Data Science for Biologists, COVID-19 Version

Updated on 3/18/20 for Remote Instruction beginning 3/30/20

Course Instructor	Stephanie J. Spielman, PhD (Dr. Spielman or Professor Spielman)
Email	spielman@rowan.edu
Link to join class Slack	CLICK ME
Course Website	https://spielmanlab.github.io/courses/datascience_for_biologists/

Course Description and Objectives:

Modern-day biological research produces volumes of data that all scientists need the skills to analyze. In this course, students will learn and apply foundational concepts in data science, including visualization, wrangling, modeling, and communicating using open-source and reproducible frameworks. This course assumes no background in any computer programming or other computational skills.

Upon completing this course, students will be able to:

- Apply the R statistical computing language (notably the "tidyverse" packages) to visualize, interpret, manage, and model data
- Communicate data analyses using professional markup languages
- Utilize version control for code and data management, communication, and reproducibility
- Become familiar with real datasets from the biological research
- Develop and address scientific questions using reproducible data science techniques

Required Texts and Materials:

- All students need a laptop or similar (i.e. Microsoft Surface) with functioning wifi. There are NO other computer requirements! Mac or Windows are both ok, the amount of memory (RAM) is ok, etc. **Students should bring their machine to EVERY class.**
    - If you have an external mouse, I highly recommended bring to class.
- All course materials are *freely available online!!* There is NO need to purchase any books or other materials. Please consult the course website for all links to materials.

Evaluation:

Your grade will be derived from the following assessments:

Assessment	Deadline	Percentage of final grade
Mid-semester Project	Due 3/29/20	25%
Weekly Assignments	Due weekly	75%

1. You are strongly encouraged to discuss assignments and the mid-semester project with your classmates. **However, all students must submit their own code and written language. *This is taken extremely seriously.***
2. You are expected to work independently on your final project, however you are still encouraged to discuss broad concepts in data science with classmates! ***Again, all code and written language must be YOUR OWN.***
3. For additional factors that may affect your grade, see "Course Policies."

Data Science for Biologists, COVID-19 Version

Updated on 3/18/20 for Remote Instruction beginning 3/30/20

Your final grade will be converted to a letter grade according to the table below. Percentages for final grades are rounded; for example, a final grade of 89.5 rounds to an A-, and a final grade of 89.4 rounds to a B+.

A = 93 - 100%	B- = 80 - 82%	D+ = 67 - 69%
A- = 90 - 92%	C+ = 77 - 79%	D = 63 - 66%
B+ = 87 - 89%	C = 73 - 76%	D- = 60 - 62%
B = 83 - 86%	C- = 70 - 72%	F = below 60%

#TidyTuesday Extra Credit

Starting in week three of the semester (Monday February 3rd, 2020) there will be *weekly* opportunities for extra credit! You can receive a bonus **1% towards your final grade** for each "Tidy Tuesday" dataset you analyze, which will be discussed more in class. To obtain the extra credit, you must do the following:

- Produce one original figure displaying the data
- Use RMarkdown and fully reproducible practices to create the figure
- Share on social media with the hashtag #tidytuesday OR provide me with all your Tidy Tuesday work (stand-alone figure and corresponding RMarkdown) so that it can be shared (with option for anonymous sharing!)
    - Once we have learned version control with git (likely early March), you can **add 0.5% to this extra credit** by submitting your TidyTuesday code and visualization to me via a GitHub pull request for a total of **1.5% bonus.**

Even though this is extra credit, any figure or code that is NOT *ENTIRELY* YOUR OWN will constitute an Academic Integrity Violation and be subject to University penalties. Do NOT try to be sneaky.

Course Policies:

1. **Assignment Policy**
    a. *Late assignments submitted will NOT be accepted without prior permission.* If you need an extension on an assignment, you must contact me at least 24 hours in advance of the deadline with a request for an extension explaining why it is merited. If you receive a positive email confirmation in response, you have been granted an extension.
    b. All students will receive a one-time "free" late assignment that must be submitted within 24 hours of the deadline. You must notify me if/when you use your late assignment no later than the original deadline.
    c. Your lowest assignment grade will be **dropped!!**

2. **Attendance Policy**
    a. No more attendance! Happy remoting!

Data Science for Biologists, COVID-19 Version

Updated on 3/18/20 for Remote Instruction beginning 3/30/20

3. **Email Policy**
   Please do not expect an email response between the hours of 7 pm - 9 am, and anytime on Saturday. Please also primarily use Slack for communication in the post-COVID times.

4. **Academic Integrity Policy**
   Academic integrity is taken **EXTREMELY SERIOUSLY** in this class. You are ALWAYS expected to submit your OWN work, both code and written language. **Any form of cheating or plagiarism, including plagiarism of code, is not permitted and WILL BE REPORTED TO THE UNIVERSITY, NO EXCEPTIONS.**

   There are *limited* circumstances when code reuse is acceptable, including:
   - Front matter in Rmarkdown documents
   - Specific phrasing of hypothesis testing results and conclusions
   - Template code that I provide explicitly for your use

   You may NOT:
   - Copy/paste your code from the internet.
   - Copy/paste your code from someone else's code.
   - Copy/paste your code from elsewhere, but change the variable names so I can't tell. (Hint: I can tell.)
   - Do anything besides write your own code!

   It is understood that there are certain circumstances when many/most ways to code a problem are exactly the same. You will never be penalized in these circumstances.

5. **Regrade Policy**
   If you feel that a re-grade of an assignment or project is merited, you must fill out and submit *on paper* a regrade form, which can be found on the course website. To be considered, this form must be submitted *within two weeks* of receiving the disputed grade. If you believe points were incorrectly added for an assignment or there is a similar minor grading error, you can speak with me directly to request a re-grade without this form.

6. **Key University Policies:**
   As a student at Rowan University, you are expected to adhere fully to all university-wide academic policies outlined here: https://confluence.rowan.edu/display/POLICY/Administrative+Policies.

   Key policies to be aware of can be found at the following links:
   - Classroom Behavior Policy: https://confluence.rowan.edu/display/POLICY/Classroom+Behavior
   - Academic Integrity Policy: https://confluence.rowan.edu/display/POLICY/Academic+Integrity+Policy

Data Science for Biologists, COVID-19 Version

Updated on 3/18/20 for Remote Instruction beginning 3/30/20

- Student Accommodation Policy: https://confluence.rowan.edu/display/POLICY/Accommodation+Policy
- University Attendance Policy: https://confluence.rowan.edu/display/POLICY/Attendance+Policy

Rowan University is required to accommodate students with documented disabilities. **Your academic success is important. If you have a documented disability that may have an impact upon your work in this class, please provide the professor with up-to-date documentation from the Academic Success Center *as soon as possible!*** Students must provide documentation of their disability to the Academic Success Center in order to receive official University services and accommodations. The Academic Success Center can be reached at 856-256-4234. The Center is located on the 3rd floor of Savitz Hall. The staff is available to answer questions regarding accommodations or assist you in your pursuit of accommodations. Additional information and resources can be found at this link: https://sites.rowan.edu/disabilityresources/.

Updated Information for `Data Science for Biologists (Spring 2020)`

READ THE CONTENTS OF THIS DOCUMENT CAREFULLY. Please post questions or clarifications in the #logistical-questions Slack channel.

If you have not yet joined the class Slack workspace, please do so ASAP:

- Link to join class workspace
- Please download the Slack app for your computer and phone - will be very helpful and is WAY better than the browser version.

Weekly class

- Every week, you can expect 1-2 lecture videos to debut **each Monday** with that week's materials. Each week's (except the last week, keep reading!) assignment will be due the following **Sundays at 11:59 pm** (more info in next section).
- As before, materials will be hosted on the class website here, and submissions will be on Blackboard.
- **Please keep all active Projects in the CLASS RSTUDIO CLOUD WORKSPACE to ensure I can help you!!** Similarly, any other code you want me to see should go there. As usual, move the completed projects

Updated Grading Scheme

Your final grade will now be determined from:

- Midterm Project (25%)
- Assignments (75%)

Details on assignments

- You will stop drop the lowest grade!!
- The Final Project will be converted to a *double assignment* (i.e. will count as TWO assignment grades - total ~14% contribution to final grade rather than 25%).
- In the end, there will be **11** assignments total that count towards your grade (remember one is dropped!). This means each assignment now counts as ~6.8% of your final grade:
  - **6** assigned before break:
    - Evaluating dataviz assignment (where we found bad figures, remember that?)
    - Intro `ggplot2`
    - More `ggplot2`
    - Wrangling with `dplyr`
    - More wrangling with `dplyr`
    - The never-ending `tidyr` assignemnt
  - **6** coming up after break:
    - Permutation testing assignment
    - Linear modeling assignment
    - Logistic regression assignment
    - Clustering/PCA assignment
    - Final assignment which counts DOUBLE
- *Please rest assured: All forthcoming assignments really will be MUCH SHORTER and based around accomplishing only 1-2 overarching tasks.*
- You may participate in `#TidyTuesday` through the bitter end aka up until grades are due to the university on Friday 5/15/20. (Note, originally this would have ended during the last week of class, 4/28/20, so you get two more chances).
  - Including *this week* (the dataset for 3/17/20), there are **NINE more** #TidyTuesday's you can do.
  - The last week considered will be for the dataset released on 5/12/20 *but ONLY IF you get it to me by 11:59 pm on Thursday 5/14* - this is necessary for the university grade entry deadline.

Communication

Virtual Meetings

Please see that other document for Google Calendar Instructions

- There will be two kinds of office hours, for which all scheduling will be done via **Google Calendar**. Join the calendar here (you must be logged into your Rowan account). Please use the GOOGLE CHROME BROWSER!! Not Internet Explorer, not Safari, not Firefox - they are not guaranteed to work cleanly with the Google services that we will use. Each meeting will either take place via Google Hangouts (likely for small one-on-one meetings), and larger class meetings will take place via Zoom.
  - Every **Thursday from 12-1:45 pm** (during normal classtime) there will be an open session for anyone to join. Links to join these meetings will be embedded in the calendar event.
  - Every **Tuesday 3-4:30 pm** and **Friday 3-4:30 pm** there will be opportunities to book individual (or team!!!) remote meetings with me, again via the class google calendar. These bi-weekly events are scheduled in the calendar. To make an appointment, click the event, and then click "Go to appointment page for this calendar." You can then book 1 or more (but please don't overbook!!) 15-minute time slots. Once you have made an appointment and I am notified, I will be able to turn on a link for the meeting. Make sure to book meetings in a timely fashion!!
  - If you are unable to make any of those times, please DM me on Slack and we will find another suitable time to meet.
  - To facilitate scheduling, I highly recommend you install the "Google Calendar" app to your Slack. See here for information.
  - **When coming to your personal meetings, please start initiating the process to join the meeting a few minutes early. It may take some time to set up. Don't take away from the next person's meeting time.**

For other types of communication (questions, emails), please default to the Slack Workspace for all communications:

- DM me on the app for indiviualized conversation. If any conversation appears important enough to merit being on email, I will transition the conversation accordingly.
- Use the `#channels` to direct flow of information and general discussions
- Protip! For each message you send to a `#channel`, there is an option to reply directly and start a thread - this keeps conversations grouped! If/when you start a thread, please be sure to click the button that shares responses with everyone in the class.

What should you NOT worry about?

- Attendance. There are no required in-person meeting times. As long as you get the work in on time, you'll be set.
- Time to complete assignments. While there are deadlines, nothing will be within the context of a *timed* Blackboard test. In other words, nothing will be a "do this in Blackboard and you get cut off in 45 minutes."
- If you get severely ill and are unable to complete assignments, please reach out to me ASAP. I may ask for a *brief* video chat to confirm your circumstances. **Extensions will be liberally granted when they are appropriate.** Please do not assume an extension is granted unless I explicitly tell you so - asking for one is not the same as receiving one.
  - If an extension is not feasible given the circumstances, there will be a possibility to exempt you from the given assignment entirely.