle Pitzer and Shelly Cooper)	Data visualization exercise	**HW4:** Git exercise assignment
5: 2/13	• Quick-R Generalized Linear Models https://www.statmethods.net/advstats/glm.html	Generalized linear modeling I	Overview of generalized linear models (Kyle Pitzer)	Generalized linear models in R exercise	**HW5:** Data visualization assignment
7: 2/20	• Long, J.S. (1997). Regression Models for Categorical Dependent Variables. Thousand Oaks, CA: Sage Publications, Chapter 8	Generalized linear modeling II	Poisson and negative binomial regression (Kim Johnson)	Poisson and negative binomial regression exercise	**HW6:** Generalized linear models assignment
6: 2/27	• https://stats.idre.ucla.edu/r/dae/multinomial-logistic-regression/ • https://stats.idre.ucla.edu/r/dae/ordinal-logistic-regression/ • Denham BE. Determinants of Anabolic-Androgenic Steroid Risk Perceptions in Youth Populations: A Multivariate Analysis. J Health Soc Behav. 2009 Sep;50(3):277-92. (focus on methods and results)	Generalized linear modeling III	Multinomial and ordinal logistic regression (Xiaoyan Wang)	Multinomial and ordinal logistic regression excercise	**HW7:** Poisson and negative binomial regression assignment
8: 3/6	• A Practical Guide to Understanding Kaplan-Meier Curves https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3932959/pdf/nihms549224.pdf	Survival analysis I	Introduction to Survival Analysis (Kim Johnson)	Kaplan Meier Curves in R exercise	**HW8:** Multinomial and ordinal regression assignment Concept proposal for project
9: 3/20	**Midterm Exam (proctored by Kyle Pitzer)**				
10: 3/27	• https://www.datacamp.com/community/tutorials/survival-analysis-R (start with Cox Proportional Hazards Models)	Survival analysis II	Cox Proportional Hazards Regression	Cox PH in R exercise	**HW9:** KM curves assignment

			(Kim Johnson)		
11: 4/3	• Suttorp MM et al. Graphical presentation of confounding in directed acyclic graphs. Nephrol Dial Transplant. 2015 Sep;30(9):1418-23. doi: 10.1093/ndt/gfu325. Epub 2014 Oct 16. (**REQUIRED**) • Westreich D and Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. Am J. Epidemiol 2013 Feb 15;177(4):292-8. PMID: 23371353 (OPTIONAL) • VanderWeele TJ and Robinson W. On causal interpretation of race in regressions adjusting for confounding and mediating variables. Epidemiology. 2014 July; 25(4): 473-484 (OPTIONAL) • Shrier I and Platt RW. Reducing bias through directed acyclic graphs. BMC Medical Research Methodology. 2008, 8:70 (OPTIONAL)	Covariate selection (DAGs)	DAGs (Kim Johnson)	DAG exercise	**HW10:** Cox PH assignment
12: 4/10	• Cummings P. Missing Data and Multiple Imputation JAMA Pediatrics. July 2013 Volume 167, Number 7 656-661.	Other critical issues	Missing data (Kim Johnson)	Sensitivity analyses for selection bias and missing data	**HW11:** DAG assignment
13: 4/17	• https://www.nature.com/articles/d41586-017-07522-z • http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1154108?needAccess=true • https://jamanetwork.com/journals/jama/article-abstract/2676503	Good science	P-values and reporting (Kim Johnson)	Working class period	**HW12:** Missing data assignment
14: 4/24	Final project presentations				
15: 5/1	Final project presentations				
5/8					Final project revisions