Introducing Huxtable

Matt Lawrence

Dcember 6, 2021

Setting Up

We'll use the `midd_survey` data for this example. Load it and the usual packages.

Introducing The Huxtable Package

The raw output of summary tables and regression output can be messy in knitted files. We have previously seen how to use `kable` to format some tables. It works well for simple tables but not as well for regression output. Fortunately, there is a different package - `huxtable` - that works well for simple tables and regression output. The package is highly customizable; we'll go over just a few options today. For more details about the package, click here.

Install `huxtable` using the package manager in the bottom right pane of R Studio. Then load it using `library()`.

```
library(huxtable)
```

Formatting Summary Tables With `huxtable()`

Summary tables use the `huxtable()` function from the `huxtable` package. Let's make a simple three-variable table summarizing GPA by gender and number of siblings.

```
gpa_table <- midd_survey |>
  group_by(gender, siblings) |>
  summarise(mean_gpa = mean(gpa, na.rm = TRUE))
```

Now wrap the table in `huxtable()`

```
huxtable(gpa_table)
```

gender	siblings	mean_gpa
Man	0	3.52
Man	1	3.48
Man	2	3.5
Man	3	3.43
Man	4	3.4
Man	5	3.4
Other	0	3.78
Other	1	3.41
Other	2	3.9
Other	5	2
Woman	0	3.59
Woman	1	3.54
Woman	2	3.51
Woman	3	3.52
Woman	4	3.49
Woman	5	3.36

We can clean up the table by changing the column labels, adjusting the row alignment, adding a title, and changing the theme.

- The easiest way to change the column titles without altering the underlying data is to use the `set_contents()` option. The syntax here follows the pattern of (row numbers, column numbers, c(column titles)).
- Change the alignment with the `set_align()` option. The possible positions are left, center, right, and decimal. The syntax follows a similar pattern as above (row numbers, column numbers, position). If you want the change to affect all rows, use `everywhere` as the row number.
- Change the title with `set_caption()`. Note that all tables titled with this function will be numbered consecutively in the final report.
- The huxtable package has several different themes. The `theme_article()` and `theme_compact()` are the two I use the most. Try `theme_article()` first; it should work well with short tables. If your table is too long for a single page (or single slide), switch to `theme_compact()`.

```
huxtable(gpa_table) |>
  set_contents(1, 1:3, c("Gender", "Siblings", "Mean GPA")) |>
  set_align(everywhere, 2:3, "center") |>
  set_caption("Mean GPA by Gender and Number of Siblings") |>
  theme_article()
```

Table 1: Mean GPA by Gender and Number of Siblings

Gender	Siblings	Mean GPA
Man	0	3.52
Man	1	3.48
Man	2	3.5
Man	3	3.43
Man	4	3.4
Man	5	3.4
Other	0	3.78
Other	1	3.41
Other	2	3.9
Other	5	2
Woman	0	3.59
Woman	1	3.54
Woman	2	3.51
Woman	3	3.52
Woman	4	3.49
Woman	5	3.36

Formatting Regression Tables With `huxreg()`

Summaries of linear regression models have a lot of information and can be messy in knitted files. One option to clean them up is to use the `huxreg()` function in the `huxtable` package.

To see how huxreg compares to the output we are used to, use the `midd_survey` data to regress `gpa` on `gender`, save the model as `model1`, and look at the summary (remember `echo = FALSE` suppresses the code in the knitted file):

```
##
## Call:
## lm(formula = gpa ~ siblings, data = midd_survey)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39967 -0.15695  0.04305  0.20487  0.60033
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  3.558767   0.017163 207.348  < 2e-16 ***
## siblings    -0.031819   0.009198  -3.459 0.000564 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3003 on 983 degrees of freedom
## Multiple R-squared:  0.01203,    Adjusted R-squared:  0.01102
## F-statistic: 11.97 on 1 and 983 DF,  p-value: 0.0005644
```

Now that you are comfortable interpreting summary outputs, you should continue using the traditional summary outputs to get a sense of your models. But replace `summary()` with `huxreg()` when you want to knit your file.

```
huxreg(model1)
```

	(1)
(Intercept)	3.559 ***
	(0.017)
siblings	-0.032 ***
	(0.009)
N	985
R2	0.012
logLik	-211.646
AIC	429.291

*** p < 0.001; ** p < 0.01; * p < 0.05.

We already know how to set the title and the theme. There are a couple other options that are good to add to regression tables:

- We don't need all the model statistics at the bottom of the table. Let's keep the number of observations only which we can do with the `statistics = c()` option. The package saves the number of observations as an object called `nobs`; we'll rename it `N. obs.` in the table.
- We can also rename the coefficient names using the `coefs = c()` option. Each coefficient in the model (including the Intercept) needs to be listed here or else it will be omitted in the table.

Here's the full code chunk with all the edits. Knit the file to see the formatted regression table on page 5.

```
huxreg(model1,
       statistics = c("Number of Observations" = "nobs"),
       coefs = c("(Intercept)" = "(Intercept)",
                 "Number of Siblings" = "siblings")) |>
  set_caption("A Better Title") |>
  theme_article()
```

Table 2: A Better Title

	(1)
(Intercept)	3.559 ***
	(0.017)
Number of Siblings	-0.032 ***
	(0.009)
Number of Observations	985
***** p < 0.001; ** p < 0.01; * p < 0.05.**	

We can include more than one model in a table too. Let's save a model with a control variable and a model with an interaction.

And then include our three models in a huxreg function that also includes several of the changes described above.

```
huxreg(model1, model2, model3,
       statistics = c("Number of Observations" = "nobs"),
       coefs = c("(Intercept)" = "(Intercept)",
                 "Number of Siblings" = "siblings",
                 "Gender = Other" = "genderOther",
                 "Gender = Woman" = "genderWoman",
                 "Siblings X Other" = "siblings:genderOther",
                 "Siblings X Woman" = "siblings:genderWoman")) |>
  set_caption("GPA on Siblings and Gender") |>
  theme_compact()
```

Table 3: GPA on Siblings and Gender

	(1)	**(2)**	**(3)**
(Intercept)	3.559 ***	3.532 ***	3.519 ***
	(0.017)	(0.021)	(0.027)
Number of Siblings	-0.032 ***	-0.031 ***	-0.023
	(0.009)	(0.009)	(0.014)
Gender = Other		-0.058	0.296 *
		(0.085)	(0.125)
Gender = Woman		0.046 *	0.056
		(0.020)	(0.035)
Siblings X Other			-0.269 ***
			(0.070)
Siblings X Woman			-0.006
			(0.019)
Number of Observations	985	985	985

*** p < 0.001; ** p < 0.01; * p < 0.05.