

# Reproducible Research

Roger D. Peng

## Course Description

In this course you will learn the ideas of reproducible research and reporting of statistical analyses. Topics covered include literate programming tools, evidence-based data analysis, and organizing data analyses. In this course you will learn to write a document using R markdown, integrate live R code into a literate statistical program, compile R markdown documents using knitr and related tools, publish reproducible documents to the web, and organize a data analysis so that it is reproducible and accessible to others.

## Course Content

- Structuring and organizing a data analysis
- Markdown and R Markdown
- knitr / R Pubs
- Reproducible research check list
- Evidence-based data analysis
- Case studies in air pollution epidemiology and high-throughput biology

## Course Book

The book *Report Writing for Data Science in R* serves as the official course reference textbook. It comes in electronic form (PDF, ePub, and mobi) and is available from Leanpub. The book contains the lecture material for the course and can serve as a reference for you once the course is completed.

## Assessments

### Quizzes

- There are two quizzes (Weeks 1 and 2).
- You must earn a grade of at least 80% to pass a quiz.
- You may attempt each quiz up to 3 times in 8 hours.
- The score from your most successful attempt will count toward your final grade.

### Course Projects

The Course Projects focus on plotting and will be assessed via peer assessment. In these assignments, you will be asked to conduct data analysis and you will be evaluated by your classmates on the work and code that you write. Assignments evaluated via peer assessment will make use of your GitHub account or the R Pubs web site.

### Points and Scoring

- Quiz 1: 20%
- Quiz 2: 20%
- Course Project 1: 25%
- Course Project 2: 35%

Learners must pass every graded assessment to pass the course, regardless of their final grade.

## Differences of opinion

Keep in mind that currently data analysis is as much art as it is science - so we may have a difference of opinion - and that is ok! Please refrain from angry, sarcastic, or abusive comments on the message boards. Our goal is to create a supportive community that helps the learning of all students, from the most advanced to those who are just seeing this material for the first time.

## Plagiarism

Johns Hopkins University defines plagiarism as "...taking for one's own use the words, ideas, concepts or data of another without proper attribution. Plagiarism includes both direct use or paraphrasing of the words, thoughts, or concepts of another without proper attribution." We take plagiarism very seriously, as does Johns Hopkins University.

We recognize that many students may not have a clear understanding of what plagiarism is or why it is wrong. Please see the following guide for more information on plagiarism:

<http://www.jhsph.edu/academics/degree-programs/master-of-public-health/current-students/JHSPH-ReferencingHandbook.pdf>

It is critically important that you give people/sources credit when you use their words or ideas. If you do not give proper credit -- particularly when quoting directly from a source -- you violate the trust of your fellow students.

The Coursera Honor code includes an explicit statement about plagiarism:

*I will register for only one account. My answers to homework, quizzes and exams will be my own work (except for assignments that explicitly permit collaboration). I will not make solutions to homework, quizzes or exams available to anyone else. This includes both solutions written by me, as well as any official solutions provided by the course staff. I will not engage in any other activities that will dishonestly improve my results or dishonestly improve/hurt the results of others.*

## Reporting plagiarism on course projects

One of the criteria in the project rubric focuses on plagiarism. Keep in mind that some components of the projects will be very similar across terms and so answers that appear similar may be honest coincidences. However, we would appreciate if you do a basic check for obvious plagiarism and report it during your peer assessment phase.

It is currently very difficult to prove or disprove a charge of plagiarism in the MOOC peer assessment setting. We are not in a position to evaluate whether or not a submission actually constitutes plagiarism, and we will not be able to entertain appeals or to alter any grades that have been assigned through the peer evaluation system.

But if you take the time to report suspected plagiarism, this will help us to understand the extent of the problem and work with Coursera to address critical issues with the current system.

## Technical Information

Regardless of your platform (Windows or Mac) you will need a high-speed Internet connection in order to watch the videos on the Coursera web site. It is possible to download the video files and watch them on your computer rather than stream them from Coursera and this may be preferable for some of you.

Here is some platform-specific information:

### *Windows*

The Coursera web site seems to work best with either the Chrome or the Firefox web browsers. In particular, you may run into trouble if you use Internet Explorer. The Chrome and Firefox browsers can be downloaded from:

- Chrome: <http://www.google.com/chrome>
- Firefox: <http://www.mozilla.org>

### *Mac*

The Coursera site appears to work well with Safari, Chrome, or Firefox, so any of these browsers should be fine.