# Week 1 Handout

Gov 50 Data Science for the Social Sciences

John Koo

September 2, 2025

## Contact Info

**John Koo**
PhD Candidate, Department of Government

**Email** (preferred): johnkoo@fas.harvard.edu
**Slack**: @johnkoo (in the Harvard University workspace)
**Office hours**: Tuesdays, 1:30pm to 3:30pm @ CGIS Café (sign-up: jkoo.nl/meet)

**Section materials**: https://github.com/tanxpyox/gov50-sections-jk

## Where to get help?

- For tech assistance or help with homework: **Course Assistant-led Study Halls**

  Study halls are a combination of office hours and drop-in tutoring sessions. Course assistants will hold a table usually at the CGIS Fisher Family Commons and help students with assignments and course material. Study halls work best if you come as a group and work on the assignments on your own while you are there and ask for help from the CAs when you get stuck.

  Schedule: Mondays and Wednesdays, 5pm to 9pm (see Syllabus for location)

- For help with course content: ask on course Slack (accessible via Canvas sidebar) or sign up for my office hours
- For perspective and inspiration: sign up for Scott's office hours: https://calendly.com/causalinf/office-hours

## Getting the most (grades) out of this class

- Podcast and Article Responses (5%) [one two-page doc every other week]
- Problem Sets (20%) [seven to eight in total]
- Mid-term exam (25%) [in-class, written, closed book]
- Final exam (25%) [in-class, written, closed book]
- *Final project* (25%)

Generic advice for maximising grades and efficiency

- Everything you learn should help you work towards the final project (which is the biggest chunk of your grades)
  - Take the project milestones seriously and do not wait until the last minute
- Problem sets
  - Low hanging fruit - don't miss them; Can reuse code in your projects
  - It's OK to make mistakes (each p-set is 2–3% of your final grade)
  - Work with your study group; but write up your p-sets individually
  - Start early, so you have time to get help
  - Set aside a "focus session" every week to do the problem sets; do not mull over the p-set for the whole week
- AI allowed, except in exams. Start learning how to use AI to debug your code.
- Keep your code organised in your GitHub repository - you may need them in your final project

## Getting Started

- Software: R, RStudio
- Version control: Git and GitHub

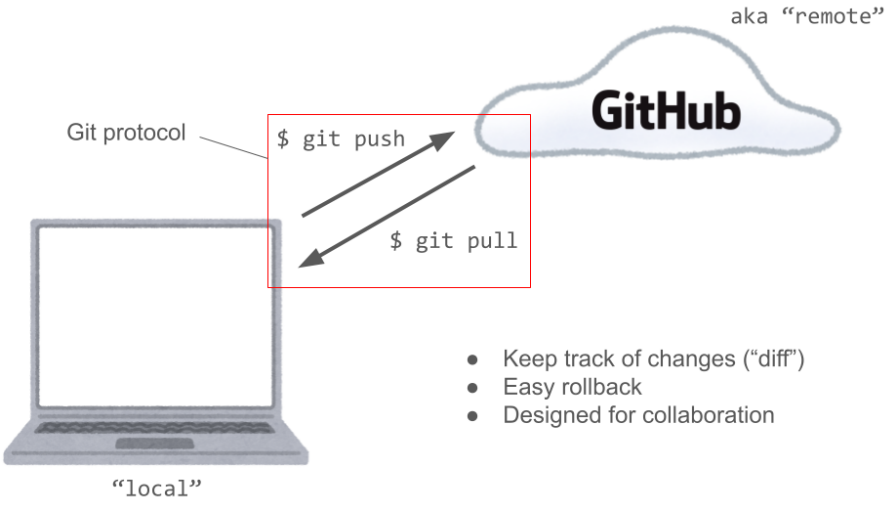


Figure 1: *Git and GitHub*

# Problem Set 0: Getting Started with R, R Studio, and Slack

**Due: Wednesday, September 10, 2025 at 11:59 PM**

## Installing R and RStudio

In this problem set, we're going to get R, RStudio, and R Markdown set up on your computer. To get started, follow these steps:

1. Download and install the most recent version of R (click me). There are versions available for the Windows, Mac, and Linux operating systems. On a Windows machine, you will want to install using the `R-x.y.z-win.exe` file where `x.y.z` is a version number. On a Mac, you will want to install using the `R-x.y.z.pkg` file that is notarized and signed.

2. With R installed, download and install RStudio (click me). RStudio is a type of "integrated development environment" or IDE designed for R. It makes working with R considerably easier and is available for most platforms. It is also free.

3. Open the .Rmd version of this file in RStudio. Install the packages we will use throughout the semester. To do this, either type or copy and paste each of the following lines of code into the "Console" in RStudio (lower left panel by default). Make sure you do this separately for each line. If you are asked if you want to install any packages from source, type "no". Note that the symbols next to `my_package` are a less than sign `<` followed by a minus sign `-` with no space between them. (Don't be worried if you see some red text here. Those are usually just messages telling you information about the packages you are installing. Unless you see the word `Error` you should be fine.)

```
my_packages <- c("tidyverse", "usethis", "devtools", "learnr",
                 "gitcreds")
install.packages(my_packages, repos = "http://cran.rstudio.com")
```

4. For many things in the course, we'll need produce PDFs from R and that requires something called LaTeX. If you've never heard of that, it's completely fine and you should just run the following two lines of R code:

```
install.packages('tinytex')
tinytex::install_tinytex()  # install TinyTeX
```

## Joining the Course Slack

5. In the left hand menu on Canvas click the item called Slack. Slack is the main way forum we will use to send out announcements and answer questions for this course. Post a short introduction message in the #general channel (e.g. `Hi all, my name is Noah and I'm looking forward to this class`).

## Submitting

6. At the top of this file where it says `author:` replace Gov 50 with your name. Click the `Knit` button near the top of the RStudio window, this should create a pdf version of this document in the same place you are storing this .Rmd file (we strongly recommend having a dedicated folder for Gov50 assignments). If you encounter any issues knitting seek help from a TF or

CA. Upload the pdf in Gradescope (can be reached via the left hand menu on Canvas) under the assignment `Problem Set 0: Getting Started`. This along with your written introduction in Slack constitutes the assignment. This assignment is graded on completion only.