

# Statistical Programming Camp Spring 2018

Sunday, January 28 – Friday, February 2

Morning Session: 10:00 AM – 12:00 PM

Afternoon Session: 1:00 – 3:00 PM

Location (TBA)

---

<b>Instructors:</b>	Naoki Egami & Diana Stanescu & Erik H. Wang
Office:	Corwin 026 (Naoki, Diana), Corwin 025 (Erik)
Email:	Naoki: negami@princeton.edu Diana: stanescu@princeton.edu Erik: haixiaow@princeton.edu
Office Hours:	3:30 – 4:30 PM (Corwin 023)
<b>Faculty advisor:</b>	Kosuke Imai

---

**Description:** This camp will prepare students for POL 572 and other quantitative analysis courses offered in the Politics department and elsewhere. Although participation in this camp is completely voluntary, the materials covered in this camp are a pre-requisite for POL 572. Students will learn the basics of statistical programming using **R**, an open-source computing environment. Using data from published journal articles, students will learn how to manipulate data, create graphs and tables, and conduct basic statistical analysis. This camp assumes knowledge of probability and statistics as covered in POL 571.

**Structure:** The camp will meet for five and a half days starting with an afternoon session on Sunday and continuing with two daily sessions (morning and afternoon) each following day. Each session lasts two hours with an hour break between. The first hour of the morning session is lecture-based; the other half will be devoted to instructors and students working on programming exercises together. In the afternoon session, students will assemble in their assigned groups to solve a daily problem set, and instructors will be present to help with questions. Students who finish early are free to leave the day's afternoon session. After the lecture, students are required to go over the assigned readings for the following day. We may devote more or less time to lecture depending on the day's materials.

We will start with the afternoon Sunday session and will finish on Friday afternoon. The content is organized into 6 modules, each containing sessions that cover material students will need for the module's problem set. In the afternoon session, we will also take a short amount of time to review the graded problem set from the previous module.

**Discussion Board:** In addition to precepts and office hours, please use the **Piazza** Discussion Board at <https://piazza.com/> when asking questions about lectures, problem sets, and other course materials. This allows all students to benefit from the discussion and to help each other understand the materials. Both students and instructors are encouraged to participate in discussions and answer any questions that are posted.

To join the Piazza site, click on "Search Your Classes" from the Piazza homepage. After specifying Princeton University as your school, search for "Statistical Programming Camp."

You will then be prompted to enter your `princeton.edu` email address to confirm your registration. Piazza can also be accessed from within Blackboard by going to the Programming Camp course page and clicking on the link to “Piazza Messageboard.” In addition, all class announcements will be made through Piazza. Blackboard will still be used for hosting all class materials.

Some useful tips for Piazza include:

- Piazza has apps available for the iOS and Android platforms. The apps are free downloads and provide complete access to all of Piazza’s message board features.
- To insert  $\LaTeX$ -formatted text in a post, place a double dollar sign ( $\$$ ) on both ends of the relevant text, or click the  $fx$  button in the Details toolbar above your post.
- To add formatted **R** code to a post, click the “pre” button in the Details toolbar above your post. A grey text box will open up where you can paste code from **R**.
- You can classify a post using pre-selected tags, or you can generate your own by prepending a hash ( $\#$ ) to your chosen label. Posts can then be sorted by these tags using the search bar in the left-hand column.

**Materials and Website:** Students are encouraged to bring their personal laptop to each session. The required textbook for the course is:

Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*, Princeton University Press (QSS).

**Assignments:** The only way to learn statistics is by doing. To ensure steady and efficient learning, we assign daily problem sets and a final exam. The final exam and problems sets will be assessed and will count towards a final grade with the following weights:

Problem Sets: 70 % (4 equally weighted assignments, completed with assigned group)  
Final Exam: 30 % (Individual, no collaboration)

We ask you to submit your solutions to the problem sets in the appropriate folders at Statistical Programming Camp Blackboard by 12 AM. The final exam will be due electronically by 10 pm on Saturday, February 3<sup>th</sup>.

Students must complete the required reading before each day’s session. For the first module, students must have read Chapter 1 in the required textbook, completed exercise 1.4 at the end of the chapter, and registered for Piazza. The first chapter provides students with a guide to install **R**, and helps them get familiarized with the programming language.

**Group-based Learning:** To promote learning and collaboration, students are assigned to a group of three students. Groups are required to work together on in-class problem sets, and all group members should contribute equally to all assignments. Students hand in problem sets, as a group, with the name of each group member on the assignment.

Groups have two options when completing problem sets. Group members can work together, from start to finish, or work individually on the entire problem set, then meet as a group to

discuss the solutions and write the final submission. Dividing the work among group members is not allowed, since the goal of problem sets is to ensure every individual learns the material. The final grade is based on both the group problem set score and your individual performance on the final exam. There is to be no collaboration between groups, aside from public posts on Piazza.

All exercises are available at the end of each of the module's required chapter. Additional class materials, including solutions to graded exercises, will be made available through the Statistical Programming Camp Blackboard site under the Course Materials link.

### **Camp Outline:**

#### **Module 1** (Sunday Afternoon)

Topic: Introduction to **R**

Required Reading: Chapter 1

#### **Module 2** (Monday Morning/Afternoon)

Topic: Causality

Required Reading: Chapter 2

#### **Module 3** (Tuesday Morning/Afternoon)

Topic: Measurement

Required Reading: Chapter 3

#### **Module 4** (Wednesday Morning/Afternoon)

Topic: Prediction

Required Reading: Chapter 4

#### **Module 5** (Thursday Morning/Afternoon)

Topic: Probability

Required Reading: Chapter 6

#### **Module 6** (Friday Morning/Afternoon)

Topic: Uncertainty

Required Reading: Chapter 7

Lunch Party: 12.00–1.00 pm