# Syllabus: The Pre-prefresher

Gary King
Harvard University*

May 2, 2020

## 1 What's it about?

Welcome to the graduate program at the Harvard Government Department. We want to make sure you have all the skills to thrive in our methods sequence and in graduate school, starting with Gov2001 this Fall. This pre-prefresher course is voluntary: no attendance taken, no grades issued. But mastering or refresh basic concepts now will be beneficial in your methods classes this fall, for the rest of graduate school, and likely throughout your professional career. Your main task in graduate school is to learn how to do research, and teaching that is the goal of the pre-prefresher, the prefresher, and the whole political methodology sequence. Even if you're already familiar with most of the concepts listed on this syllabus, it will still be useful for you to join to get to know your cohort and us, and to make sure you're up to speed on all relevant topics.

This pre-prefresher course leads into the *Math Prefresher for Political Scientists*, which we (intend to!) teach in person just before classes begin.

## 2 Prerequisites

Just hard work. Beyond that, we'll assume familiarity with basic arithmetic, high-school algebra, and a working knowledge of computers. If your high school math is rusty, Khan Academy provides useful open-source learning resources. Previous experience with statistical methods and computing is of course helpful, but not required. We've designed this experience so that you would not need to spend more than about 5–10 hours per week.

## 3 Getting Help

The primary platform for this course is Perusall.com, where we will read and annotate the learning material. All questions about the readings should be asked (only) in Perusall by annotating the readings themselves. This way, everyone receives the same information

*Albert J. Weatherhead III University Professor, Director of the Institute for Quantitative Social Science, 1737 Cambridge Street, Harvard University, Cambridge MA 02138; GaryKing.org, King@Harvard.edu, (617) 500-7570.

and benefits equally. If you think you may know an answer to a query another student posted, or have a suggestion, please try to answer it. It doesn't matter at all if you say something wrong, or get the wrong answer in these discussions, but it does matter a lot that you be engaged with the material and your fellow students. You may also post questions about course content in 'General Discussion.' Perusall also allows you to communicate in chat channels privately with instructors or with other students. To access the Perusall course site, you will need to create a free account and enter the course access code that we will email you.

There will also be weekly virtual office hours. Office hours will take place in Zoom meeting rooms hosted by the TFs. No need to sign up – simply drop in and chat with us. To join, just click on the URL provided for office hours in the Gov2001 Google calendar. If the regular office hours times don't work for you, please don't hesitate to send an email to one of the TFs, Meredith Dost (mdost@g.harvard.edu) or Sascha Riaz (riaz@g.harvard.edu), to set up an individual meeting.

# 4 Resources

This course mainly draws on material from the methods faculty in the Government Department. The textbook for the course is *Quantitative Social Science: An Introduction* (QSS) by Prof. Kosuke Imai – a former student and TF of Gov2001! The book covers data analysis, linear regression, and probability using the programming language R. We will give you free access to the book for a period of six months using an access code that we will send to you via email.

We have assigned some additional material including hands-on programming exercises and video lectures to help you understand the material. We recommend that you watch the lectures by Prof. Matt Blackwell that accompany the book. These lectures were originally recorded for Gov50, Harvard's introductory course in political methodology. You may of course discard all references to assignments and deadlines that relate to Gov50, not the pre-prefresher. To give you access to the video lectures, we will need your Harvard ID and email address, which you should obtain within 2 weeks after April 15. To share this information with us, please fill out this form at your earliest convenience.

Finally, we have assigned a number of hands-on programming exercises on Data Camp. Programming in R can seem challenging at first, but completing these exercises will help you become familiar with the syntax fast. You will receive an invitation to our data camp group via email after we have received your Harvard information through the form cited above. On Data Camp, you will find all assigned exercises including their due dates. The exercises labeled 'QSS Data Camp' are an exception to this rule. You will have to manually access these exercises by clicking on the hyperlinks provided below. These Data Camp modules were created by Prof. Matt Blackwell to accompany the QSS textbook.

# 5 Problem Sets

We strongly encourage you to complete the assigned exercises. Reading through a chapter of the textbook and actually applying this knowledge to solve a new problem are entirely

different things. The best way to learn this material is to use it!

**How Problem Sets Work**

1. You will prepare problem sets in R Markdown. We recorded two video tutorials to explain how this works here and here. Please reach out on Perusall with any questions about RMarkdown.

2. Please DO NOT include your name in the author field. The TFs will assign an animal codename to each of you (Sascha uses the name "hamster" in the tutorial). You will receive your codename through a private message on Perusall, so nobody else will know who you are.

3. When you're done, please compile a PDF document to submit your answers (not html). Then, name the file as follows: hamster_pset1.pdf where you replace "hamster" with your assigned animal. Again, your PDF document should not include any identifying information like your name or email address. Finally, submit the file using the Google form link provided on Perusall.

4. The problem sets will be posted on Perusall. Clarifying questions must be asked by annotating the problem set so all can see. (Obviously, you should not give away the answer even if you know it at this point.)

5. Turn your assignment in by the date and time indicated in the Gov2001 calendar: Sundays at 11:59PM EST.

6. By 12:00PM EST the following day (Monday), the solution set will be released on Perusall. Any questions about the problem set solutions must be asked in Perusall by annotating the solution set distributed there.

7. You will be randomly and anonymously assigned the problem set written by another class member. By Fridays at 11:59PM EST, your job is to annotate this problem set by explaining as clearly, helpfully, and concisely as you can how each answer could have been improved, using the distributed solution set as guidance. Your goal is to teach your classmate so they truly and deeply understand how to solve this problem and problems like it. When you find a mistake, you should try to go beyond referring your classmate to the solution set and identify where your classmate made a mistake. Answers that are wrong obviously require more explanation from you, but may also be able to help your classmate learn how to improve a question, or connect it to related information you know or on the web, even if they got it right.

8. The instructors will then, if possible, improve the explanations in your annotations. You should read annotations on your completed assignments and try to understand how you might have improved your answers. Any questions you might have can be sent privately to us through a private Perusall chat channel.

To complete the assigned problem sets from QSS, we suggest that you use RStudio Cloud by making a free account. You can of course also try to set up R and RStudio on your local computer. You find instructions on how to do this here.

# 6 Schedule

## 6.1 Module 1: Introduction to R

*Timeline: April 27 – May 3*

- Reading: QSS, chapter 1
- Video Lecture 2
- Data Camp
  - QSS Data Camp: Introduction
  - Introduction to R

## 6.2 Module 2: Causality

*Timeline: May 4 – May 17*

- Reading: QSS, chapter 2
- Video Lectures 3 – 5
- Data Camp
  - QSS Data Camp: Causality
- Get Started with RStudio Cloud
- Problem Set 1: 'Exclusionary Attitudes'

## 6.3 Module 3: Measurement

*Timeline: May 18 – June 7*

- Zoom meeting: May 26, 12:30-1:30PM EST
- Reading: QSS, chapter 3
- Video Lectures 6 – 9
- Data Camp
  - QSS Data Camp: Measurement
  - Intermediate R (excluding the chapter on 'Utilities')
  - Optional: Intermediate R Practice
- Problem Set 2: 'Sources of Empathy in the Circuit Courts'

## 6.4 Module 4: Linear Regression

*Timeline: June 8 – July 5*

- Reading: QSS, chapter 4
- Video Lectures 11 – 15
- Data Camp
  - Chapter 3: Exploring Data
  - Chapter 4: Correlation and Regression
- Problem Set 3: 'Immigration attitudes: the role of economic and cultural threat'

### 6.4.1 Optional: Linear Regression in Matrix Form

- Math Prefresher Booklet, Chapter 6 on Linear Algebra
- Video Series: videos 1–17
- Here's a series of videos that provides a foundational yet intuitive understanding of the geometric interpretation of matrices, vectors, etc.

## 6.5 Module 5: Probability

*Timeline: July 6 – July 26*

- Reading: QSS, chapter 6
- Video Lectures 16 – 18
- Data Camp
  - Chapters 5 – 7: Probability, Probability distributions, Sampling distributions
  - Optional: Probability Puzzles in R
- Problem Set 4: 'A Probability Model for Betting Market Election Prediction'

## 6.6 Module 6: Uncertainty

*Timeline: July 27 – August 9*

- Reading: QSS, chapter 7
- Video Lectures 19 – 24
- Data Camp
  - Chapters 8 – 9: Confidence Intervals, Hypothesis testing