73-265: Economics & Data Science

Professor John Gasper

Fall 2018
Updated: August 27, 2018

Course Time / Location:
    Tues/Thur 3:00-4:20pm;
    Room: TPR 2612

Instructor: John Gasper
    Office: TPR 4202
    Email: gasper@cmu.edu
    Office Hours: Tues/Thur: 4:30-5:30pm; Wednesday 2 - 3pm; by appointment

In general, I have an open door policy: if my door is open and I'm not meeting with someone, you are welcome to come in and meet with me. If my door is closed, I am not available (out of the office, working, etc). I know finding office hour times that work for everyone is difficult. Consider these a suggested meeting time, but I highly encourage you to schedule an appointment. Please, don't interpret a closed door as me not wanting to meet with you. I do, but just send me an email to schedule some time.

Teaching Assistants: Mason Paccione

    Office hours: TBA

Required materials:

- Clicker(either the device or app for your phone)

- *R for Data Science*, Grolemund and Wickham http://r4ds.had.co.nz/

- *Storytelling with Data*, Knaflic website for book

- *Mastering Metrics: the Path from Cause to Effect*, Angrist and Pischke

Suggested Supplemental Texts:

- *Introductory Econometrics: A Modern Approach*, Wooldridge.

- *Mostly Harmless Econometrics: An Empiricist's Companion*, Angrist and Pischke

Course Description and Goals

This course is at the intersection of economic analysis, computing, and statistics. It develops foundational skills in these areas and provides students with hands-on experience in identifying, analyzing, and using data to solve real-world problems in economics and business.

In this course we will be using the statistical programming language R, which is a powerful open-source language that is widely used. In this very hands-on, data-centric, course, students will learn the basics of data manipulation, how to visualize, present and interpret data related to economic and business activity by employing statistical analysis and various visualization techniques.

Through many interactive exercises, students will develop a foundation for data-driven decision making. This course also provides a solid base for future courses such as Econometrics I and II (73-274/275) that provide a rigorous treatment of more advanced methods used for business, economics, and public policy. The exact data sets we will use will vary and depend on student interest, but the underlying objectives of the course will remain:

- Develop competence in the statistical programming language R and manipulating data.

- Develop competence in visualizing data using R.

- Perform various statistical analyses used in economic analysis

- Interpret the estimates from these statistical analyses.

- Develop communication skills to effectively describe and visualize findings for technical and non-technical audiences.

Attendance and participation

It is easy to take the attitude that your job (and mine) is accomplished with your mastery of the material of the course, and consequently that I need not bother with whether you show up for class. Realistically, we know that in general the vast majority of students who feel they don't need to come to class are mistaken, but only find that out, to their shock, as they do poorly on exams. This course is a course with lots of material to cover. The pace of the course will be fast. If you miss a class, you will be substantially behind. I will expect you in class and I expect you on time. This not an "easy" course and a large part of the lecture material will not come from the text. The exams will cover both sets of material.

More importantly, your class participation also provides important feedback to me regarding how well topics are getting across. If something that isn't clear to you, please let me know. You are probably not the only one. Attendance is necessary, but not sufficient for effective participation, which also requires actively engaging the material. That said, if you feel uncomfortable about class participation either in general or for specific topics, see come talk to me and we'll find some way around it.

This semester we will experimenting with "clickers" in the course. Once these are setup, **you are required to bring your clicker to every class**. I will be using these to take attendance for the course, but primarily to have you real-time survey questions. If you forget

to bring your clicker and we use it that day, you will be counted as absent. In addition, I might occasionally give unannounced short quizzes.

I will also make an effort to get to know you. It is also fair to say that those who sit near the front and participate in class will get the benefit of the doubt when their grades are below a borderline. During the first week of class, I ask that you schedule an appointment with me and come by my office. Early in the first week of the semester I will pass around a sheet with available times. These meetings will probably only last about 10 minutes but I find them incredibly valuable. I feel that I can best present material to you only after I know you.

Cell phones, laptops, and other technology

I would typically teach a course like this in a computer lab. Unfortunately that isn't possible this semester. So I'm going to ask that you **bring your laptop to every class.** If you don't have a laptop, come see me ASAP and we'll see if we can find a solution.

You need to bring your laptop because we'll be running code and performing analyses pretty much every day in class. You should be using your computer to follow along and take notes. In addition, I am also willing to venture a guess that many of you will be tempted to check your email, the news, etc. This is not fine. It's distracting to me and more importantly your fellow students. Please be respectful of their learning environment and postpone all non-class activities until after class.

I also ask that you turn off your cell phone during class. If there is an emergency and you might need to be contacted, please talk to me before class. Otherwise there should be no reason to hear a phone ring or see someone send a text. Text messages and phone calls during class are very distracting and disrespectful to me and your other students. If you are surfing the internet, texting, etc, during class you will be asked to leave.

Finally, I understand that many of you might use alternative note taking methodologies, but this has its limits. I do not claim to be a riveting lecturer. So I ask that student record or tape any classroom activity without my express written consent.

Course Logistics

This course has a Canvas site. The sites should set up and functioning. Our class page can be accessed via the following URL:

<div align="center">

https://canvas.cmu.edu/courses/4292

</div>

Handouts, problem sets, updated syllabi and announcements will be posted to Canvas and you are responsible for checking the site regularly. The TA and I will also maintain the Canvas/Piazza discussion board. If you have any questions about the techniques, problem sets, etc, ask them on the discussion board. It has been my experience that one of the best ways to learn something is to try to explain it to someone else. So *we will expect you to try to answer the questions that other students ask*; doing so will aid the participation element of your grade.

I welcome questions during class: if you have a question or a comment, please let us know. I will generally pause after each slide and ask if there are any questions. *Please feel encouraged to raise questions during class.* I am also fairly accessible via email, but you should not expect a reply immediately (within 24 hours).

Accommodations for Students with Disabilities

Carnegie Mellon University is committed to providing reasonable accommodations for all persons with disabilities. I would ask any student needing a learning accommodation to let me know at the beginning of the term so that we can work out necessary alternative assessment options. All information will be considered confidential and only released to appropriate persons on a need to know basis.

Health & Well-being

Take Care of Yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, I strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call412.268.2922 to make an appointment. Consider reaching out to a friend, faculty or family member you trust for help.

Academic Integrity

You should feel encouraged to talk with your class mates about the problems on the problem sets, but do not copy even parts of someone else's work. Absolutely no copying of code, figures, or any exposition is allowed.

Take note of the following from the CMU policy on cheating and plagiarism: *In all academic work to be graded, the citation of all sources is required. When collaboration or assistance is permitted by the course instructor(s) or when a students the services provided by Academic Development, the Global Communication Center, and the Academic Resource Center (CMU-Q), the acknowledgement of any collaboration or assistance is likewise required. This citation and acknowledgement must be incorporated into the work submitted and not separately or at a later point in time. Failure to do so is dishonest and is subject to disciplinary action.*

I am very sensitive to cheating and plagiarism;cheating of any kind will not be tolerated. **If I suspect an academic integrity violation, I will report it to the CMU administration and you will be penalized one letter grade off of your final grade.** If you have any doubt about your actions, please ask me. I strongly encourage you to review Carnegie Mellon's policies regarding academic integrity:

http://www.cmu.edu/academic-integrity/index.html

Grades

Each student's grade for the course will be based on the following:

1. *Participation* 5%

2. *Problem Sets* 10%

3. *Visualization Challenge* 10%

4. *Final Project* 25%

5. *Quizzes* total 25%

6. *Final Exam* 25%

Given the breakdown above, I will assign final grades based on your cumulative score. Grade cut-offs are set as follows:

A: 90% and above

B: 80 - 90%

C: 70 - 80 %

D: 60 - 0%

R: Below 60 %

The above allocation of points is my contract with you. Please do not ask if you can raise your grade by doing extra work or if I can modify the above formula to accommodate your circumstances. In the interests of fairness, I apply the same rules to everyone. Some students will inevitably fall close to, but below a cutoff when I assign final grades. Such students sometimes ask for cutoffs to be reduced. Please do not do this. If you are just below a cutoff you may be close to the lowest A or B, but still some way from the average A or B and we consider these when assigning cutoffs. All of that said, I do reserve the right to lower the cutoffs if I deem it most fair to the class overall, but I will not raise them. I.e., if your final percentage in the course is a 92%, you will receive an A no matter what anyone else in the course receives.

The only way to learn the material is to do it. There will be multiple problem sets distributed, but the problem sets will be graded on a "check-minus / check-plus" system where credit will be given for completing the problem set, rather than on correctness per se. A check will mean that you've reasonably attempted the problems; a check-plus is awarded for exemplary work (i.e., I could use it as a solution set next year) and a check-minus for a poor and deficient attempt. Solution sets will be posted and *you will be responsible for checking that your work is correct.* Students sometimes find this frustrating, but checking your own work is an excellent way of reinforcing your understanding of the material. There will be no personal extensions to the homework deadlines because I will post the solutions right after the homework deadline. On the top of the first page of every assignment, you

must list everyone you worked with on the assignment. Absolutely no copying of code, figures, or exposition is allowed.

In addition to the problem sets distributed, mastery of the material will be gauged via in-class quizzes, presentations, and a final exam. There will be multiple quizzes throughout the semester. Some of these quizzes will be announced and some could be unannounced and "pop" quizzes.

Finally, this course is relatively heavy on presentations for being a "methods" course. Fundamentally, visualizing and presenting your analysis is incredibly important. If you conduct the best, most thoughtful, analysis possible but can't present it, then the quality of your work won't matter. Therefore the course will culminate with a final group project and presentation. We will also have a visualization challenge where you will be pitted against other students in the class to redesign an visualize an existing figure.

As faculty, we know that your schedule during the semester can be hectic. We also know that various events can happen during the semester that make finishing projects on time difficult. We also, however, expect you to know these things as well. There is a no make-up policy for any missed work, quiz, or presentation in this course, unless it is a university approved absence.

Re-grading Requests: Everyone makes mistakes, including me. If you feel that there was an error in grading your assignment, you must submit re-grading requests within one week from the date that the assessments are returned to the class. Re-grading requests must be submitted to me in writing with the assessment. When reviewing the request, the entire assessment will be reviewed. It is the responsibility of each student to check that Canvas accurately reflects grades received on the assignment.

Canvas: Do not use Canvas' final grade calculation to calculate your final grade. Given the check system of the homework assignments, I must make each assignment worth three possible points (0 = no home work, 1 = check minus, 2 = check, 3 = check plus), but it will not be graded out of three. If you see that you have received a two out of three on the assignment, that implies that you have received 100% for that homework, however Canvas will calculate this as a 66.6% and it will influence your grade calculation. Moreover, each quiz will be worth the same final percentage, but they might vary in the number of points for grading on each quiz. This will give various quizzes undue weight in the final grade calculation. Do not use the automatic grade calculation in Canvas. You should, however, be able to take the scores provided on Canvas and calculate your correct grade in the class.

Tentative course schedule

The course schedule is very tentative and will almost surely change a little (but hopefully not much). When there are changes, I'll update the syllabus and make an announcement in class. You are responsible for keeping current on the latest version and paying attention to announcements made in class.

Class	Day	Date	Topic	Reading	Assignment
1	T	28-Aug	Intro		
2	Th	30-Aug	Viz Principles		PS 0 Due
3	T	4-Sep	Beginning R		
4	Th	6-Sep	Visualizing Data / R		PS1 due
5	T	11-Sep	Visualizing Data / R		
6	Th	13-Sep	Visualizing Data / R		PS2 due
7	T	18-Sep	Data wrangling		Quiz
8	Th	20-Sep	Data wrangling		
9	T	25-Sep	Data wrangling		PS 3 due
10	Th	27-Sep	Viz Challenge		Viz Challenge
11	T	2-Oct	Regression		
12	Th	4-Oct	Diagnostics and Inference		Quiz
13	T	9-Oct	Transformations		
14	Th	11-Oct	Multiple Regression		PS 4 due
15	T	16-Oct	Categorical Data		Quiz
16	Th	18-Oct	Interactions		
17	T	23-Oct	Diagnostics and Inference		PS 5 due
18	Th	25-Oct	Visualzing Results		Quiz
19	T	30-Oct	Causality		
20	Th	1-Nov	Experimental data		PS 6 due
21	T	6-Nov	Instrumental Variables		
22	Th	8-Nov	Instrumental Variables		Quiz
23	T	13-Nov	Instrumental Variables		
24	Th	15-Nov	Classification		
25	T	20-Nov	Logit		PS 7 due
26	Th	22-Nov	BREAK		BREAK
27	T	27-Nov	Logit		Quiz
28	Th	29-Nov	Applications		
29	T	4-Dec	Presentations		
30	Th	6-Dec	Presentations		
		TBA	Final Exam		

All items are an approximation. Due dates, quiz dates, and presentation dates subject to change.