GAPS Workshop: Introduction to R

Description and Learning Objectives

This workshop focuses on developing basic skills in R. By the end of the course, students will have enough familiarity with R that they can concentrate on learning substantive methodological subject matter in their coursework instead of struggling with their software program.

The learning objective of this workshop include:

- How to install R, the RStudio working environment, and commonly-used R packages;
- How R is both similar and different to other analysis software;
- How to import and save datasets;
- How to conduct basic operations such as re-coding, subsetting/filtering, construction of composite variables, etc.;
- How to create basic data visualizations;
- Troubleshooting basic coding mistakes; and
- Good workflow practices such as setting a working directory, documenting/commenting code syntax, not overwriting existing variables, etc..

Note, this workshop is NOT designed to teach income students and substantive statistical material (e.g. how to conduct statistical tests, how to run a regression, etc.). Rather, it is meant to get new students to a level where instructors can focus on teaching substantive material without the unfamiliarity with statistical software getting in the way.

Rationale

The rationale for this workshop is two-fold. First, it will give incoming graduate students (MA and PhD) the opportunity to become comfortable using R before they encounter it in their methods courses. Second, it will give senior PhD students the opportunity to gain teaching experience.

The department has made significant changes in its methods training. One of the most challenging experiences for some new students is learning a computing language alongside statistical concepts. This workshop sets up incoming graduate students for success by familiarizing them with the software to allow them to focus on learning substantive material in their courses.

Many graduate students in the department are already well-equipped to facilitate the proposed workshop, and doing so would give them valuable teaching experience. Incoming students who participate can provide teaching evaluations, which the student instructors can include in their teaching dossier—something routinely requested in tenure-track job (and some postdoctoral) applications.

An additional benefit is this workshop provides an opportunity for senior and incoming graduate students to get to know each other before the school year begins. This kick-starts the socialization process, which–hopefully–encourages new graduate students to seek support from their peer sooner rather than later in the very stressful experience that is the first term of graduate school..

Timeline, Format, and Expectations

The workshop will total four hours of instruction over two days plus an hour after each day of "open lab"/"office hours" where students can do more self-directed practice with R or get one-on-one coaching with the student instructor(s).

The teaching would be conducted by one or two senior PhD students. If there are two instructors, both will be available during the workshops and hte labs to help students with technical issues they encounter. Alvaro Periera and John Santo are interested in facilitating the workshop. Shanaya van Hooren and Daniel Mosannef are interested in helping with the open lab/office hours component.

The sessions will be held in hybrid format across two days in August (tentatively, the 28th and 29th) . This format has been used to great success by the ICPSR Summer Program. This allows us to accommodate more students and—with recorded sessions—an online resource new students can consult to help them troubleshoot any coding problems they encounter.

In addition to the technical training, we will hold a social activity after each session to help facilitate comradery amongst students.

Continuity Plans

The department has many graduate students who are able to run this workshop and will likely continue to produce students with this level of knowledge. Pending the success of the 2022 workshop, this could be a yearly event in August that GAPS runs.

Ideally, a partnership with the department would be formed in an effort to collaborate on the topics covered and on instructor compensation. The proposed compensation is a $300 research grant (like the RTDF) to the instructor, or if there are two instructors, $150 split between them. This would be an incentive to be the instructor and help offset research costs for the student. The breakdown for that amount is roughly 12 hours at $25/hour. A more detailed breakdown is as follows:

Item	**Hours Allocated**	**Total Hours**
Workshop Instruction	2 hours x John Santos 2 hours x Alvaro Pereira	4 hours
Prep Time	2 hours x John Santos 2 hours x Alvaro Pereira	6 hours
Office Hours/Responding to Questions	1 hour x John Santos 1 hour x Alvaro Pereira	2 hours

Draft Outline

Date	Focus	Main Commands/Cheat Sheet	Data

DAY ONE
Mon. Aug. 29
10:00-12:00
	<u>"GETTING TO KNOW R"</u>
	● What is R? Why do we use it?
	● How is R different from other software? (object oriented approach, package system, etc.)
	● Downloading R + R Studio
	● Navigating and using R
	● Creating a Project
	● Loading and Discussing Initial Packages
	● Importing Data
	● Looking at the Data
	● Object types
		`install.packages()`
		`library()`
		`read.csv()`
		`write.csv()`
		`read.dta13()`
		`table()`
		`summary()`
			Canadian Election Study (CES)
			
			Varieties of Democracy (V-DEM)

DAY TWO
Tue. Aug. 30
12:00-14:00
	<u>"STARTING TO WORK IN R"</u>
	● Operations - different way to do the same thing
	● The "tidyverse" convention of R operations
	● Visualization using ggplot2
		`$ [] [[]]` *(selection operators)*
		`<-` *(assignment operator)*
		`+ - * / max min mean median sum` *(arithmetic operators)*
		`NA NULL` *(missingness operators)*
		`TRUE FALSE` *(boolean operators)*
		`recode()`
		`select()`
		`filter()`
		`pivot_longer()`
		`pivot_wider()`
		`ggplot()`
			Canadian Election Study (CES)
			
			Varieties of Democracy (V-DEM)