Introduction to R markdown

Marianne Huebner, Michigan State University

2015-05-01

In Rstudio click on File > New File. One of the options is R markdown. You can give this project a title, your name and choose whether you want the output saved in Word or html or pdf format. Only choose PDF output, if you have Miktex installed (also free software).

In Rstudio an "untitled" page will appear with a template. First save this file to the folder you created, e.g "Rnotes". It will then be saved as "Rnotes.Rmd". The extension ".Rmd" stands for R markdown and can be opened with R studio.

## Create a document

In the toolbar "ABC" is your spell checker. Always run your spell checker after you finished.

In the toolbar the question mark has a "Markdown Quick Reference". If you click on that it will open in the lower right window with tips how to get italic, bold face, or larger text size, and more.

In the toolbar, click on "Knit Word" (or "Knit html" or Knit PDF"). This will start a process that converts the current template file into a Word file (or html or pdf) which opens automatically. Do this now.

## Add R commands

Click on "Chunks" and then "Insert Chunk". This will result in a gray area starting with three backslashes followed by {r} and ending with three backslashes. In this grey area you can type R commands. A comment can be inserted following a # sign. If you click anywhere in the grey area and then click on Chunks > Run Current Chunk, all R commands will be automatically pasted into the "Console" window and evaluated. If you only want to run one line, highlight this line, and click on Run.

## Things to know about the R chunk

* echo = FALSE in {r, echo=FALSE} will not show the R commands in the final document.
* include=FALSE will evaluate the R chunk, but no output will be shown in the document.
* Your document will stop compiling, if you have an error somewhere in the R code. It is recommended that you evaluate each R chunk separately to make sure it works.
* Do a spell check before clicking on "Knit Word" or "Knit html" or "Knit pdf" for finalizing your document.
* You can use output from the R chunk in your text. For example, the Rchunk reads a dataset (use header=T, sep="")

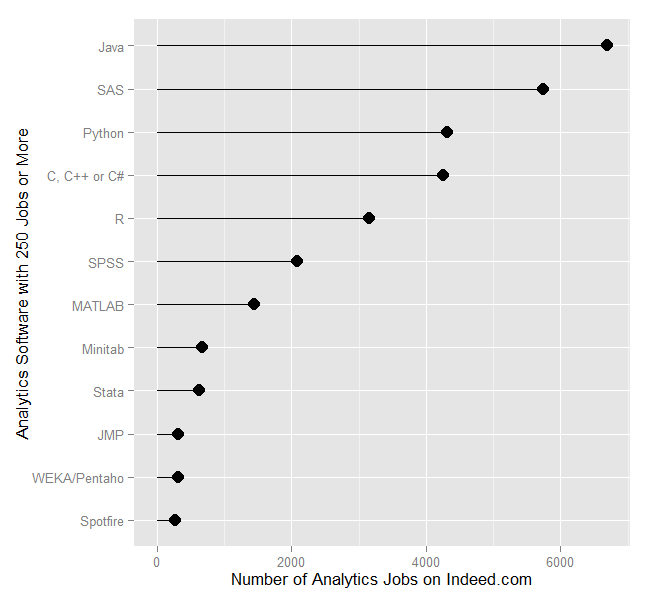
body=read.table("http://stt.msu.edu/Academics/ClassPages/uploads/FS14/200-1/Ch08\_Body\_fatv2.txt", header=T, sep="")

In order to include R output in the text, use a left apostrophe followed by "r" followed by a function or R variable followed by a left apostrophe. Example (without apostrophes): r cor(body$Weight, body$BodyFat) In the text this looks like (with apostrophes): The correlation of weight and body fat content of the dataset "body" is 0.6966328.

## Finish your Rstudio session

Before you finish your Rstudio session save your R markdown file. If it is saved, the name is shown in black, if you made changes, the name is in red and has an asterisk.

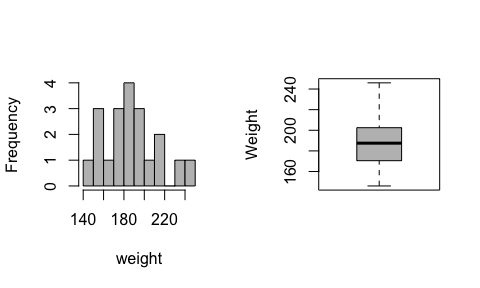
## Other convenient features for R markdown documents

Here is a feature of R markdown: you can add links to a text or to figures from the internet. For example, see [A short introduction to R] (<http://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf>) by Paul Torfs and Claudia Bauer. In order to do so you choose square brackets: [Here is the text for the link] followed by round brackets with the URL (<http://>...). If you want to insert a figure from the internet you do the same but with a ! in front of the square brackets, e.g. 

## Visualizing data

In your document you can include figures. For example, to see both a histogram and a boxplot in the same figure, we can tell R how to organize it: 1 row for 2 figures side-by-side. Try it different ways, too! There are also many ways to modify plots with labels of the axes, title, colors, lines, points, text, etc.

par(mfrow=c(1,2))  
hist(body$Weight, nclass=10, col="grey", main="", xlab="weight")  
boxplot(body$Weight, col="grey", ylab="Weight")



par(mfrow=c(1,1))

You can also embed plots, for example making a scatter plot of weight versus body fat. The plot function has the format "plot(x,y)" where x is the variable for the horizontal axis and y is the variable for the vertical axis.

plot(body$Weight, body$BodyFat, xlab="weight (lb)", ylab="body fat (%)")

