

Course Title: Data Scientist's Toolbox**Course Instructor:** Jeff Leek**Course Description:**

In this course you will get an introduction to the main tools and ideas in the data scientists toolbox. The course gives an overview of the data, questions, and tools that data analysts and data scientists work with. There are two components to this course. The first is a conceptual introduction to the ideas behind turning data into actionable knowledge. The second is a practical introduction to the tools that will be used in the program like version control, markdown, git, Github, R, and Rstudio.

This course focuses primarily on getting you set up with the appropriate tools and accounts you will need for the rest of the specialization and on giving you a solid grounding in the key conceptual ideas. If you feel like the material is basic, that is ok, you will see much more in depth treatment of each topic in the subsequent courses in the track.

Course Content:

- Track motivation
- Getting help
- Introduction to basic tools
- R
- Rstudio
- Git
- Github
- Types of data questions
- Steps in a data analysis
- Putting the science in data science

Weekly quizzes

- There are three weekly quizzes.
- 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.

The Course Project

In the Course Project, you will demonstrate that you've set up all of the necessary accounts for the tools we'll be using.

You are required to evaluate and grade at least four of your classmates' projects. For this course, the project can be evaluated with a series of yes/no answers to determine whether people completed the required installations.

Part of the course project includes submitting a screenshot to demonstrate you have installed the relevant software. Be sure not to take a screenshot with other applications open that may reveal personal information or anything else you don't want others to see.

Grading policy

You must score at least 80% on all assignments (Quizzes & Project) to pass the course.

Your final grade will be calculated as follows:

- Quiz 1 = 20%
- Quiz 2 = 20%
- Quiz 3 = 20%
- Course project = 40%

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 JHU referencing guide for more information on plagiarism.

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.