# Course number: BS 730 Introduction to R: Software for Statistical Computing

Fall 2020

4 Credits

Friday, 10am – 12:50pm

Room R107, 780 Harrison Avenue – Housman Med Research Center

## Instructors information

Sarah Leavitt, PhD

[sv1205@bu.edu](mailto:sv1205@bu.edu)

801 Massachusetts Ave. Crosstown Building, #348

Online Office hours: Monday, 12-1pm

Chathurangi Pathiravasan, PhD

[chathu@bu.edu](mailto:chathu@bu.edu)

801 Massachusetts Ave. Crosstown Building, #335

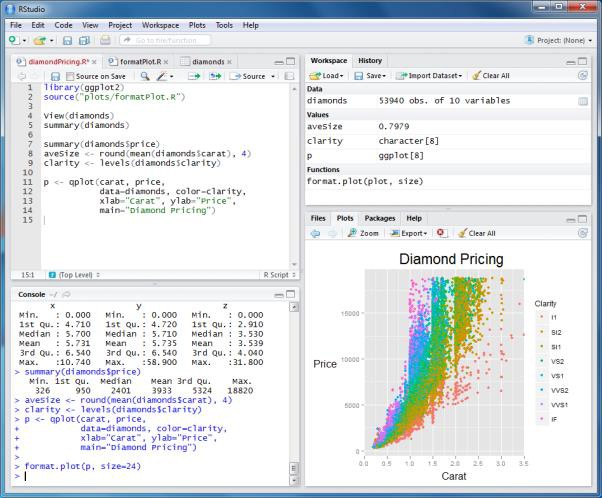
Online Office hours: Thursday, 4-5pm

## Course description

How to draw formal conclusions from data? In this course, we will address this central question to the justification of claims across scientific fields. The overall goal is for you to learn the conceptual tools and practical skills of statistical analysis. Together, we will use a quantitative focus and

* define statistical inference as inferring facts about a population of interest, by distinguishing the signal from noise in the data
* examine descriptive analysis, statistical hypothesis tests, and regression modeling
* understand how to interpret the results adequately
* learn how to identify appropriate statistical methods for the data at hand.

Statistical analysis goes hand in hand with computational tools. Together, we will learn to use R, a comprehensive software package. In particular, we will use [RStudio](https://www.rstudio.com/), an environment to use R. We will focus on manipulating data, analyzing them with the aforementioned methods, reporting and presenting the results.



This course is a prerequisite for BS704 or PH717. Please see the instructor if you have not fulfilled this requirement.

## Course learning objectives

Upon completion of this course, students will be able to:

1. Create and modify datasets; Import and export data
2. Install, load and use R packages
3. Conduct basic data manipulations and deal with missing values in R
4. Apply numerical, tabular, and graphical descriptive techniques to characterize and summarize public health data
5. Identify appropriate statistical methods to be applied in a given public health setting
6. Describe the principles and practical importance of key concepts in statistical inference and regression modeling
7. Apply the statistical methods for hypothesis testing and regression modeling using R
8. Evaluate R outputs and interpret them in the public health context of interest

# Learning resources

### Textbooks, software, materials:

RStudio (R) is available on public computers in the lecture room.

You can access RStudio (or R) by clicking on **Start**, **Programs**, **RStudio** or **R**.

While not necessary for successful completion of the course, it may be convenient to run RStudio (R) on your own laptop. RStudio and R can be downloaded from the following websites:

To download R,

<https://cran.r-project.org/>

To download RStudio,

<https://www.rstudio.com/products/rstudio/download/>

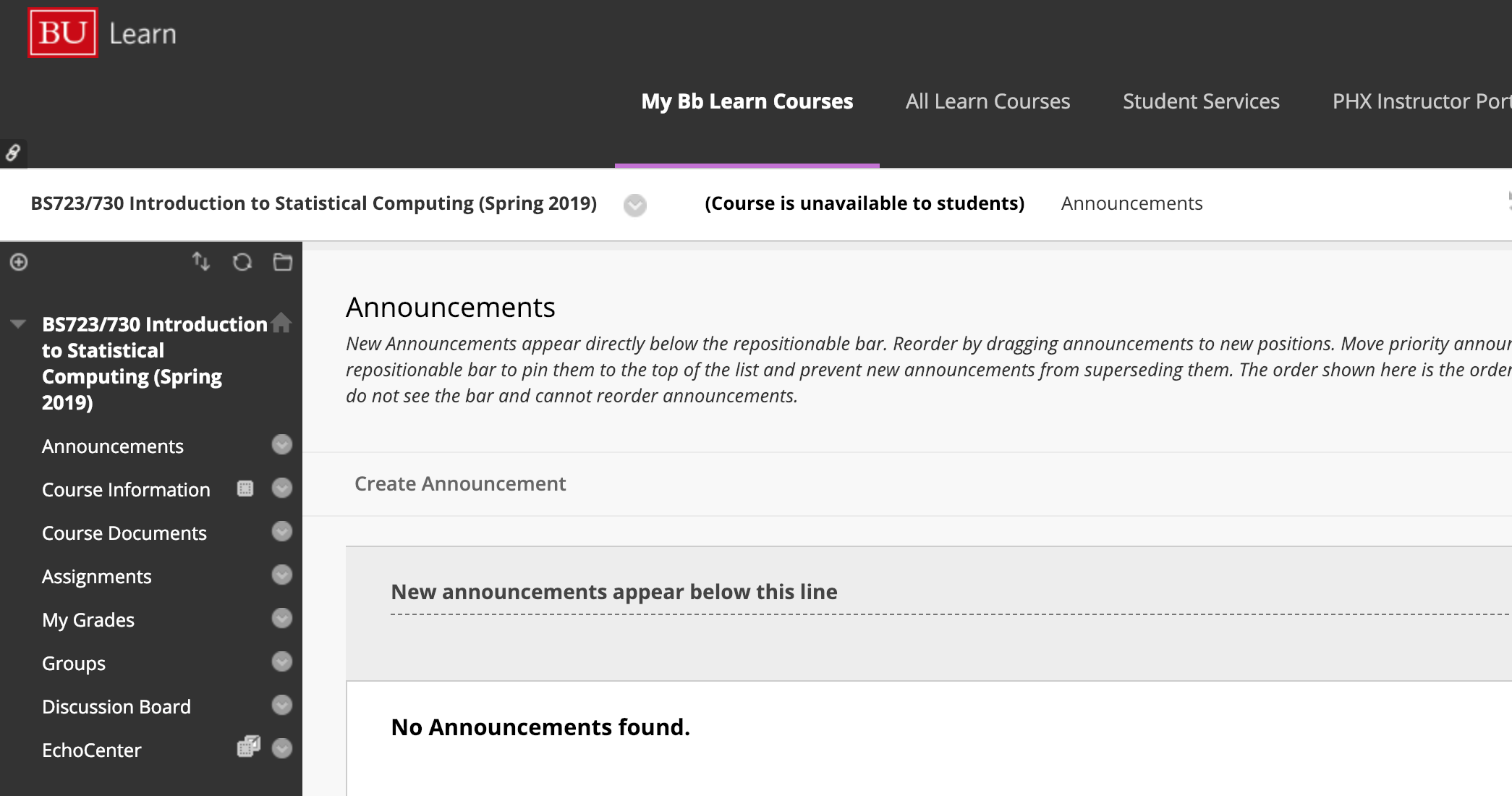
### Recommended textbooks, software, materials:

1. W. N. Venables, D.M. Smith and the R Development Core Team. An Introduction to R, Notes on R: A Programming Environment for Data Analysis and Graphics. 2016. URL: https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf
2. P. Dalgaard. Introductory Statistics with R, 2nd Edition. Springer 2008
3. R. Kabacoff. R in Action: Data Analysis and Graphics with R. Manning Publications; 2nd Edition, 2015. URL of companion website: <http://www.statmethods.net/>

### Blackboard:

Lecture slides, lecture notes, syllabi, homework assignments, and other supporting materials will be posted on Blackboard at <http://learn.bu.edu>

You may print the lecture slides or notes before class, or view them on the lab computer during the lecture.



**Announcements** - This area is used to display announcements, updates, and reminders. This area appears in the main course window each time you enter the course.

**Course Information** - This area is used to display general information about the course. Typically, this area contains course Syllabi for each section.

**Course Documents** - This content-specific area is used to hold the majority of information that will be delivered online such as course outlines, handouts, lecture materials, and related readings.

**Assignments** - This area holds course assignments, tests, quizzes, and surveys.

**Lecture Recordings –** This area will be where all classroom recordings will be made available.

# Assessments

**1. Weekly Assignments (25% of final grade)**

* Weekly homework assignments (20 points each) will be designed to allow students to apply their understanding of the concepts covered during the lecture.
* Late submissions will be penalized 3 points for the first day late (by 10am the next day after the due date) and an additional point per day for each additional day late unless prior arrangements have been made with your course instructor(s). No homework assignment will be accepted after 3 days past the due date.
* As part of the first homework assignment, all students must also complete the CITI training in “Human Subjects Protection” and “HIPAA” if they have not completed this training in the previous 3 years.
  + The purpose of CITI training is to comprehend the basic ethical and legal principles pertaining to the collection, maintenance, analysis, and dissemination of epidemiologic and public health information.

**2. Three exams (45% of final grade – 15% each)**

* The exams are to be done remotely and are open note and open book.

**4. Two Take-Home Projects (30% - 15% each)**

* Students must work independently on the take-home projects.
* Late submissions are penalized 5 points per day unless prior arrangements have been made with your course instructor(s).

**Homework Submission**

Homework assignments are due by the **beginning** of each class **before** lecture begins.

How to submit homework assignments:

* Sign into Blackboard and go to *Assignments* on the left-hand side and click on the *BS730 Assignments* folder. Then click on the appropriate homework link (for example, to submit homework 1 click on the homework 1 link).
* Submit your **write-up** (preferably embedded in the assignment document) with a copy of your **R program** (code) at the bottom of your assignment on Blackboard before class begins. **Please be sure to attach your homework assignment using the "Browse my Computer" link.**
* Do not submit your R output. All necessary information from the output should be included in the text or in tables of your formal write-up.
* Homeworks will be graded directly in Blackboard and you will be able to view and download the graded assignment.

## Grading policy

Students will be assigned the following final letter grades based on course assessments. All SPH candidates must have a minimum 3.0 GPA at SPH to graduate. Please review the full Boston University School of Public Health [grading policy](http://www.bu.edu/academics/sph/policies/grading/).

|  |  |
| --- | --- |
| Course Average | Final Grade |
| 93‐100 | A |
| 90‐92 | A‐ |
| 87‐89 | B+ |
| 83‐86 | B |
| 80‐82 | B‐ |
| 77-79 | C+ |
| 73‐76 | C |
| 70-72 | C- |
| 60‐69 | D |
| <60 | F |

# Boston University School of Public Health: Standards of Academic Honesty

Students in the School of Public Health are expected to adhere to the highest standards of academic honesty. Academic honesty is essential for students to attain the competencies the School expects of its graduates and to enable the faculty to adequately assess student performance. Any action by a student that subverts these goals seriously undermines the integrity of the educational programs of the School.

Academic misconduct is any intentional act or omission by a student which misrepresents his or her academic achievements, or any attempt to misrepresent his or her academic achievements. The following acts constitute academic misconduct. This is not an exhaustive list.

* + *Cheating on examinations:* The use or attempted use of any unauthorized books, notes or other materials in order to enhance the student’s performance in the examination, copying or attempting to copy from another student’s examination, permitting another student to copy from an examination or otherwise assisting another student during an examination, or any other violation of the examination’s stated or commonly understood ground rules.
  + *Plagiarism:* Any representation of the work of another as one’s own constitutes plagiarism. This includes copying or substantially restating the work of another person without the use of quotation marks or other indication that the words of another have been copied, the use of any written or oral work from which the student has obtained ideas or data without citing the source, or collaborating with another person in an academic endeavor without acknowledging that person’s contribution.
  + *Submitting the same work in more than one course without the consent of all the instructors*
  + *Misrepresentation or falsification of data*
  + *Allowing another student to represent your work as his or her own*
  + *Violating the rules of an examination or assignment*

A student who is found guilty of academic misconduct may be subject to disciplinary action, up to and including dismissal from the School. For more details, please review the [full policy](http://www.bu.edu/academics/sph/policies/standards-of-academic-honesty-and-disciplinary-procedures/).

Be sure to complete the [plagiarism tutorial](http://medlib.bu.edu/tutorials-secure/SPHPlagiarism/) and review [tips for avoiding academic dishonesty](https://ori.hhs.gov/avoiding-plagiarism-self-plagiarism-and-other-questionable-writing-practices-guide-ethical-writing).

# Course Schedule

|  |  |  |
| --- | --- | --- |
| **Date** | **Topic** | **Assignment** |
| **Unit 1: Introduction to R and Descriptive Analysis** | | |
| Sep 4 | Class 1: Getting started, descriptive analysis |  |
| Sep 11 | Class 2: Data importing/exporting, data exploration | HW1 due |
| Sep 18 | Class 3: Basic Data manipulation | HW2 due |
| **Unit 2: Univariate and bivariate** | | |
| Sep 25 | Class 4: One-sample test of means | HW3 due  *Pass out Project 1 (Units 1&2)* |
| Oct 2 | Class 5: Two-sample tests of means  *Quiz #1 on Unit 1* | HW4 due |
| Oct 9 | Class 6: Tabular data and analysis of categorical data | HW5 due |
| **Unit 3: Graphs** | | |
| Oct 16 | Class 7: Graphs | HW6 due |
| **Unit 4: ANOVA, correlation, simple linear and logistic regression** | | |
| Oct 23 | Class 8: ANOVA  *Quiz # 2 on Unit 2* | HW7 due |
| Oct 30 | Class 9: Correlation and Univariable linear regression | HW8 due  *Turn in Project 1*  *Pass out Project 2 (Units 3-5)* |
| Nov 6 | Class 10: Linear regression assumptions and Univariable logistic regression | HW9 due |
| **Unit 5: Multivariable regression, confounding, effect modification** | | |
| Nov 13 | Class 11: Multiple Linear and Logistic Regression, Part I | HW10 due |
| Nov 20 | Class 12: Multiple Linear and Logistic Regression, Part II | HW11 due |
| **Unit 6: More on R and Nonparametric** | | |
| Dec 4 | Class 13: Advanced data management: dplyr  *Quiz # 3 on Unit 4 and 5* | *Turn in Project 2* |
| Dec 11 | Class 14: Nonparametric Statistics | HW13 due |

# BUSPH Academic Support Resources

There are many support resources available to BUSPH students, including [communication resources](http://www.bu.edu/sph/students/resources/educational-resources/communication-resources/), a [writing guide](http://www.bu.edu/sph/students/resources/educational-resources/writing-guide/), [academic support](http://www.bu.edu/sph/students/resources/educational-resources/academic-support/) and a [core course tutoring program](http://www.bu.edu/sph/students/resources/educational-resources/core-course-tutoring-program/). For more information, contact Mahogany Price at [sphtutor@bu.edu](mailto:sphtutor@bu.edu).

## Public Health Writing Program

The Public Health Writing Program is available to SPH degree candidates who would like to discuss planning a paper, organizing a paper, writing clearly, or other aspects of the writing process. The program is not an editing service and does not guarantee that the assistant will be knowledgeable about content of the paper.

For more informationor to schedule an appointment, please visit the program’s website: [bu.edu/sph/writing](http://www.bu.edu/sph/writing). If you have any questions, please contact the Program Manager, Mahogany Price at [sphwrite@bu.edu](mailto:sphwrite@bu.edu).

## Presentation Skills Appointments

Presentation skill appointments are available to SPH degree candidates looking to practice presentations and receive feedback from a peer coach on slides, poster presentations, speech outlines, and/or oral communication skills!

For more information or to schedule an appointment, please visit the program’s website: [bu.edu/sph/present](http://www.bu.edu/sph/present). If you have any questions, please contact the Program Manager, Mahogany Price at [sphwrite@bu.edu](mailto:sphwrite@bu.edu).

## BUSPH Writing Guide BUSPH Writing Guide ([bu.edu/sph/writing-guide](http://bu.edu/sph/writing-guide)). The Guide includes components like writing strategies, communicating data, word choice, writing as a team, resources for non-native English speakers, and finding and using resources. In addition, there are explanations of several specific types of public health writing, including literature reviews, policy memos, reflections, and critiques. This Guide is designed to be a starting point for students and save you from having to answer basic questions about style and formatting.

## BUSPH Library Tutorials

Librarians from the BU Alumni Medical Library created [BUSPH library tutorials](http://www.bu.edu/sph/students/student-services/student-resources/academic-support/writing-guide/finding-using-references/library-resources/), for students in the School of Public Health. The first is a brief overview of library resources followed by guidance on advanced searches using including PubMed, Web of Science, POPLINE, Google, etc. One this site you will also find tutorials on Mendelay and Zotero, which are free citation management programs, as well as on properly citing sources and avoiding plagiarism.

If you would like to make an appointment to meet with a librarian in person to get personalized assistance with a search you can contact them directly by email ([refquest@bu.edu](mailto:)), phone (617 638-4228), or stopping by the reference desk on the 12th floor of the med school (Building L).