

# **Lec 01 - Welcome**

## **Statistical Computing and Computation (Sta 663)**

**Spring 2022**

**Dr. Colin Rundel**

# Course Details

# Course Team

## Instructor

- Dr. Colin Rundel - [colin.rundel@duke.edu](mailto:colin.rundel@duke.edu) / [rundel@gmail.com](mailto:rundel@gmail.com)

## TAs

- Samuel Rosen
- Michael Sarkis
- Congwei Yang

# Course website(s)

- GitHub pages - <https://sta663-sp22.github.io>
  - HTML, PDF, and Rmds of Slides
  - Lecture screencasts (youtube)
  - Readings
  - Links to course tools
- Sakai - <https://sakai.duke.edu>
  - Announcements
  - Gradebook
  - Zoom recordings

# Course Timetable

- Lectures (weekly)
  - Wednesdays, 1:45 - 3:00 pm - Old Chemistry 116
  - Fridays, 1:45 - 3:00 pm - Old Chemistry 116
- Labs (weekly)
  - Lab 01 - Mondays, 1:45 to 3:00 pm - Old Chemistry 116

# Lectures

- Week 0 & 1 will be online (zoom)
  - screencast posted to Youtube
  - Zoom recordings on Sakai
- Week 2 onward will be in person (tentatively)
  - screencast posted to Youtube
  - No live zoom sessions
- Traditional lecture, live coding / coding demos, and short exercises + solution discussion

# Labs

- Attendance is expected
- Opportunity to work on course assignments with TA support
- Labs will begin in Week 3 (January 24th) - no lab Week 1 or Week 2
- Monday Week 1 - Optional introductory session
  - Basic usage of Jupyter Lab
  - Git and GitHub configuration and usage

# Announcements

Will be posted on Sakai (Announcements tool) and sent via email, be sure to check both regularly.



# Grading

This course is assessed 100% on your coursework (there is no exam). We will be assessing you based on the following assignments,

Assignment	Type	Value	n	Assigned
Homeworks	Team	50%	~7	~ Every other week
Midterms	Individual	40%	2	~ Week 6 and 14
Project	Team	10%	1	~ Week 10

# Teams

- Team assignments
  - Roughly biweekly assignments
  - Open ended
  - 5 - 20 hours of work
  - Peer evaluation after completion
- Expectations and roles
  - Everyone is expected to contribute equal effort
  - Everyone is expected to understand all code turned in
  - Individual contribution evaluated by peer evaluation, commits, etc.

# Collaboration policy

- Only work that is clearly assigned as team work should be completed collaboratively (Homeworks + Project).
- Individual assignments (Midterms) must be completed individually, you may not directly share or discuss answers / code with anyone other than the myself and the TAs.
- On Homeworks you should not directly share answers / code with other teams in this class, however you are welcome to discuss the problems in general and ask for advice.

# Sharing / reusing code policy

- We are aware that a huge volume of code is available on the web, and many tasks may have solutions posted.
- Unless explicitly stated otherwise, this course's policy is that you may make use of any online resources (e.g. Google, StackOverflow, etc.) but you must explicitly cite where you obtained any code you directly use or use as inspiration in your solution(s).
- Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism, regardless of source.

# Academic integrity

To uphold the Duke Community Standard:

- I will not lie, cheat, or steal in my academic endeavors;
- I will conduct myself honorably in all my endeavors; and
- I will act if the Standard is compromised.

# Course Tools

# Jupyter Lab

Dept RStudio Workbench - <http://rstudio.stat.duke.edu:8787>

OIT Container - <https://cmgr.oit.duke.edu/containers>

- Browser based + Provides consistency in hardware and software environments
- Local Python / Jupyter installations are fine but we will not guarantee support
- Common issues:
  - If This site can't provide a secure connection change https to http in the url.
  - If This site can't be reached make sure you are on a Duke network and are not use an alternative DNS service.

# Local Python + Jupyter

If working locally you should make sure that your environment meets the following requirements:

- Recent Python (3.9 or newer) with working pip (or equivalent)
- Recent jupyterlab (3.2 or newer)
- working git installation (`jupyterlab-git` recommended)
- ability to create ssh keys (for GitHub authentication)
- Packages updated to their latest version (assignments will include `requirements.txt` when needed)



# GitHub

- We will be using an organization specifically to this course  
<https://github.com/sta663-sp22>
- All assignments will be distributed and collected via GitHub
- All of your work and your membership (enrollment) in the organization is private
- We will be distributing a survey this weekend to collection your account names
  - Next week you will be invited to the course organization.

# Username advice

Some brief advice about selecting your account names (particularly for GitHub),

- Incorporate your actual name! People like to know who they're dealing with. Also makes your username easier for people to guess or remember.
- Reuse your username from other contexts, e.g., Twitter or Slack.
- Pick a username you will be comfortable revealing to your future boss.
- Shorter is better than longer, but be as unique as possible.
- Make it timeless. Avoid highlighting your current university, employer, or place of residence.

# Before next Wednesday

- Create a GitHub account if you don't have one
- Complete the course survey (you will receive before next Monday)
- make sure you can login in to the Department's RStudio server
  - <http://rstudio.stat.duke.edu:8787>
- and OIT's Sta 663 container
  - <https://cmgr.oit.duke.edu/containers>