Statistical Programming Camp
Spring 2014

Monday, January 27 – Friday, January 31
Morning Session: 10:00 AM – 12:00 PM
Afternoon Session: 1:30 – 3:30 PM
Wallace 300

Instructors:	Ben Fifield & Katie McCabe
Office:	Robertson 322 (Ben); Robertson 224 (Katie)
Email:	Ben: bfifield@princeton.edu
	Katie: ktmccabe@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 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. In addition to traditional data sets and models, **R** can also handle many kinds of new data including GIS, network, 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 knowledge of probability and statistics as covered in POL 571.

Structure: The camp will meet for five days with two daily sessions (morning and afternoon) each day. Each session lasts two hours with an hour and a half break between. The first hour of each session is lecture-based. After the lecture, students work individually through practice exercises with the guidance of the instructors. We devote the final ten minutes of each session to review solutions to the practice exercises.

We will start with the morning Monday session and will finish on Friday afternoon. The content is organized into 5 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: 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 where you can 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 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 1^{st}.

In addition, we will be distributing a "Getting Started" problem set due by the start of class on January 27^{th}. 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.

Group-based Learning: To promote 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.

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.

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 (Monday Morning/Afternoon)
Topic: Summarizing Univariate and Bivariate Data
Getting Started Assignment due 10 AM, Monday
Homework 1 due 12 AM, Tuesday

Module 2 (Tuesday Morning/Afternoon)
Topic: Loops and Conditional Statements
Homework 2 due 12 AM, Wednesday

Module 3 (Wednesday Morning/Afternoon)
Topic: Probability and Simulations
Homework 3 due 12 AM, Thursday

Module 4 (Thursday Morning/Afternoon)
Topic: Functions, Point Estimation, and Confidence Intervals
Homework 4 due 12 AM, Friday

Module 5 (Friday Morning/Afternoon)
Topic: Statistical Tests, Analysis of Experimental Data, and Regression
Final Exam due 10 PM, Saturday