# 11.220/Quantitative Reasoning and Statistical Methods for Planning
# Syllabus and Orientation Notes

Ezra Haber Glenn, AICP

Spring 2011

## 1 Contact Information

| name | email | office | office hours |
|---|---|---|---|
| Ezra Glenn | eglenn@mit.edu | 7-337 | W9:30-10:30 |
| Marie Victoria del Campo | delcampo@mit.edu | 9-569 | F11:00-12:00 |
| Lauren Lambie-Hanson | lslh@mit.edu | 9-545 | T1:00-2:00 |
| Ning Wu | ningwu@mit.edu | 9-536 | F3:00-4:00 |

Stellar site:
https://stellar.mit.edu/S/course/11/sp11/11.220/

## 2 Overview and Prefatory Remarks

Planners use numbers, and planners use reasoning. The overarching goal of this class—the one thing we absolutely *need* to accomplish together—is to make sure that each and every MCP that graduates from the MIT program is comfortable and skilled at using quantitative information and sound reasoning to address the problems and questions they encounter in planning, design, and policy-making contexts. As with your other classes, we fully expect you to approach this course with the characteristic blend of ambition and skepticism that defines the DUSP approach to planning: that is, we expect you to be *energetic and creative* in your application of the skills you will learn in quantitative reasoning, statistical methods, and the presentation and visualization of complex information, but also to be *critical* of these methods where appropriate,[1] questioning whether the the Modern Age's confidence in statistics—the prevailing faith in "hard numbers," "scientific accuracy," and "dispassionate logic"—may at times be overstated or unjustified.

[1] When this is will be up to you...

## 2.1 A few thoughts on *quantitative* and *qualitative* methods

Much is made of the distinction between quantitative and qualitative approaches. Fortunately, planning is a field where you will be able (and expected) to master both, so we don't need to waste a lot of space here with rhetorical debates about which is "truer", or better, or more persuasive. But for the time being, for this class, we need to concentrate on the quantitative aspects of the field (hence the "Q" in this particular "QR"); we will be concerned with things that can be measured, compared, and analyzed with regard to scale, size, variation, frequency and distribution, degree, and proportion. We will also be concentrating on the differences between observed samples and entire populations, and using statistical tools to distinguish between meaningful differences and random noise.

That said, it may be worth meditating on the possibility that from an *existential* perspective, things in the world are not really either quantitative or qualitative—these words refer to ways we *approach* these things (or events, phenomena, ideas—whatever). To use a hackneyed old example, consider a tree in the forest: it is neither quantitative nor qualitative—it is just there, being treeey. These aspects of the tree only come out in relation (or perhaps in reaction) to our observation and discussion of it: "how tall is it?" "how old?" "is it healthy?" "is it pretty?" "what's it good for?" Depending on how we want to answer these questions, we may chose more quantitative or more qualitative methods. Interestingly, which approach will be more helpful for which questions is not always obvious.

Most of the implications of this line of thinking are far beyond our purposes here, but it does help to point to the artificiality of the distinction, and may help us see the these two types of approaches more as a continuum and less as two unreconcilable world views.

## 2.2 Putting the "R" back in "QR"

The second half of "QR" stands for "Reasoning." As you begin to prepare your mind for the course, please remember that this is not just a statistics course. The quantitative part—dealing with gathering, analysis, and presentation of numbers—is certainly a key aspect of the class. But equally important will be developing skills in *reasoning*: making and critiquing arguments; stating and investigating hypotheses; avoiding bias in your own work and identifying it in the work of others; struggling over the vagaries of cause and effect; learning how to simplify complexity without doing violence to truth; and a host of other challenges that have more to do with logic and clear-thinking than with numbers and data *per se.*

## 2.3 A note about knowledge and belief

In the first few weeks of the class, we will touch on the difference between *knowledge* and *belief*: a belief may be true, but only when it is *justified and*

*explained* can one be said to possess true knowledge.[2] Importantly, the need to insist upon this higher standard is all the more crucial when we are working to develop knowledge of tools (such as logic, quantitative reasoning, research design, and statistical methods), as these can then form a foundation to build further (justified and explained) knowledge; if the foundations are shaky, you will never be able to trust the upper floors.

I would like to challenge all of you to develop actual knowledge concerning the material we cover—that is, to learn the methods that we think work for this or that purpose *and* to understand *why* we use them. Not everyone agrees with this level of intellectual rigor, and to be honest, others may be right: it might be a waste of your time, and you can probably get by fine in planning or even academic research by just treating quantitative methods as a series of recipes to use as dictated by the textbooks and other experts. (Of course, it is *extremely* difficult for an outside observer to know the difference between the two, as rote memorization of facts and blind acceptance of the formulas in the text will still get you the right answers; therefore, whether you choose to pursue true knowledge or just faith in the experts is largely up to you. . . .)

### 2.4 A sappy comment to inspire you

Finally, I should mention that in addition to being useful towards meeting the MCP core requirement and your professional goals, Quantitative Reasoning is also very fun, and can be meaningful from a philosophical perspective. We'll talk about this a little in class, and I hope you find yourself inspired you see the beauty and wonder in it all.

## 3 Schedule and Logistics

### 3.1 Lectures

Lectures take place on Tuesdays and Thursdays, from 9:30-11:00 AM, in Room 3-370. Although it should go without saying, you are expected to attend these lectures. If for some reason you are *unable* to attend, it is your responsibility to figure out what you missed—including changes to the syllabus and assignments, discussions of statistics in current events, or other important announcements.

Beyond this, you are expected to *participate.* Rarely (I hope never) will the entire hour-and-a-half be spent as a lecture—we will have a group discussion, pursue interesting or meaningful sidetracks, listen to guest speakers and student presentations, and even occasionally play games. All of this is part of the course material, and you won't necessarily be able to just get it later from the readings or the web.

[2] Empirical purists such as Karl Popper (Popper, 2002) will be quick to point out that even with "justified and explained" beliefs we might be overstepping our bounds here—all "knowledge" is tentative at best and subject to revision based on future evidence.

## 3.2 Recitation Sections

In addition to the lectures, the course includes a one-hour weekly recitation section, either on Thursday or Friday. Like the regular classes, these too are mandatory (and your active participation will be even more essential in these smaller groups). Some weeks will be spent in review, especially as we get closer to the exams or when we are dealing with the more computational material; other weeks will be more discussion-oriented, as we drill down into closer readings of particular studies or articles.

Sections will meet at follows:

| Section | TA | Day | Time | Room |
|---|---|---|---|---|
| R01 | del Campo | Friday | 10:00-11:00 | 9-450B |
| R02 | Lambie-Hanson | Friday | 11:00-12:00 | 9-450B |
| R03 | Wu | Friday | 1:30-2:30 | 9-450B |

## 3.3 Lab Sections

Four times throughout the semester, we have computer labs instead of the normal weekly lecture, to become familiar working with a statistical software package (either Stata or `R`—see section 4.6 on page 8). We will use software to explore concepts and methods we covered in class a week or two earlier, so you will already be familiar with the basics (and hopefully the introduction of the computers will illuminate—not complicate or obscure—your grasp of the fundamentals).

For students interested in learning STATA, these sections will meet in Room 1-115, and will be taught by staff from the Harvard-MIT Data Center; for `R`, we'll meet in 3-370, and Ezra will teach the labs.

## 3.4 Office Hours

All of us have office hours, as listed on the first page of this handout. If these times do not work for you, we can probably find another time to meet by appointment.

## 3.5 Feedback

Somewhere in here we wanted to be sure to mention that we actually care about your feedback, and ideally would want it during the course so we can identify problems (or opportunities) early and make changes as necessary. *Please* feel free to contact us with issues as they arise, either in person or through email (or even anonymous notes).

### 3.6 Some Required Elements

Although all of this should go without saying, the Institute requires us to say the following:

#### 3.6.1 Accommodation for Disabilities

If you have a documented disability, or any other problem you think may affect your ability to perform in class, please see me early in the semester so that arrangements may be made to accommodate you.

#### 3.6.2 Academic Misconduct

Plagiarism and cheating are both academic crimes. Never (1) turn in an assignment that you did not write yourself, (2) turn in an assignment for this class that you previously turned in for another class, or (3) cheat on an exam. If you do so, it may result in a failing grade for the class, and possibly even suspension from the college. Please see me if you have any questions about what constitutes plagiarism. Anyone caught cheating on an exam will be reported to the provost in line with recognized university procedures.

## 4 Requirements

You need to (a) read all the required readings, (b) complete all the listed written assignments and problems, (c) attend classes and participate in the discussions, (d) pass the exam, and (e) otherwise think about, remember, and learn all the material.

### 4.1 Books and Readings

For every week of the Outline (below) we have listed certain readings. Please do these readings as early as possible in the week for which they are assigned, and come prepared to discuss them.

These readings are all listed in the "References" section (starting on page 18). In general, they can be grouped into the following categories:

#### 4.1.1 Required books

For most of the basic material in the class, we will be using a textbook and two short monographs, all of which are available at the Coop and on reserve at Rotch Library. Please note that the first book covers most of the general "statistical methods" part of the material, and is not that different from many other introductory textbooks. Depending on your interests and learning style you may find it helpful to rely on some other book as your "primary text," although you should be sure to look over this one as well to see how we are ordering and pacing the topics.

1. *Introduction to the Practice of Statistics*, Moore et al. 2009: This is just what it sounds like: a straightforward text covering statistical methods, with many good examples and problems to work through the basics. We will rely on this book for formulas and computational problems, and in class I will generally try to use the same notation and terms that you find there.

   The book is available at the MIT Coop, as well as various on-line vendors; we will be using the *6th edition*, which is not the latest, so you may be able to find used copies. However, please note that the text comes with very extensive on-line support through the publisher's "StatsPortal" site, which is not strictly required for this class, but is necessary if you want to access the chapter quizes (see section 4.3 on the next page) and many other resources. If you buy a used copy, you may want to be sure that you get a valid (and *unused*) access code. (You can also subscribe to StatsPortal online by itself, but it costs almost as much as just buying the book). If you don't want the hard-copy of the text at all, the publisher also provides the option of purchasing an ebook (with full StatsPortal access) for about half the cost of the full book. (We'll go over this more in class, to be sure you understand the options.)

2. *Exploratory Data Analysis*, Hartwig and Dearing 1979: This short book provides a classic introduction to the world of Exploratory Data Analysis (EDA), as pioneered by statistician John Tukey. Rather than rushing to statistical tests and regression analysis, EDA encourages the statistician to explore data first through descriptive summaries, plots, and other visualization techniques, to suggest both problems and relationships. This is a classic, used in many introductory courses, and there are probably many cheap copies on Amazon and elsewhere.

3. *Visual and Statistical Thinking*, Tufte, 1997b: This small pamphlet reprints chapter two from Tufte's larger work, *Visual Explanations: Images and Quantities, Evidence and Narrative*(Tufte, 1997a), so if you already have that, don't bother to buy this. Due to the importance of the visual quality of the material, we are including this as a "required book" (as opposed to simply including it in the reader, which would probably cost you just as much and result in a substandard copy). It is only \$7.00,[3] and is a great intro to the endlessly-fascinating and graphically-stunning world of Edward R. Tufte.

#### 4.1.2 Articles and individual chapters

Beyond these books, some weeks will include additional readings from other sources, handed out in class or provided on Stellar. Sometimes these will be a chapter in a book that covers a topic particularly well, or adds some interesting wrinkles to the standard treatment. Other times it will be an article or news

[3] even cheaper used on Amazon

item that demonstrates a particular concept, or gives a case for us to discuss. Remember: planners (and other professionals) spend very little time reading textbooks—most planning knowledge comes in the form of journal articles, case studies, publications of research findings, government reports, and other such sources. Typically these are less dry than textbooks, but they are also often written with a particular agenda or bias. Learning to read these sources in an open-minded but critical way is a real art, and an important part of a good planning.

#### 4.1.3 Recommended books

The "References" section also includes some recommended books that may be helpful. Some are more basic than the texts we are using (*e.g.* Moore and Notz, 2009; Gonick and Smith, 1993), and some go into more depth on particular topics (*e.g.*, graphs and charts; statistical software packages). Also, many of the required readings represent individual chapters from longer books—these can be good sources for further study (Zeisel, 1985 and the works by Tufte in particular are worth looking at.)

### 4.2 Homework/Problem Sets/Written Assignments

The Outline includes six homework assignments, due roughly every other week. These provide a blend of problems from the textbook, more involved examples using real-world data, and some longer written assignments to challenge you to communicate clearly about data. Please pay close attention to the deadlines, as we will often use them as a starting point for the weekly sections; come prepared to discuss your solutions or ask questions about places you got stuck.

### 4.3 Chapter Tests

Beyond the assignments which must be submitted (and will count for a combined 40% of your final grade), the StatsPortal site provides excellent chapter tests to help you make sure you have mastered the material as you encounter it and help to focus your review on spots you are still confused about. Although these are not required, they are *highly recommended*. In order to understand the material later in the course you will need to build a strong foundation from the basic techniques in the early chapters. These chapter quizzes can help you be sure you know what you need before you move on.

### 4.4 Exam

There will be a midterm exam on 3/31/11. There will not be a final exam.

### 4.5 Final Project: Data Analysis Exercise

For the final project you will pose your own research or policy question, and then gather and analyze data to answer it. Prior to starting you must submit a

clear and thoughtful proposal for approval (due 4/14/11), indicating your topic, the data you intend to use, and a basic outline of the steps you intend to follow. The final paper (6-8 pages, plus a well-documented dataset and scripts/logs from your analysis) is due on 5/12/11.

## 4.6 Computers and Software

Part of working with quantitative data is becoming comfortable using computers for data analysis. In this class and in the readings you will learn the theories and methods behind statistical analysis, and we will demonstrate how one can do these computations "by hand" with small samples; in an actual planning, policy, and social science situations, these computations—even the creation of simple tables and exploratory plots, to say nothing of statistical tests and regression analysis—would be extremely time consuming without the use of a computer and a statistical software package. As part of the class—although not the the primary focus—you are expected to become familiar with the use of a statistical software package.

Our objective in requiring you to use a statistical package stems from the hands-on nature of the class (and the MCP program in general). By the time you finish the class, you should be able to manage large and small datasets, import them into a software package, and produce descriptive summaries, plots, and inferential analyses (including regression). Beyond this, I hope you will apply and hone the more general purpose skills of a planner in regard to computing: striving to be resourceful in identifying both data and technical assistance; becoming comfortable across platforms and settings; understanding and questioning the assumptions inherent in any analysis you encounter or conduct; and interpreting and presenting the results of your work to a general audience.

In order to encourage and reward this resourcefulness, creativity, and skills for critical assessment, we are giving you a choice as to which statistical package to use. Choosing a statistical program is similar to choosing an operating system or word processor (or bicycle, or other tool): it is as much a matter of personal preference as an evaluation of technical merit. You may choose from the following three options:

### Stata

For the past couple years, DUSP has used Stata in this class (before that it was SPSS). Stata provides a good package to learn on without sacrificing the more powerful tools required by advanced users. Nearly everything can be done either through pull-down menus and dialog boxes (helpful for beginners) or through the command line (good for scripting and advanced users who care about setting different options). Graphics and output are good, import in generally straightforward (with a few quirks around missing values). The package is well-documented and well-supported, two real strengths for beginners. Stata is already installed on all the computers in the CRN, and you can buy your own copy for $95 for a one-year academic license or $155 for the perpetual license

($48 for the one-year "Small Stata" version, but this is very limited).[4] (For more information, visit `http://www.stata.com/order/new/edu/gradplans/gp-campus.html`; the bookstore may also have copies.)

If you decide to use Stata, you will probably want to look at the online documentation, as well as the manual it comes with (Stata Corporation, 2007); Acock (2008) is a more comprehensive introduction, but it is a little pricey.

### R

`R` is the open source version of the commercial `S` and `S+` programming language used by many professional statisticians and applied scientists. It is available under the GNU license, with hundreds of additional packages and extensions for free download via the `CRAN` website (`http://cran.r-project.org/`). It provides a fully-featured statistical package for data management and manipulation, statistical analysis, scripting, simulation, and graphing (the latter being a real strength). The interface is a little different if you are not used to using the command line for this sort of thing, although there are graphical user interfaces available as well ("GUIs"—windows with menus and that sort of thing). `R` is also installed on all the CRN computers, if you want to check it out.[5]

If you decide to use `R`, you will probably want to look at the online documentation; Verzani (2004) provides a more thorough introduction, ideal for our context.

### Something else

Many social scientists do not use either Stata or `R`—these have been selected as good (and affordable) options to learn on. If you have access to SPSS or SAS, and want to try to work on that, you are free to do so. (Note that Excel is *not* a suitable alternative—see section 5.2 on the following page.) The only caveat is that we cannot promise to support all of these options; the lab sections will only deal with Stata and `R`, and these programs are the only ones installed in the CRN. Different TAs (or people around the department) may be able to help, but we can't promise it.

## 4.7 Grading

Your grade for the class will based on the following allocation:

[4] The real downside of learning on Stata is that it is less common in the workplace of planning, and you may have difficulty convincing a new employer to purchase a significantly-more-expensive government or private license.

[5] `R` is not very common in government planning offices either, but it is free, so you can bring it with you wherever you go.

| Component | % of final grade |
|---|---|
| Participation (lectures, sections, and timely preparation of readings and assignments) | 10% |
| Assignment #1 | 5% |
| Assignment #2 | 5% |
| Assignment #3 | 10% |
| Assignment #4 | 5% |
| Assignment #5 | 5% |
| Assignment #6 | 10% |
| Midterm Exam | 25% |
| Final Project | 25% |
| Total | 100% |

# 5 FAQs

## 5.1 What if I can't make [a deadline/an exam/etc.]

We are handing out this course schedule now, at the very beginning of the term. Please look at it and note the dates for classes, labs, and the exam, as well as deadlines for the assignments and the final project. If for some reason you cannot meet these dates and deadlines, it is your responsibility to contact us as far in advance as possible to make alternate arrangements. Extensions are generally fairly easy to grant with good cause and advance warning; extensions at the last minute (or after the deadline) are much harder to justify.

That said, we know that sometimes life is unpredictable, and things happen; in order to accommodate this uncertainty, each student will be allowed one free unpenalized late submission of up to three calendar days (this may not be used for the exam or the final project). Beyond this, assignments submitted after the deadline without prior approval will suffer a 5% cumulative penalty for every day they are late. Also, please note that one component of your participation grade is for "timely preparation of readings and assignments," so late homework may affect your grade in more ways than one.

## 5.2 Why can't I just use Excel®?

This question always comes up. Excel is a great program for assembling, viewing, and limited manipulation of spreadsheets.[6] Where Excel fails is where it is pushed too far; this tends to be Microsoft's biggest mistake, in my book: they want to make a single giant tool that does everything for everyone and controls 100% of your time (and the market). In contrast, statistical packages (such as those listed in section 4.6 on page 8) have been designed around functions

[6] Although open source options are of course available for free that are just as good and leave a better taste in your mouth about being part of an open and free society....

specific to the tasks at hand: managing large data sets; conducting complex statistical analyses (with options to control for different assumptions and technical corrections); producing high quality graphical and tabular output; and logging all inputs, commands, options, settings, manipulations, and outputs (necessary to replicate results for true scientific peer review, as well as useful for scripting and "batch" processing of routine projects).

In the same way that most of you would probably choke if asked to use Word to create websites or Publisher to create GIS maps, trying to use a spreadsheet for statistical analysis is overextending the technology, resulting in more work for you and a worse product in the end. You came here to expose yourself to new things and learn about planning technology; here is your chance.

# 6 Course Outline

**Week 1: Introductions (to Each Other and to Core Concepts)**

**2/1/11** Course overview; epistemological foundations; math review

**2/3/11** What's in a number?; basic numeracy; measurement; numbers for planning

Readings:[7]

- "On Number Numbness", Hofstadter, 1985
- "Use and Mis-Use of Measurement Scales in City Planning", Hodge, 1963
- "Sagas of three indicators", Innis, 1995

Problems: Short reflection assignment (ungraded; due 2/4/11).

**Week 2: Descriptive Statistics; Numerical Summaries**

**2/8/11** Measures of center and spread

**2/10/11** Computer Lab #1: Introduction to STATA/Introduction to `R`

Readings:

- Moore et al., ch. 1 and 2
- "The Political Properties of Crystalline $H_2O$: Planning for Snow Emergencies in New York", Savas, 1973

Problems: Assignment #1 (due 2/10/11).

**Week 3: Special Focus: The Census**

**2/15/11** Census I

**2/17/11** Census II

- Moore et al., ch. 3
- "A Compass for Understanding and Using American Community Survey Data: What State and Local Governments Need to Know". U.S. Census Bureau, 2009
- "The American Community Survey: Warmer (More Current), but Fuzzier (Less Precise) Than the Decennial Census", MacDonald, 2006
- "In which direction should percents be run?", Zeisel, 1985, ch. 3

Problems: None.

**Week 4: Probability**

**2/22/11** No class (Monday schedule)

**2/24/11** Probability

Readings:

- Moore et al., ch. 4 and 5

Problems: Assignment #2 (due 2/22/11).

**Week 5: Introduction to Inference**

**3/1/11** Sampling Distributions

**3/3/11** Hypotheses and Statistical Tests

Readings:

- Moore et al., ch. 6

Problems: None.

**Week 6: Estimating from a Sample**

**3/8/11** Estimates and Confidence Intervals

**3/10/11** Computer Lab #2: Data Management

Readings:

- Moore et al., ch. 7
- "How to Handle the 'Don't Knows' and 'No Answers"', Zeisel, 1985, ch. 4

Problems: Assignment #3 (due 3/8/11).

**Week 7: Tests of Means and Proportion**

**3/15/11** Tests of Means; Guest Speaker Katherine McNeill on Data Sources

**3/17/11** Tests of Proportions

Readings:

- Moore et al., ch. 8

Problems: Assignment #4 (due 3/17/11).

**Week 8: Spring Break**

**3/22/11** No class

**3/24/11** No class

Readings:

- None.

Problems: Nothing due, but maybe review for exam, and think some upcoming proposal for final project (due 4/14/11).

**Week 9: Review and Exam**

**3/29/11** Review

**3/31/11** Midterm Exam

Readings:

- None.

Problems: Just review, take the exam, and rest.

**Week 10: More Than One Variable**

**4/5/11** $\chi^2$ tests

**4/7/11** Computer Lab #3: Graphics

Readings:

- Moore et al., ch. 9
- "Accidents on Route 2: Two-Way Structures for Data", (Fairley, 1977)
- By this time you should be working through one of the following:
  - Stata Corporation, 2007 (free with purchase)
  - Acock, 2008 (better, more expensive)
  - Verzani, 2004 (for using `R`)
  - One of the on-line tutorials found on the course website/Stellar

Problems: None.

**Week 11: Introduction to Regression**

**4/12/11** Regression

**4/14/11** Guest speaker (Joe Flood, author of *The Fires: How a Computer Formula, Big Ideas, and the Best of Intentions Burned Down New York City-and Determined the Future of Cities*

Readings:

- Moore et al., ch. 10

Problems: Proposals for final project due (4/14/11).

**Week 12: More Regression**

**4/19/11** More regression

**4/21/11** Computer Lab #4: Regression

Readings:

- Moore et al., ch. 11

Problems: Assignment #5 (due 4/21/11).

**Week 13: Brief Introduction to Multivariate Regression**

**4/26/11** Patriot's Day Holiday ("Go, Patriots!")

**4/28/11** Multivariate regression; examples

Readings:

- Skim all articles for Assignment #6

Problems: None.

**Week 14: Special Focus: Advanced Techniques**

**5/3/11** Visualization and exploratory data analysis

**5/5/11** Projection and simulation

Readings:

- *Exploratory Data Analysis*, Hartwig and Dearing, 1979 (O.K. to skim)
- *Visual and Statistical Thinking*, Tufte, 1997b
- "John Wilder Tukey: The Father of Twenty-First-Century Graphical Display," Wainer, 2005, ch. 19

Problems: Assignment #6 (due 5/3/11).

**Week 15: Special Focus: Dollars; Decisions**

**5/10/11** Talking about money

**5/12/11** Decision-making under conditions of uncertainty

Readings:

- "Land Values and the Measurement of Highway Benefits", Mohring, 1961
- "Creating a Transparent Budget", `http://www.massbudget.org/file_storage/documents/Creating_a_Transparent_Budget.pdf`
- *Rational Choice in an Uncertain World*, Dawes (1988, ch. 2, 6)

Problems: Final Project (due 5/12/11).

# References


Alan C. Acock. *A Gentle Introduction to Stata*. Stata Press books. StataCorp LP, 2nd edition, 2008. Recommended only; on reserve at Rotch. If you are using Stata, you can probably get by with just the *Getting Started with Stata* book that ships with the program, or one of the on-line tutorials, but if you want a little more, or you expect to continue using Stata for other classes or work, then this book would not be a bad thing to buy.

William D. Berry and Mitchell S. Sanders. *Understanding Multivariate Research: a primer for beginning social scientists*. Westview, 2000. Unfortunately this book costs $30 for a pretty short intro, but I think it is very good, and it presents some nice complex examples to demonstrate the points it is trying to get across. This book was written with exactly this sort of audience in mind—students in the first year of a program who may or may not go on to more advanced courses, but need to get a firm grasp of multivariate regression fast, to be able to make sense of research presented in other substantive courses.

James A. Davis. *The Logic Of Causal Order*. Sage Publications, Inc., November 1985. Part of the excellent Sage monograph series in Quantitative Applications in the Social Sciences.

Robyn M. Dawes. *Rational Choice in an Uncertain World*, chapter 2, 6, pages 21–31, 91–125. Harcourt Brace Jovanovich, 1988. In course reader.

William B. Fairley. Accidents on Route 2: Two-way structures for data. In Willaim B. Fairley and Frederick Mosteller, editors, *Statistics and Public Policy*, pages 23–50. Addison-Welsley, 1977. In course reader.

Stephen Few. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, 2009.

Larry Gonick and Woollcott Smith. *The Cartoon Guide to Statistics*. Collins Reference, 1993. Recommended only.

Frederick Hartwig and Brian E. Dearing. *Exploratory Data Analysis*. Sage Publications, Inc., 1979. Part of the excellent Sage monograph series in Quantitative Applications in the Social Sciences; required.

Gerald Hodge. Use and mis-use of measurement scales in city planning. *Journal of the American Planning Association*, 29(2):112–121, 1963. URL `http://www.informaworld.com/10.1080/01944366308978048`. On Stellar.

Douglas R. Hofstadter. On number numbness. In *Metamagical Themas: Questing for the Essence of Mind and Pattern*, chapter 6, pages 115–135. Basic Books, 1985. In course reader.

Lucy Horwitz and Lou Ferleger. *Statistics for Social Change*. South End Press, 1980.

Judith Innis. Sagas of three indicators. In *Knowledge and Public Policy: The Search for Meaningful Indicators*, chapter 6, pages 122–171. Transaction Publishers, 1995. In course reader.

Heather MacDonald. The American Community Survey: Warmer (more current), but fuzzier (less precise) than the decennial census. *Journal of the American Planning Association*, 72(4):491–504, 2006. URL `http://proquest.umi.com/pqdweb?did=1174434591&Fmt=7&clientId=5482&RQT=309&VName=PQD`. On Stellar.

Kenneth J. Meier, Jeffrey L. Brudney, and John Bohte. *Applied Statistics for Public and Nonprofit Administration*. Thomson Wadsworth, 7th edition, 2009.

Herbert Mohring. Land values and the measurement of highway benefits. *The Journal of Political Economy*, 69(3):236–249, June 1961. ISSN 00223808. URL `http://www.jstor.org.libproxy.mit.edu/stable/1829265`. On Stellar.

David S. Moore and William I. Notz. *Statistics: Concepts and Controversies*. W. H. Freeman, 7th edition, 2009. Recommended only, but a really fun introduction at a slightly lower level than we needed.

David S. Moore, George P. MCCabe, and Bruce A. Craig. *Introduction to the Practice of Statistics*. W. H. Freeman and Company, 6th edition, 2009. Required; also online: see `http://courses.bfwpub.com/ips6e.php`.

Karl R. Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*, chapter 10, pages 291–341. Routledge, 5th edition, 2002. In course reader.

E. S. Savas. The political properties of crystalline $H_2O$: Planning for snow emergencies in New York. *Management Science*, 20(2):137–145, October 1973. doi: 10.1287/mnsc.20.2.137. URL `http://mansci.journal.informs.org/cgi/content/abstract/20/2/137`. On Stellar.

Stata Corporation. *Getting Started with Stata for [Unix/Windows/Macintosh]*. Stata Press, 2007. Recommended only.

Edward R. Tufte. *Visual Explanations: Images and Quantities, Evidence and Narrative*, chapter 2, pages 27–53. Graphics Press, 1997a. Recommended only; chapter two is reproduced in *Visual and Statistical Thinking: Displays of Evidence for Making Decisions*, which is requird.

Edward R Tufte. *Visual and Statistical Thinking: Displays of Evidence for Making Decisions*. Graphic Press, Cheshire, CT, 1997b. Required.

U.S. Census Bureau. *A Compass for Understanding and Using American Community Survey Data: What State and Local Governments Need to Know*, 2009. URL `http://www.census.gov/acs/www/Downloads/handbooks/ACSstateLocal.pdf`.

John Verzani. *Using R for Introductory Statistics*. Chapman & Hall/CRC, 1st edition, 2004. Recommended only, for those students interested in using `R` instead on Stata.

Howard Wainer. *Graphic Discovery: A Trout in the Milk and Other Visual Adventures*. Princeton University Press, 2005.

Hans Zeisel. *Say It With Figures*, chapter 3–4, pages 34–64. Harper & Row, 6th edition, 1985. In course reader; note: this is a truly great classic in the field, and definitely worth owning, even if we don't get to read the whole thing. You can probably find an old copy on Amazon or elsewhere, but be warned that not all editions have the chapter on Regression Analysis (the 1985/sixth edition does).