

Short example of R Markdown

Garret Christensen

Thursday, April 23, 2015

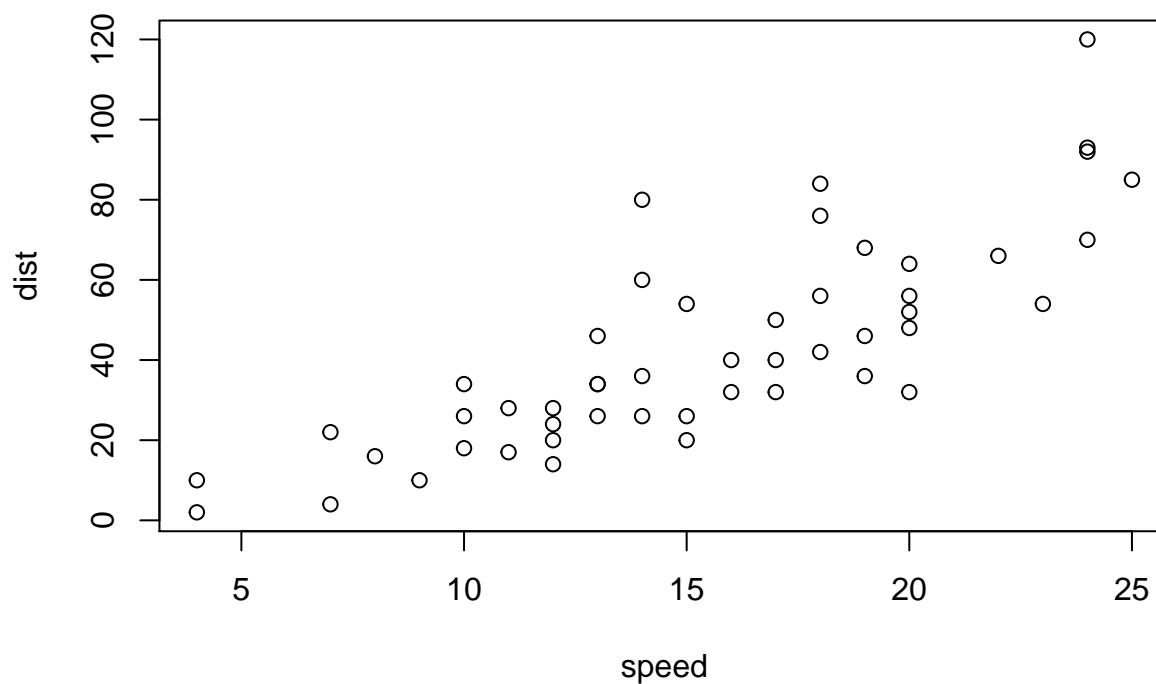
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

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:

```
summary(cars)
```

```
##      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
```

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Loading the data

You can bring in the Stata data directly with the ‘foreign’ package. You only have to install a package once, but you have to load the library every time.

```
#install.packages("foreign")
library(foreign)
setwd ("C:/Users/garret/Documents/Research/BITSS/WorkshopSlides/Riverside")
WASHB<-read.dta("WASHBpublic_mock.dta")
```

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

Running Analysis

```
#simple model
model1<-lm(free_chl_yn~treatw, data=WASHB)
#with hetero-robust se's
#install.packages("sandwich")
library(sandwich)
model2<-lm(free_chl_yn~treatw, data=WASHB)
#cluster standard errors by village
model3<-lm(free_chl_yn~treatw, data=WASHB)
```

Simple Output

You can just use built in R functionality.

```
summary(model1)
```

```
##
## 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
```

```
summary(model2)
```

```
##
## 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
```

```
summary(model3)
```

```
##
## 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
```

Fancier Output

Or you can use the new package `stargazer` to get really nice output, either in combination with LaTeX, or directly in R Markdown. First I'll save output to a .tex file, and we can use that in TeX later. (There are complicated ways to compile this using the great document conversion tool `pandoc`, but so far I haven't mastered it and don't think it's optimal.)

```
library(stargazer)
```

```
##
## Please cite as:
##
## Hlavac, Marek (2014). stargazer: LaTeX code and ASCII text for well-formatted regression and summary
## R package version 5.1. http://CRAN.R-project.org/package=stargazer
```

```
stargazer(model1, model2, model3, title="Results", out="outputR.tex", align=TRUE)
```

```
##
## % Table created by stargazer v.5.1 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard
## % Date and time: Fri, Apr 24, 2015 - 12:11:20 AM
## % Requires LaTeX packages: dcolumn
## \begin{table}[!htbp] \centering
##   \caption{Results}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lD{.}{.}{-3} D{.}{.}{-3} D{.}{.}{-3} }
## \hline
## \hline \hline
## & \multicolumn{3}{c}{\textit{Dependent variable:}} \\
## \cline{2-4}
## \hline & \multicolumn{3}{c}{free\_chl\_yn} \\
## \hline & \multicolumn{1}{c}{(1)} & \multicolumn{1}{c}{(2)} & \multicolumn{1}{c}{(3)} \\
## \hline
## treatw & 0.364^{***} & 0.364^{***} & 0.364^{***} \\
## & (0.040) & (0.040) & (0.040) \\
## & & & \\
## Constant & 0.013 & 0.013 & 0.013 \\
## & (0.027) & (0.027) & (0.027) \\
## & & & \\
## \hline
## Observations & \multicolumn{1}{c}{284} & \multicolumn{1}{c}{284} & \multicolumn{1}{c}{284} \\
## R^2 & \multicolumn{1}{c}{0.223} & \multicolumn{1}{c}{0.223} & \multicolumn{1}{c}{0.223} \\
## Adjusted R^2 & \multicolumn{1}{c}{0.220} & \multicolumn{1}{c}{0.220} & \multicolumn{1}{c}{0.220} \\
## Residual Std. Error (df = 282) & \multicolumn{1}{c}{0.340} & \multicolumn{1}{c}{0.340} & \multicolumn{1}{c}{0.340} \\
## F Statistic (df = 1; 282) & \multicolumn{1}{c}{81.002^{***}} & \multicolumn{1}{c}{81.002^{***}} & \multicolumn{1}{c}{81.002^{***}} \\
## \hline
## \hline
## \textit{Note:} & \multicolumn{3}{c}{*p<$0.1; **p<$0.05; ***p<$0.01} \\
## \end{tabular}
## \end{table}
```

Second I'll do output as text. It opens nicely in Word or Notepad, or here in Markdown.

```
stargazer(model1, model2, model3, type="text", out="outputR.txt", title="Results", align=TRUE)
```

```
##
## Results
## =====
##                               Dependent variable:
```

```
##
##
##               free_chl_yn
##              (1)      (2)      (3)
## -----
## treatw          0.364***  0.364***  0.364***
##                (0.040)  (0.040)  (0.040)
##
## Constant        0.013     0.013     0.013
##                (0.027)  (0.027)  (0.027)
##
## -----
## Observations      284      284      284
## R2                0.223     0.223     0.223
## Adjusted R2       0.220     0.220     0.220
## Residual Std. Error (df = 282) 0.340     0.340     0.340
## F Statistic (df = 1; 282)   81.002*** 81.002*** 81.002***
## =====
## Note:              *p<0.1; **p<0.05; ***p<0.01
```

Third, I'll do output as html. We can knit right to html here, and it looks great.

```
stargazer(model1, model2, model3, type="html", out="outputR.html", title="Results", align=TRUE)
```

Results

Dependent variable:

free_chl_yn

(1)

(2)

(3)

treatw

0.364***

0.364***

0.364***

(0.040)

(0.040)

(0.040)

Constant

0.013

0.013

0.013

(0.027)

(0.027)

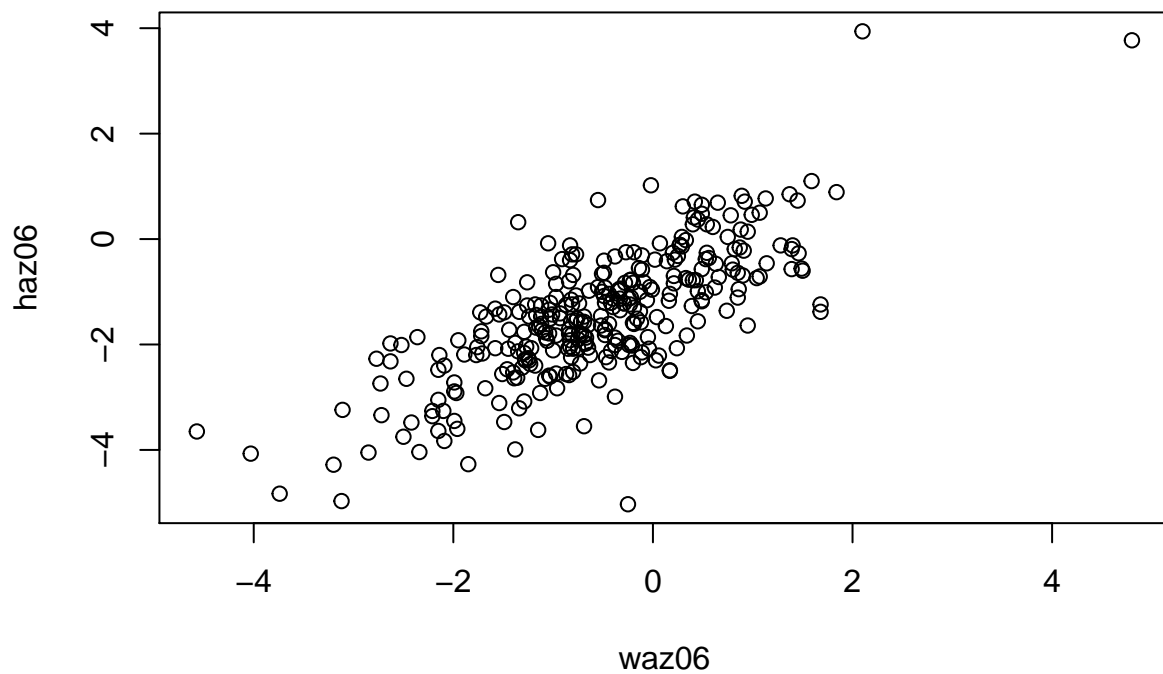
(0.027)

Observations

284
 284
 284
 R2
 0.223
 0.223
 0.223
 Adjusted R2
 0.220
 0.220
 0.220
 Residual Std. Error (df = 282)
 0.340
 0.340
 0.340
 F Statistic (df = 1; 282)
 81.002***
 81.002***
 81.002***
 Note:
 $p < 0.1$; $p < \mathbf{0.05}$; $p < 0.01$

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 work, so you can write short reports all in one file.

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