Lab 1: Intro to R

David Sichinava

October 4, 2017

Topics

- Intro to R
- R GUI: R-Studio
- Creating R markdown documents
- Literate programming

Instruction:

Follow the assignment step-by-step. Name your *.rmd* file in a following format: *your_surname_lab*1*.rmd*. For example,

```
sichinava_lab1.Rmd
```

Assignments:

File system and navigation:

First of all, create a working directory for your work in your computer. As we are going to have assignment for each lab, create a separate folder for this class *intro_stats_r*4, add a subfolder *labs* and create a folder for the first lab: *lab_*1.

Open 'R-studio' and create new *R*-notebook. Assign a name and save it into *lab_*1 folder. Saving is simple: you just click a corresponding icon in a menu and select necessary format (See picture. 1). Cool! Now in the notebook, after the preamble (see picture 2) type a short text about your impressions on *R − Studio*. Two or three sentences are enough.

Insert code chunk in your notebook. Be sure that your code is active. In order to check that look on a green button at the right side of the chunk (see picture 3). Write down a code between the accent marks which will navigate you to your working directory.
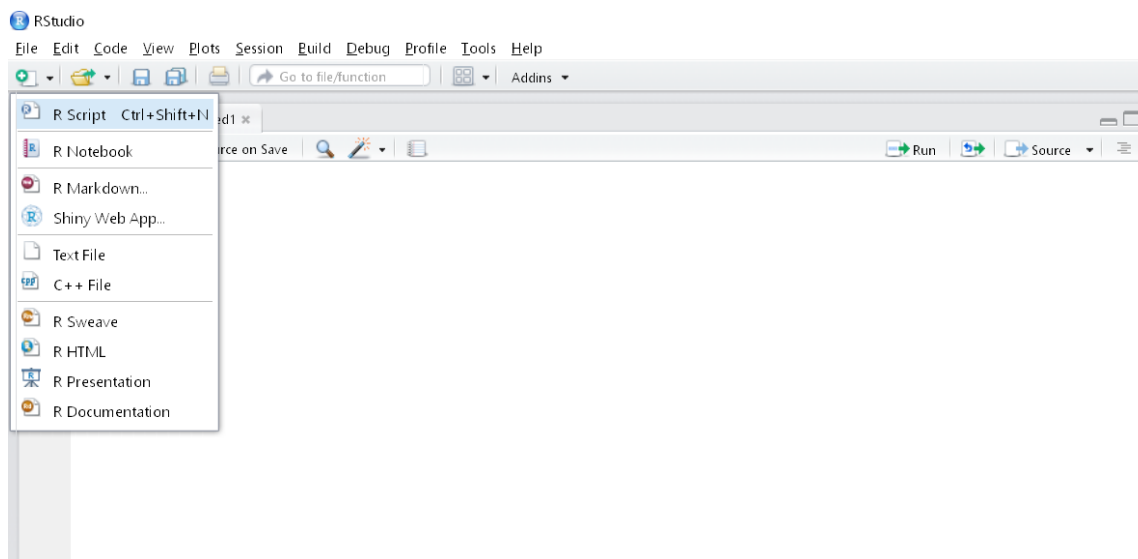
Figure 1: Creating a new file

Installing and activating libraries

Insert new code chunk in your notebook. Write down a code which will install *ggplot*2 library. Run the code and make sure it works.

In a new line write a code which will activate *ggplot*2. Run the code and again make sure it runs for you.

Creating a script file and sourcing

Click a button for creating new file and select *R* script file. R script is just a text file where we type our code, which can be opened in R environment and executed (sourced) In difference with markdown document, we cannot have a *richtext* in script files. Usually we don't include text in the script, file, however you can have one by *commenting* it. (see figure 4). The program will ignore script lines with leadign hashtags.

In the script file, write a code which will install *tidyr* library and activate it. Create a new subfolder *source* in your *lab*_1 folder and save your script file there. Give a reasonable name and extension .*R* to the script file.

According to the principles of literate programming, we should run (or, in R-speak, *source*) the script from the notebook. In order to do that, Go back to the notebook, add a code chunk, type code which will run the script and make sure that it runs correctly.

Processing the document

In order to follow the principles of literatre programming, in the notebook write a short description to each code chunk what they are supposed to do. Do not limit yourself to the plain text. Use wide formatting possibilities of markdown, for example, mark sections, insert links, pictures, bold
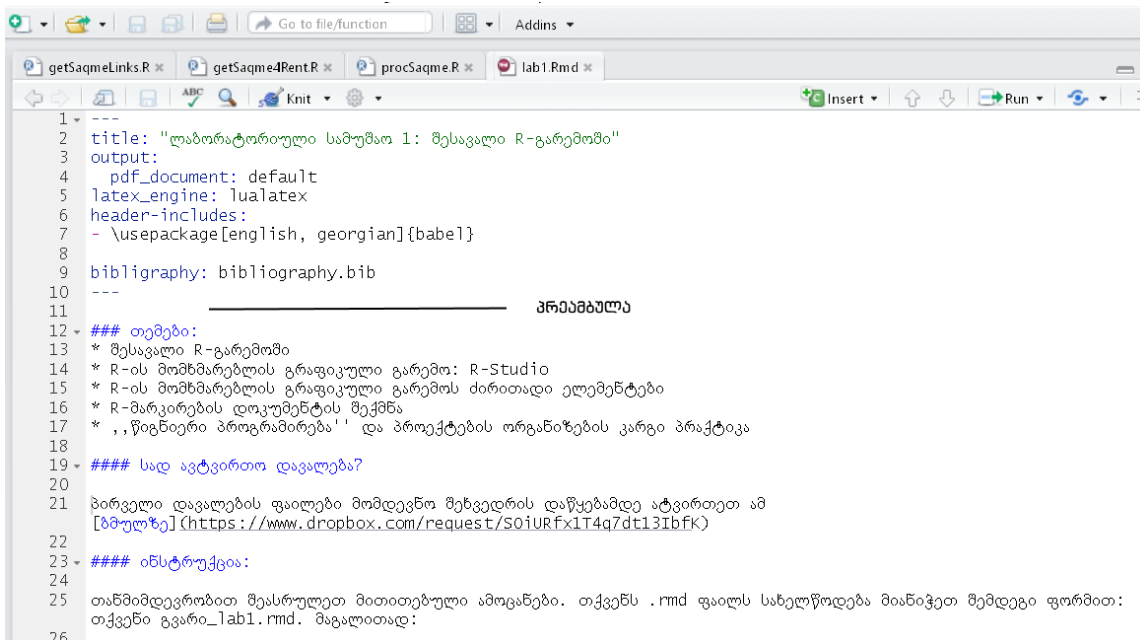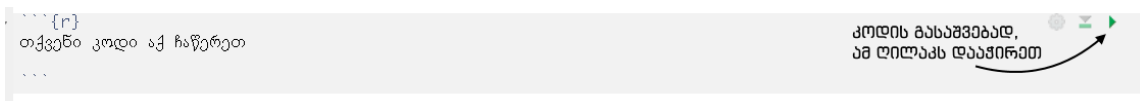
Figure 2: Document preamble

Figure 3: Selecting and running code chunk

and italic texts. Each of this will earn you extra points. Detailed documentation for the markdown could be found here and here.

Compiling the document

Before you compile the document, that is, transfer it into *.html* format, save your notebook (just click *save* button or hit *ctrl + S*). *.html* or HyperText Markup Language is a standard language for creating websites. That said, your compiled document could be opened in any relatively modern web browser.

In order to compile the document, hit the knit button and select *knit to HTML* (see figure 5. If you did everything correctly, your compiled notebook will load automatically. Remember that when compiling, R-Studio executes all the code written in the document.

Annotated bibliography

You can type the text for the annotated bibliography directly into R-notebook. Use APA standards and the following example to prepare the bibliography:

```
 9
10    # ეს ტექსტი გაკომენტარებულია, მომდევნო ხაზი კი - არა
11      init_url<-c("http://realestate.saqme.ge/ka/uZravi_qoneba-5566/cat_listing.html?alias=reale
12
13      urlstart<-Sys.time()
14 ▾    for(i in 2:373) {
15        propURLs<-GET(paste0(init_url, i, sep=""), user_agent("Mozilla/5.0 (Windows NT 6.1; WOW6
16        read_html() %>%
17        html_nodes(".udzravi_gancxadebis_title a") %>% # get the CSS nodes
18        html_attr("href") # extract the URLs
19        propURLs<-na.omit(as.data.frame(propURLs))
20        pageURL<-rbind(pageURL, propURLs)
21      }
22
23      urlend<-Sys.time()
```
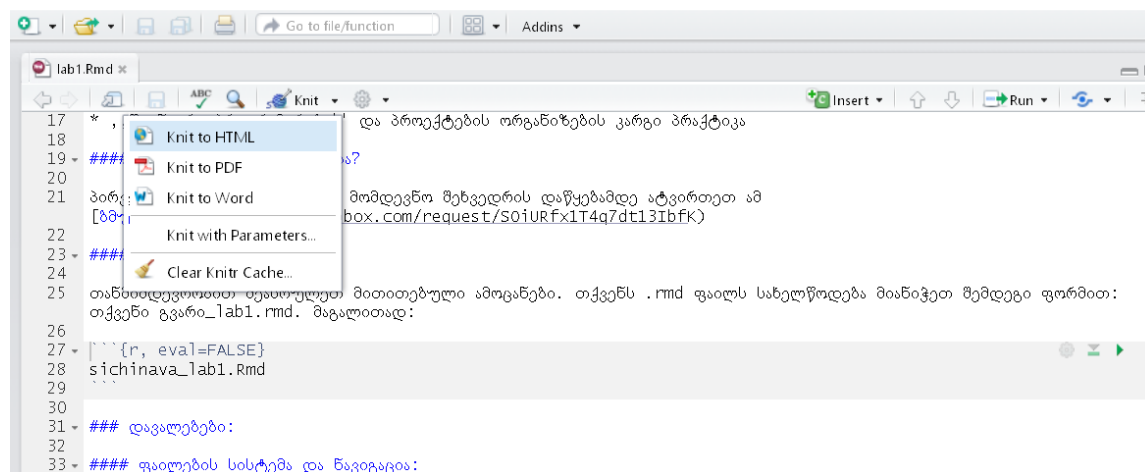
Figure 4: Commenting the code

Figure 5: Compiling the document

Waite, L. J., Goldschneider, F. K., Witsberger, C. (1986). Nonfamily living and the erosion of traditional family orientations among young adults. American Sociological Review, 51, 541-554

The authors, researchers at the Rand Corporation and Brown University, use data from the National Longitudinal Surveys of Young Women and Young Men to test their hypothesis that nonfamily living by young adults alters their attitudes, values, plans, and expectations, moving them away from their belief in traditional sex roles. They find their hypothesis strongly supported in young females, while the effects were fewer in studies of young males. Increasing the time away from parents before marrying increased individualism, self-sufficiency, and changes in attitudes about families. In contrast, an earlier study by Williams cited below shows no significant gender differences in sex role attitudes as a result of nonfamily living.

Example source: http://guides.library.cornell.edu/annotatedbibliography)

Here are the three articles which you have to read and write bibliography:

- Goodman, S. N., Fanelli, D., Ioannidis, J. P. (2016). What does research reproducibility mean?. *Science translational medicine*, 8(341).

- Iqbal, S. A., Wallach, J. D., Khoury, M. J., Schully, S. D., Ioannidis, J. P. (2016). Reproducible research practices and transparency across the biomedical literature. *PLoS Biol*, 14(1).

- Stodden, V. (2010). Data sharing in social science repositories: facilitating reproducible computational research. *NIPS* 2010.

After you are done with writing the text. Compile the document. Remember that previous *.html* file will be overwritten.

Done. How to submit my assignment?

Zip the whole folder for the first lab. Name the file according to the following format: *surname_lab*1.*zip* like this:

```
sichinava_lab1.zip
```

Upload the file to this link before the start of our next meeting.
Good luck!