

# API-201Z

## Quantitative Analysis & Empirical Methods

Tuesdays and Thursdays 1:15–2:30pm  
Review Section: Fridays 10:15-11:30am (L280)

**Maya Sen**, HKS

Email: [maya\\_sen@hks.harvard.edu](mailto:maya_sen@hks.harvard.edu)

Office Hours: Tuesdays, after class to 4:30pm (Taubman-356)

**Pamela Ban**, HKS/Department of Government

Email: [pban@fas.harvard.edu](mailto:pban@fas.harvard.edu)

Office Hours: Fridays, 11:30am-1:30pm (location TBA)

## Overview of API 201

API-201 introduces a range of quantitative tools that are commonly used to inform public policy issues. It provides an introduction to descriptive statistics, probability theory, statistical inference, and decision analysis, with an emphasis on the ways in which they are applied to practical policy questions.

Our goal is that by the end of this course you will be able to:

1. Take a data set and a broad descriptive policy question (such as “what has happened to incomes in the U.S. in the last 30 years?”), figure out what statistical analysis would be most appropriate to answer the question, conduct such an analysis, identify what are the most salient findings/patterns that emerge from the data, and present the findings in a way that is accessible to policymakers.
2. Identify real world policy situations in which the tools of probability can be used, identify which tools are most relevant to inform courses of action in those real world situations, and critically consume policy analysis in which probability is used.
3. Critically consume policy studies/papers/reports in which statistical analysis is used.

4. Use the decision analysis framework as one tool to make personal and professional decisions, and to think about policy problems.

The course content is divided into four broad units: (1) Descriptive Statistics, (2) Probability, (3) Statistical Inference, and (4) Regression. Section Z of the course also provides you with an opportunity to become proficient in the use of **Stata** (or **R**, if you so choose) as tools to analyze quantitative data.

API-201 is open to all Kennedy School students and is required for first-year MPPs. Students from other graduate programs may enroll with the permission of the instructor.

**How does the Z Section differ from the A-B-C Sections?** Section Z will follow a similar syllabus to Sections A, B, and C, but assumes a greater comfort level with mathematics and a stronger interest in learning how to work with statistical software. It will proceed somewhat faster than the other sections, allowing more time for applications, in-depth discussions, and a number of advanced topics. The problem sets will be longer and more difficult than those of the other sections. Section Z's final exercise is also open ended, allowing student groups to write on a topic of their choice.

Section Z will also give you the option of using two statistical software packages, **Stata** or **R**, to conduct data analysis throughout the semester, whereas the A-B-C Sections will use Excel. Please note that in order to enroll in Section Z of API-202 in Spring, either an A or A- will be required in either Sections A, B, C, or Z of API-201. Students are encouraged to talk with the API 201 course head to determine which section is the best fit.

First-year MPP students with prior coursework in statistics are strongly encouraged to place out of the course entirely by taking an exemption exam during orientation week. Students may also switch from the Z Section to the A-B-C Sections any time from the beginning of the semester *up to no later than one week after the first midterm*, after consultation with faculty.

## Logistics & Requirements

**Prerequisites.** There are no prerequisites for this course. Some of the problem sets will be time consuming and be quite challenging, particularly if you have never worked with statistics before. Please give yourself ample time to work through them and consider working in a group. Statistical expertise is not a prerequisite, but you will still be expected to come to class ready to learn and challenge yourself.

**Class Attendance.** We expect that you will arrive on time to class and do your best to attend every class. If you need to miss a class due to an emergency, it is your responsibility to obtain missed notes and course announcements from another student; there is no need to email the instructors. All lecture notes will be posted on the course website.

**Laptop and Cell Phone Policy.** Responsible use of laptops is permitted in class. Please be considerate and silence and put away cell or mobile phones and smartphones during class.

**Suggested Textbooks.** The study of statistics is rapidly moving to online platforms—many of which provide free and excellent overviews of all of the material we cover in this class. In addition, all course lecture notes will be posted to the course website. However, along with API 201 A-B-C, there is one strongly suggested textbook:

- *Introduction to the Practice of Statistics.* Moore, McCabe and Craig. W.H. Freeman, 8th Edition. ISBN: 1464158932.

This text can easily be rented, bought used, or bought new, all online. Unlike students in the A-B-C Sections, *there is no need for Z Section students to purchase any of the online supplementary subscriptions.*

In addition, we strongly recommend the purchase of:

- *The Cartoon Guide to Statistics.* Gonick and Smith. Harper Perennial. ISBN: 0062731025.

The *Cartoon Guide* provides an intuitive, easy-to-follow overview of every single topic we cover in this course. It is strongly recommended, particularly for those new to statistics.

**Statistical Software Packages.** One of the objectives of the course is to help you gain proficiency using statistical software packages. Section Z will use **Stata** (and **R**, as an option) to conduct statistical analysis throughout the course. Problem sets will contain exercises designed to get you to practice the basics of **Stata** (or **R**, if you so choose). Tutorials will be offered during the first weeks of class to introduce Stata, and assistance provided throughout the semester by the teaching fellow and course assistants. Only **Stata** will be officially supported by the teaching fellow and course assistants, although both the teaching fellow and head course instructor can answer questions about the use of **R**. Students concerned about or unable to invest the time and effort needed to learn a statistical software language (such as **Stata** or **R**) should consider opting into the A-B-C Sections instead.

## Grading

Your grade in this class will be composed of

- 10% - Problem sets
- 15% - Class participation and engagement – see below
- 15% - Final Exercise
- 15% - Midterm #1

- 15% - Midterm #2
- 30% - Final exam

Final letter grades will be determined on a curve using the Dean's Recommended Grade Distribution (available at [goo.gl/HdLE76](http://goo.gl/HdLE76)).

**Problem Sets.** The best way to learn statistics is to practice. Thus, problem sets will be assigned almost every week and *due Tuesdays at the beginning of class*. There will be no credit for late assignments. You should plan to spend approximately 6-8 hours on each problem set. Problem sets will be posted on the course website approximately 10 days before they are due.

Under the Kennedy School Academic Code ([goo.gl/jpQIoS](http://goo.gl/jpQIoS)), the problem sets for this course are "Type II" assignments. *You are encouraged to work in a study group, but must submit your own hand- or type-written solutions. It is not acceptable to work on one electronic document as a group and submit identical, or nearly identical versions.* Examples of assignments that are not in accordance with the HKS academic code include photocopies or reprints of substantially identical assignments, printouts of substantially identical Stata or R tables or graphs, and copies of solutions from previous years. Problem sets are due at the beginning of class on the date they are due. Please indicate on the first page the names of the students you worked with.

The lowest problem set grade will be automatically dropped in the calculation of your overall problem set grade.

**Class participation and engagement.** Student participation can substantially enrich the learning experience for both the students and the instructor. In this spirit, class participation is definitely encouraged. However, class participation is notoriously difficult to assess subjectively and subject to potential bias. Therefore, class participation will only help, rather than hurt, students, and will be used primarily to bump up students between two grade cut-offs. For example, a student with demonstrated strong participation who is otherwise on the bubble between a B+ and an A- will be bumped up to an A-, while a student who lacks that participation will remain at a B+.

**Final Exercise.** The final exercise is a group-based project that engages all of the skills acquired in the class by asking you to take the statistical skills learned and apply them to the analysis of some original, interesting, or otherwise professionally relevant data set.

Groups should be 3-4 students total. The components of the final exercise include (1) a short write-up of your analyses and results, and (2) a short power-point presentation of the key points.

The final exercise is due on December 1, 2015. More details and deadlines will be provided later in the course. Students are encouraged to start thinking about their groups and also about potential data sources as early in the term as possible.

**Midterm Exams.** There will be two midterms, and a final exam. These will be closed book/notes and a formula sheet will be provided. Calculators may be used, but statistical functions on them may not. Calculators that allow text storage are not permitted.

*The two midterms will be given in class on Thursday, October 8, and Thursday, November 12.*

Please schedule your travel plans accordingly and arrange to be in the classroom on time. We will adhere to the HKS Registrar policy regarding rescheduling of exams (i.e., only to be done in case of documented health-related or personal emergencies).

**Final Exam.** Our exam has been scheduled by the HKS Registrar on **Friday, December 11, from 3 to 6pm**. We have no discretion to change the time or date of the final exam. Schedule your travel plans accordingly.

**Exam Regrade Policy.** Requests for regrades will be accepted only in writing, with a clear statement of what has been mis-graded, and why, and within one week of return of your graded exam. Note that we reserve the right to review all of your answers, and you might end up having more points deducted than if you hadn't requested a regrade.

All course activities, including class meetings, homework assignments, and exams are subject to the HKS Academic Code and Code of Conduct.

## Tentative Schedule

This schedule is subject to change, depending on how far we get in each class and how much discussion the topics generate.

**September 3 (Lecture 1):** Introduction and Class Overview

**September 8 (Lecture 2):** Descriptive Statistics

- MMC 1.3, 2.3, 2.7

**September 10 (Lecture 3):** Probability I

- MMC 4.1, 4.2, 4.5

**September 15 (Lecture 4):** Probability II

- MMC 4.5
- Problem Set #1 Due

**September 17 (Lecture 5):** Bayes' Rule

- Article on monograms

**September 22 (Lecture 6):** Application: Public Pensions in Mexico

- Case study on Public Pensions in Mexico
- Problem Set #2 Due

**September 24 (Lecture 7):** Discrete Distributions

- MMC 4.3 (through page 255), 4.4, 5.2

**September 29 (Lecture 8):** Continuous Distributions

- MMC 1.3, 4.3 (from page 255 onward)
- Problem Set #3 Due

**October 1 (Lecture 9):** Introduction to Decision Analysis

- Case study

**October 6 (Lecture 10):** Sampling

- Readings TBA
- Problem Set #4 Due

**October 8** Midterm Review/Midterm

**October 13 (Lecture 11):** Introduction to Inference and the Central Limit Theorem

- MMC 3.3, 5.1
- Problem Set #5 Due

**October 15 (Lecture 12):** Inference with Means I

- MMC 6.1, 6.2, 7.1

**October 20 (Lecture 13):** Inference for Means II

- MMC 7.2, 8.2, 6.3
- Problem Set #6 Due

**October 22 (Lecture 14):** Inference for Proportions I

- MMC 8.1

**October 27 (Lecture 15):** Inference for Proportions II

- MMC 8.2, 6.4
- Problem Set #7 Due

**October 29 (Lecture 16):** Paired Data

- MMC 7.1, 8.2 (review)

**November 3 (Lecture 17):** One-ANOVA

- MMC 12.1, 12.2
- Problem Set #8 Due

**November 5 (Lecture 18):** Application: Oregon Health Care Experiment

- Oregon Health Care experiment case study

**November 10 (Lecture 19):** Chi-Square Tests

- MMC 9.1, 9.2, 9.3
- Problem Set #9 Due

**November 12 (Lecture 19):** Midterm

**November 17 (Lecture 20):** Regression Analysis I

- MMC 2.3, 2.4, 10.1 (p. 563–580)
- Problem Set #10 Due

**November 19 (Lecture 21):** Regression Analysis II

- MMC 11.1, 11.2

**November 24 (Lecture 22):** Regression Analysis III

- 6.3
- Problem Set #11 Due

**November 26:** THANKSGIVING HOLIDAY RECESS – NO CLASS

**December 1:** Final Presentations/Final Exercises Due

**December 3:** Final Exam Review

**December 11:** Final Exam