# Brown School

# Washington University in St. Louis

# Fall 2018

## Foundations of Public Health: Biostatistics

## S55 MPH 5003

##### **credit hours:** 3

###### **grade:** L/G

###### **room:** Hillman 120

###### **day/time:** Tuesdays 1pm - 4pm

###### **section:** 03

###### **github:** <https://github.com/jenineharris/Foundations-Biostatistics-Fall2018>

##### **instructor**: Jenine Harris

###### **email**: [harrisj@wustl.edu](mailto:harrisj@wustl.edu)

###### **phone**: 935-3522

###### **office location**: Goldfarb 357

###### **office hours**: Thursdays 11am - 1pm

##### **teaching assistant:** Kyle Pitzer

###### **teaching assistant office hours:** Fridays 11am-noon

###### **teaching assistant office hours location:** Brown Hall LL03

###### **teaching assistant email:** [kyleapitzer@wustl.edu](mailto:kyleapitzer@wustl.edu)

### I. COURSE DOMAIN AND BOUNDARIES

The purpose of this course is to introduce the basic principles and methods of biostatistics, providing a sound methodological foundation for public health and/or social work research and practice. This course will cover descriptive and inferential statistics with applications in health care, medicine, public health, social work, and epidemiology.

### II. MPH COMPETENCIES

#### Foundational Knowledge:

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

#### Foundational Competencies:

* Select quantitative and qualitative data collection methods appropriate for a given public health context.
* Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate.
* Interpret results of data analysis for public health research, policy, or practice.

### 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](https://insidebrown.gwb.wustl.edu/People/students/studenthandbook/Pages/default.aspx)

#### Accommodations

If you have a learning disability, sensory, or physical disability or other impairment, and you may need special assistance in lectures, reading, written assignments, and/or exam taking, please contact the Brown School Director of Student Affairs who can provide coordination of accommodations at Washington University and the Brown School. The [Disability Resource Center](http://cornerstone.wustl.edu/disability-resources/), a University-wide resource, provides diagnostic and academic accommodations support and referrals.

#### 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)](http://oiss.wustl.edu/english-language-programs/). 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](http://oisshome.wustl.edu/students/) 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 ([kim\_webb@wustl.edu](mailto:kim_webb@wustl.edu)), Director of the Relationship and Sexual Violence Prevention Center. Ms. Webb is a confidential resource; 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 Coordinator. If you would like to speak with the Title IX Coordinator directly, Ms. Kennedy can be reached at (314) 935-3118, [jwkennedy@wustl.edu](mailto:jwkennedy@wustl.edu), or by visiting her office in Umrath Hall. Additionally, you can report incidents or complaints to Tamara King, Associate Dean for Students and Director of Student Conduct, or by contacting WUPD at (314) 935-5555 or your local law enforcement agency. You can also speak confidentially and learn more about available resources at the Relationship and Sexual Violence Prevention Center by calling (314) 935-8761 or visiting the 4th floor of Seigle Hall.

#### 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](file:///C:\Users\jenine\Box\teaching\Teaching\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. See: [shs.wustl.edu/MentalHealth](file:///C:\Users\jenine\Box\teaching\Teaching\shs.wustl.edu\MentalHealth).

#### 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](mailto:erochman@wustl.edu). She can assist you in navigating a myriad of concerns. Her office in Brown Hall, room 320.

### VI. READINGS

There are 3 books used in this course. The first two (Fischetti, Dalgaard) are available for free electronically through the Wash U library system. The Fischetti book has 8 WUSTL licenses so 8 people can have the electronic copy of this book open at any given time. The Dalgaard has unlimited user access. While unlikely, there is a small risk that more than 8 students might try to access the Fischetti book at the same time, so if you want to ensure you always have access to this text, you may wish to buy a hard copy. Multiple copies of the book club book (Salsburg) will be available on reserve in the Brown School library.

1. Fischetti, Tony. 2018. Data Analysis with R - Second Edition. Packt Publishing.
2. Dalgaard, Peter. 2008. Introductory Statistics with R. Springer.
3. Salsburg, David. 2002. The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century. Holt Paperbacks.

### VII. ORGANIZATION OF COURSE

Most in-class meetings will consist of a peer-review activity, a hands-on activity, a tutorial/demo in R, and time to work on the weekly challenge. On other weeks we will have a course book club, a course review, and a final exam. The course work is designed to introduce you to statistical approaches common in the social sciences. Most weeks you will have time to start the weekly challenge in class. In addition to the hours you spend in class each week, expect to spend 3-5 hours each week reading, reviewing course materials, and working on challenges.

If you are working on coursework outside of class time and your books and other materials do not contain enough information for you to complete the work, there are multiple options for you to get help:

1. Check the UCLA statistics website, youtube, and other online resources
2. Email the instructor or TA
3. Go to the instructor or TA office hours
4. Make an appointment with the StatLab
5. Make a one-on-one appointment outside office hours with the TA or instructor

The instructor and TA will attempt to provide you with a response within 24 hours during the course, although it may be longer if you request help during the weekend.

### VIII. ROLE OF FACULTY AND STUDENT

The instructor will facilitate the student’s learning experience through demonstrations, activities, exercises, exams, and outside consultation with students. The instructor will provide timely feedback on student performance. Students are expected to attend class on time and be prepared; complete all required readings and assignments in a timely manner; and participate actively in class. If any student has problems with attendance, meeting deadlines, or completing work on/by a given date, it is important that these difficulties be discussed promptly with the instructor.

### IX. COURSE OUTLINE AND GRADING

During the course you will earn points by completing peer reviews, individual challenges, and a final exam. Five percent of your grade is also based on professionalism which includes overall professionalism and participation in the group work for book club. Your book club teammates and TA will be consulted in the calculation of the professionalism grade. At the end of the course, the percent you have earned will translate into a letter grade. Note that your final score is not rounded up, so you have to reach at least the percentage shown to earn the grade.

#### A typical day:

1. Peer review (20-30 minutes)
2. Hands-on warm-up activity (20-40 minutes)
3. Tutorial (60-90 minutes)
4. Individual work time for peer review and challenges (20-80 minutes)

#### Weighting:

* 9% peer reviews (1% per review *session* which may include reviewing 1 or 2 classmates)
* 10% book club
* 10% statistics in the wild
* 6% DataCamp
* 5% professionalism
* 30% challenges (3% per challenge)
* 30%\* final exam

\* **Late work is not accepted for any reason.** If you are unable to submit something on time, the percentage available for the item will be *added to the weight of your final exam*. For example, if you do not submit a challenge on time and it is worth 3% of your grade, the final exam will be re-weighted to be 33% of your total grade. This is true for the first 3 things you do not submit on time. After three items are late or missing, additional late or missing items will receive no credit. *Please consider the possibility that you might need to miss something late in the semester due to personal or family reasons.*

##### The highest threshold you reach will be your earned grade:

###### 95% A

###### 90% A-

###### 88% B+

###### 85% B

###### 80% B-

###### 78% C+

###### 75% C

###### 70% C-

###### 0% F

#### Graded components of the course

**Peer review (9%)**

At the beginning of the class period on ten days of class each student will be assigned one or two peer-reviews to complete. During a peer review, you will have access to the assignment(s) submitted by one or two of your classmates. You will run their code and provide detailed feedback on its clarity and on whether it worked to answer the questions assigned.

Reviews contribute to meeting all three *Foundational competencies* from section II.

**Challenges (30%)**

There will be ten challenges throughout the semester. Most challenges include a *coder* version and a *hacker* version. The two versions are worth the same amount of points. The *coder* version focuses on practicing the skills learned in the class for the week. The *hacker* version typically will include practice of the skills from class but also go beyond the material from class to present a new related coding problem for you to solve. Challenges will be reviewed by one or two classmates and the instructor.

Challenges will be graded as complete (100), partial (70), or incomplete (0). A score of *complete* is earned when your challenge includes code and text to answer all the questions in the challenge, uses good coding practices, and includes few (if any) errors. A score of *partial* is earned when your challenge answers all or most questions but does not use good coding practices or includes a moderate amount of error. A score of *incomplete* is earned when your challenge does not answer most questions or answers most/all questions but demonstrates poor coding practices and is mostly or entirely incorrect.

Each challenge will contribute to meeting all of the competencies for the course listed in section II.

**DataCamp (6%)**

Completing chapters from DataCamp courses will be among your reading/preparatory work for class on several different days throughout the semester. Complete the DataCamp chapter(s) due *before* class on the week assigned with a score of 50% or higher in the course earned to collect 2% for the week. Do this 3 times to earn the 6% for your course grade. DataCamp scores will be either complete (1) or incomplete (0). There are 6 opportunities to complete DataCamp chapters. Completing more than 3 will not result in additional points, but you may gain some extra knowledge that will help you more efficiently complete challenges and the final exam!

Note that, once I enter your WUSTL email into the DataCamp interface, you will have free access to ALL DataCamp courses for 6 months.

DataCamp chapters contribute to meeting all *Foundational competencies*.

**Statistics in the Wild (10%)**

Choose ONE of the following projects to complete outside of class ON YOUR OWN. The projects are designed to take 12-16 hours total. This may vary based on how quickly (or slowly) you read, write, and code.

1. Read *Rigor Mortis* by Harris and write a memo of *up to 500 words*. The memo should have two parts: (1) a summary of the book, and (2) a discussion of the relevance (or lack of relevance) of the topics to you and/or your profession.
2. Attend an [R-Users](https://www.meetup.com/Saint-Louis-RUG/) or [R-Ladies](https://www.meetup.com/R-Ladies-St-Louis/) Meetup and create a brief tutorial to teach others about something you learned at the Meetup. Note that R-Users meets *monthly* and R-Ladies meets *less frequently*, so plan ahead! The tutorial should be an R-Markdown file or a video (5-7 minute video length). Assume that your audience is the other students in your *Foundations of Public Health: Biostatistics* course.
3. Identify a quantitative published article that uses a publicly available data set like NHANES or BRFSS, or an article you are able to obtain the data for in some other way. Reproduce a table or figure from the article as closely as you can. Submit an R-Markdown file with a link to the original article, your annotated R commands, the final table or figure you reproduced, and comments about how you reproduced the table or figure including any challenges you faced or anything you were not able to figure out during the process.
4. Complete at least four of the courses in the [Intro to Statistics with R](https://www.datacamp.com/tracks/learn-statistics-with-r) skill track on DataCamp with 50% or more of the available points in each course.

Projects will be graded as complete (100), partial (70), or incomplete (0). See schedule for due date and **plan ahead**!

Option 1 meets the *Foundational knowledge* statement in section II, while options 2, 3, and 4 meet all three *Foundational competencies* from section II.

**Book club (10%)**

For book club you will be assigned to a team by the topic you are most interested in and your team will be responsible for leading a 20-minute class discussion or activity on the topic (No 20-minute presentations! The activity or discussion should include opportunities for participation by other students in the class). Additional details will become available on GitHub when book club gets closer. Like the weekly challenges, book club team scores will be complete (100), partial (70), or incomplete (0). Contributing to your book club team is part of the professionalism score in class and your team will be asked to evaluate your contribution. It is possible that students in the same team receive different scores if peer-evaluations indicate a lack of contribution by one or more team members.

**Final exam (30%)**

The final exam is a comprehensive written test that includes 2 parts: (1) data analysis and interpretation, and (2) multiple choice and short answer questions. For the first part, students will be provided with data sets and research questions at the end of class on the week before the last class meeting and will have a full week to complete the work using whatever materials they have. The second part of the exam will be an in-class multiple choice and short answer test during the class period on the last day of class; no materials will be permitted during this second part. Both parts must be submitted by the end of the class period on the last day of class. The first part of the exam must be submitted before the second part will be distributed to each student, so having the first part complete or nearly complete before class is advised; it is a long exam. All topics in the course may be on either part of the exam. The first part of the exam is worth 75% of your final exam score; the second part is worth 25% of your final exam score. You MUST PASS the final exam with a total of 70% or higher to pass the class. In working on the first part of the exam, you may not ask for any sort of help (including electronic and in-person) from anyone other than the instructor; to do so will result in a 0 on the exam and a referral to the administration for academic dishonesty.

The final exam will contribute to meeting the *Foundational knowledge* statement and covers all three *Foundational competencies* from section II.

**Professionalism (5%)**

Coming to class on-time and prepared, participating in in-class activities and team work, providing feedback to the instructor via course evaluations, and treating your fellow students, teaching assistants, and instructors (regular and guest) in a professional manner in person and electronically, etc. are required and will be translated into the professionalism score.

##### **Schedule**

Week 1 (Aug 28): Course intro & data preparation

BEFORE CLASS:

###### \* Install R and RStudio

###### \* (optional) Try the [Introduction to R](https://www.datacamp.com/courses/free-introduction-to-r) DataCamp course

IN CLASS:

###### \* Course overview

###### \* Workshop on importing data and preparing it for analysis

###### \* Watch peer review video: <https://www.youtube.com/watch?v=Pv3cDy9gIp0>

Week 2 (Sept 4): Descriptive statistics

BEFORE CLASS:

###### \* Submit Challenge 1

###### \* Complete Chapter 1 of [Introduction to Data](https://www.datacamp.com/courses/introduction-to-data)

###### \* Read Fischetti The Shape of Data first four sections (stop after Spread)

###### \* Read Dalgaard sections 4.1-4.2

###### \* Read [Descriptive Statistics: Reporting the Answers to the 5 Basic Questions of Who, What, Why, When, Where, and a Sixth, So What?](http://www.medschool.umaryland.edu/media/SOM/Departments/Anesthesiology/Documents/Faculty-Development/Descriptive_Statistics___Reporting_the_Answers_AA.PDF)

IN CLASS:

###### \* Challenge 1 peer review

###### \* Descriptive statistics workshop

Week 3 (Sept 11): Bivariate descriptive statistics

BEFORE CLASS:

###### \* Submit Challenge 2

###### \* Read Fischetti Describing Relationships

###### \* Read Dalgaard sections 4.3-4.4

###### \* Complete Chapters 1 and 2 of [Exploratory Data Analysis](https://www.datacamp.com/courses/exploratory-data-analysis)

IN CLASS:

###### \* Peer review Challenge 2

###### \* Bivariate descriptive statistics workshop

Week 4 (Sept 18): Graphs and tables for descriptives

BEFORE CLASS:

###### \* Submit Challenge 3

###### \* Read <http://r4ds.had.co.nz/data-visualisation.html> and try the exercises as you go

###### \* Read Dalgaard 4.5-4.6

###### \* Complete Chapters 3 and 4 of [Exploratory Data Analysis](https://www.datacamp.com/courses/exploratory-data-analysis)

IN CLASS:

###### \* Peer review Challenge 3

###### \* Graphs and tables workshop

Week 5 (Sept 25): Probability & Sampling

BEFORE CLASS:

###### \* Submit Challenge 4

###### \* Read Vetter TR. Fundamentals of research data and variables: the devil is in the details. Anesth Analg. 2017;125:1375-1380.

###### \* Read Fischetti Using Data to Reason About the World (Probability chapter)

###### \* Read Dalgaard Chapter 3

IN CLASS:

###### \* Peer review Challenge 4

###### \* Probability and sampling workshop

Week 6 (Oct 2): Bivariate for two categorical variables

BEFORE CLASS:

###### \* Complete Challenge 5

###### \* Read Fischetti Testing Hypotheses chapter

###### \* Read Dalgaard sections 8.1-8.2

IN CLASS:

###### \* Chi-squared workshop

Week 7 (Oct 9): Bivariate for one categorical and one continuous variable

BEFORE CLASS:

###### \* Submit Challenge 6

###### \* Review Fischetti Describing Relationships chapter

###### \* Read Dalgaard sections 5.1-5.3, 7.1-7.2

IN CLASS:

###### \* Peer review Challenge 6

###### \* t-tests and other bivariate workshop

Week 8 (Oct 16): Bivariate for two continuous variables

BEFORE CLASS:

###### \* Submit Challenge 7

###### \* Complete chapters 1 and 2 from the [Correlation and Regression](https://campus.datacamp.com/courses/correlation-and-regression/) DataCamp course

###### \* Review Fischetti Describing Relationships chapter

###### \* Read Dalgaard section 6.4

IN CLASS:

###### \* Peer review Challenge 7

###### \* Correlation and other bivariate workshop

Week 9 (Oct 23): Linear regression

BEFORE CLASS:

###### \* Submit Challenge 8

###### \* Complete chapters 3 and 4 from the [Correlation and Regression](https://campus.datacamp.com/courses/correlation-and-regression/) DataCamp course

###### \* Read Fischetti Predicting Continuous Variables

###### \* Read Dalgaard sections 6.1-6.3

IN CLASS:

###### \* Peer review Challenge 8

###### \* Linear regression workshop

Week 10 (Oct 30): Logistic regression

BEFORE CLASS:

###### \* Submit Challenge 9

###### \* Complete chapter 4 from the [Multiple and Logistic Regression](https://campus.datacamp.com/courses/multiple-and-logistic-regression/) DataCamp course

###### \* Read <https://stats.idre.ucla.edu/r/dae/logit-regression/>

###### \* Read Fischetti Predicting Categorical Variables section on Logistic regression only

###### \* Read Dalgaard sections 13.1-13.5

###### \* Read Tolles, Juliana, and William J. Meurer. “Logistic Regression: Relating Patient Characteristics to Outcomes.” JAMA 316.5 (2016): 533-534.

IN CLASS:

###### \* Peer review Challenge 9

###### \* Logistic regression workshop

Week 11 (Nov 6): Book Club

BEFORE CLASS:

###### \* Complete [Multiple and Logistic Regression Chapter 4]<https://campus.datacamp.com/courses/multiple-and-logistic-regression/> (This counts as Challenge 10 and not toward the DataCamp 6%, Complete with >75% correct for 100 points, completed with 50%-75% for 70 points, not completed or completed <50% correct for 0 points)

###### \* Read book for book club

###### \* Work with team to prepare activity/discussion on part of book

IN CLASS:

###### \* Lead the class in a discussion or activity related to your book club question. Discussion or activity should be active and encourage participation by students (no presentations!). As all good book clubs are, it’s a potluck! Feel free to bring a favorite snack to share with the class.

Week 12 (Nov 13): Data Management (Guest lecturer: Kyle Pitzer)

BEFORE CLASS:

###### \* Read <http://r4ds.had.co.nz/exploratory-data-analysis.html>

###### \* Read <https://coding2share.github.io/ReproducibilityToolkit/Mod3Code.html>

###### \* Read Fischetti Dealing with Missing Data and Dealing with Messy Data in the checking unsanitized data section

###### \* Read Dalgaard 10.1-10.4

IN CLASS:

###### \* Data management workshop

Week 13 (Nov 20): Statistics in the Wild work day

IN CLASS:

###### \* Open time to work on remaining coursework with help from instructor as needed

Week 14 (Nov 27): Course Review & Pick Up Final Exam Part I

BEFORE CLASS:

###### \* Submit Statistics in the Wild project

###### \* Review notes and make a list of topics/questions to ask in review

IN CLASS:

###### \* Review course material

###### \* Distribute Final Exam Part I

Week 15 (Dec 4): Final Exam Part II

IN CLASS:

###### \* Submit Final Exam Part I

###### \* Complete and submit Final Exam Part II

*NOTE: There may be minor revisions to this syllabus, usually to correct typos or errors. If a major change is needed to the structure of the course or the course requirements, students will be notified in advance. This is a very rare occurrence typically reserved for extreme situations.*