ECON 103: Econometric Data Science, Fall 2021

University of Pennsylvania

Syllabus

(Tentative: August 26, 2021)

Instructor:

Wayne Gao

Email: waynegao@upenn.edu
Zoom Office Hours: Monday 3-4pm and Thursday 12:30-1:30pm
Zoom Meeting Room ID: waynegao
My office hours will be held via Zoom. If the office hours do not fit into your schedule, or if you feel the necessity to meet in-person, please contact me for an appointment.

Lectures:

Time: Tuesday & Thursday, 1:45pm-3:15pm
Location: ANNS 110

Teaching Assistants:

Sajad Ghorbani: sajadgh@sas.upenn.edu
David Mao: davidmao@sas.upenn.edu
TA Office Hours: to be announced.

Recitation Sections:

201: Friday, 8:30-9:30am, Sajad Ghorbani
202: Monday, 10:15-11:15am, David Mao
203: Friday, 10:15-11:15am, Sajad Ghorbani
204: Monday, 8:30-9:30am, David Mao

Course Description:

This course focuses on data description, probability, and statistics, as relevant for economics. Topics include economic data sources, descriptive statistics, probability distributions and population moments, sampling distributions and sample moments, statistical estimation, confidence intervals, hypothesis testing, and an introduction to linear regression. Focus is on both theoretical and practical issues involved in the substantive interpretation of economic data using

statistical and econometric techniques. Empirical case studies are discussed throughout. Students will be guided to learn the statistical computing language R, and will be required to carry out various statistical analyses in R.

This course fulfills the College's Quantitative Data Analysis requirement.

Prerequisites:

ECON 001 and ECON 002 and MATH 104 and (MATH 114 or MATH 115).

To do well in this course you will need to be comfortable with algebra, manipulating sums, differentiation and partial differentiation, solving unconstrained optimization problems, and integration.

Notes: This course is intended primarily for economics majors. ECON 103 cannot be taken by any student who has already completed Statistics at the level of STAT 430 (including the sequence STAT 430 and 431) or higher. Such students must take an additional 200-level course to satisfy course requirements of the major.

Textbook:

The official textbook for this course is "*Introductory Statistics* (4th edition)", by Sheldon M. Ross, Elsevier. Digital copies of the textbook are freely available online via the Penn Library or the following webpage (PennKey login required): https://www.sciencedirect.com/book/9780128043172/introductory-statistics . While I suggest that you complete the assigned readings, my lecture slides, which will be posted online at the start of each week, are the final authority on course material. In particular, you are not responsible for material in the textbook unless it is also covered in lecture, but you are responsible for material from lecture even if it is not covered in the textbook.

Required Software: R

We will use the statistical package R via a front-end called RStudio throughout the course. Both R and RStudio are open source and free. First, download and install R from http://cran.r-project.org/. Second, download and install RStudio by visiting https://rstudio.com/products/rstudio/download/ and clicking "Download" under the free RStudio Desktop version. You might need to further choose the right version for your operating system (Windows/Mac). Make sure you download and start using R as the homework assignments will involve coding and running some data analysis on R .

You will be *guided to learn* R *by yourself*. The lectures, recitations and homework assignments will provide the necessary guidance, tutorials or instructions about the R basics and the specific R commands used in this course. However, this course will *not* provide a comprehensive and systematic coverage of computer programming with R, for the following reasons: (1) The focus

of this course is on probability theory and statistical method, and we have a limited amount of time for instruction. (2) The heterogeneity in the students' past exposure to coding also makes it inefficient to teach coding in lectures. (3) The best way to learn coding (in R) is "learning by doing", and the process of figuring out how to code on your own is in many ways more important and useful than the "correct codes" per se.

That said, *students are highly encouraged to seek help from the Instructor, the TAs, and other students on coding*. Often times, the questions and problems students encountered in coding tend to be very specific and detailed. Thus "one-on-one tutoring" during office hours, along with discussions among students (e.g. on Piazza), are often much more effective and efficient than lectures.

In addition, an invitation link for *free* access to *DataCamp*, an online learning platform for computer programming and data analysis, will also be sent out on Canvas. The platform allows you to learn and practice R interactively on your web browser or mobile device. Those students who are interested in learning computer prorgamming in R more systematically are encouraged to take advantage of DataCamp. Homework assignment may contain R exercises on DataCamp.

Some additional R resources:
Optional textbook:*"The R Student Companion*", by Brian Dennis
http://cran.r-project.org/other-docs.html
http://www.twotorials.com/
https://www.r-bloggers.com/how-to-learn-r-2/
http://cran.r-project.org/doc/contrib/Farnsworth-EconometricsInR.pdf
https://stats.idre.ucla.edu/r/

Course Website: Canvas

We will use *Canvas* to make course announcements, post course material, answer questions about course material and respond to private messages from individual students regarding personal issues. All written communication for ECON 103 should be directed to *Canvas*, *not* to the instructors' personal email accounts.

Discussion Board: Piazza

You are highly encouraged to the discussion board, *Piazza*, which is accessible via *Canvas* for Q&A about course material. By asking your questions and answering others' questions on *Piazza*, you create a positive externality: other students benefit from your questions/answers and you benefit from theirs. The instructor and TA's will actively moderate *Discussion*/*Piazza* on *Canvas* both to answer questions and approve (or correct) answers written by your fellow-students. In order to encourage participation, and reward the benefit that students bring to the discussion, I will award bonus points for participation in Piazza. This could be up to 5% of the

final grade. Bonus points will be awarded for constructive questions, answers, and notes on Piazza, as well as in-class/online student presentations about assigned Piazza discussion questions during lectures. Even if you post anonymously to other students (but not to instructors: if you post anonymously to the instructors as well, I will not be able to recover your identity and award you bonus points) on Piazza, we can still award you bonus points for online participation. Bonus points are discretionary.

Departmental Course Policies:

All Economics Department course policies are in force in ECON 103 even if not explicitly listed on this syllabus. See https://economics.sas.upenn.edu/undergraduate/course-information/course-policies for full details.

Academic Integrity:

All suspected violations of the code of academic integrity as set forth in the Pennbook will be reported to the Office of Student Conduct. Confirmed violations will result in a failing grade for the course.

Public Health Policies

Please adhere to all then-effective public health policies announced by the University and the SAS. See the following webpages for the most up-to-date policies and information.
https://coronavirus.upenn.edu/content/public-health-guidance
https://provost.upenn.edu/fall-2021-covid-19-faqs-students
https://provost.upenn.edu/fall-2021-covid-19-faqs-instructors
https://www.sas.upenn.edu/covid-19-announcements-resources

In particular, please pay special attention to the following SAS guidelines:

- ***Mask Mandate:***

  "Per the University's August 25 Message to the Penn Community on the Start of the Fall Semester, masks covering the nose and mouth must be worn at all times in all public indoor spaces, including classrooms, by all persons. Students who refuse to wear masks in the classroom should be referred by their instructor of record in the course to the Office of Student Conduct for disciplinary action; masking non-compliance by any Penn community member can also be reported anonymously through the University's Masking Violations page. Instructors, including TAs and LAs, should remind students of their obligation to wear masks in the classroom and the consequences for non-compliance. Should a student refuse to wear a mask during a particular class meeting, the instructor should first ask the student to comply. If the student still refuses to wear a mask, the instructor should ask the student to leave

the class meeting. If the student refuses to leave, the instructor should announce that the class meeting is canceled and ask the class to vacate the classroom for the sake of health and safety."

- **What if a student or students in a class test positive for COVID-19?**

  "Students who test positive will be quarantined away from the classroom and other public spaces for 10 days and their close contacts will be traced and tested. The remainder of the class will continue meeting in person in the event of an isolated positive test or tests among students in the class. It is important to note that for reasons of privacy, the identity of a student testing positive will not be reported to instructors. Individual students will utilize the Course Absence Report (CAR) system to notify instructors of their absence from class. Instructors should familiarize themselves with the CAR system, urge students to use it to report absences for any reason, and be on the alert for CAR messages from students concerning absences."

Grading:

Grades for this course will be determined based on 10 homework assignments, 7 quizzes, 2 midterms, a comprehensive take-home final exam, and Piazza bonus points:

$$\begin{aligned}\text{Course Score} = \ & (10\% \times \text{Homework}) + (10\% \times \text{Quizzes}) \\ & + (20\% \times \text{Midterm 1}) + (20\% \times \text{Midterm 2}) + (40\% \times \text{Final}) \\ & + (\text{up to } 5\% \text{ of Piazza Bonus})\end{aligned}$$

Homework:

Homework assignments will be posted on Canvas each Tuesday, starting from the second week of the semester. Homework will be collected and graded. Group work is encouraged, but you will have to submit your own answers. When calculating your homework average, I will drop your two lowest scores and weigh the remaining homeworks evenly.

Quizzes:

There will be seven short in-class quizzes over the course of the semester. Unless otherwise indicated, each quiz will cover all the material before, and including, the last lecture with higher weight given to material that was not tested in previous quizzes. When calculating your quiz average, I will drop your two lowest scores and weight the remaining quizzes evenly. There will be no makeup quizzes so be sure to use your two "free skips" carefully.

Exams:

There will be two 80-minute in-class midterm exams and a 2-hour final exam during the exam period. Each midterm is worth 20% and the final is worth 40% of your grade. Neither midterm is comprehensive, but the final is: it will focus on the final third of the course but also include several questions on earlier material. There will be no makeup midterms: if you miss one midterm, your final exam will be worth 60% to compensate; if you miss two midterms, it will be worth 80%. The makeup final will take place at the beginning of next semester and is outside of the instructor's control: eligibility as well as the time and date are determined by the Department of Economics. Cheat sheets are not permitted on exams. Scientific calculators are allowed but graphing calculators are not. You may write in pencil or pen on your exam as it will be photocopied before being returned to you. We will check ID cards at each exam. Also, please note that it is not advisable to make early flight arrangements before knowing when your final exam will be held. Accommodation for sudden emergencies should be discussed with and will be determined by the Undergraduate Chair.

Regrade Requests:

Exam regrade requests must be made in writing within a week of receiving your graded exam. As we re-grade the entire exam, your score could rise or fall. You may not discuss your answers with the Teaching Assistants or the instructor before submitting a regrade request. To deter regrade-related cheating, we will photo-copy the exams of all or a random subset of students before returning the exams.

Course Curve:

We typically try to target an average GPA in the range between 3.0 and 3.2, or slightly above a B average. In a nutshell, I will give about 30% As and A-s, 40-50% Bs and 20-30% Cs. If necessary, I will curve overall course scores (not individual assignments) so that they approximately fall into these ranges. I reserve grades below a C-minus for those cases in which a student fails to attain a minimum level of basic competence in statistics, an absolute rather than relative standard.

Course Schedule (Tentative)

Date	Day	Lecture	HW Posted	Quiz
Aug. 31	Tue	Introduction		
Sep. 02	Thu	Summary Statistics: I		
Sep. 07	Tue	Summary Statistics: II	HW1	
Sep. 09	Thu	Basic Probability I		Q1
Sep. 14	Tue	Basic Probability II	HW2	
Sep. 16	Thu	Basic Probability III		
Sep. 21	Tue	Discrete Random Variables I	HW3	Q2
Sep. 23	Thu	Discrete Random Variables II		
Sep. 28	Tue	Discrete Random Variables III	HW4	
Sep. 30	Thu	Discrete Random Variables IV		Q3
Oct. 05	Tue	**In-Class Midterm 1**		
Oct. 07	Thu	Continuous Random Variables		
Oct. 12	Tue	Normal Random Variables	HW5	
Oct. 14	Thu	**Fall Break: No Lecture**		
Oct. 19	Tue	Sampling Distribution		
Oct. 21	Thu	Estimation I		Q4
Oct. 26	Tue	Estimation II	HW6	
Oct. 28	Thu	Confidence Intervals I		
Nov. 02	Tue	Confidence Intervals II	HW7	Q5
Nov. 04	Thu	Hypothesis Testing I		
Nov. 09	Tue	Hypothesis Testing II	HW8	
Nov. 11	Thu	Reserve Lecture		Q6
Nov. 16	Tue	**In-Class Midterm 2**		
Nov. 18	Thu	Linear Regression I		
Nov. 23	Tue	Linear Regression II	HW9	
Nov. 25	Thu	**Thanksgiving Break: No Class**		
Nov. 30	Tue	Linear Regression III		
Dec. 02	Thu	Linear Regression IV	HW10	Q7
Dec. 07	Tue	Linear Regression V		
Dec. 09	Thu	Concluding Remarks		
		Final Exam: 12pm-2pm, Dec. 16 (Tentative)[1]		

See the College's Academic Calendar for important dates such as deadlines for course selection, course drop, grade type change and course withdrawal.

[1] Check https://srfs.upenn.edu/registration-catalog-calendar/final-exams for updates.