

# Using statweave for fun and profit

Nicholas J. Horton  
Smith College and University of Auckland

June 10, 2008

Copies of this file (`statweave.swv`), formatted output (`statweave.pdf`) and presentation (`talkjun2008.pdf`) can be found in <http://www.math.smith.edu/~nhorton/statweave>. The Statweave distribution and documentation can be found at

<http://www.stat.uiowa.edu/~rlenth/StatWeave>

We begin by reading in the dataset.

```
R> options(digits=3)
R> options(show.signif.stars=FALSE)
R> ds <- read.csv("http://www.math.smith.edu/~nhorton/statweave/fev.csv")
R> attach(ds)
R> i <- 4
R> ds[i,]
```

```
id height age initht initage logfev loght loginiht
4 1 1.42 12.5 1.2 9.34 0.751 0.351 0.182
```

At this point we can access particular values, for example, the age of the subject in row 4 is 12.46. We can fit a linear model and display the results.

```
R> lmres <- lm(logfev~age+loght+initage+loginiht, data=ds)
R> summary(lmres)
```

Call:

```
lm(formula = logfev ~ age + loght + initage + loginiht, data = ds)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-0.46501 -0.07267  0.00192  0.07922  0.37385
```

Coefficients:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.33289    0.02071  -16.07 < 2e-16
age          0.02868    0.00208   13.81 < 2e-16
loght       2.04343    0.06880   29.70 < 2e-16
initage     -0.01498    0.00396   -3.78 0.00016
loginiht    0.39223    0.08263    4.75 2.2e-06
```

Residual standard error: 0.114 on 1988 degrees of freedom

Multiple R-squared: 0.88, Adjusted R-squared: 0.88

F-statistic: 3.64e+03 on 4 and 1988 DF, p-value: <2e-16

Or we can use the `xtable()` function (in the `xtable` package) to create a  $\LaTeX$  table:

```
R> library(xtable)
R> lmtab <- xtable(lmres,digits=c(0,3,3,2,4),caption="Better formatted results")
R> print(lmtab)
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.333	0.021	-16.07	0.0000
age	0.029	0.002	13.81	0.0000
loght	2.043	0.069	29.70	0.0000
initage	-0.015	0.004	-3.78	0.0002
lognitht	0.392	0.083	4.75	0.0000

Table 1: Better formatted results

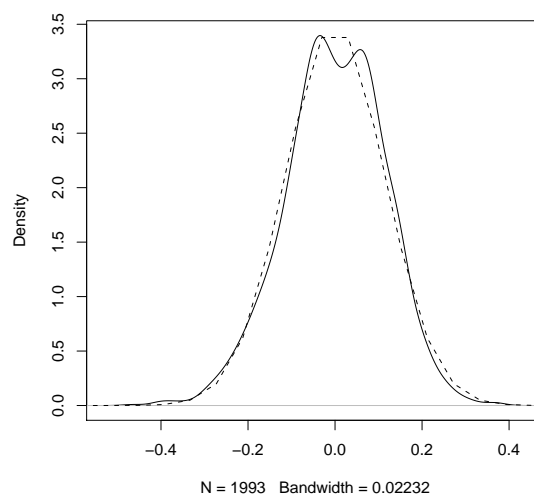
We can and should take a look at the residuals. Figure 1 displays the empirical density of residuals, while Figure 2 displays the scatterplot of normalized residuals by predicted values.

```
R> resid <- residuals(lmres)
R> pred <- predict(lmres)
R> fivenum(resid)
```

```
1729      778      1759      1895      85
-0.46501 -0.07267  0.00192  0.07922  0.37385
```

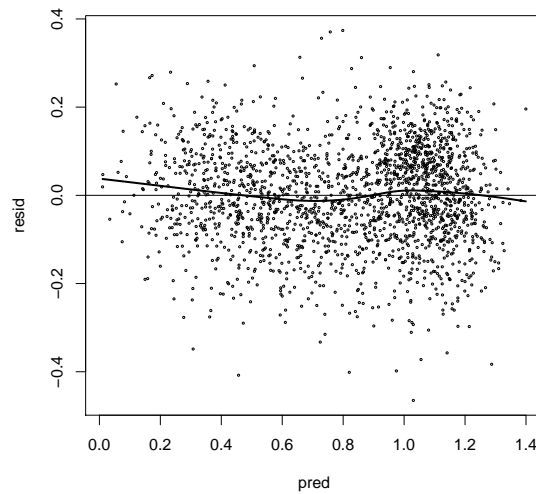
Figure 1: Empirical density of residuals

```
R> plot(density(resid),main="")
R> x <- seq(from=-3,to=3,length=100)
R> lines(x,dnorm(x,0,sd(resid)),lty=2)
```



```
R> fivenum(resid)
```

Figure 2: Scatterplot of residual by predicted value



```

      1729      778      1759      1895      85
-0.46501 -0.07267  0.00192  0.07922  0.37385

```

```

R> resid[resid< -.4]
R> pred[resid< -.4]
R> ds[resid< -.4,]

```

```

      214      232      1729
-0.401 -0.408 -0.465
      214      232      1729
0.820 0.457 1.030
      id height   age initht  initage logfev loght loginitht
214   30   1.50 11.73   1.30   7.67 0.4187 0.405   0.262
232   32   1.32  8.27   1.32   8.27 0.0488 0.278   0.278
1729 246   1.62 14.58   1.29   9.39 0.5653 0.482   0.255

```

This output was created by running `statweave statweave.swv`, which ran R, snarfed the output, then created the file `statweave.pdf`. At the University of Auckland, StatWeave is installed on `stat12.stat.auckland.ac.nz`. At Smith College, StatWeave is installed on the Scinix cluster and scheduled to be installed on Science Center classroom images for Macintosh and Windows.

While this example created a  $\text{\LaTeX}$  file (which was automatically turned into a pdf), StatWeave also supports Open Office (`.odt`) files, which can then be saved as Microsoft Word (`.doc`) files. In the present release StatWeave supports the following languages: R, Splus, SAS, IML, Stata and Maple and runs under Windows, Linux and Mac OS X.

Intrigued? I hope that you are. These static documents are a step in the right direction for reproducible research (check out Duncan Temple Lang's dynamic StatDocs projects for the holy grail). More information about StatWeave (documentation, sample files, tests, and installation instructions) can be found at <http://www.stat.uiowa.edu/~rlenth/StatWeave>.