Statistical Programming Camp
Spring 2013

Sunday, January 27 – Friday, February 1
Morning Session: 10:00 AM – 12:00 PM
Afternoon Session: 1:30 – 3:30 PM
Wallace 300

Instructors:	Alex Ruder & Ben Fifield
Office:	Robertson 318 (Alex); Robertson 322 (Ben)
Email:	Alex: aruder@princeton.edu
	Ben: bfifield@princeton.edu
Office Hours:	5:00 – 6:00 PM (Corwin 126)
Faculty adviser:	Kosuke Imai

Description: This camp will prepare students for POL 572 and other quantitative analysis courses offered in the politics department and elsewhere. Although the 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 the open-source computing environment called **R**. In addition to the traditional data sets and models, **R** can also handle many kinds of new data including GIS and text data. 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 the knowledge of probability and statistics covered in POL 571.

Structure: The camp will meet for six days with single sessions on Sunday and Friday and two daily sessions (morning and afternoon) on Monday through Thursday. Each session lasts two hours with an hour and a half break between. We will start with the afternoon Sunday session and will finish with the Friday morning session (followed by the pizza lunch party!). The content is organized into 6 modules, each containing sessions for coverage of 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 module's graded problem set.

Discussion Board: We will be using the Piazza discussion board (`https://piazza.com/`) to facilitate discussions and questions throughout the Statistical Programming Camp. Piazza provides an interactive environment in which to both ask questions and answer those of others. To join the Programming Camp 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 and for submitting assignments.

Some tips and tricks 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 messageboard 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.

- We encourage you to mark helpful contributions (particularly those from classmates) using the "Thanks!" button at the bottom of each post.

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 count towards a final grade with the following weights:

Problem Sets: 70 % (5 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 Software Camp Blackboard by 12 AM. The final exam will be due electronically by 5 pm on Saturday February 2nd. We will also issue a preliminary exercise due Sunday January 27th at 12 AM to guide you in installing **R** and loading a data set.

To promote group learning and collaboration, students are assigned to a group of three students. Groups are required to work together on take-home 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. The final grade is based on the group problem set score and the individual performance on the final exam. There is to be no collaboration between groups, aside from public posts on Piazza. Furthermore, all in-class exercises are to be worked on individually.

In addition, we will be distributing a "Getting Started" problem set due by the start of class on January 27th. This problem set will guide you in installing **R** and introduce you to the Piazza messageboard. This assignment is to be completed and submitted individually.

At the end of the course, as a reward for hard work and mastery of the materials, the top three groups will be awarded a **prize**, which is provided by the Program for Quantitative and Analytical Political Science (`http://q-aps.princeton.edu`). The Department has kindly agreed to buy pizza for the whole class to celebrate the end of camp on Friday, and we will announce the winning groups then.

Materials and Website: Students are encouraged to bring their personal laptop to each session. There are no required textbooks for the course. As an optional textbook, we recommend the following book, which can be purchased at Labyrinth on Nassau Street or elsewhere,

Fox, J. and Weisberg, S. (2010). *An* **R** *Companion to Applied Regression.* 2nd ed. Sage.

Handouts, practice exercises, problem sets, and other course materials will be made available through the Statistical Programming Camp Blackboard site under the Course Materials link.

Camp Outline:

Module 1 (Sunday Afternoon)
        Leader: Ben
        Topic: Introduction to **R**
        Homework 1 due 12 AM, Monday

Module 2 (Monday Morning/Afternoon)
        Leader: Alex
        Topic: Data Manipulation and Summarizing Univariate Data
        Homework 2 due 12 AM, Tuesday

Module 3 (Tuesday Morning/Afternoon)
        Leader: Ben
        Topic: Summarizing Bivariate Data, Loops and Conditional Statements
        Homework 3 due 12 AM, Wednesday

Module 4 (Wednesday Morning/Afternoon)
        Leader: Ben
        Topic: Loops and Conditional Statements, Functions, Probability and Simulations
        Homework 4 due 12 AM, Thursday

Module 5 (Thursday Morning/Afternoon)
        Leader: Alex
        Topic: Point Estimation, Confidence Intervals, and Statistical Tests
        Homework 5 due 12 AM, Friday

Module 6 (Friday Morning)
        Leader: Alex
        Topic: Analysis of Experimental Data and Regression
        Pizza party after class and prize winners announced
        Final exam due 5 PM, Saturday