

**PUBHLTH 390R**  
**Introduction to Data Science Using R**  
**Fall 2019**

Contact Me 	Meet with Me 	Location and Time 	Course Materials 
Ben Rogers, PhD Pronouns: he, him, his bwrogers@umass.edu	Please drop-in on Wednesday 11am- 12pm, or email me for an appointment at a different time	Goodell Room 608 - TBL  Tuesday/Thursday 2:30pm-3:45pm	Wickham, Hadley & Golemund, Garrett (2017). R for data science: import, tidy, transform, visualize , and model data. O'Reilly Media, Inc. link: <a href="https://r4ds.had.co.nz/index.html">https://r4ds.had.co.nz/index.html</a>  Golemund, Garrett. (2014) Hands on programming with R. O'Reilly Media, Inc. link: <a href="https://rstudio-education.github.io/hopr/">https://rstudio- education.github.io/hopr/</a>  Access to a device capable of running R and RStudio.

Data science a rapidly growing discipline which combines programming skills, statistical knowledge and applied subject expertise to use data to answer questions. In this course we will learn which kinds of questions to ask, how to visualize and model data to answer those questions, and how to communicate our findings. You will gain hands-on experience through in-class coding activities as well as homework assignments, and the course will culminate in an original data analysis project. All coding will be done in R, a popular free open-source data analysis program.

**Course Objectives and Learning Outcomes:**

By the end of the semester you will be able to:

- 1) Formulate appropriate questions which can be investigated with a given data set.
- 2) Clean and transform data to prepare it for visualization and analysis;
- 3) Generate clear and informative plots to understand data;
- 4) Form and communicate results of data exploration.

**Prerequisites:** There are no prerequisites expected for this course. We will do in class coding, so you should bring a device to class that can run R and RStudio.

**Grading**

Grading is based on participation, homework assignments and a final project.

**Participation: 5%**

Participation will be determined by the activity on the Moodle discussion board, in order to encourage you to ask and answer questions.

**Homework assignments: 60%**

There will be 6 homework assignments. Late submissions will be penalized. Homework must be turned in within 1 week of due date to receive credit. Homework with the lowest grade will be dropped. Students are encouraged to collaborate with classmates to solve homework problems, but they must write their own answers and hand in homework assignments individually. Under no circumstances may students copy each other's answer including computer code.

**Project:35%**

Students will conduct an exploratory data analysis for a real dataset.

**Grading scale**

Percentage	Grade
[90, 100]	A
[85, 90)	A-
[80, 85)	B+
[75, 80)	B
[70, 75)	B-
[65, 70)	C+
[62, 65)	C
[59, 62)	C-
[56, 59)	D+
[53, 56)	D
[50, 53)	D-
[0,50)	F

**Policies on Grading**

There will be no extension without a documented, legitimate reason that is outside of your control. Social conflicts or a heavy workload are unacceptable excuses. It is required to contact the instructor no less than 14 days before the due date for any submissions, except for emergencies. No exception can be guaranteed and decisions will be made on a case-by-case basis.

Any grading dispute must be submitted in writing to the instructor within one week after the grade is available. No change will be made before the grade is posted or after this deadline. Reasons irrelevant to your academic performances will not be considered or responded. These reasons include but not limited to ``I need

to pass this class to graduate," ``I will be in academic probation if I failed this class," ``It's not uncommon that professors would consider this situation when assigning grades," etc.

**Tentative weekly schedule (subject to change):**

- **Week 1: Introduction to data science and R basics**
  - a. Introduction to data science
  - b. Introduction to R/R-studio
  - c. Using R as a calculator
  - d. R Scripts and R markdown
- **Week 2-5: data visualization using *ggplot***
- **Week 6-9: Data transformation and wrangling**
- **Week 10: Modeling**
- **Week 11: Exploratory data analysis**
- **Week 12: Other topics in R including functions, iterations**
- **Weeks 13: introduction to causal inference**

**Course materials:**

The following are two Textbooks for general reference:

1. Wickham, Hadley & Golemund, Garrett (2017). R for data science: import, tidy, transform, visualize , and model data. O'Reilly Media, Inc. link: <https://r4ds.had.co.nz/index.html>
2. Golemund, Garrett. (2014) Hands on programming with R. O'Reilly Media, Inc. link: <https://rstudio-education.github.io/hopr/>

Software: R and RStudio can be downloaded here

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

<http://www.rstudio.com/download>

**Copyright protection :**

Many of the materials created for this course are the intellectual property of the instructor. This includes, but is not limited to, the syllabus, lectures, and course notes. Except to the extent not protected by copyright law, any use, distribution or sale of such materials requires the permission of the instructor. Please be aware that it is a violation of university policy to reproduce, for distribution or sale, class lectures or class notes, unless the faculty member has explicitly waived copyright.

**Accommodation statement:**

The University of Massachusetts Amherst is committed to making reasonable, effective and appropriate accommodations to meet the needs of students with disabilities and help create a barrier-free campus. If you have a disability and require accommodations, please register with Disability Services (161 Whitmore Administration building; phone 413-545-0892) to have an accommodation letter sent to your faculty. Information on services and materials for registering are also available on their website <http://www.umass.edu/disability>.

**Academic honesty statement:**

Intellectual honesty requires that students demonstrate their own learning during examinations and other academic exercises, and that other sources of information or knowledge be appropriately credited. Scholarship depends upon the reliability of information and reference in the work of others. Student work at the University may be analyzed for originality of content. Such analysis may be done electronically or by other means. Student work may also be included in a database for the purpose of checking for possible plagiarized content in future student submissions. No form of cheating, plagiarism, fabrication, or facilitating of dishonesty will be condoned in the University community. Please see <https://www.umass.edu/honesty/> for more information.

**Valuing, recognizing, and encouraging diversity:**

Promoting and valuing diversity in the classroom enriches learning and broadens everyone's perspectives. Inclusion and tolerance can lead to respect for others and their opinions and is critical to maximizing the learning that we expect in this course. Our own closely held ideas and personal comfort zones may be challenged. The results, however, create a sense of community and promote excellence in the learning environment. Diversity includes consideration of (1) the variety of life experiences others have had, and (2) factors related to "diversity of presence," including age, economic circumstances, ethnic identification, disability, gender, geographic origin, race, religion, sexual orientation, social position. This class will follow principles of inclusion, respect, tolerance, and acceptance that support the values of diversity.

**Title IX statement:**

The University of Massachusetts Amherst is committed to fostering a safe, productive learning environment. Title IX and our school policy prohibits discrimination on the basis of sex. Sexual misconduct — including harassment, domestic and dating violence, sexual assault, and stalking — is also prohibited at our school. UMass Amherst encourages anyone experiencing sexual misconduct to talk to someone about what happened, so they can get the support they need and our school can respond appropriately. If you wish to

Speak confidentially about an incident of sexual misconduct, want more information about filing a report, or have questions about school policies and procedures, please contact our Title IX Coordinator, Débora D. Ferreira, Equal Opportunity Office (EO), 413-545-3464, [equalopportunity@admin.umass.edu](mailto:equalopportunity@admin.umass.edu).

UMass Amherst is legally obligated to investigate reports of sexual misconduct, and therefore it cannot guarantee the confidentiality of a report, but it will consider a request for confidentiality and respect it to the extent possible. If you want to talk with someone who is not a mandated reporter, you can contact the Center for Women and Community, (<https://www.umass.edu/cwc/>, 413-545-0883, or 24-hour hotline 413-545-0800), the Center for Counseling and Psychological Help (<https://www.umass.edu/counseling/>, 413-545-2337), or University Health

Services SANE program (<https://www.umass.edu/uhs/services/sane>, 413-5775000). As an instructor, I am also required by our school to report incidents of sexual misconduct and thus cannot guarantee confidentiality. I must provide our Title IX coordinator with relevant details such as the names of those involved in the incident.