Short PDF example of R Markdown

Garret Christensen

Thursday February 11, 2016

REMEMBER TO SET THE WORKING DIRECTORY AND IN-STALL/LOAD PACKAGES!

You only have to install a package once, but you have to load the library every time you want to use it. You'll see that a bunch of ugly commands and output gets displayed. You probably wouldn't really want that in your actual paper.

```r
#install.packages("foreign")
library(foreign)
#install.packages("stargazer")
library(stargazer)
```

```
##
## Please cite as:

##  Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.

##  R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

```r
#install.packages("sandwich")
library(sandwich)
setwd ("C:/Users/garret/Box Sync/BITSS/1_Events/3_Workshops-Seminars/UCSD/UCSDWorkshop/Rmarkdown")
```

Which is why you can turn off code (with echo=FALSE) and/or results (with results='hide'). You're not going to see anything from this chunk here:

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
##     speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Loading the data

You can bring in the Stata data directly with the 'foreign' package.

```
## Warning in `levels<-`(`*tmp*`, value = if (nl == nL) as.character(labels)
## else paste0(labels, : duplicated levels in factors are deprecated
```

Running Analysis

I ran some regression analysis. The results are good.

Graphics

Graphics can be easily inlaid. Here, I'll make a map of the number of US military recruits to a 16 year period by county.
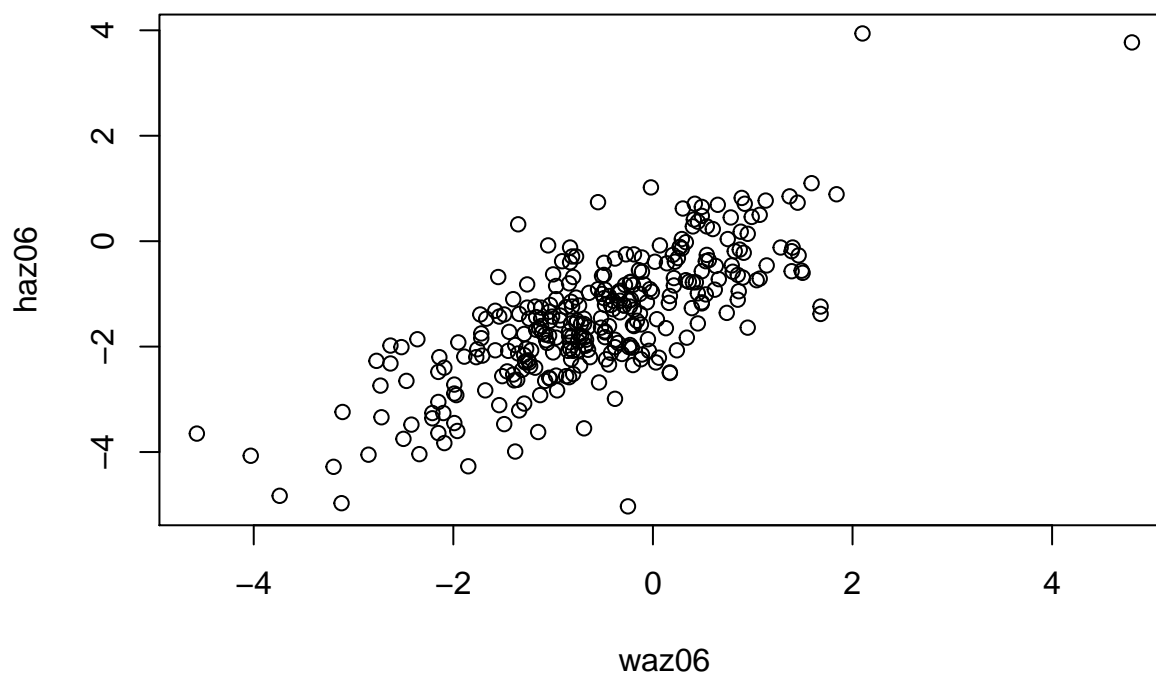
Equations

Equations written with LaTeX syntax works, so you can write short reports all in one file.

$$\frac{dN}{dt} = r * N * (1 - \frac{N}{K})$$

Refer to Values

You can refer to values calculated in R by just surrounding "r" and the code with single accent marks. For example, the mean frequency is 0.4822888.

Simple Output

You can just use built in R functionality.

```
## 
## Call:
## lm(formula = free_chl_yn ~ treatw, data = WASHB)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.37692 -0.37692 -0.01299 -0.01299  0.98701
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.01299    0.02736   0.475    0.635
## treatw       0.36394    0.04044   9.000   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.3395 on 282 degrees of freedom
##   (83 observations deleted due to missingness)
## Multiple R-squared:  0.2231, Adjusted R-squared:  0.2204
## F-statistic:    81 on 1 and 282 DF,  p-value: < 2.2e-16
```

```
## 
## Call:
## lm(formula = free_chl_yn ~ treatw + kiswahili + english, data = WASHB)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.39602 -0.35122 -0.02021  0.00334  0.97979
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.003344   0.052899  -0.063    0.950
## treatw       0.365188   0.040614   8.992   <2e-16 ***
## kiswahili   -0.010624   0.075772  -0.140    0.889
## english      0.034176   0.063663   0.537    0.592
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.3405 on 280 degrees of freedom
##   (83 observations deleted due to missingness)
## Multiple R-squared:  0.2243, Adjusted R-squared:  0.216
## F-statistic: 26.98 on 3 and 280 DF,  p-value: 2.33e-15
```

Fancier Output

Markdown is designed to be simple and also readable by humans in marked-up form. Like I said, mark*down*, not mark*up*. But you can still get really nicely formatted regression output with a couple of R packages, xtable or stargazer. (Very similar to estout or outreg2 in Stata.)

Stargazer has three types of output (text, html, and LaTeX).

TeX->PDF

When we Knit a Markdown as a PDF, it actually makes that PDF using LaTeX. (See here.) So you can use the .tex output option from stargazer and get nice PDF documentation.

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, Feb 11, 2016 - 1:12:22 PM

Table 1: Made Automatically in R

	Dependent variable:		
	free_chl_yn		
	(1)	(2)	(3)
treatw	0.364***	0.364***	0.365***
	(0.040)	(0.043)	(0.041)
kiswahili			−0.011
			(0.076)
english			0.034
			(0.064)
Constant	0.013	0.013	−0.003
	(0.027)	(0.009)	(0.053)
Observations	284	284	284
R^2	0.223	0.223	0.224
Adjusted R^2	0.220	0.220	0.216
Residual Std. Error	0.340 (df = 282)	0.340 (df = 282)	0.340 (df = 280)
F Statistic	81.002*** (df = 1; 282)	81.002*** (df = 1; 282)	26.982*** (df = 3; 280)

Note: *p<0.1; **p<0.05; ***p<0.01

Everything All in One Place?

You can do citations. Plots, graphs, and citations, what else do you need for a research paper?

Complicated, time consuming for very long articles.

Maybe try Sweave (direct combo of LaTeX and R).

Send your output to .tex files, include those in your master paper file.