Homework 0: Configuring Your Environment

CS109b Staff

Jan 22, 2018

Welcome to CS109b!

Homework 0 is due Wednesday, January 31, 11:59PM

Homework 0 is intended to setup and test your environment to work out any kinks before Homework 1 is due. Homework 0 must be submitted (see Step 6 below), but it will not count towards your homework grade.

In CS109b, the R programming language and the RStudio IDE will be used for the first two thirds of the class, including most of the homework.

The final third of the class will focus on Deep Learning using Python. A separate setup document will be available for that portion later on.

If you encounter problems, please use Piazza: piazza.com/harvard/spring2018/cs109b

Step 1: Install R

Download and install the R runtime for your machine at cran.rstudio.com

Step 2: Install RStudio

Download and install RStudio for your machine at: www.rstudio.com/products/rstudio/download/

Choose the appropriate installer for RStudio Desktop, Open Source License.

OPTIONAL Step 3: Install TeX for PDF Generation

This step is optional.

Assignments will be submitted as PDF files. If you do not perform this step, the default option is to generate HTML output from RStudio and then convert the HTML file to PDF. However, an alternative is to have RStudio generate PDF directly via TeX or LaTeX

If you already TeX or a LaTeX editor installed, you likely do not need to do any additional steps.

Installing TeX takes approximately 2 GB of disk space. Having TeX will save a step of converting HTML output to PDF for submission and allow slightly cleaner output.

For more information:

- Customizing LaTeX in RStudio
- Install LaTeX for R

Step 4: Install a Package

R gets significant leverage from thousands of packages. Install one of the packages that will be used in Homework 1 and ensure that it loads correctly.

Open RStudio.

In the left hand pane marked Console is an interactive interpreter for R (a REPL).

At the `>` prompt, enter:

```
install.packages("ggplot2")
```

If no errors are encountered, it should say `* DONE (ggplot2)`

Next, load the library into memory

```
library(ggplot2)
```

If there are no errors, it will return silently.

Step 5: Test the installation

1. In RStudio, select `File`, `New`, `R Markdown...`
2. Fill in the Title, such as "CS109b Homework 0", your name
3. If you performed Step 3 or have TeX installed, select PDF as the output format. If you did not perform Step 3, choose HTML as the output format.
4. Save the R Markdown file locally
5. Generate the output: select `File`, `Knit Document`

A nicely formatted output preview should appear in a new window.

If the output is HTML, convert the HTML to PDF. The easiest way to do this is typically by printing to PDF from your browser.

Homework requires PDF output so grading comments can be made inline. Canvas does not support inline HTML comments.

Step 6: Submit to Canvas

Go the Canvas, select the `Rmd` and generated `PDF` files and upload both to the Canvas Homework 0 assignment.

Step 7: Sign up for github

github will used for the Deep Learning project to allow collaboration and publishing of your results.

If you do not yet have an account on github:

Create a free educational account by going to education.github.com and sign up using your `.edu` email address.

This will provide you with unlimited free private repositories (and other goodies!). If you have an existing github account, you can link your `.edu` email address and get the same benefit. It will take only minutes to get approval, not weeks as they may claim.

Step 8: Sign up for Amazon Web Services

Students will have access to JupyterHub (as in CS109A). AWS will be used for the Scalability and Deep Learning Modules. For the projects, teams can decide which combination of tools is most appropriate. Sign up for an account now.

Visit wiki.harvard.edu/confluence/display/USERDOCS/AWS+Educate for detailed instructions.