

Quantitative Analysis for Public Policy
UPADM-GP 111, Fall 2017

Lecture: J. Andrew Sinclair

J.Andrew.Sinclair@nyu.edu

Tuesday, 12:30 pm – 3:15 pm

Room: 60FA 125

Lab Section: Xiaodi Li

xl1060@nyu.edu

Friday, 12:30 pm – 1:45 pm

Room: WAVE 668

(Formal) Course Description: *Introduces students to basic statistical methods and their application to management, policy, and financial decision-making. Covers the essential elements of descriptive statistics and univariate and bivariate statistical inference and introduces multivariate analysis. Emphasizes applied statistics and data analysis in addition to statistical theory. Encourages a critical approach to reviewing statistical findings and using statistical reasoning in decision-making.*

Translation: In this class, you will learn about how to use data to think about social science questions. This course is introductory only in the sense that we assume no prior knowledge of statistics or advanced mathematics. We plan to move fast and help you acquire the tools you need to be successful in a rapidly evolving field, to be able to perform your own independent analysis in a variety of contexts, to understand top-tier research, and to be a sophisticated consumer of quantitative information. You *will* leave this class with new technical skills that you can use and a better ability to understand the world around you. This is powerful and awesome stuff.

Textbooks & Materials: Readings for this class will come from two main texts, listed below, and from other handouts and materials that I will produce and post on the NYU Classes website.

Required:

Wheelan, Charles. 2013. *Naked Statistics: Stripping the Dread from the Data*. New York, NY: Norton & Company.¹

Angrist, Joshua D. and Jorn-Steffen Pischke. 2015. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton, NJ: Princeton University Press.²

¹ You can get a used copy on Amazon for around \$10. There tend to be lots of these available. You can still find copies at regular bookstores too (try the Strand bookstore). Also, ask around with other policy majors.

² You can get a used copy on Amazon for about \$20. New copies are \$25-35.

For the exam, you will want to have a non-programmable calculator – a Casio FX-300ES Natural Display Scientific Calculator or an approved equivalent.³ It might be quite helpful to bring this to class as well.

To successfully complete your assignments, you will also need access to the computer software for the course, Stata 14. There are several different flavors of Stata; for the applications in the course, Stata IC is fine (I personally use Stata SE).⁴ You may be tempted to buy “small stata” (cheaper) but some of the data applications for our course will exceed its capabilities. You have a couple of options to get access to Stata. I recommend looking up:

<http://www.nyu.edu/life/information-technology/getting-started/software.html>

And:

<http://guides.nyu.edu/stata>

You can follow the links to buy a copy of Stata 14, find out where on NYU’s campus you can get access to a computer with it, and find out how to use the Virtual Computer Lab (VCL) to use it for free. It seems that students find the free version through the VCL more annoying to use than a purchased copy – but it is free.⁵

Optional:

You may benefit from additional perspectives on the course material. I will distribute specific readings, including small sections of some of these, via NYU Classes. These are **optional** but useful supplemental materials.

Diez, David M, Barr Christopher, and Cetinkaya-Rundel Mine. 2012. *OpenIntro Statistics: Second Edition*. URL: <http://www.openintro.org/stat/textbook.php>. [Free]

Acock, Alan C. (Various versions): *A Gentle Introduction to Stata*. I’d get at least the 4th edition if you plan to get a copy of this book.⁶

³ Amazon link: www.amazon.com/Casio-fx-300ES-Scientific-Calculator-Black/dp/B007HJ89VE. It’s about \$12. You can also get something similar from the NYU Bookstore. Again, ask around; somebody can lend you one.

⁴ The difference between the two is really about what size projects they can handle (very large vs. huge). It won’t matter for this class.

⁵ The 6-month license at \$75 is roughly the equivalent of buying a traditional course textbook. If you are going to do future quantitative class, or at some point do independent quantitative research (as part of another course or in a summer job, etc.), I’d encourage you to buy the perpetual license Stata/IC version for \$198. If cost is an issue, check out the VCL.

⁶ There is a fifth edition, which is more up-to-date and “better” for a more recent version of stata. Not much changes, though, so used versions of the 4th edition on Amazon are currently selling for about 30% of the 5th Ed. price. I have the 4th Ed. and it’s fine. You should know that you *can* get most of this information for free online if you are patient in your google searches. This might be helpful, though, and a good investment if you are struggling in the first week or two with STATA.

Mitchell, Michael N. (Various versions). *A Visual Guide to Stata Graphics*. I'd buy a used copy of the third edition.⁷

Reinhard, Alex. 2015. *Statistics Done Wrong: The Woefully Complete Guide*. San Francisco: No Starch Press. Lots here for free: <https://www.statisticsdonewrong.com/>

Seife, Charles. 2010. *Proofiness: How You're Being Fooled by the Numbers*. New York: Penguin Group.

Klein, Grady and Dabney, Alan. 2013. *The Cartoon Introduction to Statistics*. New York: Hill and Wang.

Look for further recommendations to be posted on NYU Classes as we move through the course. You'll also encounter bits and pieces of other books in your HW.

Course Requirements & Grading: Your performance in this class will be assessed through multiple assignment types. These components are weighted:

Attendance	5%
Participation	5%
Pre-Class Quizzes	5%
Homework Assignments	20%
In-Class Exam:	35%
Final Memo:	30%

Attendance (Completion, In-Class Exercises): Most lectures will feature in-class exercises. You will complete these as part of a group. Showing up to class and doing these is an important part of preparing you to manage the homework. You can miss class twice without a penalty.

Participation: There is a participation component to your grade so you *must* engage with the course readings, lecture materials, homework questions, and assignments on the Forum on NYU Classes. To obtain a full score you are expected to make at least seven contributions over the term: posting a question, answering someone else's question, posting a topical article that may be of general interest, posting review materials that would be of general use. You are absolutely expected to engage online in a manner that is respectful and gracious towards your fellow classmates.

Pre-class quizzes: These are available online in advance of class via the NYU Classes Tests & Quizzes feature. These are multiple-choice questions based on the reading for that week as well as the past material for the course. These questions focus on major themes to help you

⁷ This is indeed what I have. It's expensive --- \$60-\$80 on Amazon --- but it is a really great reference book, especially if you intend to go on to do quantitative things.

identify and review critical concepts, prepare for class each week, and prepare for the exam. These are for you to do on your own, not as part of a group, because they are graded on the basis of how many you answer correctly. These should only take about 2-15 minutes to complete: *only open the quiz when you are ready; they have a 15-minute time-limit* and may only be submitted once.

Homework Assignments: Your homework assignments will become available on NYU Classes after each lecture and are due *online* on NYU Classes before the next lecture. Your homework assignments will be graded based in part on how well you answer the questions. While you must turn in your own answers, it is acceptable to work on these with other people in the course. The Friday lab section will address *only the most difficult questions, so attempt your homework in advance.*

- **You must turn in your HW electronically.** There are a couple of ways to do this: most of the time, students simply type answers in Word or some other standard text editing software. It can be time consuming to type math this way, so you can scan in something that is written sufficiently neatly for us to read it. I recommend typing.
- There is also a solution to this called LaTex. This is how I produce all of the assignments and other materials for the course. It takes a little bit of effort to learn but would save you time over the course of the term. And, of course, if you are considering going to a variety of graduate programs, you may want to invest in learning how to do this now (my PhD program, for example, required that our HW in all of our courses be turned in as tex documents).

In-Class Exam: We will have an exam in the 11th lecture period. This is an in-class exam, closed book, and entirely on paper (no computers). You will want to bring a non-graphing calculator of the type mentioned above under “materials.”

Final Memo: You will provide a statistical analysis and write up the results in the form of a professional memo. This will be due at the end of the term (instead of having a final exam). Details for this project will be released later in the course via NYU Classes. This will also be turned in via NYU Classes online.

Administrative Details: This section provides a few other guidelines for the course.

Course Organization: This course has two regular components – a lecture and a lab section.

The lectures will be based on, but not limited to, the readings listed for that lecture period: do the reading in advance of class! In addition, you should review the past week’s homework solutions and make sure you remember the main points from the prior lecture as well. You will really need to do the reading to have success on the concept quizzes. Most lectures will be broken into parts: Lecture A, Practice Problems A, Solutions A, Break, Lecture B, Practice

Problems B, Solutions B, done. You should bring your computer to class or arrange to share a laptop with somebody in class.

Xiaodi Li will teach the lab section; this focuses mostly on broad discussions on the points raised by the more difficult homework problems or key skills students seem to be struggling with overall. Before you come to the lab section, try your homework. There will not be time in the lab sections to cover every homework question; you will get the most out of these by preparing in advance.

Anticipated Plan:

Mondays: Read, Do the Concept Questions (Quiz)
Tuesdays: Lecture, Do the In-Class Questions
Wednesdays: Do the Homework
Thursdays: Do the Homework
Fridays: Lab Section, Ask Homework Questions, Finish HW.
Saturdays/Sundays: Do something else. It's a big, interesting world.

Office Hours/Additional Help: I will post via NYU Classes specific times when I will be in my office. Also: use the Forum to ask questions --- I know you often work late at night; I do too, and that's often the fastest way to get an answer.

Email: "Should we email you or email Xiaodi?" In general, you should copy both of us. All emails should have "Spring 2017 Quantitative Analysis" in the subject-line: we get a lot of email every day and I want to make sure I manage to catch the emails about this course. If you do not hear back from me within about 48 hours, I do not consider it rude to re-send your message; I mean to answer every email I get from a student in my course. If I haven't answered yours, it is a mistake. In general, email should be restricted to personal matters -- content questions should go to the Forum. If you have a question about something, others almost assuredly do too, and they will appreciate that you asked.

Late or Missing Assignments: Major assignments to be turned in via NYU Classes will have some provision for accepting late assignments, although these may be penalized. You must complete the concept question quizzes by their deadline and submit the homework within the time-window set up on NYU Classes.

Special Requests: Excused absences, extensions on assignments, and other accommodations should be requested and discussed in advance. A student with a qualified disability requesting a reasonable accommodation should do so through the Moses Center:

<https://www.nyu.edu/life/safety-health-wellness/students-with-disabilities.html>

As general advice: it is always easier to solve a problem earlier rather than later. If you are worried about anything related to the course, it does not take very long to get in touch and start a discussion.

One common request to which the answer is “No”: “I already know how to use R / SPSS / Excel / Python... can I just do my work in the program I already know?” – No. It will never hurt you to have learned another program, especially since Stata is widely used in policy-related professions.

Changes to the Syllabus: Expect changes in the syllabus; I will post any updated versions on NYU Classes.

The Course Website: You should regularly check the NYU Classes website and, of course, pay attention to course-related emails sent to you through it. These will often contain important announcements.

Technology: This is a high-use course for technology – computers, cell phones, etc. are often useful in class for doing computations and accessing the course website. Two words of caution: (1) please remember to put your phone on silent and (2) it really is true that I can tell who is paying attention and who is playing *Angry Birds*.⁸ At least observationally, students who take notes by hand also seem to have a better experience with the course – but I haven’t exactly tested that hypothesis.

⁸ If that is even a thing, these days.

Schedule

Lecture 1 (9/5): Quantitative Methods in Social Science, Descriptive Statistics

Reading	NO
Quiz Due	NO
Homework	NO – <i>but note the amount of reading for Lecture 2.</i>
LAB	YES

This week we broadly discuss the role of quantitative methods in public policy and other social sciences. We review the outline of this course before moving on to descriptive statistics and their use in policy argument. The Friday lab focuses on making sure you have stata loaded correctly on your computer and will briefly touch on key points in the reading that may be helpful for the first quiz.

Lecture 2 (9/12): Central Tendency & Dispersion, Point Estimates, & An Introduction to Distributions

Reading	YES
Required:	Wheelan pp. 1-57, 68-109
Quiz Due	YES
Homework	YES, DUE 9/19 (L2 HW #1)
LAB	YES

This week includes foundational statistical material, including an introduction to the properties of the normal distribution that will be useful throughout this course.

Lecture 3 (9/19): The Sampling Distribution, The Central Limit Theorem & Confidence Intervals

Reading	YES
Required:	Wheelan pp. 110-142
Quiz Due	YES
Homework	YES, DUE 9/26 (L3 HW #2)
LAB	YES

This week we focus on building your statistical intuition. This theoretical material is important for the rest of the course. The bulk of this material is not an end all unto itself (although you should find it interesting); this is the build-up to the material in Lecture 4.

Lecture 4 (9/26): Hypothesis Testing & the t-Test.

Reading	YES
	Required: Wheelan pp. 143-168
	Required: Mastering 'Metrics pp. 1-46
Quiz Due	YES
Homework	YES, DUE 10/3 (L4 HW #3)
LAB	YES

The lecture this week provides a statistical foundation for the most important question a policy analyst will ask: does this policy work? More generally, the hypothesis testing framework provides guidelines for deciding if the characteristics of one group are different than other. The t-test is one type of test (and the one most commonly used as introductory material). More to come starting in the next week.

Lecture 5 (10/3): ANOVA, Chi-Square, and Correlation

Reading	YES
	Required: Wheelan pp. 58-67
	Required: Selection from <i>The Victory Lab</i> .
Quiz Due	YES
Homework	YES, DUE BY 10/10 (L5 HW #4)
LAB	YES

Last week covered one commonly used test, the t-test, which is appropriate for some but not all circumstances. This week more formally categorizes the kinds of relationships we can study and then describes three types of tests appropriate, for differing types of data, for analyzing "bivariate" (between two variables) relationships. Don't worry! We cover more complex relationships beginning next week.

Lecture 6 (10/10): Introduction to Linear Regression

Reading	YES
	Required: Wheelan pp. 185-211
Quiz Due	YES
Homework	YES, DUE BY 10/17 (L6 HW #5)
LAB	NO – it's a grueling semester. Take a day off.

Linear regression is a commonplace tool for understanding the relationship between one thing, or one thing controlling for other things, and some outcome we care about. This week covers the basics: bivariate regression, the problem of omitted variable bias, and multivariate regression. The next week continues the discussion.

Lecture 7 (10/17): Regression Options (Choices About Functional Form)

Reading	YES
Required: Mastering Metrics 47-97.	
Quiz Due	YES
Homework	YES, DUE BY 10/24 (L7 HW #6)
LAB	YES

This continues the discussion on linear regression. You have choices for the way you specify the relationships between the potential explanations and the outcome you have in mind. This lecture reviews some of those options.

Lecture 8 (10/24): Multivariate Regression – Implementation Advice

Reading	YES
Required: Wheelan 212-224.	
Required: Mastering Metrics Ch. 3, pp. 98-146 (IV)	
Quiz Due	YES
Homework	YES, DUE BY 10/31 (L8 HW #7)
LAB	YES

Actually making the decision about what to include in a model, which model to use, and how much faith to put in the results can be quite difficult in practice. This week reviews some of those challenges and introduces the concept of “instrumental variables,” an elegant approach for some types of problems. The weeks that follow also focus on the question of how you appropriately design a real-world study.

Lecture 9 (10/31): Regression Discontinuity Designs

Reading	YES
Required: Mastering Metrics Ch. 4, pp. 147-177 (RDD)	
Quiz Due	YES
Homework	YES, DUE BY 11/7 (L9 HW #8)
LAB	YES

Regression discontinuity designs (RDD) is one commonly-used option to make inferences in less-than-ideal circumstances.

Lecture 10 (11/7): Differences in Differences

Reading	YES
Required: Mastering 'Metrics Ch. 5, pp. 178-208 (DID)	
Quiz Due	YES
Homework	NO
LAB	YES

Similarly to RDD, Diff-in-Diff (DID) is one way to think about setting up your analysis. This is an extremely common approach in policy settings. There is no homework this week – but study hard for the exam!

Lecture 11 (11/14): EXAM

Reading	NO
Quiz Due	NO
Homework	NO
LAB	NO

Your exam will be taken in-class, use the full-time, be “closed book,” and allow you to use only a non-graphing calculator. Students frequently use the entire period to take the exam.

Lecture 12 (11/21): Stata Programming Workshop

Note: Thanksgiving Week

Reading	NO
Quiz Due	NO
Homework	YES, DUE BY 11/28 (L12 HW #9)
LAB	NO

This week focuses on providing you with some additional stata skills that you will need to complete the final project.

Lecture 13 (11/28): Polling & Public Opinion

Reading YES

Required: Articles on polling in the 2016 election on NYU Classes

Quiz Due YES

Homework YES, DUE BY 12/5 (L13 HW #10)

LAB YES (Focused on Projects)

The content for this week will focus on polling and public opinion.

Lecture 14 (12/5): Logit/Probit Models

Note: on 12/12, NYU makes Tuesday into Monday – this is our last class.

Reading YES

Required: Handout on Logit/Probit on NYU Classes

Quiz Due YES

Homework NO

This final lecture will focus on a different type of regression model: models for binary outcomes. This is actually an introduction to a whole class of models for different types of outcomes that are in two or more categories. These models are both *very useful* and *very important* since, as many questions have categorical answers, these are widely used models in social science.

Finishing the Term:

Your final projects are due **12/19/2017 at 11:55pm on NYU Classes**. Note that you can turn these in remotely; I'll plan to hold some office hours on 12/14, if you wish to attend, but otherwise there is no reason you cannot leave town after the 12/5 class (unless you are relying on campus-only resources for Stata).