

Tutorial 5: Producing professional-looking tables

Pierre Chardaire

July 2, 2017

Perfection is finally attained not when there is no longer anything to add, but when there is no longer anything to take away.

Antoine de Saint-Exupéry (2000)

Now that you know how to produce basic tables, as presented in Oetiker et al. (2011) and in tutorial 4, you will learn how to create professional-looking tables in your reports. To help you, the `cmreport.cls` class file, which was used to produce this document, defines the `cmptable` and `cmptablenotes` environments and includes macro-commands to create sophisticated tables more easily. From now on, to try the examples provided in this tutorial, you must use the `cmreport.cls` class file with `tutorial` option. To be more precise your `tex` file should start with `\documentclass[tutorial]{cmreport}`.

1 The aesthetics of tables

Examine table 1. What is wrong with it? This table is unnecessarily large. It contains long similar statements in each of the row headings, as well as the repetition of the word *Awareness* in the column headings. Tables should contain succinct information¹. In academic writing you should avoid the sort of style you see in marketing brochures, such as shading every other row — which you can do with `LATEX 2ε` but will not learn from me!

¹Do not use a table with only two columns; either use a list or some form of chart or graph depending on the nature of the information presented.

Table 1: τ_b coefficients for awareness and usage

τ_b coefficients of. . .	Awareness of Textual Terminology	Awareness of N-Mark logo	Awareness of Contactless logo
Current use of credit/debit cards for contactless payment	.016	.067	.062
Current use of mobile phones for contactless payment	-.016	-.039	-.071
Future use of credit/debit cards for contactless payment	-.114	-.070	-.099
Future use of mobile phones for contactless payment	.091	.093	.068

Table 2: τ_b coefficients for awareness (columns) and **P**resent and **F**uture usage of contactless payment (rows)

	Textual terminology	N-Mark logo	Contactless logo
Credit/debit cards (P)	0.016	0.067	0.062
Mobile phones (P)	−0.016	−0.039	−0.071
Credit/debit cards (F)	−0.114	−0.070	−0.099
Mobile phones (F)	0.091	0.093	0.068

Besides making the table more compact you may choose to observe some guidelines (Fear, 2005; Dewey, 2013) when redesigning it:

- Do not use vertical rules² — In British typesetting a *rule* is a line, and I will use the word *rule* from now on.
- Do not use double rules.
- Do not use unnecessary horizontal rules (to separate similar rows).
- Place the units in the column headings. Avoid changing the units of measurement within a column.
- Always precede a decimal point by a digit: use 0.1 instead of .1, unless the number is a statistic that cannot be larger than one, for example a correlation $r = .55$, or a probability $p < .01$
- Make sure decimal numbers are aligned to the decimal point.
- Do not change the number of decimal places within a column.
- If necessary add notes to explain the table contents.
- For statistics tables use asterisks to indicate the value for which the null hypothesis is rejected with probability p specified in the notes, for example $*p < .05$ and $**p < .01$. Consistently use the same number of asterisks for a given alpha level throughout your paper.

Table 2 provides the same information as table 1 and uses far less space. As a consequence *it is easier to read*. The caption is used to give all the information needed to interpret the table even though this is not the right thing to do³: in section 2 you will see how to use table notes for adding explanations. I have also followed the above guidelines by avoiding unnecessary rules and using some heavier rules to separate the main sections. Compare the two tables. Which one looks more professional?

The code for producing table 2 is given in figure 1 and is of the form

```
\begin{cmptable}[ht]
{
% This is the table caption.
% It should not contain blank lines.
%
 $\tau_{\text{b}}$  coefficients for awareness (columns)
and \textbf{P}resent and \textbf{F}uture usage
of contactless payment (rows)\label{tab3}
}
%%%%%%%% CODE FOR THE ACTUAL TABLE FOLLOWS.
\end{cmptable}
```

²The APA style guide, amongst others, prohibits the use of vertical rules in tables (APA, 2004).

³According to The University of Chicago Press (2013) titles ‘should not contain explanatory material; this should be placed in a “note” at the bottom of the table.’

```

\begin{cmptable}[ht]
{
% This is the table caption.
% It should not contain blank lines.
%
$ \tau_{\text{b}}$ coefficients for awareness (columns)
and \textbf{P}resent and \textbf{F}uture usage
of contactless payment (rows)\label{tab3}
}
\begin{tabular}{L{0.30}C{0.25}C{0.25}C{0.20}} \toprule
& Textual\newline terminology & N-Mark\newline logo & Contactless\newline logo \\ \uhrrule

Credit/debit cards (\textbf{P})
& $\phantom{-}0.016$ & $\phantom{-}0.067$ & $\phantom{-}0.062$ \\

Mobile phones (\textbf{P})
& $-0.016$ & $-0.039$ & $-0.071$ \\

Credit/debit cards (\textbf{F})
& $-9.114$ & $-0.070$ & $-0.099$ \\

Mobile phones (\textbf{F})
& $\phantom{-}0.091$ & $\phantom{-}0.093$ & $\phantom{-}0.068$ \\ \bottomrule

\end{tabular}
\end{cmptable}

```

Figure 1: Code for producing table 2

This structure includes the passing of the caption as *compulsory* parameter of the table⁴. Also the `\label` command used to refer to the table must be included in the caption.

The code for the actual table starts with

```
\begin{tabular}{L{0.30}C{0.25}C{0.25}C{0.20}} \toprule
```

`L{0.30}C{0.25}C{0.25}C{0.20}` indicates that the table has 4 columns. The first column has specification `L(0.30)` which means it is written in paragraph mode (long lines will be wrapped) and is *Left*-justified; its width is equal to 30% of the length of a line of text. Similarly the second column has specification `C(0.25)`. Its text is *centred* in paragraph mode and its width is 25% of the length of a line of text. The width parameters add up to 1, so that the table width is exactly the width of a line of text in the document. If I had added vertical rule specification (for instance `|L{0.30}|C{0.25}|C{0.25}|C{0.20}|`) the space for the vertical rules would be automatically deducted from the width of the columns, so that the table would have the same width with or without vertical rules. Another specification for column formatting is `R` for a right aligned column.

In the table headings `\newline` is used to break lines inside cells. The code also contains a number of rule commands. `\toprule`, `\uhrule` and `\bottomrule` produce heavy horizontal rules. `\toprule` and `\bottomrule` have a wider space below and above respectively than a normal line of text, while `\uhrule` has above and below space of the same width.

The rest of the code is standard, except that I use the command `\phantom` which creates an invisible character. This is a useful trick that enables to align the content of cells in a centred column. Numbers have to be surrounded by `$` signs (i.e. typed in maths mode) to produce the correct minus sign. All this is a bit cumbersome. You will see a much better ways of producing aligned numbers in the next section.

2 Professional-looking tables

Table 3, which is adapted from an example in the instructions of The University of Chicago Press (2013) for building and formatting tables for publication, illustrates most of the components used in the writing of tables: table notes, headings associated with several columns, rules that span a subset of the columns, and alignment of numbers to decimal points.

2.1 Table notes

The code that produces the table is shown in figure 2. It comprises a `cmptablenotes` environment after a `tabular` environment. The former gives explanations associ-

⁴The `cmreport.cls` file also provides the `cmpfigure` environment that has the same requirement.

ated with the note marks included in the latter and in the caption⁵. Marks are produced with the command `\tnote`, except when they are in the middle of a tabular cell text where you must use `\mtnote`. The non-indexed notes that appear before the `\end{cmptablenotes}` are produced using the command `\note`.

2.2 Number alignment

`L{0.05}L{0.35}SSS` is the column format specification of the `tabular` environment. The table has five columns. The first two are used to make the row headings. They produce the sectioning effect seen in the table through the use of `\mcc` explained below. The last three have the format `S` that specifies columns containing numbers to be aligned. The `S` format also instructs $\text{\LaTeX 2}_{\epsilon}$ to perform some processing of the input numbers; for example the output 0.123×10^2 would be produced from the input `0.123e2`.

2.3 Sizing and grouping columns

The format `S` does not give a width for the corresponding columns. The width of the columns is given when specifying the format of the column heading cells by using the `\mcc` command. Remember the definition of the `\multicolumn` command given in tutorial 4. The macro `\mcc` is specific to the `cmpreport` class and used to simplify writing

`\mcc{hello}` is the same as:

`\multicolumn{1}{c}{hello}`

`\mcc[3]{hello}` is the same as:

`\multicolumn{3}{c}{hello}`

`\mcc(0.20){hello}` is the same as:

`\multicolumn{1}{C{0.20}}{hello}`

`\mcc[3](0.20){hello}` is the same as:

`\multicolumn{3}{C{0.20}}{hello}`

As you can see `\mcc` has two optional parameters that are distinguished by using different types of brackets: square brackets for the first one, round brackets for the second one. When optional parameters are omitted they take default values: 1 for the first parameter and `c` for the second one. The class also provides commands `\mcl` and `\mcr` for left-justified and right-justified cells respectively.

⁵When you create a list of tables each table is listed using a title obtained from its caption by removing all table note marks from the caption.

Table 3: Rates of DHS^a shelter use by selected characteristics

	Shelter system		
	Either ^b (%)	Family (%)	Single Adult (%)
<i>History of out-of-home care:</i>			
Yes	22.4	17.0	8.9
No	10.8	9.4	2.5
<i>Type of final exit from ACS^c:</i>			
Reunification	19.4	14.7	7.6
Independent living	25.6	18.8	10.7
Absconding from care	33.6	22.4	15.6
Preventive services	12.4	11.0	3.0
<i>Race and ethnicity:</i>			
African American (non-Hispanic)	24.3	18.8	8.9
Hispanic	14.1	11.0	5.2
Caucasian (non-Hispanic)	6.0	4.0	3.0
<i>Gender:</i>			
Female	25.1	23.7	4.8
Male	11.1	3.8	9.3
Total	18.7	14.5	6.9

^a Department of Homeless Services.

^b This category reflects the unduplicated sum of the other two columns.

^c The New-York City Administration for Children's Services.

All relationships are statistically significant for χ^2 test ($p < .001$).

Source: The University of Chicago Press (2013).

```

\begin{cmptable}[t]
{
Rates of DHS\tnote{a} shelter use by selected characteristics
\label{tab4}
}
\begin{tabular}{L{0.05}L{0.35}SSS} \toprule
&& \mcc[3]{Shelter system}\\ \cmidrule{3-5}

&& \mcc(0.20){Either\tnote{b}\newline (\%)} \\
& \mcc(0.20){Family\newline (\%)} \\
& \mcc(0.20){Single Adult\newline (\%)} \\ \uhrule

\mcl[2]{\emph{History of out-of-home care:}} \\
& Yes & 22.4 & 17.0 & 8.9 \\
& No & 10.8 & 9.4 & 2.5 \\ \addlinespace

\mcl[2]{\emph{Type of final exit from ACS\mtnote{c}:}} \\
& Reunification & 19.4 & 14.7 & 7.6 \\
& Independent living & 25.6 & 18.8 & 10.7 \\
& Absconding from care & 33.6 & 22.4 & 15.6 \\
& Preventive services & 12.4 & 11.0 & 3.0 \\ \addlinespace

\mcl[2]{\emph{Race and ethnicity:}} \\
& African American (non-Hispanic) & 24.3 & 18.8 & 8.9 \\
& Hispanic & 14.1 & 11.0 & 5.2 \\
& Caucasian (non-Hispanic) & 6.0 & 4.0 & 3.0 \\ \addlinespace

\mcl[2]{\emph{Gender:}} \\
& Female & 25.1 & 23.7 & 4.8 \\
& Male & 11.1 & 3.8 & 9.3 \\ \midrule
& Total & 18.7 & 14.5 & 6.9 \\ \bottomrule
\end{tabular}
%
\begin{cmptablenotes}
\item[a] Department of Homeless Services.
\item[b] This category reflects the unduplicated sum of the other
two columns.
\item[c] The New-York City Administration for Children's Services.
% some general notes
\note{All relationships are statistically significant
for  $\chi^2$  test ( $p<.001$ ).}
\note{Source: \cite{UCP}.}
\end{cmptablenotes}
%
\end{cmptable}

```

Figure 2: Code for producing table 3

Table 4: Effect of parameters on performance^a

Parameter	Data set			
	SINGLE		MULTIPLE	
	CPU (msec)	Effective (%)	CPU (msec)	Effective (%)
<i>n</i> (<i>k</i> = 10, <i>p</i> = 100)				
2	75.5	55.5	174.2	22.2
3	21.5	50.4	79.4	19.9
4	16.9	47.5	66.1	16.3
<i>k</i> (<i>n</i> = 2, <i>p</i> = 100)				
10	57.5	51.3	171.4	21.7
100	60.0	56.1	163.1	21.3
1000	111.3	55.9	228.8	21.4
<i>p</i> (<i>n</i> = 2, <i>k</i> = 10)				
100	3.3	5.5	6.1	1.2
1000	13.8	12.6	19.8	2.1
10000	84.5	56.0	126.4	6.3
100000 ^b	—	—	290.7	21.8

^a Results on processing time and effectiveness for various combinations of the three parameters for both data sets. Default parameters are shown in parentheses.

^b Value not meaningful for the data set SINGLE.

Adapted from Zobel (1997).

The top part of the banner is produced using `\mcc[3]{Shelter system}` which indicates that three columns are grouped as a single column with format `c`, and that the content of the resulting cell is `Shelter system`. The three sub-header columns are given sizes of 20% of the text width. The commands that start with `\mcl[2]{\emph{` group the first two cells of a row and include a left-justified text that forms an italic row heading. Note that the headings for the other rows start in column 2. Using two columns for the row-headings enables to produce the layered effect in the table stub column.

2.4 Rules that span a subset of columns

To create the rule under the header “Shelter system” you use `\cmidrule{3-5}` that specifies a light weight rule that spans column three to five. (Remember that what appears as the table stub column is in fact made of two columns.)

Now, an interesting example is shown in table 4 with the corresponding code in figure 3.

```

\begin{cmptable}[t]
{Effect of parameters on performance\tnote{a}\label{tab5}}
\begin{tabular}{S[table-number-alignment = center,
                 table-format=6,
                 table-space-text-post = \mtnote{b}]
               SSSS}
\toprule
\mcc(0.3){Parameter} & \mcc[4]{Data set} \\ \cmidrule{2-5}

& \mcc[2]\textsc{Single} & \mcc[2]\textsc{Multiple} \\ \cmidrule[r]{2-3} \cmidrule[l]{4-5}

& \mcc(0.175){CPU \newline (msec)} & \\
& \mcc(0.175){Effective\newline (\%)} & \\
& \mcc(0.175){CPU\newline (msec)} & \\
& \mcc(0.175){Effective\newline (\%)} \\ \uhrule

\mcc{$n$ (k=10, p=100)$} \\
2 & 75.5 & 55.5 & 174.2 & 22.2 \\
3 & 21.5 & 50.4 & 79.4 & 19.9 \\
4 & 16.9 & 47.5 & 66.1 & 16.3 \\ \midrule

\mcc{$k$ (n=2, p=100)$} \\
10 & 57.5 & 51.3 & 171.4 & 21.7 \\
100 & 60.0 & 56.1 & 163.1 & 21.3 \\
1000 & 111.3 & 55.9 & 228.8 & 21.4 \\ \midrule

\mcc{$p$ (n=2, k=10)$} \\
100 & 3.3 & 5.5 & 6.1 & 1.2 \\
1000 & 13.8 & 12.6 & 19.8 & 2.1 \\
10000 & 84.5 & 56.0 & 126.4 & 6.3 \\
100000 \mtnote{b} & \mcc{---} & \mcc{---} & 290.7 & 21.8 \\ \bottomrule
\end{tabular}
\begin{cmptablenotes}
\item[a] Results on processing time and effectiveness for
various combinations of the three parameters for both data sets.
Default parameters are shown in parentheses.
\item[b] Value not meaningful for the data set \textsc{Single}.
\note{Adapted from \cite{zobel}.}
\end{cmptablenotes}
\end{cmptable}

```

Figure 3: Code for producing table 4

This example is from Zobel (1997) who remarks that “even this rather complex table does not require vertical rules”. If you do not use vertical rules then you need a way of shortening rules that group headings. You will notice the gap between the two horizontal rules that group the sub-headings for SINGLE and MULTIPLE . It is obtained by the commands `\cmidrule[r]{2-3}` that trims the right side of the rule that spans columns 2 and 3 and the command `\cmidrule[l]{4-5}` that trims the left side of the rule that spans columns 4 and 5. The first parameter of `cmidrule` can also be `[lr]` to shorten both ends of the rule⁶. The specification for the first column format involves adding options to the *S* format. This is rather complicated and is necessary only because of the note mark associated with the last integer in the first column. If we were to use a simple *S* without additional options the mark would not be displayed next to the number. Moreover, the note mark has to be produced using `\mtnote` rather than `\tnote`. The *S* format options are presented in the `siunitx` package documentation. Note also that to introduce cell contents that are not numbers in a column whose format specifier is *S* (with possible extra parameters) you have to use the command `\mcc`.

3 More on formatting numbers in tables

Table 5: Homicides by race of the victim in 1993 in the USA^a.

Race	Population ^b	Homicides	Rate ^c
Black ^d	29986	12114	40.5*
White	199686	12153	6.1
Other ^e	19038	635	3.3**
Total	248710	24932	10.0

^a Data developped from the Statistical Abstract of the United States – 1993.

^b Population in 1000s.

^c Rate per 100000 in the population.

^d The rate for *Black* males was 69.2, for *White* males it was 9.0.

^e About one-half of the *Other* category was composed of Asian Americans and Native Americans.

* $p < .001$, two-tailed test. ** $p < 0.05$, two-tailed test.

Adapted from Dewey (2013).

Table 5 presents an example adapted from Dewey (2013). In the corresponding code, shown in figure 4, `table-format=6` means integer numbers with up to six digits, while `table-format=2.1` means decimal numbers with up to two digits before the decimal point and one after.

⁶The `\cmidrule` command has a slightly different syntax to that of the `booktabs` package and is specific to `cmreport.cls`.

```

\begin{cmptable}[t]
{Homicides by race of the victim in 1993 in the USA\tnote{a}.
\label{tab6}}
\begin{tabular}{L{0.10}
                S[table-format=6]
                S[table-format=5]
                S[table-format=2.1,
                table-space-text-post = {**}]
                }
\toprule
Race
& \mcc(0.20){Population\tnote{b}}
& \mcc(0.20){Homicides}
& \mcc(0.10){Rate\tnote{c}} \\ \uhrrule

Black\tnote{d} & 29986      & 12114 & 40.5* \\
White          & 199686    & 12153 & 6.1   \\
Other\tnote{e} & 19038     & 635   & 3.3** \\
Total          & 248710    & 24932 & 10.0  \\
\bottomrule
\end{tabular}
\begin{cmptablenotes}
\item[a] Data developped from the Statistical Abstract of the
United States -- 1993.
\item[b] Population in $1000$s.
\item[c] Rate per $100\,000$ in the population.
\item[d] The rate for \emph{Black} males was 69.2, for \emph{White}
males it was 9.0.
\item[e] About one-half of the \emph{Other} category was composed
of Asian Americans and Native Americans.
\note{$*p<.001$, two-tailed test. $*!p<0.05$, two-tailed test.}
\note{Example adapted from \citet{Dewey}.}
\end{cmptablenotes}
\end{cmptable}

```

Figure 4: Code for producing table 5

4 Conclusion

No more excuses for creating poorly presented tables!

References

- APA (2004). *Publication manual of the American psychological association*. American Psychological Association.
- Dewey, R. (2013). Apa research style crib sheet. <http://www.psywww.com/resource/APA%20Research%20Style%20Crib%20Sheet.htm#Tables>.
- Fear, S. (2005). Publication quality tables in L^AT_EX. <http://www.ctan.org/tex-archive/macros/latex/contrib/booktabs/booktabs.pdf>.
- Oetiker, T., Partl, H., Hyna, I., and Schlegl, E. (2011). The not so short introduction to L^AT_EX 2_ε. <http://tobi.oetiker.ch/lshort/lshort.pdf>.
- Saint-Exupéry, A. (2000). *Wind, Sand and Stars*. Penguin Books Ltd, United Kingdom.
- The University of Chicago Press (2013). Manuscript preparation - tables. <http://www.press.uchicago.edu/infoServices/prep-table.html>. Instructions for building and formatting tables according to University of Chicago Press guidelines.
- Zobel, J. (1997). *Writing for computer science*. Springer-Verlag New York Incorporated.