STAT 412/612 Statistical Programming in R, Fall 2019

Time: 5:30 – 8:00 pm Wednesdays East Quad Building (EQB) 205

Instructor: Dr. David Gerard
Email: dgerard@american.edu   **Office:** DMTI 106E   **Phone:** 202-885-3022
Office Hours: Monday 3pm – 4pm, Thursday 4pm – 5pm

Overview of Topics and Course Objectives

The basics of programming using the open source statistical program R. Imputing data, performing basic analyses, graphing, data types, control structures and functions in R. Most of the course will use the tidyverse packages in R and will use the RStudio IDE (Integrated Development Environment).

Learning Objectives:
- Reproducible research using R
- Use R as a powerful calculator
- Import data from external sources
- Perform analyses including hypothesis testing and regression
- Write basic R programs using control and data structures
- Install and use packages for specific applications
- Use graphical tools to visualize and understand data

Materials:
- **Required: Laptop computer**
- **Books: R for Data Science by Wickham and Grolemund (O'Reilly) http://r4ds.had.co.nz/**
    and **style.tidyverse.org**

Class Structure This class will be a blend of lecture, class discussion and labs. I want you all to be involved during class and please do not hesitate to ask questions whenever something is unclear to you. You are expected to attend all class meetings, as I believe that attending class regularly contributes greatly to your performance in the course. It is understandable that you may have to miss class on the rare occasion. You are responsible for any assignments or papers given out during any missed class. Please obtain these materials from a colleague BEFORE the next class meeting.

I will use Blackboard to post any supplementary materials, suggested readings/practice exercises, assignments, and announcements.

Data scientists must learn to discover solutions for themselves. You should expect to have to research (use Google, stackoverflow, etc) to do your assignments. All you need to do the assignment will **NOT** have been provided to you in the lectures and course book. This is an essential part of becoming a data scientist!

GRADED WORK

Assignments: During the semester I will assign, collect, and grade assignments. There will be **approximately 10 formal assignments** throughout the semester.

You may receive assistance from other students in the class and me, but your submissions must be composed of **your own thoughts, coding and words**. I expect you to get ideas from online resources such as stackoverflow or github when you get stuck. Please cite your source when you do so and be specific about what you have added to it. You should be able to redo the code "cold" when you do this. **Failure to do so is a violation of AU's *Academic Integrity Code*.**

I will **not accept** late assignments. To accommodate extraordinary circumstances and illness, I will drop the lowest assignment grade. You should not use this as an excuse to skip an assignment, but rather to handle a situation where you are unable to turn in your assignment in a timely manner.

Exams: We will have approximately three in-class exams. Any material covered in class, reading assigned, or on assignments is "fair game." **No make-up exams** will be given unless you have an extremely compelling excuse such as observance of a religious holiday (in which case you need to let me know in advance) or a documented medical emergency.

Project: Students taking the course for graduate credit will prepare a final project using the tools learned in the class. Work with me to get your topic approved. Undergraduate students may complete a final project at their option. Your project should involve working with a fairly large real-world dataset to answer some question of interest to you. It should be reproducible and include graphical representations of your data.

Grading
You should be able to explain your work on assignments, exams, and project and your rationale. Based on your explanation (or lack thereof), I may modify your grade.
Your final grade will be determined by:

Graduate students:	Undergraduate Students:
Assignments (40%)	Assignments (50%)
Exams (40% composed of: Exam 1 = 12%, Exam 2 = 12%, Final Exam = 16%)	Exams (40% composed of: Exam 1 = 12%, Exam 2 = 12%, Final Exam = 16%)
Final Project (20%)	Attendance and participation (10%)

Please visit my office hours if you would like to see or discuss your grade at any point during the semester.

The Curve
We will curve the final course score, but not each individual assignment/exam score. Your final curved score will be calculated as described below.

Let S_i be the composite score (a number between 0 and 1) for student i based on the grade weighting above.

Let $C = median(S_1, S_2, \ldots, S_n)$, where n is the number of students in the class. Then student i's post-curve score is

$$\overset{\circ\sim}{S_i} = \frac{0.85 - C}{1 - C} * (1 - S_i) + S_i.$$

We will then use the following cutoffs on the ceiling of $100 * \overset{\circ}{S}_i$ for letter grades:

| Grade	Lower	Upper
A	93	100
A-	90	92
B+	88	89
B	83	87
B-	80	82
C+	78	79
C	73	77
C-	70	72
D	60	69
F	0	59

The idea is that we came up with a linear relationship such that the median student will get an 85% and any student who received a 100% will stay at 100%, post-curve.

If the median pre-curve grade is above 85%, then there will be no curve.

In context, a recent class of mine had a median score of 0.76 and a standard deviation of 0.12, which resulted in the following grade distribution

Grade	A	A-	B+	B	B-	C+	C	C-	D	F
Number	22	22	12	42	18	13	17	3	6	1

Below is a visualization of the curve.
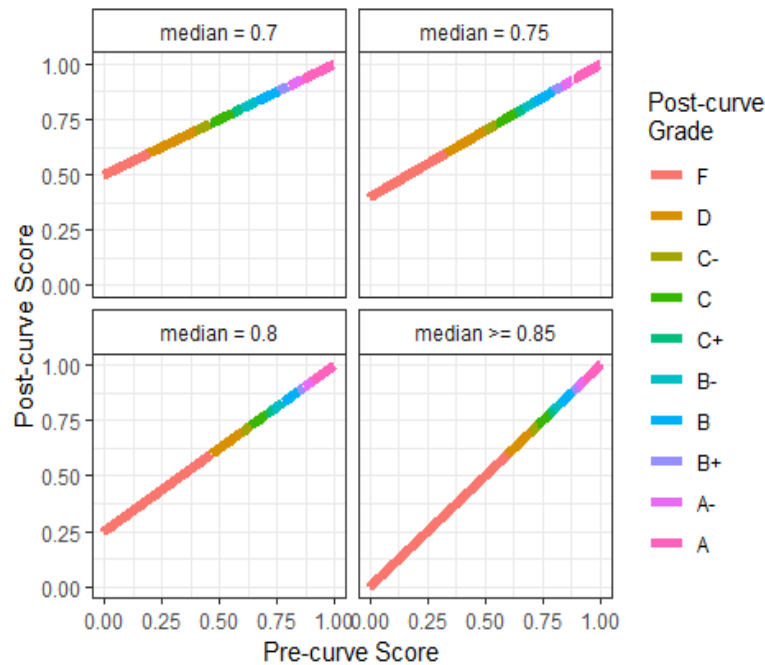
IMPORTANT DATES:

October 2	Exam 1**
November 6	Exam 2**
November 13	Final Project Progress Report Due
November 27	Thanksgiving Break (No Class)
December 4	Final Projects Due at 5:30 pm & presentations of projects
December 11	Final Exam

**Dates subject to change.

ASSISTANCE/SUPPORT: Before receiving any assistance please make sure that you have read through the class materials, and that you have made a fair attempt at the problem. You have <u>many</u> excellent resources to use for assistance outside of class:

- Always feel welcome to come visit me during my office hours. Office hours are often busy so please come prepared with specific questions.   If you are having **ANY** trouble with the class, please come see me about it as soon as possible. Do not wait until it is too late.
- I set aside a few hours each week specifically for individual meetings with students. If you need extra help, please email me to set up a time.
- Use your peers! Feel free to work with your classmates on assignments. Just make sure that you write down the solutions in your own words, just as you do in class.
- You are also encouraged to ask me questions online via email. If you are having problems with your code, be sure to attach your code to your email.

Additional support services are available on campus that may assist you in successfully completing the course requirements. Details provided by each support service's office are provided below.

- The **Academic Support and Access Center** (x3360, MGC 243) supports the academic development and educational goals of all AU students while also providing support to students with disabilities. We offer workshops on topics of interest to all students such as

time management, note taking, critical thinking, memory skills, and test taking. Additional support includes free private and group tutoring in many subjects, supplemental instruction, The Math Lab and The Writing Lab.
- The **Counseling Center** (x3500, MGC 214) is here to help students make the most of their university experience, both personally and academically. We offer individual and group counseling, urgent care, self-help resources, referrals to private care, as well as programming to help you gain the skills and insight needed to overcome adversity and thrive while you are in college. Contact the Counseling Center to make an appointment in person or by telephone, or visit the Counseling Center page on the AU website for additional information.
- **Center for Diversity & Inclusion** (X3651, MGC 201) is dedicated to enhancing LGBTQ, Multicultural, First Generation, and Women's experiences on campus and to advance AU's commitment to respecting & valuing diversity by serving as a resource and liaison to students, staff, and faculty on issues of equity through education, outreach, and advocacy.
- **OASIS: The Office of Advocacy Services for Interpersonal and Sexual Violence** (X7070) provides free and confidential advocacy services for anyone in the campus community who is impacted by sexual violence (sexual assault, dating or domestic violence, and stalking).

 A Few Additional Notes
- I expect you to be courteous to me and your fellow classmates both inside and outside of the classroom. This generally just involves a bit of common sense. **Cell phones need to be silenced and put away during class. Laptops** should be out during class time for use only on class activities. Please save texting, typing/sending emails, checking Facebook, etc. for outside of class time. Any correspondence pertaining to the course needs to be handled in a respectful manner.
- Please let me know during the first week of classes if you have any special needs that require accommodations.
- A grade of incomplete will only be given under extreme circumstances and will not be granted to any student who is failing.
- In the event of an emergency, refer to the AU information line at (202) 885-1100 and the AU Web site (http://www.american.edu/emergency) for general university-wide information. In the event that class is cancelled for **ANY** reason I will communicate with you via email and Blackboard to let you know what work you will be responsible for.
- Please be sure that you are familiar with AU's Academic Integrity Code, as I am required to report any cases of academic dishonesty to the dean of CAS. For your review: http://www.american.edu/academics/integrity/.