

Homework 1

Upmanu Lall

Due Wednesday, February 7, 2018

This homework will give you an introduction to the R programming language. We will use R throughout this course to access, read, plot, model, and analyze data from many different sources. As you are going through this assignment, please refer to Hadley Wickham's R for Data Science (free online at <http://r4ds.had.co.nz/>).

1 DATA CAMP

The fastest and easiest introduction to R is through Data Camp (<https://www.datacamp.com>). This platform presents ideas with short videos, then gives you short programming problems (1-5 lines) with interactive feedback.

Content is organized into courses, and upon completing a course you can download a certificate of completion. You will need a paid subscription to Data Camp to access their courses. Please create an account (it is preferable but not required to use your Columbia e-mail). Once you have created an account, go through the following courses *in order*:

1. Introduction to the tidyverse: <https://www.datacamp.com/courses/introduction-to-the-tidyverse>
2. Data manipulation in R with dplyr: <https://www.datacamp.com/courses/dplyr-data-manipulation-r-tutorial>
3. Data visualization in R with ggplot2: <https://www.datacamp.com/courses/data-visualization-with-ggplot2-1>

4. Reporting with Rmarkdown: <https://www.datacamp.com/courses/reporting-with-r-markdown>

Once you have completed these courses, you will need to submit your certificates of completion. To get them, go to your profile (click top right and select your name), go to the “Completed Courses” section, and click on the course name. On the right hand side is an option to download the “Statement of Accomplishment.” Please submit the 4 statements of accomplishment, as .pdf files, corresponding the the four courses above.

One feature of R is that there are many different ways to code the same thing, and each has its own benefits and disadvantages. This tutorial will get you started with one way to do things, and as you spend more time in R you will learn much more.

2 RMARKDOWN

Your second task is to install R and RStudio, and to use them to run a simple RMarkdown document. You will use RMarkdown for most of the work we will do later in this class, and so if you are having difficulties with installation we want to know as soon as possible so that we can address them.

The following instructions come from <http://r4ds.had.co.nz/introduction.html#prerequisites>, which you can visit for more detail:

1. To download R, go to CRAN, the comprehensive R archive network. CRAN is composed of a set of mirror servers distributed around the world and is used to distribute R and R packages. Don't try and pick a mirror that's close to you: instead use the cloud mirror, <https://cloud.r-project.org>, which automatically figures it out for you.
2. RStudio is an integrated development environment, or IDE, for R programming. Download and install it from <http://www.rstudio.com/download>. RStudio is updated a couple of times a year. When a new version is available, RStudio will let you know. It's a good idea to upgrade regularly so you can take advantage of the latest and greatest features. For this course, make sure you have RStudio 1.0.0 or higher.
3. Open RStudio and explore a bit. Use the commands that you have learned on Data Camp. To install the packages we will need for this homework, run

```
install.packages('tidyverse')
install.packages(c('ggthemes', 'viridis', 'knitr', 'rmarkdown'))
```

in the console.

4. Open the `template-file.Rmd` file in RStudio. Edit your name and UNI. You do not need to edit anything else. Then click “knit”.

Please turn in the .Rmd file with your edits and the .html file generated by knitting your document.

SUMMARY

The things you need to turn in as part of this assignment are:

1. 4 statements of accomplishment from Data Camp, as `.pdf` files
2. Your modified `template-file.Rmd` file
3. The `template-file.html` file generated by knitting the `.Rmd` file

If you have difficulty with this homework, particularly with using the Data Camp website or installing R and the required packages, please use the `r-computing` channel on the Slack page. It's likely that someone else has had or will have the same problem, so check if it has been answered before you post.